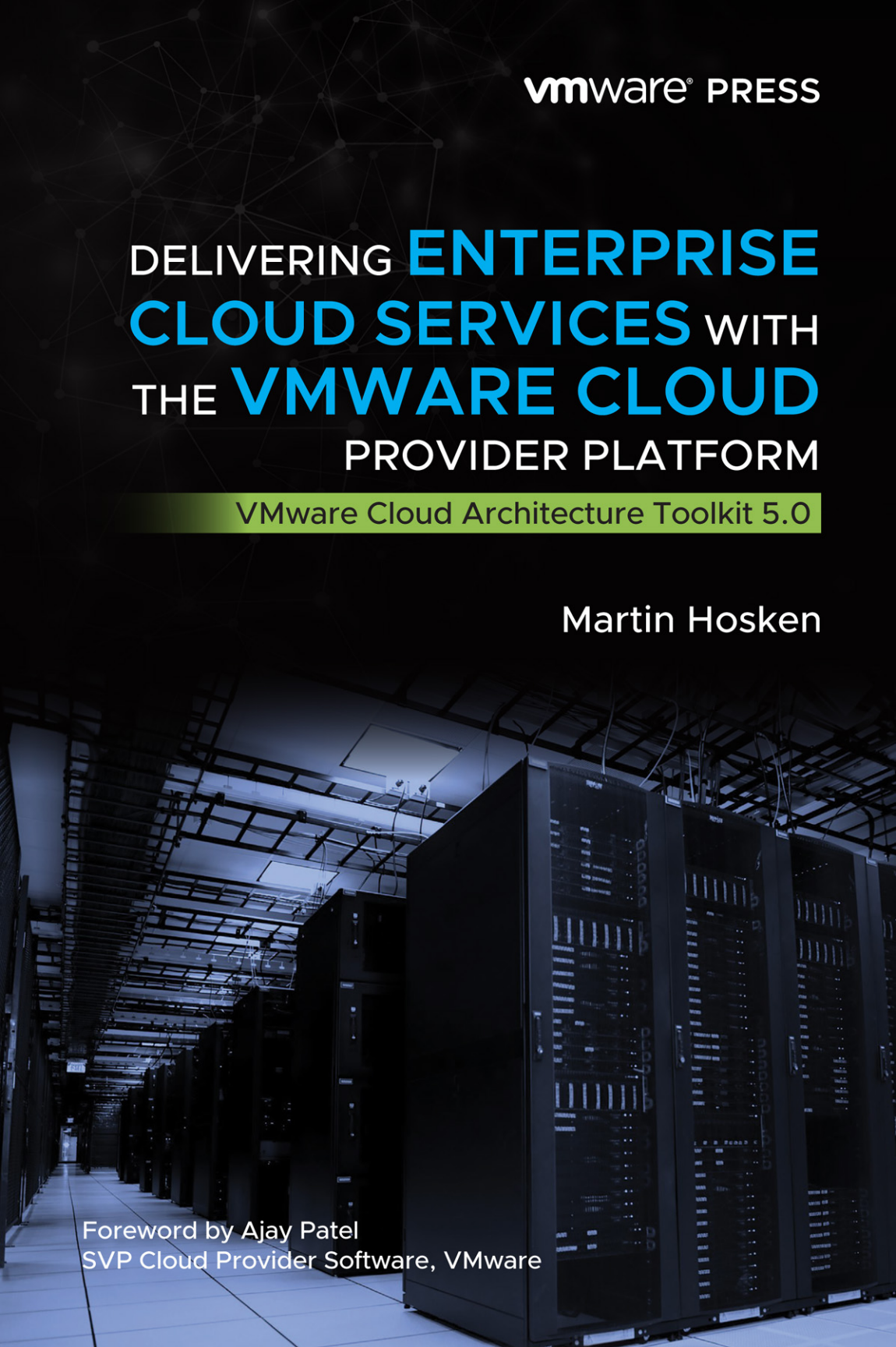




vmware® PRESS

DELIVERING **ENTERPRISE CLOUD SERVICES** WITH THE **VMWARE CLOUD** PROVIDER PLATFORM

VMware Cloud Architecture Toolkit 5.0

Martin Hosken

Foreword by Ajay Patel
SVP Cloud Provider Software, VMware

**DELIVERING ENTERPRISE
CLOUD SERVICES WITH
THE VMWARE CLOUD
PROVIDER PLATFORM**

VMware Cloud Architecture Toolkit 5.0

Martin Hosken

vmware® PRESS

Data Center Modernization with VMware Cloud on AWS™
Copyright © 2020 by Martin Hosken

Technical Reviewers:

Daniel Paluszek, Staff Solutions Engineer - Cloud Providers, VMware

John Dwyer, Director, Product Management, Cloud Provider Software, VMware

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means without written permission from the author.

Version 1.0 [March 2020]

ISBN: 978-1-5272-5969-0 print version

ISBN: 978-1-5272-5970-6 digital versions (ePub & eBook)

Warning & Disclaimer

Every effort has been made to make this book as complete and as accurate as possible, but no warranty or fitness is implied. The information provided is on an “as is” basis. The authors, VMware Press, VMware, and the publisher shall have neither liability nor responsibility to any person or entity with respect to any loss or damages arising from the information contained in this book.

The opinions expressed in this book belong to the author and are not necessarily those of VMware. VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 www.vmware.com. Copyright © 2019 VMware, Inc. All rights reserved. This product is protected by U.S. and international copyright and intellectual property laws. VMware products are covered by one or more patents listed at <http://www.vmware.com/go/patents>. VMware is a registered trademark or trademark of VMware, Inc. and its subsidiaries in the United States and/or other jurisdictions. All other marks and names mentioned herein may be trademarks of their respective companies.

Table of Contents

Figures..... ix

About the Author xvii

Foreword..... xix

Preface xxi

Chapter 1: VMware Cloud – Concept and Evolution 1

 The Next Evolution of Cloud 2

 Providing Businesses True Value for Money 3

Chapter 2: VMware Cloud Provider Platform6

 Provider Data Center – Core Capabilities and Integrated Components 8

 Cloud Provider Hub.....10

 VMware vRealize Operations and the Tenant App.....11

 VMware vCloud Usage Meter and VMware vCloud Usage Insight.....13

 Software-Defined Everything 13

 VMware Cloud Verified Program14

 VMware Validated Designs (VVD) for Cloud Providers.....15

 Building a Service Definition.....16

 Building a Service-Level Agreement.....17

 Designing an Enterprise Class Public/Hybrid.....17

 On-Premises-to-Provider Cloud Connectivity.....20

 Connectivity over the Public Internet 20

 Connectivity over Private Circuits.....23

 Evaluation of WAN Platforms24

 Hybrid Cloud Use Case Example31

 Cloud Provider Platform Core Use Cases - Moving Beyond IaaS..... 35

 VMware Cloud Provider Platform Networking and Security Overview35

 Workload Migration and Initial Onboarding36

 Data Protection and Recovery41

 Kubernetes Service42

Chapter 3: A Single Platform for a Global Cloud Business	45
Introducing VMware Cloud Director	45
Cloud Provider Use Cases	48
Hosted Private Cloud with Cloud Director.....	49
Cloud Director's Core Components	49
Platform Constructs.....	51
Personas & Role-based Access Control.....	74
Cloud Resource Design	77
Management Cluster	79
Resource Clusters	80
Scalability	81
Dedicated Hosted Cloud.....	90
Multi-Site/Multi-Geo Clouds.....	93
Multi-site Architecture with Discrete Management.....	94
Multi-site Architecture with Stretched Resource Clusters (vSAN Stretched Cluster(s)).....	94
Multi-site Architecture with Central Management Stack	96
Multi-site Architecture with Customer Associated Sites	97
Cloud Director Custom Branding	104
Custom Portal (Optional) Design Factors.....	105
 Chapter 4: Designing a VMware Cloud Verified	
Compute Platform.....	111
Cloud Verified Overview.....	111
Compute Platform Resources.....	113
Scalability and Designing Physical Resources.....	113
Infrastructure Protection & Availability.....	118
Comparing Blade Servers with Rackmount Servers	118
Compute Host Sizing – Scale-Up Compared with Scale-Out	120
Determining Host CPU and Memory Requirements	121
Determining an Appropriate Virtual CPU-to-Physical CPU Ratio	122
Performance Tuning with NUMA	124
Capacity Management for Cloud Service Providers.....	126
Automating Host Deployment.....	127
Planning a Scalable vSphere Host Clusters Architecture	129
Cloud Platform Edge Cluster.....	131
Planning for Server Failure with Resilience	133
Resource Balancing and Transparent Maintenance	136
Enhanced vMotion Compatibility	139
vCenter High Availability	140
Cloud Platform (vSphere) Patch Management	142
 Chapter 5: Designing a VMware Cloud Verified	
Storage Platform	143
VMware vSAN Overview.....	144
VMware vSAN Architecture Overview	147

vSAN Design Overview.....	160
Disk Controllers.....	164
Summary of Key vSAN Design Factors	177
vSAN Stretched Cluster Deployments	177
vSAN Storage and Cloud Director.....	180
vSAN Storage Resource Isolation.....	183
Cloud Director Storage Design with vSAN	186
Cloud Director NFS Storage Requirement	194
Designing a Highly Scalable HCI Platform	196
Consuming External Array Storage.....	196
Virtual Volumes (vVols)	201
 Chapter 6: Designing a VMware Cloud Verified	
Network Platform.....	211
VMware NSX for vSphere Verses	
NSX-Data Center	212
NSX Data Center Architecture Overview.....	214
Cloud Director and NSX Integration	216
Cloud Director Networking with NSX Data Center	217
Physical Network Architecture (NSX Underlay)	218
vSphere Network Constructs.....	220
VMware Cloud Director Tenant Networking	223
Network Use Cases	230
Distributed Firewall (Micro segmentation)	236
Bastion Connectivity Service.....	239
 Chapter 7: Platform Automation and Standardization.....	240
Cloud Automation	241
vRealize Orchestrator	242
Service Library	246
RabbitMQ	248
Blocking Tasks Overview	250
Terraform	254
VMware Cloud Foundation (VCF).....	256
Designing a VMware Cloud Foundation with Cloud Director Solution	259
Cloud Provider Pod.....	260
VMware Cloud on AWS (MSP).....	264
VMware Cloud Director Service (VCDS).....	266
VMware MaaS Solution Partners.....	267
 Chapter 8: Building and Delivering Services with Extensibility	271
VMware Cloud Director Availability	273
Architecture Overview.....	274
Disaster Recovery to Cloud (DR2C) Overview	276
Cloud to Cloud (C2C) Overview	280
Native Integration with Cloud Director.....	280

Cloud Director Availability Migration versus HCX.....	282
Planning for Application Migration with VMware HCX.....	282
Introduction to HCX.....	285
HCX Component Overview.....	285
Migration Methodology	287
HCX for Hybrid Network Extension.....	289
HCX for Disaster Recovery	290
VMware Cloud Director HCX Design Considerations.....	290
Operating System Assisted Migration (OSAM).....	292
vRealize Operations Management Pack for Cloud Director and the Tenant App	292
Tenant Metering	294
Central Point of Management (CPOM)	294
SDDC as-a-Service with CPOM.....	297
Cloud Provider Hub.....	301
Container Service Extension (CSE).....	304
Overview of Common Storage Services	307
Object Storage.....	308
Object Storage on VMware Cloud Provider Platform	315
Machine Learning-as-a-Service (MLaaS).....	316
Multi-Cloud Management-as-a-Service	320
VMware Cloud Marketplace and Bitnami for Cloud Providers.....	322
App Launchpad	323
Data Protection & ISV Service Ecosystem.....	325
Chapter 9: Operations and Security	328
Developing a Consistent Operational and Governance Model.....	330
Building a Logging Service	334
Security Monitoring	335
Logging Service.....	335
Data Archiving	335
Extending vRealize Log Insight Services	336
Management Environment Logging.....	337
Cloud Provider Platform Management Component Logging	339
Sample Syslog Design Scenario.....	342
Log Intelligence	342
Secure Multitenancy	344
VMware Cloud Director Security	347
Designing Host Security for Multi-Tenanted Clouds.....	347
Secure Communication.....	348
Certificate Configuration and Usage.....	349
Local Account Management	349
Host Active Directory Configuration Status.....	350
Transparent Page Sharing Security	350
SNMP Hardware Monitoring.....	351
Host Lockdown Mode	351

ESXi Firewall	351
Role-based Access Control	352
Change and Configuration Management	353
Patch and Update Practices	353
Final Word	355
Cloud Provider Platform Bill of Materials	357
Index	361

This page intentionally left blank

Figures

Figure 2.1	Cloud Provider Platform Overview	7
Figure 2.2	Cloud Provider Platform Component Architecture	8
Figure 2.3	VMware Cloud Director Abstraction Layers	9
Figure 2.4	VMware Offerings for VMware Partners	10
Figure 2.5	Cloud Provider Hub	11
Figure 2.6	vRealize Operations Manager Tenant App	12
Figure 2.7	Cloud Usage Meter Data Collection Workflow	14
Figure 2.8	The Cloud Provider Platform Stack	16
Figure 2.9	VMware Native Hybrid Cloud.....	18
Figure 2.10	Consistent VMware Cloud Platform	19
Figure 2.11	Consistent Hybrid Cloud Experience	19
Figure 2.12	VMware NSX L2VPN Using a VXLAN/Geneve-Based Solution	21
Figure 2.13	VMware NSX L2VPN Using a VXLAN-to-VXLAN-Based Solution	22
Figure 2.14	Dark Fibre Connectivity	25
Figure 2.15	Wave Division Multiplexing (DWDM and CWDM).....	26
Figure 2.16	Multiprotocol Label Switching (MPLS)	28
Figure 2.17	A VPLS WAN Deployment, Full Mesh of Layer 2	29
Figure 2.18	Sample Architecture Database Functions	32
Figure 2.19	SQL Server 2019 AlwaysOn Availability Group for HA and DR.....	34
Figure 2.20	Cloud Director Extender Architecture and Components.....	38
Figure 2.21	VMware HCX Overview	39
Figure 2.22	Scheduling a Failover for Bulk Migration.....	40
Figure 2.23	Dell EMC IDPA Cloud Director Integrated Solution	42

Figure 2.24	CSE - Flexible Multi-tenanted Kubernetes using Cloud Director Constructs—Native (without PKS).....	43
Figure 2.25	Container Service Extension Architecture	44
Figure 3.1	Key Pillars of the Cloud Provider Platform	47
Figure 3.2	Extension Plug-in Integration in Cloud Director	48
Figure 3.3	VMware Cloud Director Architecture	50
Figure 3.4	Cloud Director Constructs.....	50
Figure 3.5	Organization Construct Overview	52
Figure 3.6	Organization to Org. VDC Mappings.....	53
Figure 3.7	Support Organization Catalog Publishing.....	54
Figure 3.8	Self-contained LAMP Stack	55
Figure 3.9	Linked Clone and Full Clone	56
Figure 3.10	Cloud Director Cell Architecture	57
Figure 3.11	Cloud Director Appliance Database HA Cluster	58
Figure 3.12	Cloud Director Management Portal.....	59
Figure 3.13	Virtual Machine Metric Database Design	61
Figure 3.14	Default Organization Login Page	62
Figure 3.15	Securing Cloud Director	63
Figure 3.16	Remote Console	65
Figure 3.17	API Component Interactions	67
Figure 3.18	Flex Allocation Model.....	68
Figure 3.19	Island Clusters	70
Figure 3.20	Island Cluster Provisioning Workflow.....	71
Figure 3.21	Compute Policy Examples	72
Figure 3.22	Cloud Director Key Personas	75
Figure 3.23	Public Cloud Provider Operated Model for Managed Service Providers.....	78
Figure 3.24	Cloud Provider Platform Infrastructure.....	79
Figure 3.25	Cloud Provider Platform Resource Logical Infrastructure	81
Figure 3.26	Pod Architecture Resource Cluster Design.....	83
Figure 3.27	Primary Pod Architecture.....	84
Figure 3.28	Distributed Architecture	85

Figure 3.29	High-Level Pod Management Architecture	86
Figure 3.30	Flexible Pod Architecture.....	87
Figure 3.31	Building-Block Pod Architecture Evolution.....	88
Figure 3.32	Hyperscale Regional Architecture	89
Figure 3.33	Dedicated Infrastructure with Reservation Pool Allocation	91
Figure 3.34	Dedicated Resources for Different Lines of Business	92
Figure 3.35	Multi-Site, Federated Cloud Director Architecture.....	93
Figure 3.36	Cloud Provider Platform Multi-Site Architectural Options	95
Figure 3.37	Cloud Provider Platform Discrete Administration.....	95
Figure 3.38	Multi Data Center Stretched Resource Cluster Architecture.....	96
Figure 3.39	Multi-site Distributed Resource Groups.....	97
Figure 3.40	Multi-site with Customer Associated Sites	97
Figure 3.41	Organization Association Sequence	98
Figure 3.42	Association Mesh Between Multiple Member Sites.....	100
Figure 3.43	Cloud Director Federation	100
Figure 3.44	Multisite Cloud Provider Platform Load Balancing	102
Figure 3.45	Cross-Site NSX for vSphere and vCenter Server.....	103
Figure 3.46	Customizing User Interface Code Sample & User Interface Custom Links.....	104
Figure 3.47	Custom Login Portal User Experience and Journey	106
Figure 3.48	Access to Cloud Director API.....	107
Figure 3.49	Custom Portal Architecture Example.....	109
Figure 4.1	Cloud Provider Platform – Core and Non-core Components	112
Figure 4.2	Sample HCI Cluster Architecture	115
Figure 4.3	Sample Management Cluster Architecture.....	116
Figure 4.4	Planning for a Highly Scalable Deployment	117
Figure 4.5	Virtual CPU-to-Physical CPU Ratio	124
Figure 4.6	NUMA Architecture.....	125
Figure 4.7	Capacity Management Planning.....	126

Figure 4.8	Cloud Platform Edge Cluster Design and Traffic Flow	132
Figure 4.9	JSON Network Profile Definitions	134
Figure 4.10	vSphere Distributed Resource Scheduler (DRS).....	136
Figure 4.11	vSphere DRS Configuration Options.....	137
Figure 4.12	Migration Threshold Slider – Recommendation for the Cloud Provider Platform.....	139
Figure 4.13	Enhanced vMotion Compatibility	140
Figure 4.14	vCenter High Availability	141
Figure 5.1	Flexible Mixed Storage Layer.....	144
Figure 5.2	vSAN Object-Based Storage System.....	145
Figure 5.3	vSAN All-Flash Architecture	148
Figure 5.4	Enabling vSAN.....	149
Figure 5.5	Enabling Deduplication and Compression.....	149
Figure 5.6	Enabling Storage Policy-Based Management Fault Tolerance Methods.....	151
Figure 5.7	Rack Awareness (Fault Domains) Example	153
Figure 5.8	vSAN QoS Enforced by SPBM.....	155
Figure 5.9	vSAN Stretched Cluster Architecture	157
Figure 5.10	Three-Site Disaster Recovery Architecture.....	158
Figure 5.11	vSAN Encryption Framework.....	159
Figure 5.12	Network Diagram	167
Figure 5.13	Network I/O Control	169
Figure 5.14	Stretched Cluster Architecture	179
Figure 5.15	Cloud Director Provider VDC over a vSAN Stretched Cluster.....	180
Figure 5.16	Cloud Director Platform.....	181
Figure 5.17	vSAN Datastore Abstraction verses Storage Array Volumes	184
Figure 5.18	Cloud Director Storage Isolation	185
Figure 5.19	vSAN Storage Resource Isolation.....	186
Figure 5.20	Model 1: Multiple Provider VDC Storage Tiers, Supporting Multiple Org. VDCs	186
Figure 5.21	Manual Storage Tiering within a Single Organization	187

Figure 5.22	VMware Technology Solution Stack	188
Figure 5.23	Multi-Tiered Application Storage Mapping.....	189
Figure 5.24	vCloud Elastic Provider VDC vSAN Logical Architecture	190
Figure 5.25	Provider VDC with Mixed Storage Availability Policy Capabilities.....	192
Figure 5.26	Balanced Provider VDC Resource Offerings.....	194
Figure 5.27	Cloud Provider Platform—Mixed Storage Model.....	197
Figure 5.28	Logical Boundaries Separating Storage Types.....	199
Figure 5.29	Manually Tiered Storage Capabilities and Attributes Example	200
Figure 5.30	Virtual Volumes Storage Policy-based Management	203
Figure 5.31	Traditional storage LUNs Verses Virtual Volumes.....	204
Figure 5.32	Protocol Endpoint (PE) Architecture.....	205
Figure 5.33	VASA Provider Architecture.....	206
Figure 5.34	Storage Policy Based Management and VASA Provider Solution.....	208
Figure 5.35	Virtual Volume	209
Figure 6.1	NSX and Cloud Director Logical Architecture	213
Figure 6.2	NSX Data Center High-Level Architecture.....	215
Figure 6.3	NSX Data Center and Cloud Director API Integration.....	217
Figure 6.4	Cloud Director Network Layers.....	218
Figure 6.5	Host to Top of Rack (ToR) Connectivity and Layer 3 Transport.....	219
Figure 6.6	NSX Data Center N-VDS Architecture.....	221
Figure 6.7	N-VDS Teaming Policies.....	221
Figure 6.8	VMware Cloud on AWS Logical Routing.....	222
Figure 6.9	Cloud Director NSX Integration Comparison	224
Figure 6.10	Creating Tier-0 devices in NSX Manager.....	225
Figure 6.11	Cloud Director External Network Configuration	226
Figure 6.12	Customer Internet and WAN Access	227
Figure 6.13	Organization VDC Network Configuration	229
Figure 6.14	Directly Connected Org. VDC Network.....	230
Figure 6.15	Routed Org. VDC Network	231

Figure 6.16	Internal (Isolated) Network	231
Figure 6.17	Organization VDC Static Route Communication.....	233
Figure 6.18	NSX Edge Load Balancer.....	234
Figure 6.19	VMware NSX Advanced Load Balancer (Managed Service)	235
Figure 6.20	NSX Advanced Load Balancer and Cloud Director Example.....	235
Figure 6.21	Distributed Firewall Architecture	237
Figure 6.22	Cloud Provider Bastian Service	238
Figure 7.1	vRealize Orchestrator Initial Customer Provisioning Workflow	244
Figure 7.2	Request for Additional External Networks	245
Figure 7.3	Service Library	246
Figure 7.4	Access Control to the Service Workflow	247
Figure 7.5	Change Org. VDC Storage Schema	248
Figure 7.6	RabbitMQ Deployment Architecture.....	249
Figure 7.7	Distributed/Multi-Site RabbitMQ Architecture.....	249
Figure 7.8	RMQ Message Architecture.....	250
Figure 7.9	Blocking Task Orchestration Logic.....	251
Figure 7.10	Separating Business Logic and Technical Logic for Improved Scalability	253
Figure 7.11	Service Provider Onboarding a Tenant	254
Figure 7.12	Terraform Application Building.....	255
Figure 7.13	Software-Defined Cloud Provider Platform.....	257
Figure 7.14	VMware Cloud Foundation Workload Domains.....	258
Figure 7.15	VI Workload Domain to Resource Group Mapping.....	260
Figure 7.16	Cloud Director and Cloud Foundation Architecture	261
Figure 7.17	Cloud Provider Pod Workflow	262
Figure 7.18	Anatomy of VMware Cloud on AWS	263
Figure 7.19	A Hybrid Cloud that Is Based on a Common Platform	264
Figure 7.20	VMware Cloud on AWS Service Differentiators.....	265
Figure 7.21	Multi-tenancy for VMware Cloud on AWS.....	266
Figure 7.22	Multi Cloud Service Vision.....	267

Figure 7.23	Azure VMware Solutions	269
Figure 7.24	Azure VMware Solutions—Intelligent Application Delivery	270
Figure 8.1	Rubrik User Interface Extension	272
Figure 8.2	Extensible Service Framework.....	273
Figure 8.3	Cloud Director Availability Architecture.....	275
Figure 8.4	Cloud Provider Platform Component Overview	276
Figure 8.5	vSphere On-Premises Architecture	277
Figure 8.6	HTML5 vSphere Plugin Integration	278
Figure 8.7	On-Premises vSphere Architecture	279
Figure 8.8	Cloud to Cloud Architecture	281
Figure 8.9	Cloud Director Availability and Cloud Director User Experience	282
Figure 8.10	Bucketizing Workloads	284
Figure 8.11	VMware HCX Overview	286
Figure 8.12	Scheduling a Failover for Bulk Migration.....	288
Figure 8.13	VMware Cloud Director HCX Design Example	291
Figure 8.14	vRealize Operations Management Pack for Cloud Director .	293
Figure 8.15	vRealize Operations Manager Tenant App	293
Figure 8.16	Central Point of Management Architecture.....	295
Figure 8.17	Central Point of Management User View.....	296
Figure 8.18	Multi-Endpoint Resource Consumption	296
Figure 8.19	SDDC Private Cloud Service Offering—Application Stack.....	298
Figure 8.20	Provider Management & Operations.....	299
Figure 8.21	Service Packages (P&P Examples)	300
Figure 8.22	Cloud Provider Hub.....	301
Figure 8.23	VMware Cloud on AWS MSP Delivery Model	302
Figure 8.24	MSP Organization Structure	303
Figure 8.25	CSE Component Architecture and Interaction Points	305
Figure 8.26	CSE Architecture	306
Figure 8.27	Common Storage Models/Features and Costs	307
Figure 8.28	Cloud Storage Landscape.....	309
Figure 8.29	The Object Structure	310

Figure 8.30	Key Object Storage Characteristics	311
Figure 8.31	vCloud Object Storage Service	315
Figure 8.32	Distributed Machine Learning on TensorFlow for Machine Learning	317
Figure 8.33	Distributed Machine Learning on TensorFlow Component Architecture.....	318
Figure 8.34	GPU Computing	319
Figure 8.35	Multi-Cloud Application Blueprinting through Cloud Director	321
Figure 8.36	Operating vRealize Automation Cloud from Cloud Director	322
Figure 8.37	VMware Cloud Marketplace	322
Figure 8.38	App Launchpad Catalog Management—Provider View.....	323
Figure 8.39	App Launchpad High-Level Architecture.....	324
Figure 8.40	App Launchpad—Tenant View	324
Figure 8.41	Dell EMC IDPA Extension for Data Protection	326
Figure 8.42	Multi-Tenant/Multi-Tier Data Protection Service.....	327
Figure 9.1	Cloud Operational Architecture Example	329
Figure 9.2	High-Level Monitoring and Altering Architecture	333
Figure 9.3	Cloud Provider Platform Federated Deployment Architecture	334
Figure 9.4	Sample Cloud Provider Management Design	338
Figure 9.5	Multitenant Scenario 1	339
Figure 9.6	Multitenant Scenario 2	340
Figure 9.7	Edge Gateway Provider and Tenant Syslogs	341
Figure 9.8	Design Scenario	343
Figure 9.9	High-Level Secure Multitenancy Design.....	345
Figure 9.10	Network Segmentation.....	348

About the Author

Martin Hosken



Martin works at the evolution of VMware Cloud Services, the intersection where IT Architecture, Solution Architecture and Software Development meet. As a Chief Technologist at VMware, Martin helps VMware partners prepare to make the most of the opportunities presented by emerging technology and software development practices. His primary focus is on architecting solutions that create cutting edge solutions in today's complex cloud market. Martin is part of the VMware Office of the CTO, Global Field team, a double VMware Certified Design Expert

(VCDX Number 117) in Data Center Virtualization and Cloud Management and Automation, an established vExpert, and is the author of multiple papers, blogs and articles based on VMware and other technologies. He is also the sole author of the Sybex publication VMware Software-Defined Storage: A Guide to the Policy Driven, Software-Defined Storage Era and the new publication: Data Center Modernization with VMware Cloud on AWS. Follow Martin on Twitter: @hoskenm

This page intentionally left blank

Foreword

By Ajay Patel

Senior Vice President, Cloud Provider Software, VMware

It's no secret that public cloud infrastructure-as-a-service market is dominated by the top five cloud providers (Amazon, Microsoft, Alibaba, Google, and IBM), accounting for 76.8% of the \$32.4 billion IaaS cloud market share in 2108 according to Gartner. With the IaaS market predicted to double over the next 5 years, that represents over \$15 billion in infrastructure-as-a-service cloud spend up for grabs by regional and specialist cloud providers. Private clouds are growing fast, too, as companies not only move workloads to the top public clouds but also look to regional and specialist cloud providers to offer hosted private clouds as compelling alternatives to building their own. Demand for managed services across private and public clouds continues to accelerate, creating a significant market opportunity for cloud providers to the tune of \$30 to \$45 billion over the next 5 years.

But what is the definition of a cloud provider? As a leader of the VMware Cloud Provider business that has partners managing over 10,000 data centers, I have a simple definition: Anyone building public or hosted private clouds and/or offering managed services using their own cloud or one of the above public clouds is a cloud provider.

Just like the cloud has forced traditional software companies to transform to an “as-a-service” delivery model, the partner community too must also transform to help companies use the cloud to migrate, modernize, and transform to digital businesses. The lines between traditional partner segments of value-added resellers, system integrators and outsourcers, and hosting providers are starting to blur. Successful cloud providers are strategic partners who bring technical and functional expertise to customers, are adept at helping organizations make the cultural changes required, are skilled in helping accelerate the shift to the cloud, and can build and deliver innovative solutions all while establishing a foundation for operating in a highly secure, reliable, and cost-optimized manner across clouds.

A successful cloud provider is a multi-cloud specialist who smartly manages and uses the right cloud for the right workload. To successfully operate and deliver value in a multi-cloud world, cloud providers need an extensible and scalable services delivery platform. This book by Martin is about the VMware Cloud Provider Platform, the required foundation for building an innovative, highly differentiated, and customer-focused cloud business built using VMware vCloud Director, VMware vSphere, VMware NSX, and other enabling VMware solutions. Whether you are a “cloud builder” looking to

deliver a modern enterprise-ready public and hosted private cloud to the large VMware install base of customers, or a partner looking to shift your business to “managed services” using one of the many VMware Solutions available on the major public clouds, this book is for you. It is a guide to modernizing your own infrastructure to deliver a cloud that can compete with public clouds with self-service, multi-tenant infrastructure-as-a-service offering wrapped-with-your-own differentiated services and value. The VMware Cloud Provider Platform enables innovation by leveraging third-party ISV solutions, open-source certified content from VMware Bitnami, and co-opting public clouds as part of your own business transformation. It opens the door to leveraging the scale, economics, and global reach of hyperscaler cloud infrastructure and the expansive portfolio of cloud services.

Martin provides valuable insight and learnings from working with successful cloud providers across the globe. The book is not just a technical guide to building a modern cloud using VMware technologies. It also shares proven business models and leading services driving significant revenue for partners globally. You will walk away with a clear understanding of VMware’s cloud strategy for cloud providers and the benefits of adopting a single platform for your global cloud business. You will appreciate the significant investments being made to help accelerate your cloud business with the Partner Acceleration Program (PAP), which covers all aspects of building a profitable and competitive cloud business. Filled with examples, this book is a must read for both product offering owners and technologists responsible for a company’s cloud strategy and offerings. For those who don’t know Martin, he is one of our “go-to” cloud architects responsible for helping design some of the largest and most complex hybrid clouds for our partners. His experience and learnings shared in this book are recommended for partners looking to benefit from being a VMware Cloud Verified partner.

At VMworld 2019 we announced Project Path, which will build and expand on the foundational capabilities of the VMware Cloud Provider Platform. Project Path will deliver a true cloud-agnostic platform for partners to accelerate delivery of new services and effectively manage the cloud resources used across all their customers. While we have accomplished a lot in the past few years as a VMware Cloud Provider partner community, the sheer speed of change and complexity of operating in a multi-cloud environment requires us to rapidly transform our own business.

I am confident our technology and go-to-market investment in cloud providers, complemented with the guidance provided in this book, will help you capture the tremendous market opportunity driven by the shift to the cloud. We as a community represent the “neutral cloud,” harnessing the power of the ISV ecosystem while providing customers a single destination for operating across on-premise, partner, and major public clouds.

I want to thank our VMware partners who have been with us since the early innings and look forward to inviting many more on this exciting journey.

Thank you.
Ajay Patel

Preface

This book aims to reside at the intersection where VMware products, services, IT architecture, and software development all meet. For cloud service providers, the technology landscape is changing fast, driven by the increasingly digital nature of customers' businesses. In order to keep up with the fast-changing world and respond to the dynamic needs of customers, VMware has developed the VMware Cloud Provider Platform to help cloud providers modernize enterprise public cloud offerings and keep up with technology evolution and the growing popularity of new service delivery requirements for DevOps, multi-cloud, and advanced automation.

What do these trends mean for the future of the VMware Cloud Provider delivery model? Well, in order to keep up, cloud providers need a more agile and flexible approach to the delivery of services, if they want to keep pace. As we look at what is ahead for cloud providers, this new reality will drive a number of important changes, including:

1. A shift toward a more DevOps centric delivery model
2. Software-defined everything with advanced automation
3. Cross-cloud and multi-cloud expansion

Businesses today face an era of unprecedented choice of cloud services. VMware's broad range of private and public cloud solutions have traditionally given cloud service providers choice and flexibility in how to build, run, and manage public and hybrid cloud platforms. For customers who have built a best-in-class Software-Defined Data Center (SDDC)-based private cloud, extending to a VMware cloud provider partner for a hybrid cloud solution that integrates new services types, such as Kubernetes, multi-cloud, object storage, or on-demand disaster recovery, has given them the freedom to implement cloud solutions that support their cloud strategies for their businesses, with freedom and control. Now, with the VMware Cloud Provider Platform, cloud providers can take advantage of the same common model approach to provide an easy way to design, build, and manage hybrid clouds, with a common architectural approach and operating model.

With more than 4,000 partners located in more than 120 countries and managing more than 10 million workloads under management, VMware cloud providers deliver reliable services with the features and capabilities

enterprises need to stay ahead of the curve. These partners provide robust public and private VMware-validated cloud services tailored to customers' business needs while ensuring hybrid compatibility with existing VMware customer environments.

The VMware Cloud Provider Platform delivers a way for cloud service providers to easily build a multi-tenant public cloud software-defined data center-based architecture, a VMware hybrid cloud-enabled solution, with the capabilities to meet the unique needs of their customers.



Chapter 1

VMware Cloud – Concept and Evolution

In an industry that has become highly commoditized, the principal aim of a VMware public cloud service, delivered through cloud service provider partners, is to raise the bar and help the industry define a new class of enterprise public cloud computing, which brings with it high-performance, flexible service delivery, and security as standard. Effectively, defining a new enterprise-class cloud computing segment.

A VMware Public Cloud can also exist as a hybrid cloud solution, enabling organizations to extend their private cloud to the public cloud with flexibility, scalability, security, and operational efficiency. The Cloud Provider Platform model, which delivers a common platform built around VMware Cloud Director, alongside a flexible framework of solutions that integrate into it, gives cloud providers a comprehensive enterprise-grade public cloud solution. The Cloud Provider Platform delivers what many enterprise organizations have been waiting for—a common management and security model built across a platform that provides simple application portability. With this framework, VMware’s leading global cloud service providers are delivering cloud-compatible, connected, and integrated hybrid clouds.

The challenge many enterprises have to navigate is that when it comes to moving applications to the cloud, they are faced with more and more options. Public cloud, managed private cloud, and hybrid cloud solutions all have their strengths and use cases. However, when evaluating their cloud requirements, both for now and for the future, businesses must carefully consider, *what are the benefits to my organization?*

While this list will vary business to business, most businesses are looking for predictable performance, security, and required compliance standards, and high levels of availability. In addition, in many cases, they need a platform that helps

them optimize their existing investments, one that provides flexibility, extensibility, global access, and delivers a demonstrable Return on Investment (ROI).

The VMware Cloud Provider Platform is a defined ecosystem for cloud service providers that is designed from the ground up to deliver modern-day enterprises with cloud-based solutions that organizations need to grow their businesses. Built on the foundation of existing VMware technology, the Cloud Provider Platform allows cloud providers to deliver a seamless and common cloud experience.

THE NEXT EVOLUTION OF CLOUD

As any CTO knows, the decision about whether to move applications and data to the cloud marks a significant shift in the approach that a modern enterprise IT department must adapt to. While many adopters have already made that shift in approach, increasingly, more organizations are now looking for VMware cloud service providers that can fulfill their needs for a more flexible and capable approach to cloud-based services. Migration to the cloud continues to prove to be a huge risk for many organizations and demonstrating the return on investment of refactoring applications is hard to quantify to business leaders. Therefore, developing a comprehensive strategy for moving to the public cloud is becoming more important than ever before. For many organizations, a one-size-fits-all, commodity approach to cloud simply will not suffice. A forward thinking, modern enterprise needs to look beyond the hundreds of services a typical hyperscale cloud offers, which few customers ever consume, and adopt a solution that can fit their particular business and service needs, and not to try and fit a square peg into a round hole.

VMware vSphere is the world's leading server virtualization platform, used by more than 500,000 customers. Therefore, the vast majority of global organizations are already employing VMware technology in their data center, giving them a head start on their cloud journey.

The VMware Cloud Provider Platform-based hybrid cloud environment helps facilitate this by supporting the widest possible variety of powerful use cases, including regional capacity expansion, disaster recovery, application migration, data center consolidation, new application development, and burst capacity—all with complete operational consistency, seamless workload portability, and support for more than 5,000 applications and more than 90 operating systems. This makes the Cloud Provider Platform, built on VMware Cloud Director and vSphere technology, the most trusted and versatile platform for any application.

Next, we can extend that story with a VMware hybrid cloud solution, giving businesses a VMware hybrid cloud option tailored to the unique needs of that organization, which leverages enterprise-grade hybrid cloud capabilities to protect, extend, or replace on-premises data center infrastructure and run cloud-based resources on a trusted VMware platform, with hybrid

compatibility across public and private cloud environments. This model gives organizations a well understood, safe, and easy way to integrate public clouds with their on-premises infrastructure, while gaining the flexible and elastic capacity of a public cloud provider.

PROVIDING BUSINESSES TRUE VALUE FOR MONEY

VMware cloud service providers, through the Cloud Provider Platform model, can offer customers an unprecedented level of choice, extensibility, flexibility, and security, helping businesses of all sizes seamlessly migrate workloads into the cloud, whether it is migrating existing on-site virtual machines, or starting new application development, deployment, and operational oversight directly within the Cloud Provider Platform. In addition, through hybrid cloud models, with a VMware Cloud Provider Partner, businesses can move their most critical business workloads back and forth between the service provider cloud and the customers' local data center, all over the secure and capable foundation of the VMware Cloud Provider Platform architecture, which provides complete multilevel security and a multi-tenant architecture, reducing complexity and ensuring adherence to corporate security and compliance policies.

Another one of the most compelling aspects of the VMware cloud service provider ecosystem is the widespread availability of partners, and as a result, the choice of “Cloud Verified” cloud services across the globe. The number of VMware cloud provider partners is in excess of 4000 and the number of “Cloud Verified” partners is continuously expanding. These cloud partners are located in more than 120 countries and collectively manage more than 10 million¹ workloads. This cloud delivery model is unparalleled globally and gives cloud consumers an extensive choice of available services and providers.

Also, in a world where many global enterprises have data sovereignty requirements, especially companies based in Europe, and can only engage with cloud providers physically located in the country, data sovereignty delivered through a VMware cloud partner can offer the security, reliability, performance, and regulatory compliance that these organizations need. In many cases, it is only through working with a local VMware cloud service provider that these heavily regulated organizations can be offered this level of data sovereignty and security and achieve the compliance and certification required for their regulated industries. By providing these options, VMware cloud providers can better design services specifically for these types of organizations and market segments, allowing them to grow their businesses into new areas.

Many cloud providers offering services through the VMware Cloud Provider Platform also deliver multiple compliance certifications alongside managed application services and specialized certifications for various vertical

¹VMware Internal estimation, January 2019

markets, including HIPAA, PCI-DSS, ISO 27001, and/or application-specific accreditations such as SAP Cloud Platform Partner or Oracle Cloud Managed Service Provider. Through this approach, VMware cloud providers are able to offer a deeper and more integrated solution for compliance and regulated offerings, which are critical for organizations operating in industries such as financial service, healthcare, government, retail, and telecommunications. For instance, a cloud provider capable of delivering PCI-DSS-compliant services is essential for financial services and retail organizations in order to maintain data privacy and to ensure that payment card systems support all the required security layers, protocols, data protection procedures, and disaster recovery plans.

With regular news reports highlighting the reputational consequences and fines associated with credit card breaches at either major airlines or personal data loss at financial institutions, maintaining PCI compliance standards is more essential than ever before, so why not use an expert in this field? VMware cloud providers can provide this expertise to ensure the security of customer credit card data, establish protocols to prevent data breaches, and as a result, protect an organization's brand and reputation not only among its existing customer base but universally protecting itself from litigation, serious monetary losses, and loss of business credibility. Furthermore, VMware validates its cloud service providers to ensure they meet certain service and capability requirements. The "Cloud Verified" and other badges help potential customer organizations easily identify cloud service providers that can deliver the validated cloud services they need.

As we have seen, VMware cloud service provider solutions can align to every business's requirement and provide a trusted partner to deliver services that allow them to grow and expand their organization. This global ecosystem of cloud providers offers the most flexible, robust public, private, and hybrid cloud solutions for enterprises of all sizes. The VMware Cloud Provider Platform aims to provide these partners with a standardized approach and toolset to deliver control at every level and replicate on-premises network architectures, using common constructs, at scale, to create and manage intricate networking topologies using overlay networks and leveraging hybrid cloud reference designs. This approach gives customers the ability to build solutions that require no architectural changes to run in the cloud, and allows them to maintain the network, security, and access control policies configured on-premises with a hybrid cloud solution that is architecturally consistent. This gives organizations unmatched cloud flexibility and choice and removes the need, cost, and complexity of managing silos of proprietary clouds and technologies. With VMware and its partners, enterprises can take advantage of complete workload portability between vSphere on-premises or private cloud environments and public cloud.

This common approach across a single operating environment helps ensure operational consistency and the maximization of existing skill sets and tools. This allows IT staff to enjoy a common management experience across

hybrid IT environments with end-to-end business insight, automation, intelligent operations, integrated high availability, consistent disaster recovery and data protection, and live migration to improve performance and proactively avoid outages. In addition, this architectural model can help establish a foundation for future innovation to reach across customer environments and the public cloud with newer technologies such as Kubernetes or Object Storage.

Chapter 2

VMware Cloud Provider Platform

VMware's Cloud Provider Platform is built on VMware technologies and solutions to deliver enterprise class cloud computing solutions to cloud service providers. Fundamentally, it is a platform for delivering services. Taking this approach allows cloud service providers to leverage efficient pooling of on-demand, self-managed IaaS, to provide resources consumable as a service, as well as provide layers of additional service delivery, such as PaaS (Platform as a Service) or SaaS (Software as a Service) solutions.

The Cloud Provider Platform focuses on all layers of service delivery, and not simply IaaS, by extending the robust VMware Cloud Director capabilities to facilitate an ever-growing portfolio of cloud services. The platform is designed to allow cloud providers to build flexible cloud services that are truly enterprise ready through a portfolio of best-in-class private, public, or hybrid cloud solutions.

The Cloud Provider Platform is made up of several core VMware components, integrated to facilitate service delivery and provide the foundation for a VMware "Cloud Verified" platform. Some of these components illustrated in Figure 2.1 are optional and can be consumed to enhance service delivery of the overall customer experience.

As illustrated in Figure 2.1, the Cloud Provider Platform goes beyond the traditional cloud service provider use cases and forms the foundation for multi-cloud service delivery and Managed Service Provider revenue opportunities. The four core components of the Cloud Provider Platform include:

- 1. Run VMware Cloud:** Develop in-house platform for delivery of services. This is the traditional cloud service provider model where the cloud service provider develops, builds and operates a range of services from their

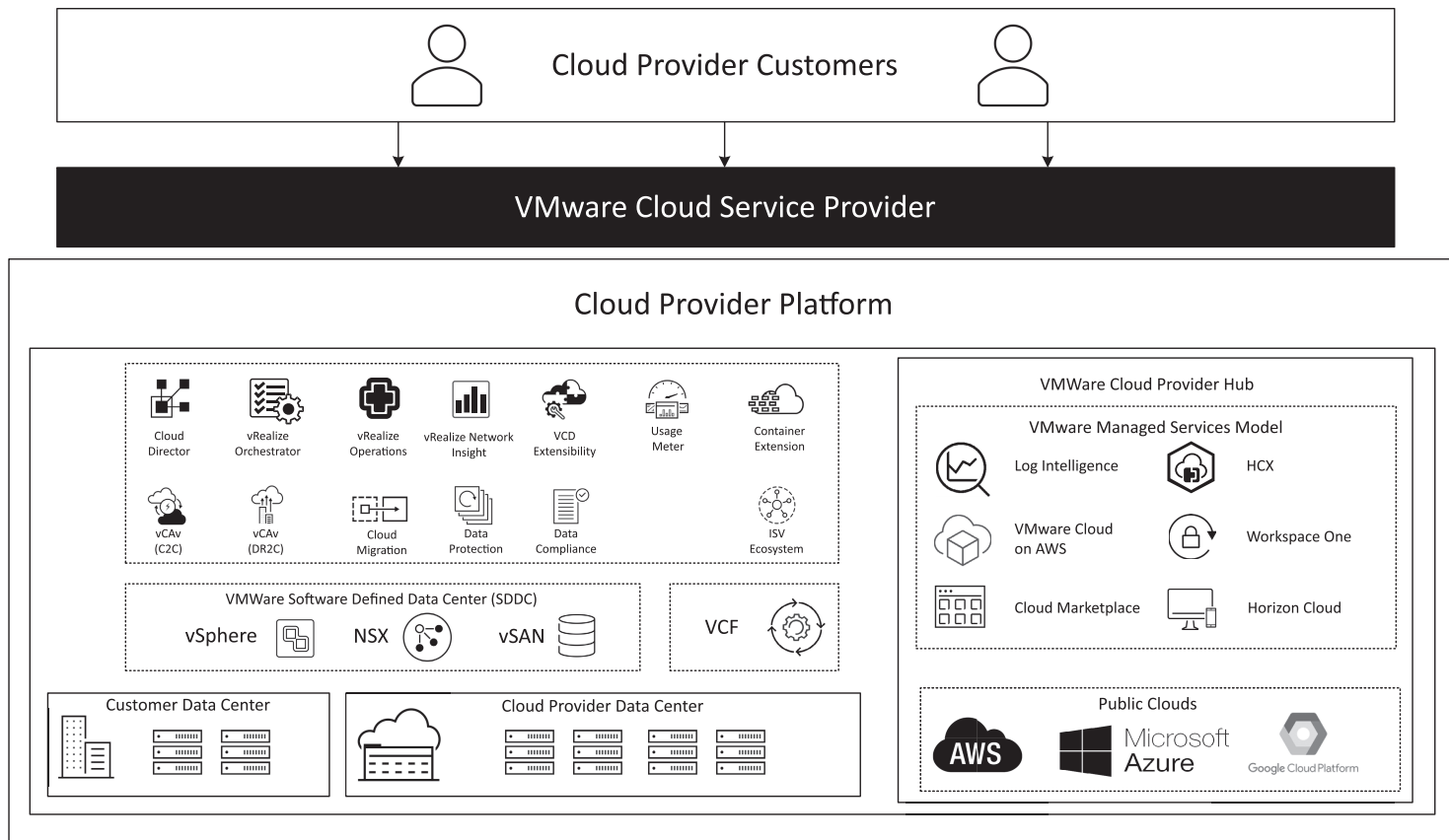


FIGURE 2.1 Cloud Provider Platform Overview

own data center facilities. The cloud provider designs, builds and operates all aspects of the cloud platform.

- 2. **Run VMware Cloud in Hyperscale Clouds:** VMware are giving new monetizing opportunities to provide services around VMware running on hyperscale cloud platforms. These include VMware Cloud on AWS, Azure VMware Solutions, Google Cloud VMware Solutions, and Oracle Cloud VMware Solutions.
- 3. **Expand Managed Services on Public Cloud:** VMware Cloud Services portfolio continues to expand through in-house developed solutions as well as acquisitions. These SaaS-based offerings can be surfaced and managed by cloud service providers through Cloud Provider Hub. This gives providers the ability to offer a range of MSP solutions alongside their own offerings, which might include solutions such as Cloud Health Technology, Cloud Assembly, or Wavefront.
- 4. **Expand Service Opportunities at Edge:** Alongside the cloud providers own data center offerings and VMware Hyperscale cloud offerings, solutions such as VMware Cloud on Dell and VMware Cloud on AWS Outposts provide monetizable opportunities for running MSP services in a customer's own data center.

PROVIDER DATA CENTER – CORE CAPABILITIES AND INTEGRATED COMPONENTS

The first of the Cloud Provider Platform components, Run VMware Cloud, has of course been the traditional cloud provider model for service delivery for many years. As illustrated in Figure 2.2, this requires, as it always has, the development of services around multiple components, which make up a cloud provider's hybrid cloud capabilities. At the center of this is VMware Cloud Director, which provides the core platform that enables a cloud service provider to deliver services for its end customers on top the VMware SDDC.

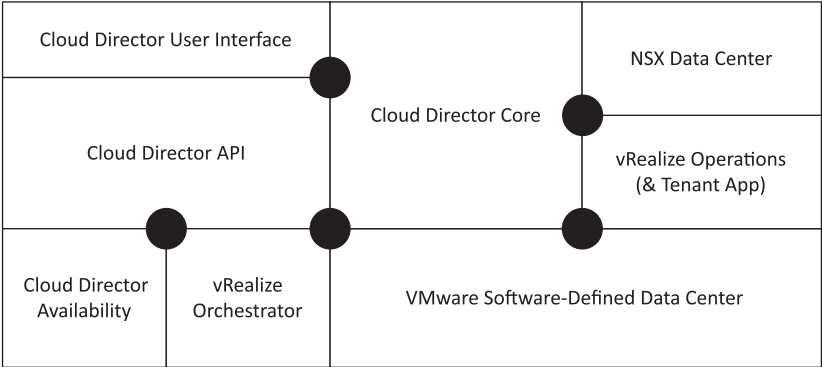


FIGURE 2.2 Cloud Provider Platform Component Architecture

VMware Cloud Director, which as we have stated at its core is a platform, enables creation of software defined virtual data centers. This allows cloud service providers to take their physical data center resources and convert them into elastic pools of compute resources that can be offered to end customers through various consumption models. It achieves this by creating a new layer of abstraction that groups all the physical data center assets, such as compute, storage, and networks into pools of virtual resources and partitions these to provide modular virtual data centers that can be allocated to individual customers, in a fully isolated and tenanted manor.

As illustrated in Figure 2.3, Cloud Director leverages VMware vCenter and VMware vSphere to convert these physical resources into logical resource pools and in turn leverages NSX and vSAN to create corresponding virtual networks and storage.

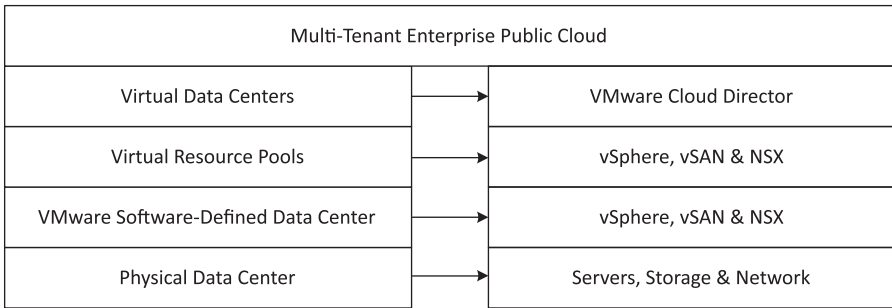


FIGURE 2.3 VMware Cloud Director Abstraction Layers

This mechanism of creating elastic resource pools allows cloud service providers to have improved management of their data centers resources. The Provider VDC creates an abstraction layer between resource and consumption, helping service providers to create, expand, manage, and delete resources without disrupting tenant services. This pooled abstraction layer makes it easy for the cloud provider to add more resources, which can be used to scale Provider VDCs. As the Provider VDC synchronizes with the vSphere inventory, the additional resources are added automatically and ready to be consumed. Another benefit of this architecture is that consolidation of data center resources and allocation pools enables oversubscription of resources, which allows cloud providers to avoid oversizing data centers and helps reduce capital expenditure costs.

Finally, in addition to providing virtual data centers that can be consumed by multiple tenants, Cloud Director also adds application catalogs and cloud provider and cloud consumer management services via a user interface and API, to facilitate application deployment and management services.

Last but not least, Cloud Director is a key component in a cloud provider obtaining VMware “Cloud Verified” status. This is the mark of the most elite

and differentiated VMware cloud providers across the world, and gives them access to multiple benefits, such as improved visibility, joint go-to market funds, and more.

CLOUD PROVIDER HUB

The purpose of VMware Cloud Provider Hub is to provide a central portal to enable cloud service provider partners to transact, deploy, and manage VMware SaaS offerings, such as VMware Cloud on AWS, Cloud Health Technology, and the VMware Marketplace. The aim is to ensure partners can expand their managed services portfolio with asset-light solutions from VMware and offer multi-cloud managed services across their customer environments, provider environments, and public cloud deployments.

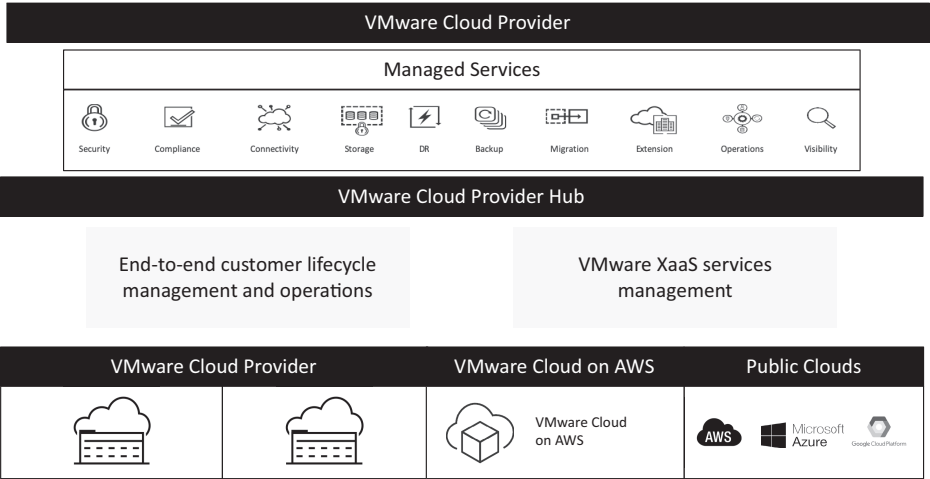


FIGURE 2.4 VMware Offerings for VMware Partners

As illustrated in Figure 2.4, with Cloud Provider Hub, cloud providers can manage customer lifecycle management end-to-end including:

- Onboarding customers
- Provisioning and managing services
- Consolidated view of their usage
- Managing support

Cloud Provider Hub simplifies and standardizes the way partners leverage VMware Cloud Services and offers them to their customers. Hub exposes a common mechanism to consume, provision, and manage customer access to these services. Partners can offer unmanaged, partially managed, or fully managed services to their customers. Hub offers a Provider console, available

via user interface and API, for the partner to onboard and manage customers, manage service provisioning, access for customers, and perform usage and support management. Hub also offers a Tenant console, shown in Figure 2.5, which again is available via user interface and API, for the partner's end customers to consume services that are provisioned for them by the partner.

Tenant Org	Org Type	Admin Contact	Tenant Company	Domain	City	State	ZipCode	Country	Tag
Beacon Prod	Customer	bianca1@harakirimail.com	Beacon Inc	beacon.com	New York	New York	10027	United States	prod_hr
Convo	Customer	charlie1@harakirimail.com	Convo LLC	convolic.com	Chicago	IL	60601	United States	
Dundie Scranton	Customer	dwright1@harakirimail.com	Dundie Paper Co	dundiepaper.com	Scranton	Pennsylvania	18507	United States	Account 8765

FIGURE 2.5 Cloud Provider Hub

More and more of the VMware Cloud Services portfolio are designed to be multi-cloud capable and partners can leverage these services on Hub to offer centralized managed services on vSphere, Cloud Director, VMware Cloud on AWS, Native AWS, Google Cloud Platform, and Azure. By leveraging and delivering VMware Cloud Services through Hub, cloud provider partners do not need to invest in and operate siloed management services for multiple clouds, whether for on-premises service delivery or in the public cloud.

VMWARE VREALIZE OPERATIONS AND THE TENANT APP

VMware vRealize Operations plays multiple roles within the Cloud Provider Platform architecture. At the core level, VMware vRealize Operations aim to deliver continuous performance optimization to drive operational efficiency and reduce cloud provider costs. At a high level, through the use of self-driving operations, capacity management, intelligent remediation, and integrated compliance checks, vRealize Operations can automate and simplify many IT operations. In addition, through the use of federated dashboards it can provide insight from the application layer to infrastructure across a cloud provider's entire asset portfolio, no matter where or in which cloud they reside in.

For cloud providers, vRealize Operations can be looked at for providing value in three keys areas:

- **Operational Efficiency:** With Cloud Directors shared resource economies and huge scalability, vRealize Operations provides operational efficiency for cloud providers themselves. Through the use of vRealize Operations, cloud provider admins are able to be more productive. It is estimated that engineers can manage **3x more workloads per admin**. This allows a more customer-centric approach.
- **Service Monetization:** The vRealize Operations Management Pack for Cloud Director in conjunction with the Tenant App provides a means to deliver tenant-facing dashboards for VCD Tenant Admins, giving enhanced workload metrics, specific to their tenant Org. VDCs. These include Orgs., Resource Pools, VDCs, vApps, VMs, Alerts, and Metrics. This feature can be enabled on a tenant-by-tenant basis, which gives the cloud provider a mechanism to upsell this service in conjunction with existing VCD resources. This solution is fully integrated into the VCD user interface via extensibility, as illustrated in Figure 2.6. In addition, the Tenant App solution delivers a tenant metering function through pricing policies, which is a critical component that enables accurate and automated usage measurement, cost calculation, analysis, and reporting to cloud providers and tenants. This feature is key to delivering transparency and accountability of billing to consumers.
- **Differentiation:** Building on the above, every cloud provider needs to find their niche in today’s all-cloud market. The Cloud Provider Platform helps cloud providers innovate to stay ahead of the competition in their area of specialty. By reducing operational complexity of maintaining cloud infrastructure, vRealize Operations in conjunction with Cloud Director allows cloud providers to focus on value additions and innovation investments.

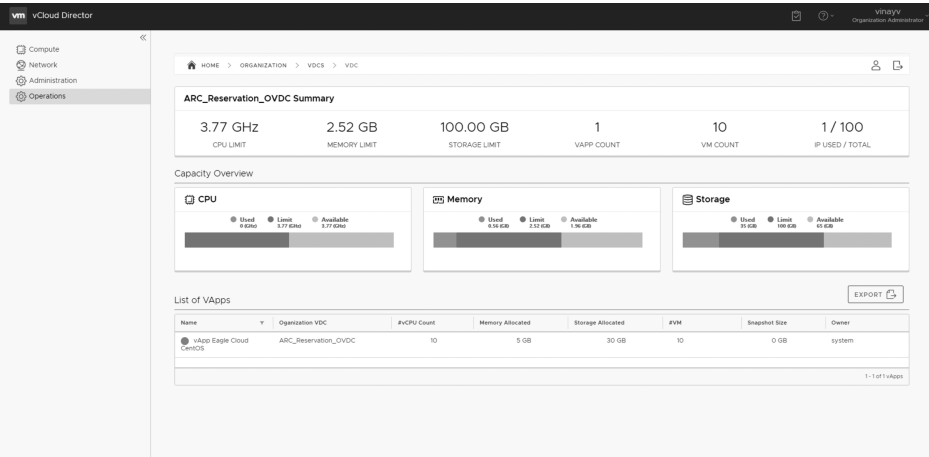


FIGURE 2.6 vRealize Operations Manager Tenant App

This solution is recommended to cloud providers who wish to have greater insight into the metrics associated with their entire Cloud Provider Platform, across multiple data center sites and regions, and provide an upsell opportunity for service users through the use of the capabilities of the Tenant App.

VMWARE VCLOUD USAGE METER AND VMWARE VCLOUD USAGE INSIGHT

The VMware Partner Connect Program offers for cloud providers a subscription-based licensing model where cloud providers consume licenses in a pay-as-you-go model. vCloud Usage Meter is a virtual appliance that collects usage data from vCenter Server and Cloud Director to provide data to VMware for license reporting purposes.

The local Usage Meter appliance securely transfers data to VMware's Usage Insight Service where usage is pre-populated into the Business Portal, making it ready for cloud providers to review before submitting the usage data to VMware.

The local vCloud Usage Meter appliance should be deployed in the management cluster along with the rest of the common management platform components in each region. As illustrated in Figure 2.7, the vCloud Usage Meter appliance requires access to metered endpoints, such as vCenter Server APIs, NSX APIs, and the Cloud Director APIs.

The speed at which vCloud Usage Meter collects data is dependent on the size of the data set and number of vCenter Server inventories in the environment. Typically, one vCloud Usage Meter appliance will be deployed per each Cloud Director instance and data center.

The local vCloud Usage Meter instances can be accessed via the following methods:

- **Virtual Machine Console**—Used for access to configuration settings.
- **Web Application**—A GUI view that provides access to manage metering settings and reports.
- **JMX Service**—The JMX service provides access to internal data such as database size, customer count, and customer rule count.
- **REST API**—vCloud Usage Meter provides a REST API interface that provides access to external clients (portals, scripts, etc.).

SOFTWARE-DEFINED EVERYTHING

Compute virtualization is the most mature category in VMware's SDDC model, which has enabled vSphere to deliver huge benefits, such as lower cost, increased flexibility and simple scalability, across the entire industry for almost two decades. In today's world, all critical data center components are also able to benefit and advance in the same way with technologies such as

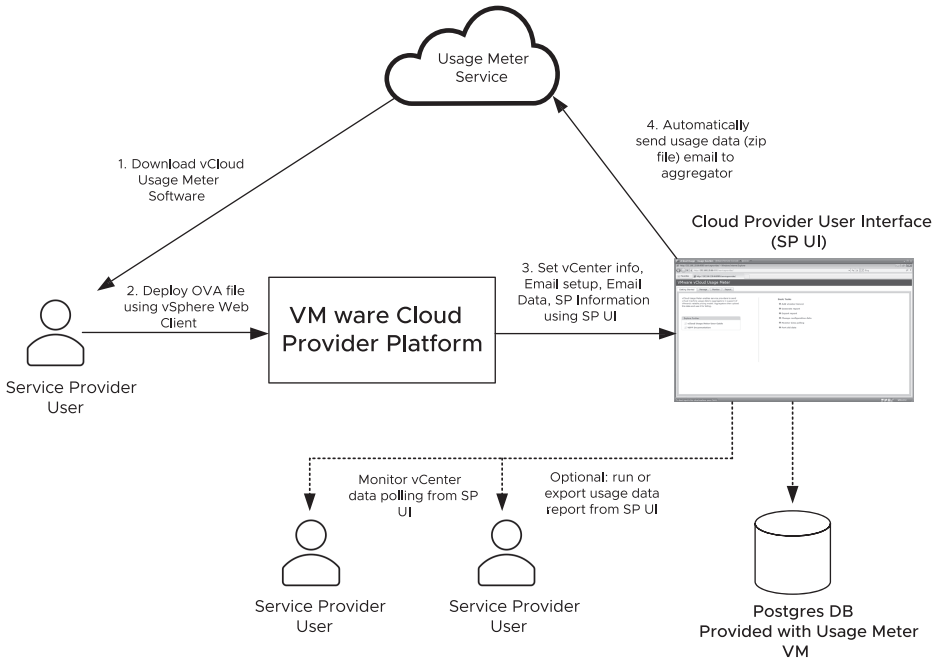


FIGURE 2.7 Cloud Usage Meter Data Collection Workflow

NSX and vSAN, allowing for dynamic scalability and software defined flexibility for software-defined networking and software-defined storage. These all pave the way for software-defined application services delivery inside the SDDC giving us a software abstraction of the entire data center.

The VMware SDDC model brings multiple benefits to cloud providers while trying to deliver services to customers. Firstly, it brings the economics of industry-standard servers to networking and storage, often eliminating the need for more expensive and complex purpose-built hardware. The SDDC architecture also improves performance and unlocks an affordable server-side flash model, to deliver greater performance with lower latencies, and allowing the cloud provider to converge storage, compute, management, and storage networking on one common platform.

The SDDC model also enables the rapid adoption of new technologies, since it's easier to upgrade software than upgrade a hardware platform. This gives the cloud provider a more efficient way of working, offering the flexibility and agility required to match the dynamic needs of customer's new and complex application architectures, as the DevOps world continues to evolve and dominate.

VMWARE CLOUD VERIFIED PROGRAM

The VMware "Cloud Verified" branding delivers the VMware Cloud Provider Partner Program's premiere status to a partner's service, and ensures the

service architecture of the cloud provider is truly built to optimize the customer's hybrid cloud experience. Customers can migrate their workloads to these partners with the confidence that the cloud provider is running a validated solution, optimized for cost, performance, security, and service delivery. The aim is to give end customers the strongest, most flexible cloud infrastructure for digital business transformation, with:

- **Integration and interoperability:** Deploy core compute, storage, and network building blocks that provide the flexibility to run apps wherever you want, in the cloud or a on-premises data center, or both.
- **Flexibility:** Cloud consumers can choose from a wide range of VMware cloud providers with a range of vertical focuses in any region worldwide. This allows cloud providers to deliver unique capabilities, and differentiated services for their customers.
- **Cost and Skill Optimization:** End customers can rely on the cloud provider's skill sets to lower TCO with consistent operations and infrastructure across their cloud delivery models.

This branding and associated logo helps cloud customers and prospects to identify and engage with VMware cloud provider partners, who VMware have validated to deliver service offerings based on the VMware Cloud Provider Platform and associated cloud infrastructure technology.



VMWARE VALIDATED DESIGNS (VVD) FOR CLOUD PROVIDERS

The VMware Validated Designs (VVD) for Cloud Providers provides a pre-validated set of software components that can be quickly deployed by a service provider and are pre-validated for a representative multitenant cloud provider environment, with multiple scaling profiles. The intent of the initiative is to document a verified stack and provide scale and performance benchmarking to help cloud providers deploy and maintain the lifecycle of the Cloud Provider Platform. This also reduces the complexity relating to interdependencies between the VMware components required for a Cloud Provider Platform-based service. The VMware Validated Designs for Cloud Providers defines which versions should be deployed, and what kind of scale and performance VMware Cloud Providers can expect to achieve.

This validated guidance gives cloud providers clarity with which version of each software component of the Cloud Provider Platform stack is recommended at any given time. Each version also comes with a predictable support time frame for all the underlying components, typically 12 to 18 months from the launch of the corresponding release. This helps reduce the time and expense involved in determining what components to upgrade to, and

when, allowing the entire associated software stack to stay inline from both a support and compatibility perspective, so that incompatible combinations can be avoided.

The other key element of the VMware Validated Designs for Cloud Providers is the guidelines for sizing hardware and software components to align with expected tenant load. The scale and performance element of this documentation cannot cover every cloud configuration and size; however, it aims to provide a sizing recommendation for a typical cloud, one that is broadly representative of the VMware cloud provider’s requirements.

Another key consideration is that it is not the intention for the scale and performance guidance to push the Cloud Provider Platform to its absolute limits. The VMware Cloud Provider Platform should be sized based on the anticipated tenant demand and follow the common sizing guidelines set out throughout this book. As illustrated in Figure 2.8, cloud providers should use good design practices and comprehensive monitoring to ensure that the performance observed by tenants and administrators is in line with the expectation of consumers, and enables the provider to meet the service levels defined for the service.

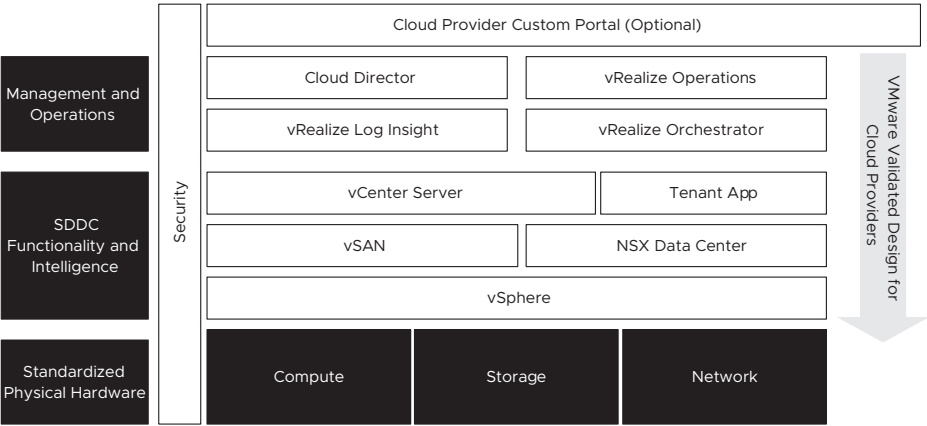


FIGURE 2.8 The Cloud Provider Platform Stack

BUILDING A SERVICE DEFINITION

Service definition is a key aspect of service design, delivery, and management. It allows both the cloud consumer and the cloud provider to know what the service will deliver, and perhaps more importantly, what the service will not deliver. Having a clearly defined service definition helps both customers to providers to understand the scope, limitations, and cost of the Cloud Provider Platform-based service offering.

The following key considerations represent some of the service design factors to take into account when developing the Cloud Provider Platform service objectives:

- Use cases
- User roles that interact with the service (managed service demarcation line)
- Consumption model
- Serviceability (time to respond, time to resolution, service credits or penalties)
- Metering, reporting, and pricing (reporting details/frequency/history)
- Service offering details (infrastructure, applications)
- Backup (schedule, restore time, data retention)
- Performance (application performance, network performance)
- Security and Compliance (regulatory compliance, logging, auditing, data retention, reporting)
- Availability (uptime) and Continuity

BUILDING A SERVICE-LEVEL AGREEMENT

The service-level agreement (SLA) is a negotiated contract or agreement between the cloud service provider and the cloud consumer. The primary aim of the SLA is that serves to document the service-level guarantees between the two parties.

VMware cloud service providers are expected to deliver enterprise grade cloud services to their customers and therefore will typically carry a comprehensive service-level agreement guaranteeing service levels, typically including an availability target that is equal to or exceeds three 9s (99.9%), and includes key considerations for overall service performance, customer support, and responsiveness. The Cloud Provider Platform SLA is typically either uniquely negotiated between the cloud provider and customer or is delivered as a highly standardized contract between the provider and customer that defines in detail responsibilities and limitations associated with the services between the two parties.

Providing detailed guidance on how to calculate the level of availability and performance for all Cloud Provider Platform service elements is beyond the scope of this book. Cloud providers typically have an SLA framework in place that can be leveraged or supplemented to support Cloud Provider Platform service delivery. However, it is important to recognize that SLA guarantees should extend to all elements of the Cloud Provider Platform, including the underlying infrastructure. Individual components, such as compute, network, storage services, and management elements that directly support the platform are all in scope, including the management delivery framework, consisting of both APIs and the user interface consumers use to access the services.

DESIGNING AN ENTERPRISE CLASS PUBLIC/HYBRID

In the enterprise business world, application downtime, especially for patching and maintenance, is no longer acceptable. The VMware Cloud Provider Platform framework allows cloud providers to deliver reliable services with the

features and capabilities enterprises need to run their businesses. Partners who provide services in this model are able to deliver robust, VMware ‘Cloud Verified’ services tailored to customers’ business needs and have years of experience bringing together public and private clouds to ensure hybrid compatibility with existing VMware customer environments.

The native VMware Hybrid Cloud model is a composite of two or more VMware instances (private and public) that remain unique entities but are also bound together by standardized VMware technology. This native VMware technology enables data and application portability with complete flexibility and even live migration for bursting and extending application scope. For instance, bursting applications for load balancing between VMware Cloud instances on-premises and in the public cloud. With a native VMware Hybrid Cloud, the enterprise business customer gets the advantages of both worlds, with the ability to burst into the public cloud when needed while maintaining critical assets in their own data center.

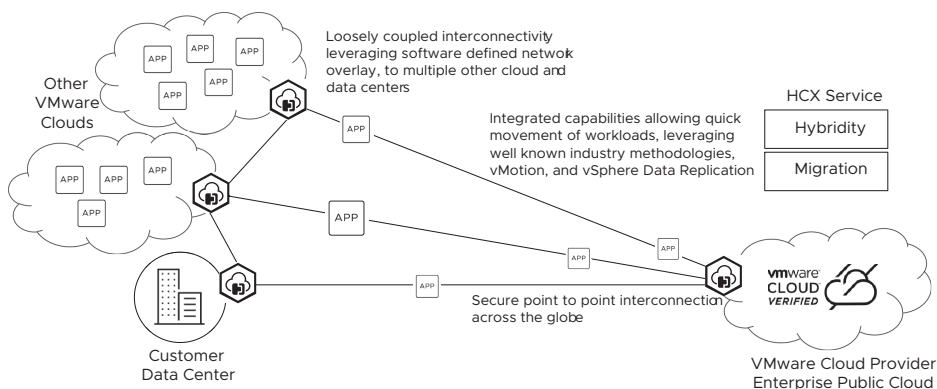


FIGURE 2.9 VMware Native Hybrid Cloud

An enterprise business customer seeking to leverage a native VMware Hybrid Cloud can benefit from an easy path to hybrid cloud with a VMware Cloud Providers standardized reference design for a repeatable hybrid cloud model between managed on-premises resources and cloud provider resources, which offers a number of key advantages:

- **Cloud Management:** Offered as a fully or semi-managed service by the VMware cloud provider and delivered as-a-service to provide automated application deployment, scaling and workflow management and vRealize Operations to ensure in-depth management of the entire infrastructure—increasing productivity.
- **On-Premises SDDC Architecture:** The customer Software-Defined Data Centre service offering provides access to a 100% dedicated cloud infrastructure with “Software-Defined” technologies to not only enhance the performance, security, and management of the infrastructure, but also to

facilitate an automated building-block, scale-up and scale-out architecture. Consistent, compatible, and simple to operate across on- and off-premises.

- **Security:** A common security model across both on-premises and cloud provider resources, delivering a secured cloud platform for delivering a wide range of business services to consumers of varying size and profile.

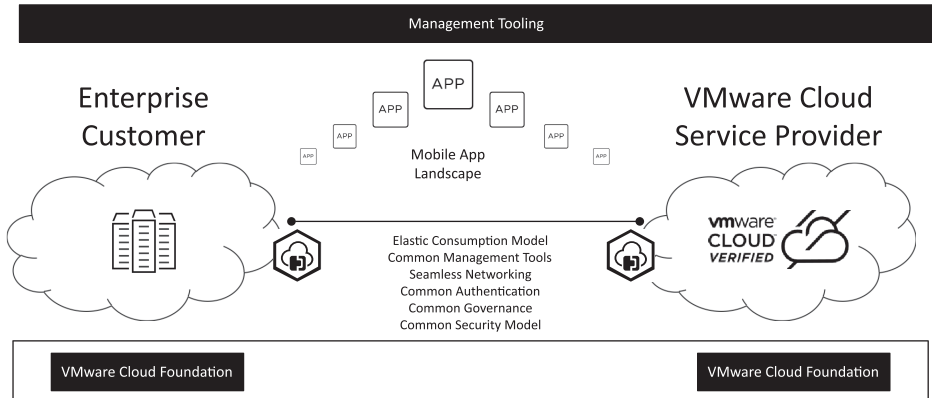


FIGURE 2.10 Consistent VMware Cloud Platform

This preservation of existing investments and choice of cloud through VMware solutions is key to VMware's public cloud strategy. As a global leader in cloud, VMware offers end customers solutions that promote freedom of choice. The same basic VMware platform deployed by enterprises is also deployed by thousands of cloud service providers across the globe, which enables a compatible bridge between a private cloud, hosted internally in a data center, and public clouds hosted remotely.

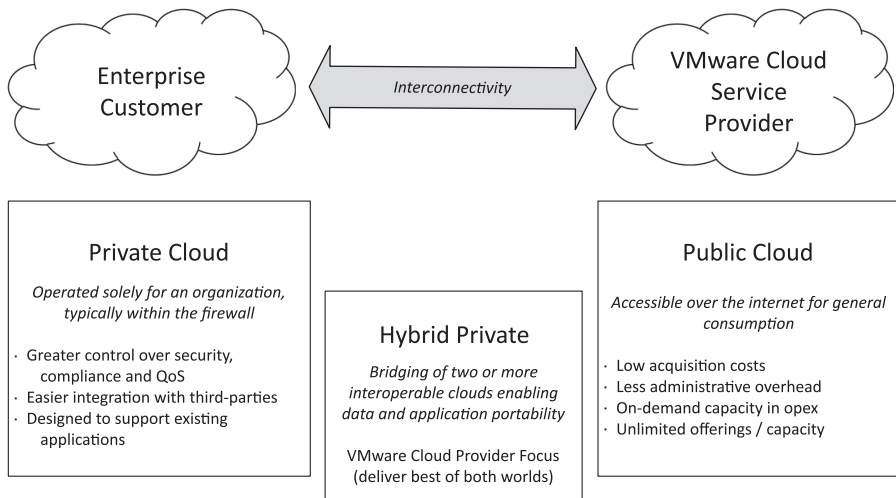


FIGURE 2.11 Consistent Hybrid Cloud Experience

This ability to bring existing applications seamlessly to the cloud and offering a consistent security, management, and performance model provides end customers with a simple and efficient method of extending into the public cloud at a significantly lower risk.

ON-PREMISES-TO-PROVIDER CLOUD CONNECTIVITY

A hybrid cloud incorporates a combination of cloud instances and can include private, external, or public resources. These resources often need to change as various projects and customer requirements evolve. The ability to dynamically shift workloads and their supporting configurations between cloud platforms allows the infrastructure to grow and adapt to these changing business requirements. Routine data center maintenance, such as software upgrades, the deployment of new hardware, or troubleshooting often requires downtime. However, a properly designed hybrid cloud infrastructure utilizing the following technologies can allow a customer to securely migrate workloads across physical hardware, with no disruption to the services.

One key cloud provider requirement is to provide services that enable their customers to connect their on-premises vSphere implementations to the Cloud Provider Platform solution. Many different approaches to architecting a VMware-based hybrid cloud exist, depending on the specific use cases, requirements and technologies employed. However, the end goal is always the same, a hybrid cloud that provides an organization with the ability to manage some resources in-house and have others provided externally.

A business of any size, with an existing private cloud, can choose or be required to provide and manage public or external cloud resources in a secure and scalable way. The connectivity between different cloud instances that enables data and application portability is key to the hybrid cloud solution, meaning the environments must have reliable high-speed connectivity with each other through the use of a data center interconnect.

Connectivity over the Public Internet

IPSec VPNs are mostly used in areas where the provider cannot enable private connectivity or the customer has specific requirements, such as a need to support remote users. IPSec solutions for site-to-site communications are typically not selected by medium or large enterprise businesses as the traffic is routed over a public IP backbone. Despite this, IPSec connectivity still has its place where the requirements define a need.

With VMware NSX, there are several options available to leverage IPSec to create a common network between the customer and provider's data centers:

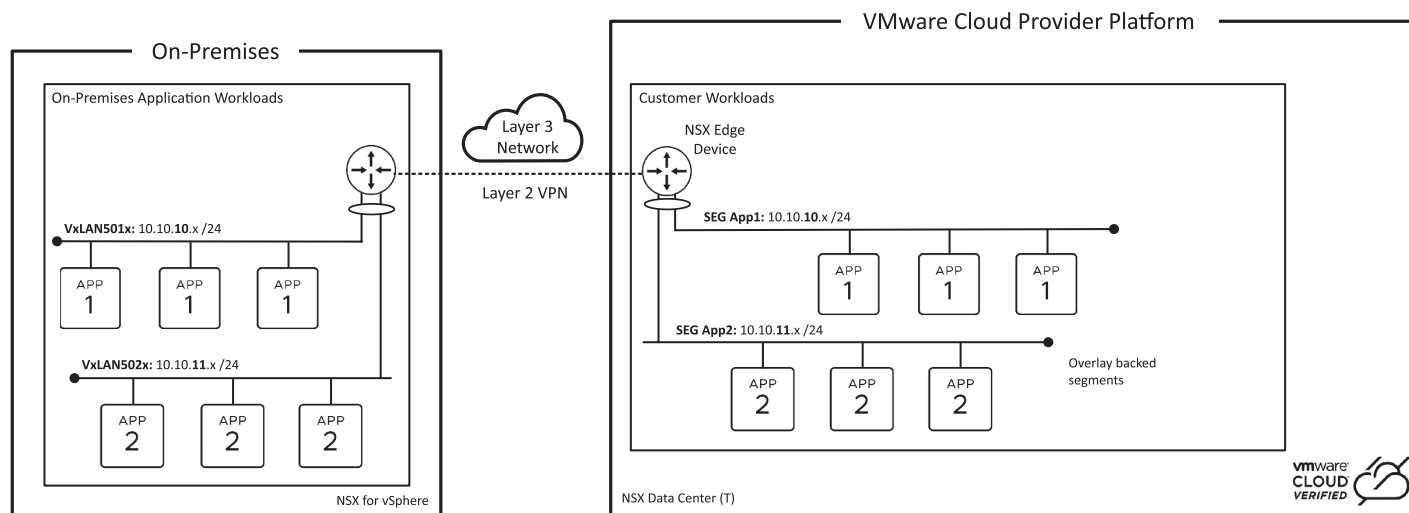


FIGURE 2.13. VMware NSX L2VPN Using a VXLAN-to-VXLAN-Based Solution

Software-Defined Wide Area Networking is a newer approach that aims to simplify Wide Area Networking while ensuring optimal performance. VMware's VeloCloud SD-WAN solution enables the use of broadband Internet, with or without traditional MPLS, to create enterprise grade wide area networks with increased bandwidth, high performance access to cloud provider resources, service insertion and extensive network visibility.

Unlike traditional WANs, SD-WAN delivers increased network agility and cost reduction. Software-Defined WAN has its roots in Software-Defined Networking (SDN), the underlying principle of which is to abstract the network hardware and transport characteristics from the applications that use the network.

This new strategy helps in many use cases where a business wishes to create their own private WAN connections by leveraging VPN over the public internet. While this certainly is a great benefit, there are still many use cases where dedicated or virtual leased lines are required in order to provide a more secure and guaranteed throughput for the customer.

Connectivity over Private Circuits

As outlined above, larger organizations typically want to set up a private circuit, an option that is highly scalable and offers increased security and reliability. Such solutions are common among public cloud providers with Microsoft Azure offering ExpressRoute circuits, delivered as a standardized service, and Amazon Web Services providing the same type of private connected circuit with their Direct Connect (DX) service.

In most regions there are several different technologies available to decide upon when choosing how to build an organization's private wide area network, although these can vary globally. Most VMware cloud service providers will offer a choice between two different circuit types:

- Private circuits (Leased Line)
- Point-to-point networks with MPLS (Multi-Protocol Label Switching)

To determine the most appropriate data center WAN interconnect solution, which meets the customer's business requirements, the features of different available WAN technology systems must be evaluated. The design decision must be centered on the technology and the business requirements that have been established. Key factors in determining the most appropriate solution are the bandwidth and latency requirements for application mobility use cases, although other factors, such as data replication and management access can also provide additional design considerations, alongside:

- Distance
- Speed (bandwidth and latency)
- Cost
- Redundancy

Evaluation of WAN Platforms

Many cloud providers will offer private circuits (Leased Line) connections using Dark Fibre from 1mb up to 100GB. This means the bandwidth of a leased line can evolve with a business, enabling you to increase capacity as the requirements grow, or even for seasonal bursting. Since a Dark Fibre line is completely dedicated to the customer, data security is guaranteed. Put simply, this is the best choice if the customer's business needs complete control of their connectivity at all times, and any loss of service would have a major financial or operational impact.

Dark Fibre: Dark fibre are strands of fiber optic cables that are deployed underground but are not currently being used. Because there is no light being transmitted across them, they are considered dark. While any type of fiber (single or multi-mode) can technically be dark, the term dark fibre typically refers to 9 μ single mode fiber. Dark fibre can be privately owned, typically across a campus-style network, or can be leased from a provider to create a metropolitan area network (MAN).

However, getting a dedicated fibre connection in this fashion is likely to be costly, and cloud service providers will only have a finite amount of fiber available. For providers, the solution is the use of Dense Wavelength Division Multiplexing (DWDM), effectively allowing multiple fiber endpoints to traverse a single fiber line, allowing more customers to utilize this private service.

Wavelength Division Multiplexing (DWDM and CWDM): Wavelength division multiplexing (WDM) is a mechanism used to simultaneously transmit multiple streams of data on a single strand of fiber (or pair of fibers). Using WDM allows the customer to overcome the limitation of needing a separate fiber pair for each type of traffic that they want to transmit (unlike dark fibre). There are two different types of WDM—Dense WDM (DWDM) and Coarse WDM (CWDM).

DWDM, as its name suggests, supports more wavelengths because they are spaced together more tightly. As the wavelength separation is so small, DWDM requires high-precision equipment that is typically very costly to purchase.

This equipment provides filters to the light emitted from the data source, such as the switch, then multiplexes it onto the fiber strands along with the light from other sources. At the far end of the fiber strand, another multiplexer separates the signals from the different sources and forwards them to the appropriate destination equipment. Because the signals are also amplified by the DWDM system, they can travel much further than with native dark fibre. With the right hardware and media components in place, DWDM can potentially be extended to thousands of kilometers. The number of different wavelengths that can be supported by DWDM varies by vendor, but 32, 64, and even 128 are possible.

In summary, DWDM has the following key design considerations:

- Most typically used over short distances, such as 100 km to 200 km, but can extend further
- Dark fibre must be available

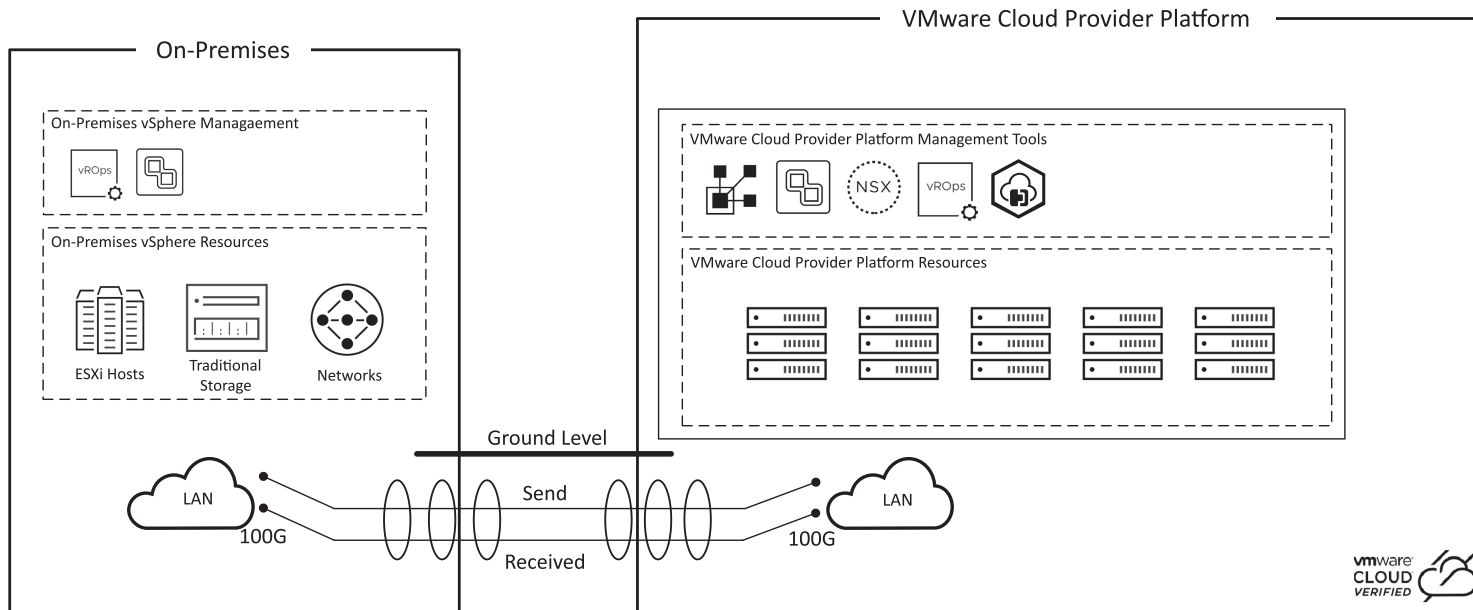


FIGURE 2.14 Dark Fibre Connectivity

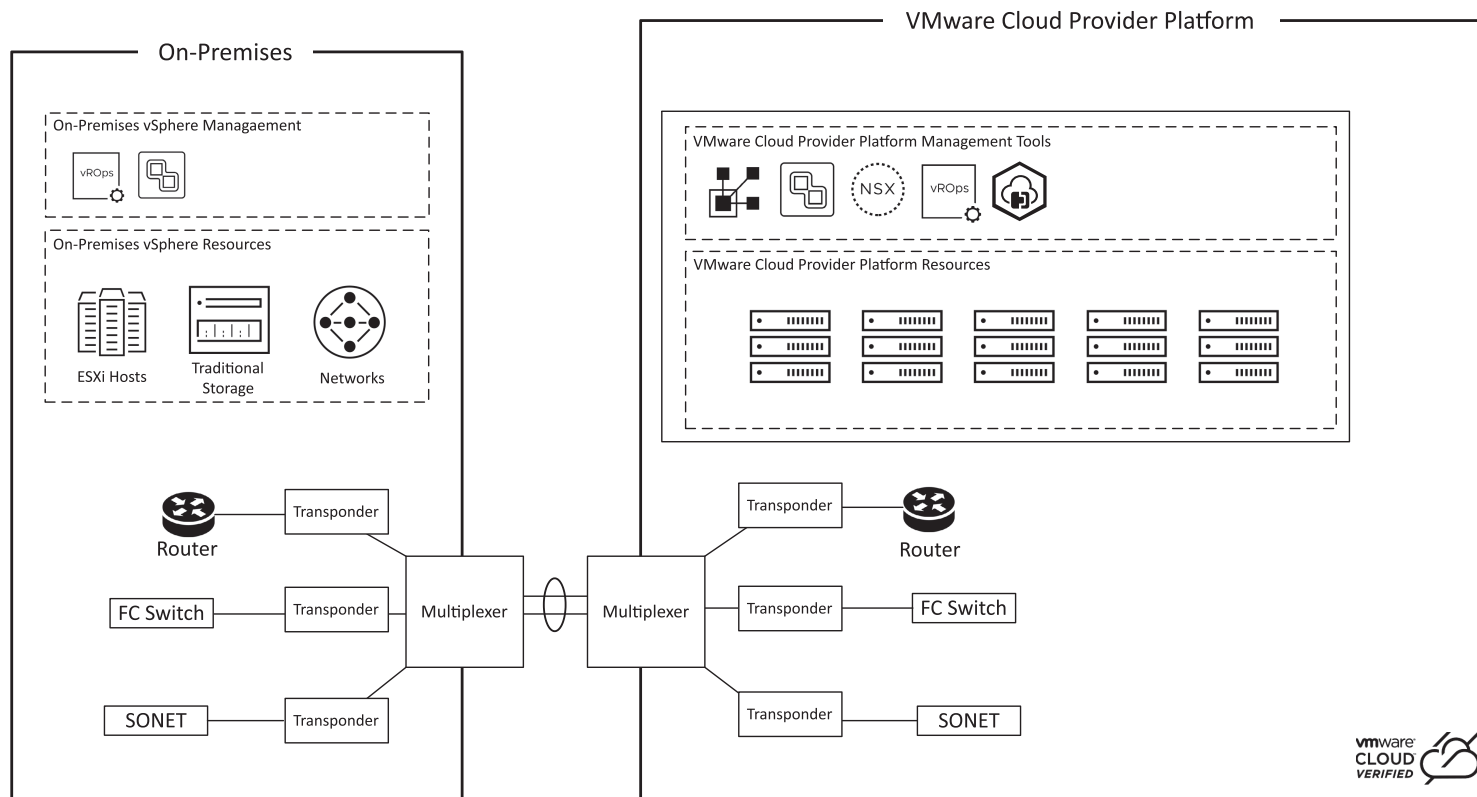


FIGURE 2.15 Wave Division Multiplexing (DWDM and CWDM)

- Divides a single beam of light into discrete wavelengths (lamdas)
- Each signal can be carried at a different rate
- Dedicated bandwidth for each multiplexed channel. Approximately 0.4 nm spacing
- DWDM transponders can support multiple protocols and speeds (LAN, SAN, and other signals)

CWDM was introduced as a lower-cost alternative to DWDM. The functionality is almost identical to DWDM, except that the wavelengths (called channels in CWDM) are spaced further apart. The result of this is that fewer channels can fit on the fiber strands (maximum of 16 channels per fiber pair). Also, CWDM does not require as many components as DWDM, and those components are significantly less costly. However, a CWDM solution does not amplify and clean the signal as DWDM does, so the distance limitations for CWDM are significantly lower. In most implementations, the maximum distance is typically less the 100 km, but multiple signals, such as Ethernet and Fibre Channel, can be carried, giving it a significant advantage over native dark fibre.

Virtual Leased Lines: Dark fibre aside, often when businesses require private leased lines, they tend to be referring to a Virtual Leased Line (VLL). A Virtual Leased Line provides an isolated private network for its customers; however, this is different from the dedicated dark fibre leased lines discussed previously, which have physical isolation from other customers.

Regardless of Layer-2 or Layer-3 capabilities, the options listed below all offer private connections between cloud provider and customer but do so on the backbone of a network that is shared among all customers in that service-provider's area. MPLS (Multi-Protocol Label Switching) offerings from service-providers, can often become muddled. For instance, a service provider can sell a product as "MPLS VPN" while other concepts, to be discussed later, are built on top of an existing service-provider MPLS backbone.

Multiprotocol Label Switching (MPLS): A MPLS Network is a flexible private wide area networking service that functions similarly to IP, but the infrastructure switches based on a predetermined path instead of calculating paths hop-to-hop.

For an MPLS or IP interconnect between data centers, the design changes from the fiber model. Because the interconnect between the data centers is a Layer 2.5 or Layer 3 link, Layer 2 connectivity is not available beyond the data center boundary. However, to perform the same actions as a fiber link, with cross-site Layer 2 adjacency (routed over a Layer 3 network) when using an IP-based WAN, a tunnel must be established between the two sites and the Ethernet data encapsulated within that tunnel structure.

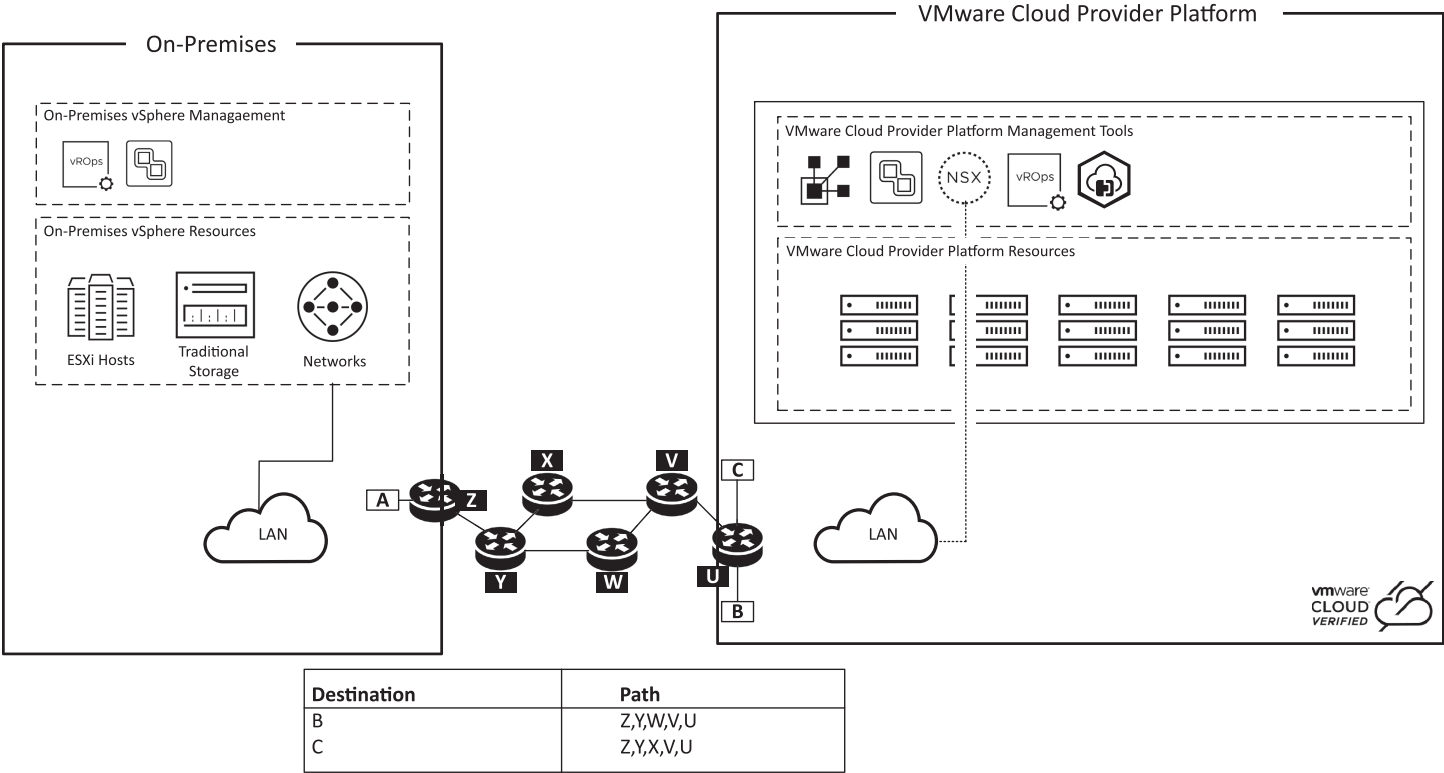


FIGURE 2.16 Multiprotocol Label Switching (MPLS)

In summary, MPLS has the following key design considerations:

- Layer 2.5 protocol
- Used over short or long distances
- Used when dark fibre is not available
- Links might be shared
- Operates over a variety of Layer 2 infrastructures
- Switches based on label
- Uses predefined paths, reducing delays from dynamically calculated paths
- MPLS performance (which affects delay, jitter, failover, and so on) is dependent on service provider SLAs

Virtual Private LAN Services (VPLS): A VPLS (Virtual Private LAN Services) WAN is an evolution of the MPLS protocol suite. VPLS is an example of an overlay WAN service that is structured on top of the cloud service provider's MPLS network. With VPLS, customers are given an Ethernet connection into the MPLS cloud, and therefore are able to connect to other WAN endpoints as with they were on the same network segment.

Compared to the MPLS VPN option, where the ISP participates in sharing customer IP routes over the VPN, VPLS allows endpoints to reside on the same subnet. VPN can still be utilized for additional security and offers point-to-point and multipoint capabilities. Since the endpoints connect at Layer-2, IGP routing protocols can be used to share routes from each end of the WAN, as illustrated in Figure 2.17.

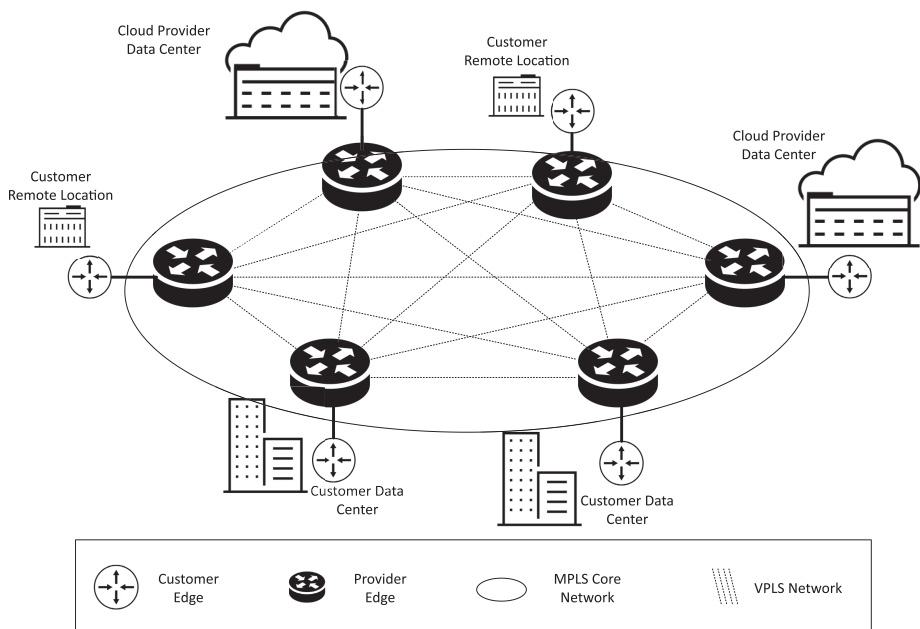


FIGURE 2.17. A VPLS WAN Deployment, Full Mesh of Layer 2

WAN Interconnect High Availability: High availability for data center interconnect WAN links is significantly more complex than it is for a LAN environment due to factors such as distance, cost, non-ownership of equipment, and connectivity. The HA model for WAN connectivity varies widely depending on the type of WAN being employed. In general, a typical WAN HA solution can be classified as either including redundant components, redundant links, or both.

Redundant components refer to having multiple devices, such as routers or multiplexers, at each site. This provides redundancy in the event of component failure and allows for non-disruptive maintenance and upgrades. Redundant links refers to having multiple connections between data centers. This can be achieved through multiple dedicated media, such as multiple strands of dark fibre, or through a combination of fiber and a shared medium, such as MPLS. You should also provide a combination of both methodologies to have a fully redundant data path for the highest level of availability.

Some service providers might also provide the option of having a standby connection. This option often has a lower total cost of ownership because you are charged by the amount of data transmitted. This might be a cost-effective way of achieving a redundant link without having to pay for unrequired bandwidth.

Before making the appropriate design decision about WAN availability, understand the available options and what needs to be protected against. For example, deploying redundant links or components does not provide any benefit if the single ISP has a large-scale outage.

A hybrid WAN Architecture: Many customers typically elect to use a combination of WAN products and capabilities, especially when you consider that most large enterprises must meet the needs of global data centres, remote users, extranet partners, and remote offices. To this end, a large customer is rarely likely to implement a single product or technology. Some locations will be better suited to layer 3 managed services, others will be better suited to LAN extension technology. How the cloud service provider goes about connecting their sites should align to the customer's specific requirements.

Securing Communication: The primary method used to secure data as it transverse between data center sites is IPsec, or Internet Protocol Security. IPsec is commonly used to encrypt portions of the IP packet. The dataflow is encrypted at the edge of one site and then decrypted at the entry point to the second site. IPsec has two different modes of operation:

- Transport mode
- Tunnel mode

Transport mode only encrypts the IP packets payload and leaves the header in an unencrypted state so it can be read by other network devices. However, in tunnel mode, the entire IP packet, including its header, is

encrypted, meaning that it must be encapsulated in another IP packet with an unencrypted header. This way the source and destination information in the header of the packet can be read by network devices. Tunnel mode is most commonly used for point-to-point or site-to-site VPNs, as discussed earlier in this chapter.

HYBRID CLOUD USE CASE EXAMPLE

Many different approaches to architecting a VMware-based hybrid cloud exist, depending on the specific use case requirements and technologies employed. However, the end goal is always the same, a hybrid cloud is a cloud computing environment in which an organization provides and manages some resources in-house and has others provided externally.

In today's dynamic business environment, which requires continuous uptime and fast recovery from disasters, ensuring business requirements and service level agreements are met, one of the most important requirements is the availability of databases. The "being prepared" approach to providing application high availability is aimed at meeting customer agreed service-level agreements, while reducing risk of revenue losses and maintaining compliance. Planning for disasters and minimizing their impact is critical for all enterprises and government agencies. VMware cloud providers can assist in providing data center resources and services to leverage application high availability solutions for Microsoft and other software vendors. Designing and deploying a database solution such as Microsoft SQL Server 2019, in conjunction with a highly available hybrid cloud solution from a VMware cloud provider, can help businesses mitigate risks and still control deployment and operational costs.

This example solution architecture focuses on deploying Microsoft SQL Server 2019, taking advantage of its AlwaysOn functionality to provide a hybrid cloud approach to application availability with a VMware Cloud Provider Platform solution. SQL Server 2019 AlwaysOn provides flexible design choices for selecting an appropriate solution for your application with protection against failures within a data center and adequate redundancy across data centers. For a further overview of the AlwaysOn design options for high availability and disaster recovery, please refer to Microsoft documentation.

The sample solution architecture shown in the following figure is built on the Windows Server 2019 Enterprise Edition operating system platform, with a four-node WSFC, consisting of a two-node cluster in the primary on-premises data center and a two-node cluster at the secondary VMware Cloud Provider Platform facility, using a stretched, active-active failover cluster configuration.

The four SQL Server 2019 instances are configured with a single AlwaysOn Availability Group. This availability group was created with one primary replica and three secondary replicas.

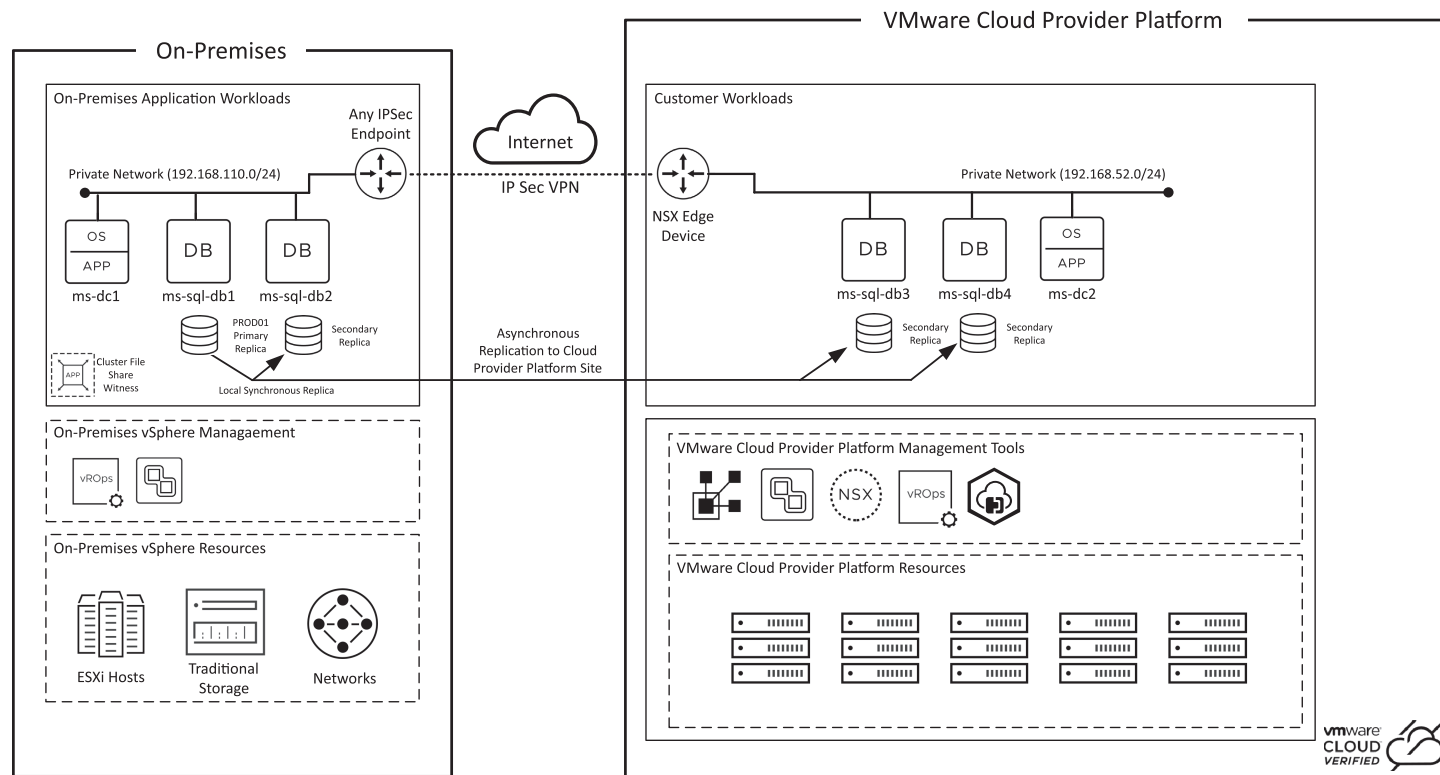


FIGURE 2.18 Sample Architecture Database Functions

In the sample design shown in Figure 2.18, the availability group is configured with a primary replica database (MS-SQL-DB1\PROD01) and one local secondary read-only replica (MS-SQL-DB2\PROD01) at the customer's on-premises data center. Two secondary replicas (MS-SQL-DB3\PROD01 and MS-SQL-DB4\PROD01) are configured in the VMware Cloud Provider Platform partner facility, for high availability and disaster recovery purposes. This architecture provides redundancy in multiple failure scenarios at both the on-premises and provider data centers.

The automatic failover option provides high availability by ensuring the database can fail over quickly and can become available after the loss of the primary replica. In the event of an unplanned outage, the failover takes place without manual intervention by operational teams or database administrators.

To configure an AlwaysOn Availability Group for automatic failover, the current primary replica of the database and one secondary replica must be set to synchronous-commit mode with automatic failover. Synchronized mode provides for little or no data loss, but requires more available network bandwidth, and therefore the environment must be within the minimum latency requirement limits set by Microsoft. In a typical architecture, to maximize performance and minimize data loss for disaster recovery purposes, configure all replicas for synchronous-commit mode with automatic failover.

In addition to the above, The Microsoft SQL Server 2019 AlwaysOn Availability Groups ability to support multiple replicas with a mixture of synchronous and asynchronous data replication between primary and multiple mirrors, makes this technology an ideal solution for building a disaster recovery solution on top of the existing high availability solution.

In the sample design shown in the following figure, the on-premises data center hosts two availability replicas in synchronous commit mode for high availability. The VMware Cloud Provider Platform facility hosts one replica, running in asynchronous mode for disaster recovery protection. In addition, each data center facility employs vSphere HA to protect virtual machines from ESXi host failure and to facilitate the rapid recovery caused by a locally contained outage.

This sample architecture allows for variations to this topology using multiple VMware Cloud Provider data centers and multiple replicas (up to five) within the solution architecture.

The AlwaysOn Availability Groups feature is the ideal solution for deploying highly available Microsoft SQL Server 2019 instances on a Cloud Provider Platform. AlwaysOn Availability Groups provide out-of-the-box protection for hardware, software, and data failure as well as additional capabilities to off-load backups and read-only reporting queries to a secondary replica. With its non-shared disk architecture, AlwaysOn Availability Groups can be used safely with vSphere vMotion, DRS, and vSphere HA to reduce downtime and improve flexibility in your application architecture, while lowering costs and minimizing the need to perform an SQL failover.

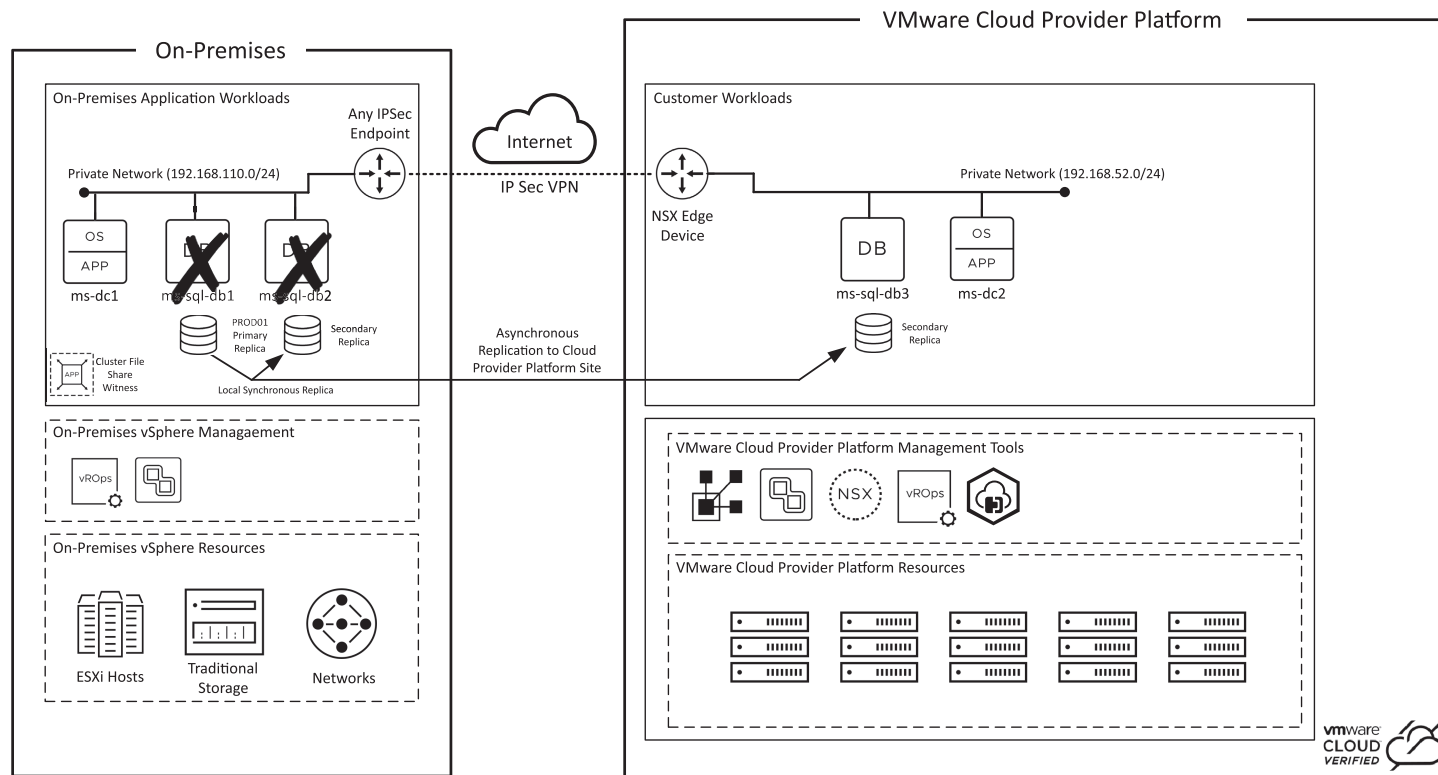


FIGURE 2.19 SQL Server 2019 AlwaysOn Availability Group for HA and DR

This sample solution can significantly reduce Microsoft SQL Server recovery time from hours or days to minutes or even seconds but is only one example of how application level availability solutions by software vendors can be leveraged for hybrid cloud use cases.

CLOUD PROVIDER PLATFORM CORE USE CASES – MOVING BEYOND IAAS

VMware Cloud Director provides a robust, multi-tenant infrastructure service platform designed to consume compute, network and storage resources from data centers and convert these resources into individual elastic units that can be provided as a service and consumed by multiple tenants.

What was once a platform for delivering purely infrastructure services has now evolved into a platform for service delivery and while Cloud Director maintains its core IaaS functionality it has expanded these capabilities to facilitate service delivery through a modern cloud delivery platform. The traditional role of cloud providers making IaaS accessible to tenants now has limited margins and little opportunity to increase profits. In order to stay relevant, cloud providers are now expected to deliver PaaS and SaaS offerings, and they need to be able to do this in a multi-tenanted aware model, which is fully isolated at the tenant level, while also delivering consistency to the user.

Throughout this book we're going to be addressing all aspects of the Cloud Provider Platform, including user interface and API extensibility, vRealize Orchestrator integration, for both pro-active and reactive service delivery and ISV integration into the Cloud Director product itself, to provide a simple and unified approach to service delivery.

VMware Cloud Provider Platform Networking and Security Overview

Traditional physical network and security solutions were rigid and complex to implement for non-specialists. In addition, they typically created a costly barrier to realizing the full agility of delivering public cloud infrastructure. These old-school limitations bound physical networking and security to inflexible dedicated hardware, and created barriers to optimizing network architecture and capacity utilization. This manual provisioning, dedicated physical hardware appliances and fragmented management interfaces reduced efficiency and limited the ability of cloud providers to rapidly deploy, move, scale, and protect customer applications and data based on their business needs.

To overcome these challenges, VMware NSX is consumed by Cloud Director to create efficient, agile, extensible logical constructs that meet the performance and scale requirements of tenant's virtual data centers. The result is dramatically simplified operations, efficient resource utilization, and greater agility to scale in response to business needs, in either a self-service or managed delivery model.

At a high-level, this is achieved by VMware NSX providing software defined networking and security which is built into the Cloud Provider Platform. Just as Cloud Director abstracts compute capacity from the server hardware to create virtual pools of resources, which can be consumed as a service, NSX abstracts networking and security into pools of capacity and separates the consumption of these services from the underlying physical infrastructure.

VMware NSX delivers network and security services with a broad range of natively consumable services through the Cloud Provider Platform and as such reduces the cost and complexity of data center operations and unlocks the operational efficiency of public cloud service delivery. The key NSX services that are consumed by customers (but not to say these are the only services available) include:

- **Firewall Services:** A stateful inspection firewall that can be applied either at the perimeter of the virtual data center (north/south traffic) or at the virtual network interface card level with the distributed firewall, directly in front of specific workloads, (east/west traffic). The firewall-rule table is designed for ease of use and automation with VMware Cloud Director objects for simple and reliable policy creation. Stateful failover of Edge devices enables high availability for business-critical applications.
- **Virtual Private Network (VPN):** Industry-standard IPsec capabilities that securely extend the virtual data center. Site-to-site VPNs link virtual data centers and enables hybrid cloud use cases at low cost.
- **Load Balancer:** A natively integrated virtual load balancer to scale application delivery, without the need for dedicated hardware, is delivered in an easy to consume model. NSX load balancers are placed at the edge of the virtual data center, the load-balancer supports Web, SSL, and TCP-based scale-out for high-volume applications. More advanced services consuming NSX Advanced Load Balancer (formally Avi Networks) capabilities are also available to consume and will be addressed in a later chapter.
- **GENEVE Overlay:** By using the Geneve overlay, NSX-T abstracts the network into a pool of capacity and separates the consumption of these services from the underlying physical infrastructure. This model is similar to the model vSphere uses to abstract compute capacity from the server hardware to create virtual pools of resources that can be consumed as a service. This is the enabling technology for network virtualization, providing network abstraction, elasticity, and scale across the data center. Geneve provides an architecture for organizations to scale applications across clusters and data center sites without any physical network reconfiguration.
- **Integrated management:** Cloud Director provides separation of duties with role-based access control (RBAC) while providing a central point of configuration and control for network and security services.

Workload Migration and Initial Onboarding

Today, it's no longer a question of whether workloads can be moved to the cloud, but which ones and how quickly. In addition, enterprise organizations

have to think about cross-cloud architecture. They must satisfy demands for on-premises private clouds and stretching workloads into the VMware public cloud with data management and governance. The Cloud Provider Platform provides the core foundation for a cross-cloud architecture, delivering a common compute, storage, and management platform for public cloud as customers are running on-premises, along with the necessary flexibility and mobility.

For end customers, the move to a software-defined architecture is just the beginning of moving to a hybrid cloud architecture. As everything becomes software-defined, it allows IT organizations to provide flexible, service-oriented capabilities, including migration, live or otherwise, to and from a public cloud platform. Having application mobility is the key to leveraging public cloud as part of a hybrid cloud strategy.

The onboarding of customer workloads is a key service that cloud providers must offer, typically through some sort of professional services engagement. Multiple options exist from VMware, third-party solutions as well as manual solutions, often used by providers to move large quantities of data into their facilities, through high capacity removable disk solutions.

The two options provided by VMware which are now most commonly used are Cloud Director Availability onetime migration and HCX. The migration factors that will identify the most appropriate solution for a specific migration come down to scale and complexity.

Cloud Director Availability Migration Overview: One of the key use cases for Cloud Director Availability, is the ability to create a hybrid cloud environment between an end user on-premises data center and a multi-tenant Cloud Director environment. This feature allows customers to simply migrate workloads from their on-premises vSphere environment to a Cloud Provider Platform-based Public Cloud at zero cost.

Cloud Director Availability is able to stretch an on-premises layer 2 network to the cloud, a feature that helps customers to migrate workloads without the hassle of having to re-IP. Furthermore, there is no requirement to have an on-premises deployment of NSX to stretch a layer 2 network to cloud, although without NSX on-premises you are limited to extending only VLANs to VXLAN on the cloud provider side.

As illustrated in Figure 2.20, on the tenant side, we only need a single on-premises appliance (which includes the replicator and tunnel functions) and the vCenter plug-in to perform migrations.

Cloud Director Availability migration offers two migration methods: cold and warm migration. Warm migration is driven through pre-staged replication, for efficient low downtime migration, effectively providing a reboot to cloud solution. Warm provides a low-downtime solution as the workload remains online during the replication process. For a cold migration the workload is powered off. This migration type will likely allow for more effective transfer with less complexity. However, it will require downtime as the system needs to be powered down.

A deeper look at Cloud Director Availability will be provided in Chapter 8, where we will address not only the migration use case but also the disaster recovery to cloud (DR2C) and cloud-to-cloud disaster recovery solutions.

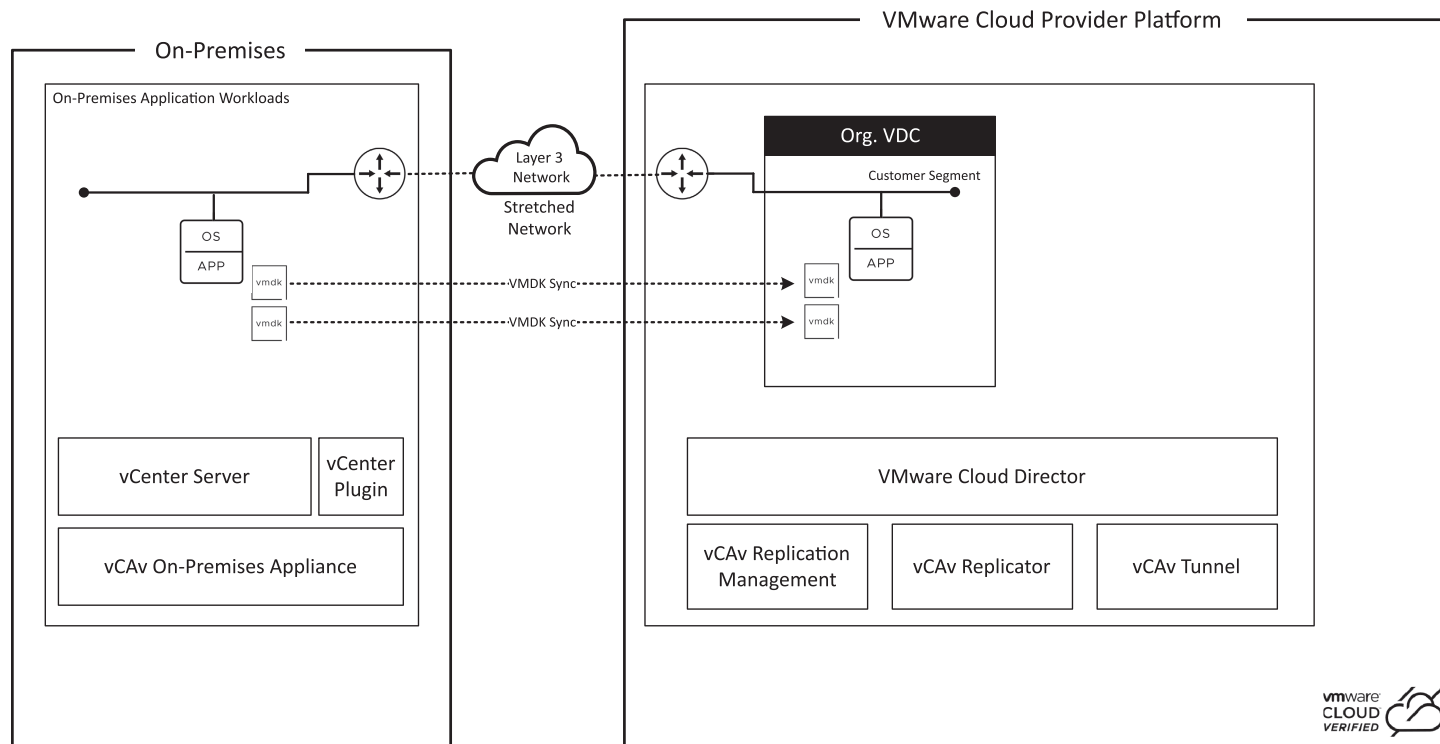


FIGURE 2.20 Cloud Director Extender Architecture and Components

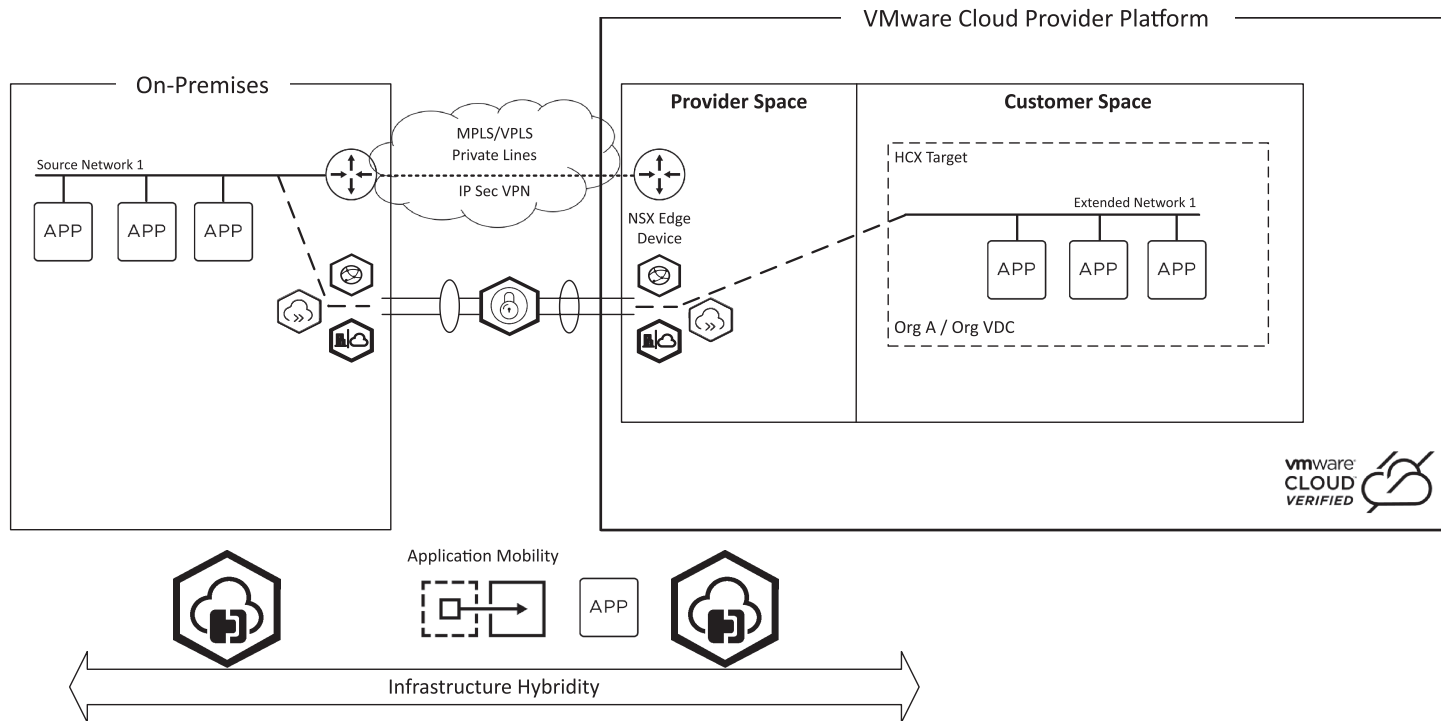


FIGURE 2.21 VMware HCX Overview

HCX Overview: The primary aim of VMware HCX is to provide application migration capabilities without application downtime, or with minimal application downtime, and with no requirement to refactor applications, or their dependencies, in any way.

VMware HCX provides the hybrid cloud capabilities to bridge between an existing on-premises data center and the VMware Cloud Provider Platform. The service offers application migration, which delivers data center mobility for live bulk migration and low downtime migration to the Cloud Provider Platform. Built on Top of SD-WAN technology, HCX provides a high-throughput network extension, with WAN optimization, traffic engineering, automated VPN with strong encryption data center interconnectivity, with built-in vSphere protocol proxies.

Through this combination of features, VMware HCX enables cloud on-boarding, without the need to upgrade the source infrastructure, supporting migration from vSphere 5.0+ to Cloud Director, with a simple low-risk migration SaaS platform.

While VMware HCX is provided as a SaaS-based platform through Cloud Provider Hub, migration-based connectivity is point-to-point, meaning no customer migration data traverses the SaaS-based components.

Depending on the application migration window, tolerable downtime, available bandwidth, latency, and other network factors, HCX offers multiple migration options to align with business and application owner requirements.

For instance, if the business plans to migrate 200 workloads over the course of a single weekend, HCX could be configured to begin replicating over to the VMware Cloud Provider Platform well in advance, allowing the initial full replication to seed from the remote site. Once complete, only delta synchronization need occur. It is noteworthy to highlight HCX utilizes hyper-visor replication, not the vSphere Replication Appliance and benefits from the

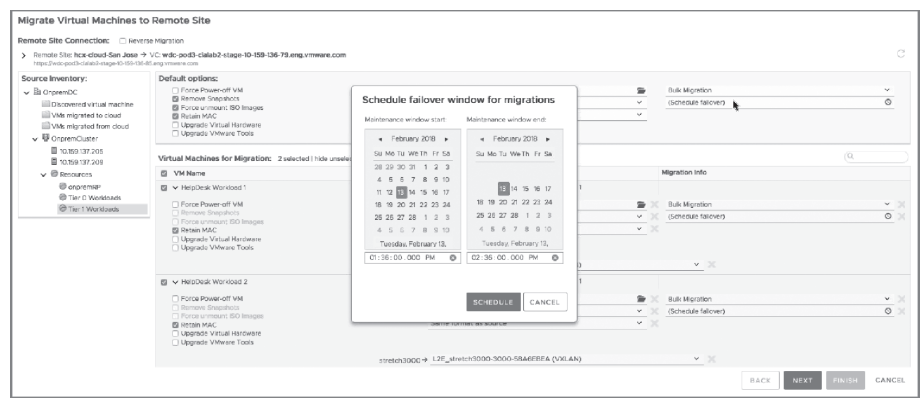


FIGURE 2.22 Scheduling a Failover for Bulk Migration

same deduplication, compression, and WAN optimization features applied to live migration.

When ready, the administrator can select the day and time they wish to schedule the switchover, a maintenance window agreed by the business, such as Sunday from midnight to 6am. During the scheduled migration time, the source workload is powered down, removed from the source inventory, and a final delta synchronization performed. The final task is for the workload to be booted on the target Cloud Provider Platform, after which the migration is completed, and the administrator can begin any required testing or validation tasks.

Like Cloud Director Availability migration, even though the Cloud Provider Platform is built on NSX technology, that does not necessarily mean that the customer must be operating NSX in their existing data centers. NSX on-premises is not a requirement, but might have additional benefits when some workloads may remain running in the customer's existing data center or operate within a hybrid application solution.

VMware HCX can help organizations bypass many challenges by providing a Layer 2 secure bridge over the internet. The HCX virtual appliance gives you everything you need to quickly set up a secure bridge, typically over the Internet, either permanently or perhaps while you're waiting for the private circuit to be enabled.

We will return to HCX and its specific implementation requirements when using Cloud Director as the target in Chapter 8.

Data Protection and Recovery

Protecting virtual machines is a critical component of any production environment. Data protection is not handled natively within the Cloud Provider Platform but is instead delivered through partner integrated extensions, developed by Independent Software Vendors (ISV).

This is achieved through Cloud Director's extensibility. In previous releases of Cloud Director, customization of the Flex user interface was not typically an option and there was no way for cloud service providers to differentiate through either the look and feel of the UX or through custom extensibility or integration with third-party ISV solutions, without building their own user interface. However, this has changed with the introduction of the new HTML5 user interface, first made available in VCD 9 and extended in every release since. Enhanced user interface extensibility promotes the integration of ISV solutions and further extends the option to build custom services, publishing those services to consumers and monetizing as appropriate—in essence, further enhancing the customization capabilities of VCD to provide XaaS.

Data protection software vendors have been among the first to take advantage of these new capabilities, through vendor-specific data protection solutions.

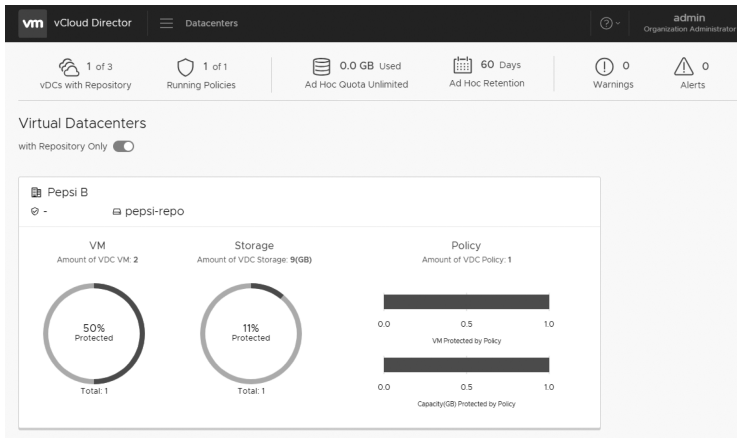


FIGURE 2.23 Dell EMC IDPA Cloud Director Integrated Solution

Data protection is the critical process of safeguarding important information from corruption, compromise, or loss, which today must be addressed alongside newer challenges, such as massive data growth, the growth of container-based workloads, and the pivot toward storing data in the cloud.

However, these are just some of the challenge's organizations are facing today. The ever-increasing costs for greater storage capacity, demands for shorter Recovery Point Objectives (RPO), the increasing expectations for the rapid provisioning of backup services, and the complexity of managing backup workflows alongside the sheer volume of data makes the data protection challenge greater than ever.

Therefore, these types of integrated solutions provide a common customer experience by extending the Cloud Director tenant portal user interface and REST API, providing tenants a single management endpoint for their virtual data centers. This in turn can help lower the total cost of data protection solutions for the customers by minimizing operating costs for Cloud Service Providers.

The VMware Cloud Provider Platform maintains full support for the vStorage APIs for Data Protection (VADP) including VDDK (Virtual Disk Development Kit) support, which provides a new set of libraries to create and access VMware virtual disk storage that can be consumed in conjunction with VADP.

Kubernetes Service

More and more modern applications are increasingly built using containers, using a microservices model, packaged with their dependencies and configurations. Kubernetes is open-source software for deploying and managing those containers at scale. Kubernetes allows developers to build, deliver, and scale containerized apps faster than ever before.

Kubernetes is an incredibly hot topic in the cloud provider space currently. Kubernetes is the industry-standard container orchestrator, used by companies of all sizes. It facilitates the deployment, resiliency, and scalability of customer applications, even in hybrid or multi-cloud infrastructures.

However, managing Kubernetes for many customers is not an easy task, therefore managed Kubernetes solutions on VMware Cloud Provider powered public cloud instances is becoming a key service offering for many of VMware's strategic cloud service providers.

In managed Kubernetes offerings, the solution's internal components are deployed, monitored and maintained by the cloud service provider's teams, and updated as required. This allows customers to focus on their business' containers and services, while enjoying all the advantages of the managed Kubernetes service.

A typical managed Kubernetes service provides intelligent scheduling, self-healing, horizontal scaling, load balancing, automated rollouts and roll-backs, and secret and configuration management. A Kubernetes service might also have advanced capabilities around simplified cluster management, container security, isolation policies, and DDoS protection.

Other customers are more interested in consuming Kubernetes in a self-service model. Giving the customer the ability to design their own cluster and integrate operational tools of their choice for consistency in deployment.

The Cloud Provider Platform provides a single tool to consume Kubernetes across multiple platforms and multiple delivery models. The Container Service Extension (CSE) is a VMware Cloud Director extension that helps tenants create and work with Kubernetes clusters.

CSE brings Kubernetes-as-a-Service to Cloud Director by creating customized VM templates (Kubernetes templates) and enabling tenant users to deploy fully functional Kubernetes clusters as self-contained vApps.

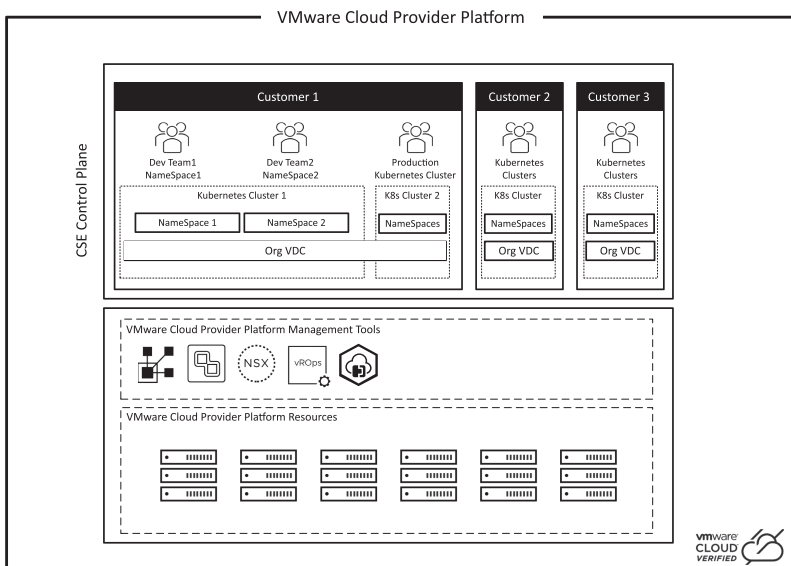


FIGURE 2.24 CSE – Flexible Multi-tenanted Kubernetes using Cloud Director Constructs—Native (without PKS)

CSE has a server component that installs as a Cloud Director extension which exposes REST API endpoints. CSE also has a client component that plugs in `vcd-cli`, to communicate with the CSE server via the exposed API endpoints. `vcd-cli` facilitates VCD users to create Kubernetes clusters on the Cloud Provider Platform.

Developers and other Kubernetes users interact with CSE Kubernetes clusters using `kubectl`. For these developer users, Kubernetes clusters work like any other Kubernetes cluster implementation. No special knowledge of Cloud Director or CSE administration is required. Developer users do not even need a Cloud Director account.

CSE Enterprise PKS Enablement: The Container Service Extension 2.0 enables orchestration of Kubernetes cluster deployments on VMware Enterprise PKS. At the same time, it maintains the CSE 1.x feature set of Native Kubernetes cluster deployments directly on VMware Cloud Director. As a result, the capabilities of CSE 2.0 allow tenants to leverage both Kubernetes Providers, Native and Enterprise PKS, for seamless Kubernetes cluster deployments, while ensuring cluster isolation between tenants. It also offers great flexibility to administrators to onboard tenants onto the Kubernetes Provider of their choice, be it Native and/or Enterprise PKS.

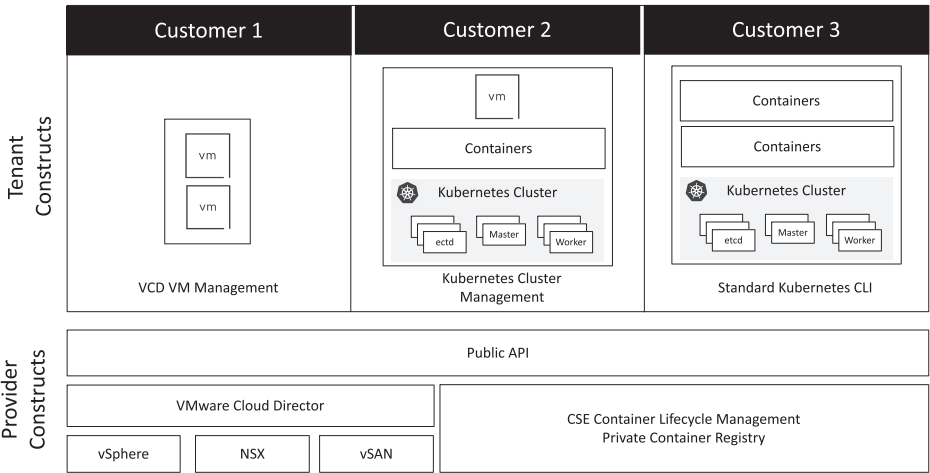


FIGURE 2.25 Container Service Extension Architecture

Enterprise PKS helps development teams with a CI/CD pipeline that is able to deploy a Kubernetes cluster on demand. PKS also comes with improved access control including integration with LDAP/AD and compliance with SAML standard. It implements RBAC in terms of who can create and manage the clusters. It also dynamically generates a Kubernetes service account with the cluster admin role.



Chapter 3

A Single Platform for a Global Cloud Business

INTRODUCING VMWARE CLOUD DIRECTOR

Formally known as Cloud Director, VMware Cloud Director (VCD) provides the core framework, workflow engine, and extensible user interface and API on which the Cloud Provider Platform is built. Cloud Director is central to VMware's cloud provider strategy and delivers the Cloud Management Platform for Public and Private Clouds delivered by partners.

VMware Cloud Director is at the core of giving cloud providers the ability to build a secure public cloud platform, which can significantly increase virtual data center efficiency for delivering resources to consumers, by pooling compute (CPU and Memory), networks, and storage, and presenting them through catalog-based services. This allows Cloud Director tenants to consume infrastructure without the operational overhead of having to manually configure and provision hardware resources. It is these resources that are the fundamental building blocks for service providers to consider when designing a solution. By employing Cloud Director, these resources can be grouped into multiple tiers, based on their specification, capacity, or performance capabilities.

Cloud Director has evolved significantly in recent years. In earlier releases, Cloud Director provided a solution that purely converted physical data center resources into elastic pools of compute, storage, and network resources. This was done through Cloud Director's ability to abstract physical resources into Virtual Data Center (VDC) resource pools, allowing the cloud provider to make them available to cloud consumers alongside catalog-based services via the Web Portal or public API. In addition, Cloud Director's policies allow the cloud provider to apply pre-determined limits on their consumers

to control resource consumption and limit access through a comprehensive tenant-facing role base access control (RBAC) framework.

However, Cloud Director has now become far more than just the engine to deliver locally controlled compute, storage, and network resources. VCD is now able to provide the platform for service delivery, which goes far beyond its traditional IaaS capabilities. As you will see as you progress through the chapters of this book, VCD provides the framework to offer Enterprise Public Cloud services across both VMware and third-party solutions. This might include delivering secure multi-tenanted Object Storage, self-service data protection solutions or Managed Kubernetes, all fully integrated into the user experience. Whatever the service the provider is looking to deliver, through VCD extensibility and integration with both VMware solutions and those of independent software vendors (ISV), the cloud provider now has flexibility and control over the customer's user experience to align their service offerings with the demands of their customer base. This means that making differentiation of service delivery from cloud providers around the globe is one of the key pillars of the Cloud Provider Platform.

Over and above vSphere, Cloud Director provides:

- A multi-tenant self-service user interface:
 - Easy Web and API-based deploy and operate models
 - Designed for the Non-IT User
 - Multi-tenanted constructs to strictly isolate customers from one another
 - Extensible tenant user interface to integrate custom services
 - Integrated vRealize Orchestrator workflow delivery
- Central libraries of templates, media and marketplace software:
 - Simple access to all the resources administrators and developers need
 - Personalization of templates and media
 - Ability to deploy multiple identical copies of applications
- Gives tenant admins control over users and developers:
 - Control user resource usage through roles/rights, quotas, and leases
 - Leverage multiple authentication sources, including Active Directory, LDAP, SAML 2.0, OAuth & Active Directory Federation Services (AD FS), so that Cloud Director tenants can federate their on-premises identity source with Cloud Director, to simplify consumer management of their Cloud Organization
- Metering for billing purposes via vRealize Operations and Tenant App through tenant dashboards and reports:
 - Data collection for billing on resource usage
 - Provide tenant-facing dashboards from vRealize Operations
- User abstraction from underlying hardware:
 - Ability to scale platform without impacting consumers or how they interact with cloud resources
 - Consumers don't know what compute and storage resources are being used on the underlying platform
 - Resource allocation decisions made by the platform, not by the consumer

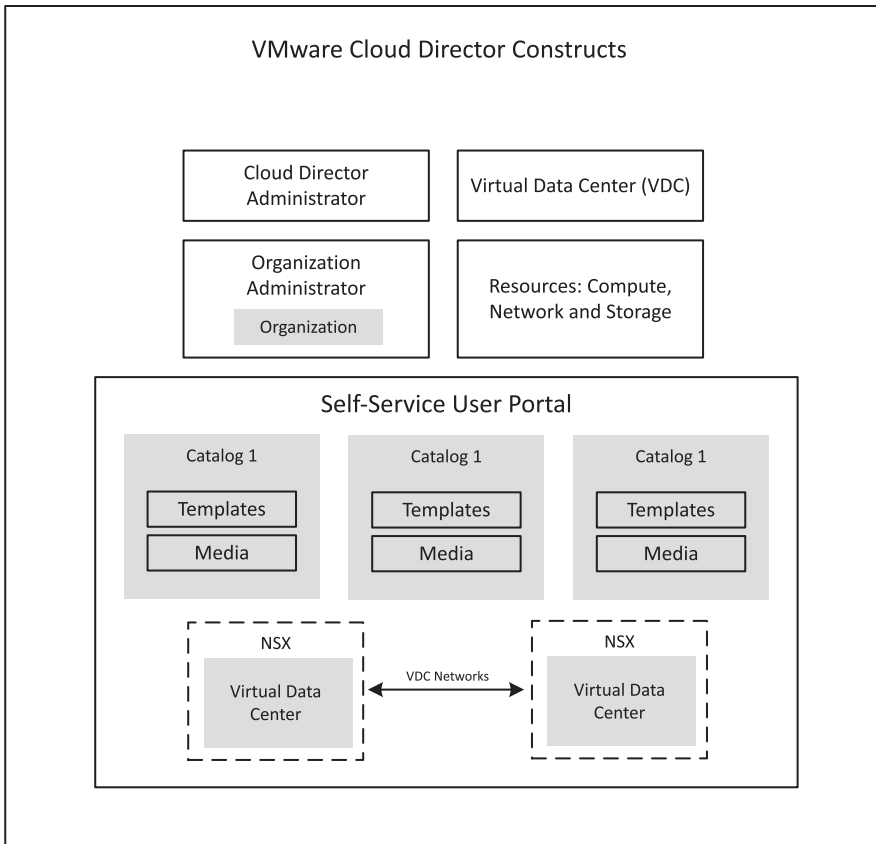


FIGURE 3.1 Key Pillars of the Cloud Provider Platform

The key construct to Cloud Director's end-to-end secure multi-tenancy model, as illustrated in Figure 3.1, is the Organization. The Organization collects together users and groups, defined by the cloud administrator, so that resources can be assigned in the form of Organizational Virtual Data Centers (Org. VDCs). The Org. VDC provides the pool of resources to the Organization, in which workloads can be deployed. Increased flexibility is provided as Organizations can be assigned multiple VDCs from resources backed by different physical characteristics. While Cloud Director operational administrators are responsible for creating an Organization and assigning VDC resources to it, the individual Organization admins can add, remove, and assign rights to users in the Organizations under their control in a self-service manner. Also, there is no limit to the number of cloud consumers that can be assigned access to the resources within the organization.

Cloud consumers, with appropriate access rights, are granted access to a self-service user interface and/or API, which can be used to interact with the Organization's compute, storage, and network resources. It is from this user interface, or provided public APIs, that workloads, Kubernetes clusters,

or other resources can be deployed and configured, either manually or from pre-configured catalog or marketplace items.

Cloud Provider Use Cases

Cloud Director is focused on Cloud Provider use cases, although NFV also provides input for telco-based subscribers. As we have seen, Cloud Director allows the cloud providers to abstract VDC resources and make them available to consumers. This includes the ability to assign advanced network resources and features to virtual machines and Kubernetes clusters in a self-service manner, rather than having to ask the cloud provider to create a new network or firewall rule on the consumers behalf. This facilitates much faster application provisioning and allows organization admins to upload and maintain their own applications and infrastructure, without intervention from the cloud provider. Therefore, as Cloud Director removes the day-to-day need for the cloud provider, it is well suited to environments where multiple consumers or developers wish to maintain greater control over their platform.

The other key factors in Cloud Director’s growing popularity is the native integration with NSX network and security services and the flexibility and extensibility offered by the new HTML5 user interface. Cloud customers are able to consume NSX services in a self-service manor, on a feature-by-feature basis. In addition, in previous releases, customization of the Flex user interface was not typically an option, meaning there was no easy way for cloud service providers to differentiate through either the look and feel of the user experience or through custom extensibility or integration with third-party software solutions, without building their own user interface. However, this changed with the introduction of the new HTML5 user interface. In addition, closer vRealize Orchestrator integration further extends the option to build simple custom services and publish those services to cloud consumers—in essence, further enhancing the customization capabilities of VCD to become a XaaS delivery platform.

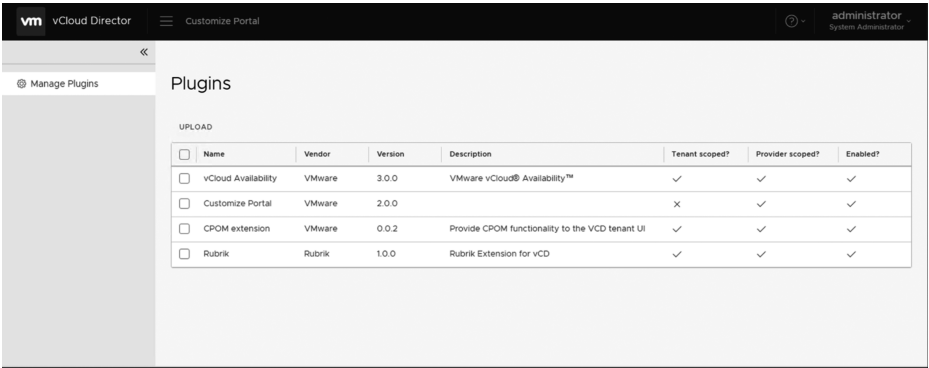


FIGURE 3.2 Extension Plug-in Integration in Cloud Director

Hosted Private Cloud with Cloud Director

It is also important to highlight that Cloud Director is also well suited for private cloud use cases, where an organization with multiple company divisions wishes to sub-divide resources and provide isolation between business units. In this model, Cloud Director is used to provide pools of isolated resources, potentially all managed but their own teams, made accessible through the HTML5 user interface or API. This solution allows the organization to behave like a cloud service provider, partitioning the private cloud into multiple sub-organizations, all with their own VDC resources. This model is often used when the separation of resources and administration is beneficial to the lines of business, from either a political or compliance perspective. In addition, this delivery model for Cloud Director can also have significant benefits from a networking and security standpoint, as well as provide the ability to the customer to cross-charge the various lines of business through accurate billing for the resources they are consuming, increasing accountability and governance over resource consumption.

CLOUD DIRECTOR'S CORE COMPONENTS

As illustrated in Figure 3.3, VMware Cloud Director is linked to one or more vCenter Server appliance, an NSX Manager and an arbitrary number of VMware ESXi hosts. The VCD cell server group (a group of servers that share a common Cloud Director database) and its embedded database manage access to vCenter resources from cloud consumers. Figure 3.3 is a schematic representation of a simple Cloud Provider Platform. All cells in the server group share a common active database. The entire cell group is connected downstream to three vCenter server appliances and the VMware ESXi hosts they manage. Each vCenter server appliance is connected to an NSX Manager, which provides network services to the cloud platform. The initial VCD installation and configuration process establishes a set of connections to each of these components. This architecture can be scaled out or up at any time.

As illustrated in Figure 3.4, VMware Cloud Director abstracts underlying virtual infrastructure and creates consumable resources through secure multi-tenanted constructs. It does this through creating a number of objects, each with attributes allowing different levels of interaction from the cloud provider and cloud consumer.

At a high level, VMware Cloud Director segregates the resources from vSphere to create a common pool called the Provider Virtual Data Center (Provider VDC). This is done in most cases by mapping the vSphere clusters to a Provider VDC, with all resources available in the vSphere cluster being allocated to the Provider VDC.

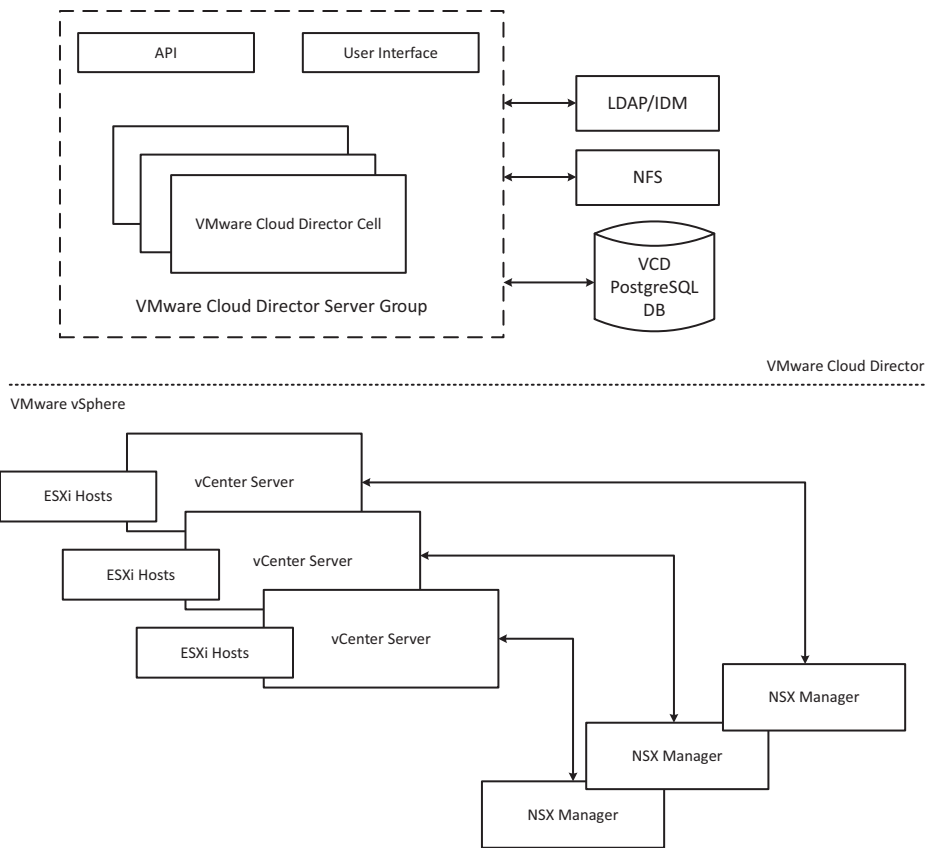


FIGURE 3.3 VMware Cloud Director Architecture

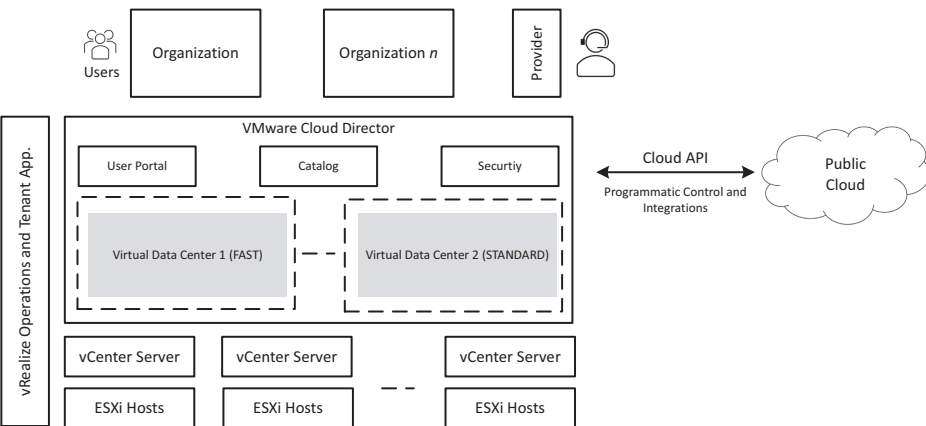


FIGURE 3.4 Cloud Director Constructs

Note that it is also possible to map a vSphere Resource Pool to a Provider VDC, which is the mechanism used by Cloud Director Service consuming resources from VMware Cloud on AWS, addressed in Chapter 8. This model might also be used in small shared vCenter Server environments or in testing scenarios.

The Provider VDC is an abstraction layer from which resources are delivered and allocated as individual units of compute and storage. This individual unit of compute and storage resource is called the Org. VDC and allocated based on a predetermined allocation model. In addition to the compute and storage resources, the Org. VDC consumes network resources from NSX.

The Org. VDC is a container for all workloads, virtual machines, or Kubernetes clusters that are being used in the cloud by a group of consumers within an Organization. The cloud provider then allocates Org. VDC's to its end consumers, ready for consumption.

Platform Constructs

Each of these object types, or constructs, outlined above is covered in more detail below. For more information on these elements and their installation and configuration, please refer to the VCD product documentation.

Provider Constructs: Provider constructs are owned and managed by the cloud provider, and in most cases, these are abstracted away from the customer experience. However, some constructs, such as the Organization, overlaps both the tenant and provider, but is ultimately owned by the provider.

Provider Virtual Data Center: The Provider Virtual Data Center represents the grouping of compute, memory, storage and network resources from a single vCenter Server. A provider virtual data center can be composed of one or more Resource Groups with VMware vSAN datastores and/or external array-based datastores and external networks. The Provider Virtual Data Center can share these resources across multiple organizations.

The fundamental principal that Cloud Director operates to create cloud services is by the pooling together of all the resources of a data center (Provider VDC) and then allocating them to the customer organizations as needed. Because of this abstraction layer, VCD is able to source resources from its pool for customers when needed and put back resources in the pool when they are no longer required. Also, if more hosts, network or storage is added to a cluster, the Provider VDC picks up the additional resources and expands its pool. This model helps cloud providers to better manage and operate their cloud infrastructure, as well as provide a strong service platform to its customers.

Organization: Multi-Tenancy is one of the intrinsic characteristics of the Cloud Provider Platform. As we have seen, this is delivered through Cloud Director, which has specific constructs built around this key characteristic to create isolated compute, storage, and network pools that can be mapped to individual tenants of the cloud. Organizations are the unit of multi-tenancy within Cloud Director and represent a single logical security boundary. Effectively, the Organization in VCD represents the container in which all of a tenant's resources reside. A Cloud Director Organization maps to an end customer or tenant. Organizations can use local Cloud Director accounts, direct LDAP integration, or integrate with a SAML2.0 compatible identity provider, allowing for flexible identity management, including multi-factor authentication.

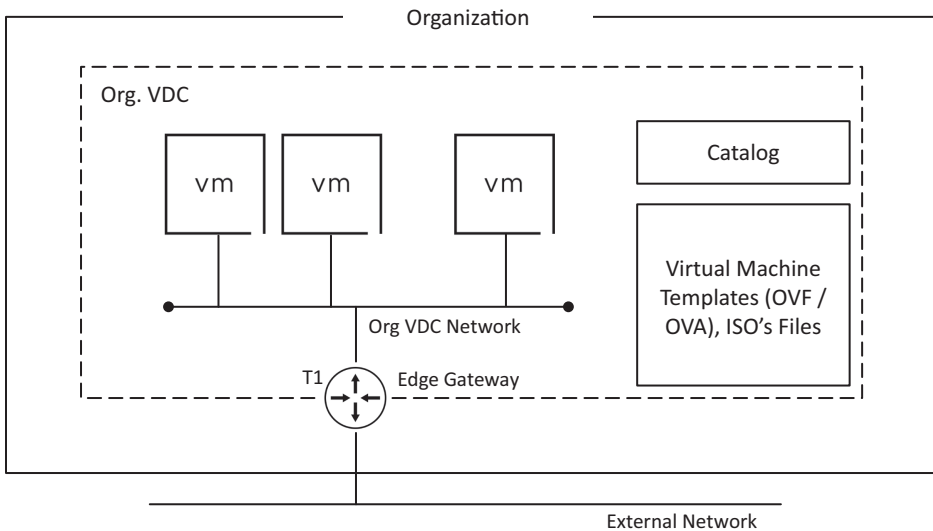


FIGURE 3.5 Organization Construct Overview

As illustrated in Figure 3.5, each organization is provided with units of resources for compute, storage, and network, which are defined by the Org. VDC. When an Org. VDC is created, a Provider VDC is used to allocate resources to the Org. VDC. The Provider VDC maps these resources into different containers in vSphere, which isolates these units of resource. For instance, the compute resources CPU and memory are defined by creating a resource pool in vSphere and storage is mapped by assigning storage profiles, which in turn map to the vSAN datastore. The Org. Networks are derived from network pools and map to port groups on the N-VDS (NSX distributed Switch) that is offered as part of the provider cluster.

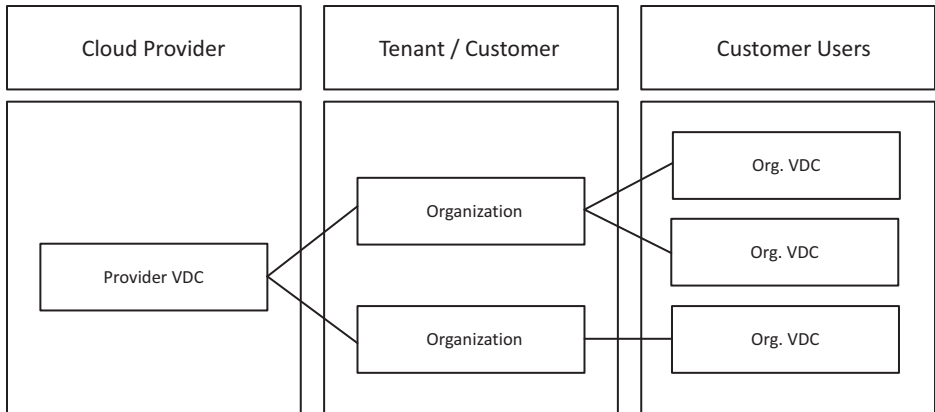


FIGURE 3.6 Organization to Org. VDC Mappings

Support Management Organization: It is common for a cloud provider to create an administrative Support Organization first. This organization owns a master catalog of templates and media that can be published and shared with some or all other customer organizations.

Operational teams are exclusively assigned to the Support Organization to take responsibility for creating official provider-based templates for placement in the master catalog, for other organizations to use. Templates in development should be stored in a separate development catalog, which is not shared with customer organizations.

Note that by default a System Organization environment is created when deploying Cloud Director. The Support Organization being created for the purposes outlined above is different from the built-in System Organization. Be sure to confirm that when you create the Support Organization you configure it to *allow publishing of catalogs*.

Provider Catalogs, vAPPs, and Virtual Machines: The use of catalogs gives cloud providers the ability to easily provide standardized application service delivery capabilities to their customers. This mechanism allows cloud providers to maintain standard builds, which are typically automated operating system installations that give customers a consistent configuration and may, in managed use cases, deliver security and operational policies by including management tools, such as backup agents, antivirus tools, and monitoring agents.

Through the use of public catalogs, cloud providers can enable customers to deliver PaaS, middleware or basic servers instances using scripts and automation as part of a standard build. The use of catalogs can also lead to significantly Faster Provisioning in the cloud by allowing users to use pre-built templates to provision virtual machines and containers, reducing the time to install and configure operating systems and applications. Guest customization

and scripts help pass configuration parameters within the guest operating system, providing users with granular control of their workloads.

The sharing of catalogs is also a key feature that allows a cloud provider to publish catalogs either within a single organization or across all customer organizations. This allows cloud providers to deliver a wide library of pre-built solutions, while also allowing customers to create and upload templates for their private use. As outlined above, this is typically achieved by cloud providers maintaining a Support Organization, from which they publish catalogs to all or a subset of tenants, as illustrated in Figure 3.7.

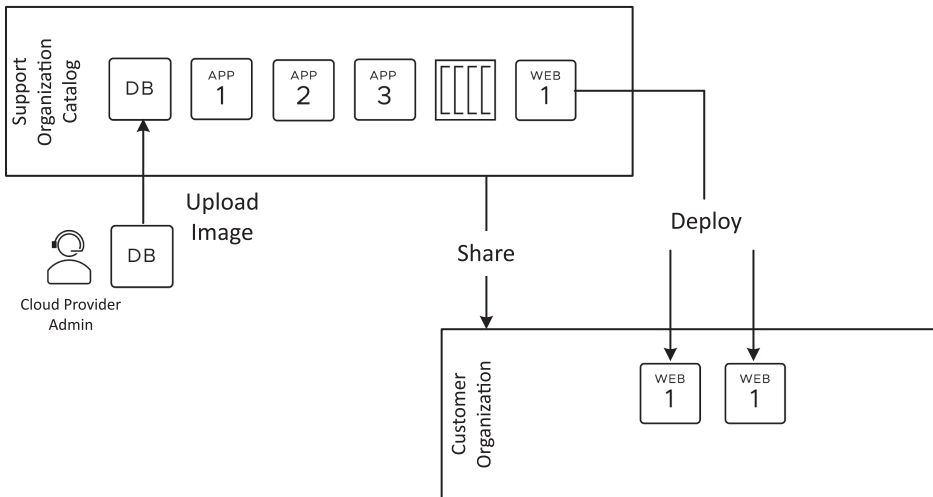


FIGURE 3.7 Support Organization Catalog Publishing

VMware Cloud Director provides two mechanisms for distributing catalogs across tenant organizations.

- **Publishing:** This is where the catalog is presented to Organizations, with no granularity.
- **Sharing:** When a catalog is shared, access can be assigned to specific users and groups within the publisher's organization.

Through these mechanisms, cloud providers can maintain catalogs of base operating systems, similar to traditional vSphere templates. Providing base operating systems templates gives customers a way to maintain standard builds, without the need to boot a blank virtual machine from a virtual disk or PXE network server. This allows the cloud provider to centrally maintain an appropriate patch level of their base images. Operating system templates such as these might typically include:

- Windows 2019
- CentOS 7

- Red Hat Enterprise Linux (RHEL) 7
- Ubuntu

Whatever the list of templates available, a cloud provider should ensure to maintain a uniform naming convention, identifying the product, version, and edition. For instance, **Windows_2019_Std_x64** or **RHEL_8_x64**.

In addition to base operating systems, cloud providers can also use this same model to add application services to public catalogs and the required metadata to create additional add-on XaaS offerings for their customers.

Alternatively, the catalog interface also allows customers to upload their own operating system and application templates or application installation ISO media. However, this approach is not a common operating model for most customers as it is far easier to consume prebuilt templates from the catalog or to leverage the VMware Marketplace and Bitnami Catalog, which will be addressed in more detail in Chapter 8.

Another important concept in Cloud Director is the vApp. The vApp is a logical construct that groups one or more virtual machines and networks, into an application blueprint, which includes CPU and memory allocations, disks, networks cards and details of any connections to organization networks. The vApp construct also allows for power actions and startup orders to be pre-defined, with their required resource allocations. For instance, when deploying a self-contained LAMP stack (Linux, Apache, MySQL, PHP), as shown in Figure 3.8, a LAMP vApp connected to an Org. Network, with an NSX Edge firewall connecting it to an external network.

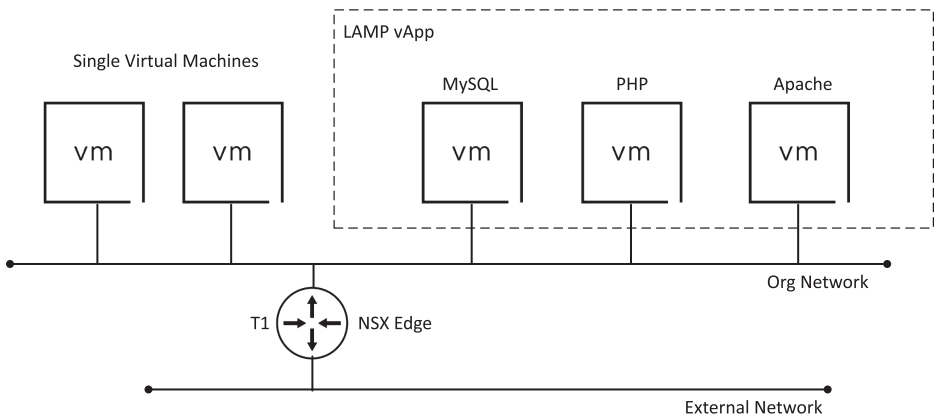


FIGURE 3.8 Self-contained LAMP Stack

As illustrated in Figure 3.8 vApps and virtual machines must be consumed in conjunction with each other and can co-exist in any application architecture that meets the solution goals. They are not exclusive entities.

Fast Provisioning (Linked Clone): Cloud Director can, if required, provision workloads quickly using linked clone technology. Linked clones utilize the vSphere redo-log linked clone implementation to provide fast virtual machine provisioning. Compared with a full clone, linked clones improve the agility on the Cloud Provider Platform by reducing provisioning time and providing near-instant provisioning of virtual machines in the customer's cloud environment.

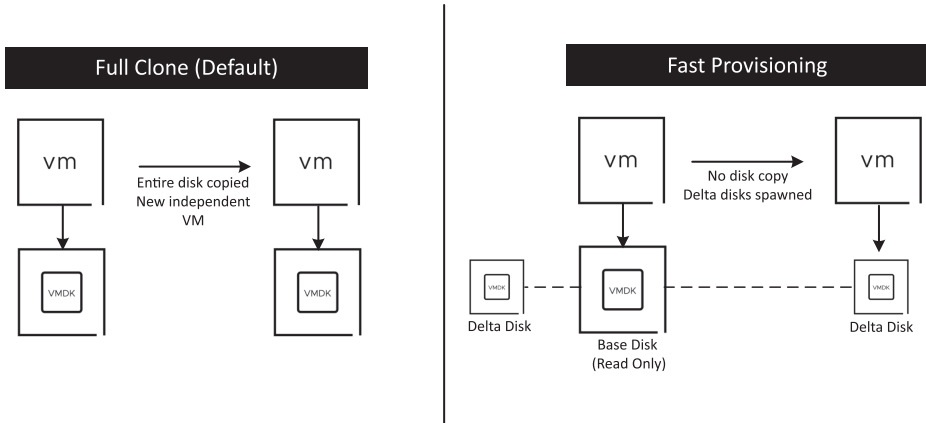


FIGURE 3.9 Linked Clone and Full Clone

As illustrated in Figure 3.9, when a cloud customer provisions a virtual machine using a full clone, the entire virtual disk is replicated, then a new independent virtual machine is created. Typically, with a full clone writes go to the VMDK and reads come from the same VMDK.

However, when a linked clone virtual machine is provisioned, a small 16MB VMDK is created. This VMDK is an empty delta disk that serves to capture disk writes for the newly created virtual machine. This takes a very short amount of time to create and only consumes a very small amount of disk space. Writes to disk from virtual machine then go to this delta disk, which grows to accommodate the writes. Disk read operations traverse up the linked clone chain until the desired block is found.

For workloads that are not generating I/O intensive operations, linked clones offer the flexibility and agility of instant provisioning. However, for workloads that generate I/O-intensive operations, customers should consider using full clones instead of linked clones. Keep in mind the additional I/O processing required by delta disks when using linked clones.

Cloud Director Cells: The Cloud Director cells provide the Cloud Management Platform components, all provided in either a stateless Photon OS-based virtual appliance or by installing the Cloud Director software on appropriately supported Linux servers.

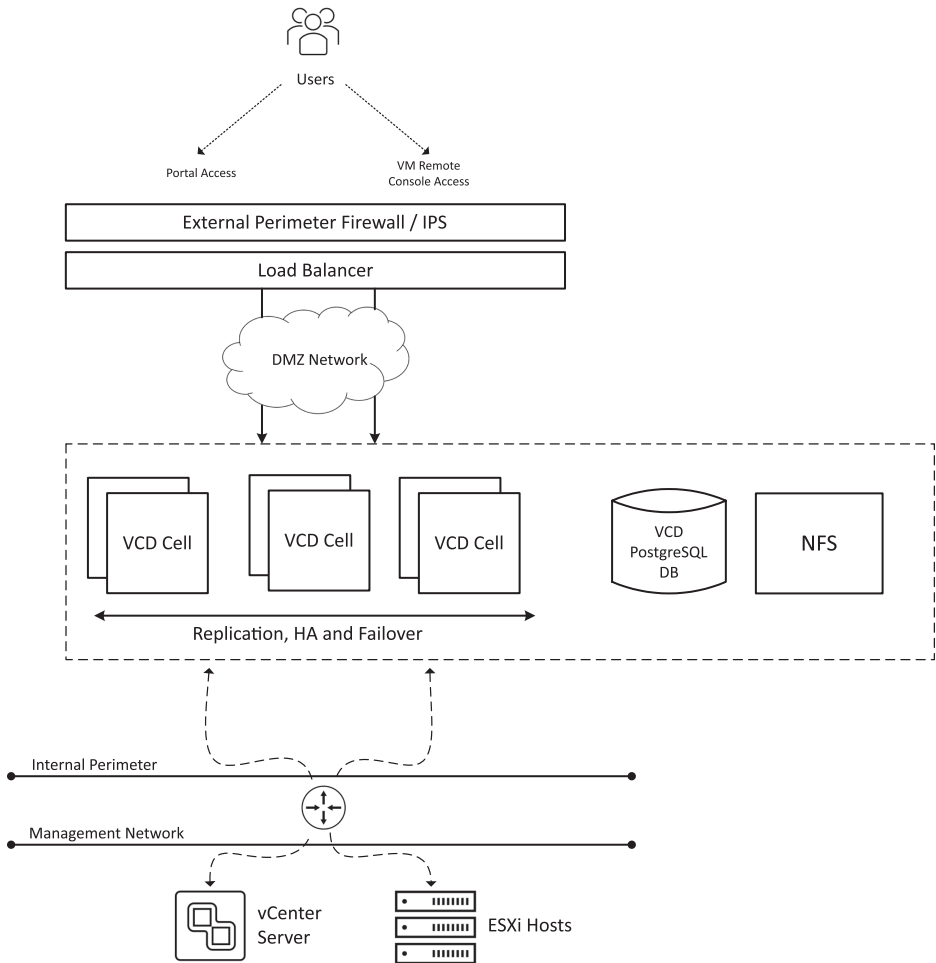


FIGURE 3.10 Cloud Director Cell Architecture

The cells provide the functionality for higher-order resource management and structure. There is some caching at the Cloud Director cell level, such as SSL session data, but all refreshes and updates are done against information stored in the PostgreSQL database.

As illustrated in Figure 3.10, cells are the main entry point of customer requests through the user interface or API layer. All cells interact with vCenter Server as well as directly with ESXi hosts through VIM API calls. One cell has an open proxy connection with the vCenter Server listening for updates at all times.

There are typically multiple Cloud Director cells deployed for load balancing and high availability reasons. The recommended minimum number is two, or $n+1$ where n is the number of Resource Group vCenter Servers. This formula is based on the number of vCenter Servers, providing sufficient resources to

allow for cell failover and maintain resources for cell maintenance or failure, for which we recommend having a spare cell. Using this methodology to size the cells allows for a level of high availability for the Server Group in case if a cell failure or for routine maintenance.

Cloud Director scalability can be achieved by adding more cells to the system. If VMware Remote Console or OVF transfer tools are extensively used by the customer users, this could impact Cloud Director cell utilization and the deployment of new instances should be considered.

A Cloud Director Server Group is the term used for one or more Cloud Director cells deployed as an appliance or through the Linux installable that share a single database, a single NFS transfer storage location and connect to the same underlying vSphere and network resources.

Cloud Director Database: Cloud Director maintains a database of all the resources and inventory items from all attached vCenter Servers by synchronizing with vSphere inventories across the entire platform. As we have seen, the Cloud Director platform has transitioned to a standard VMware appliance deployment, which offers multiple benefits to cloud providers, such as simplified deployment, scaling, and lifecycle management. As part of this evolution, Cloud Director has been required to support an open-source database, PostgreSQL. For the Linux installable deployment model, an external PostgreSQL database is supported.

As illustrated in Figure 3.11, the Cloud Director appliance uses an embedded PostgreSQL database.

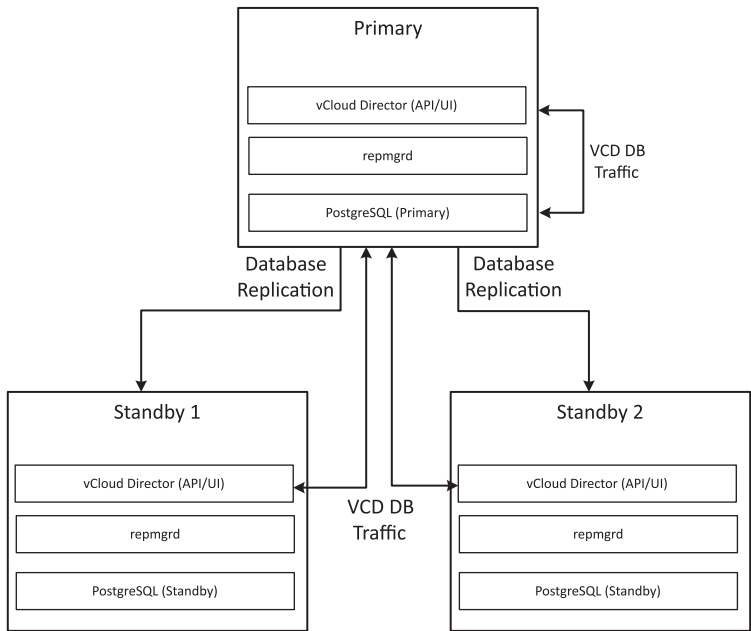


FIGURE 3.11 Cloud Director Appliance Database HA Cluster

The Cloud Director appliance embedded PostgreSQL database provides a high availability function to a cluster of PostgreSQL servers, meaning you can create an appliance deployment with a database HA cluster that provides failover capabilities to the Cloud Director database. There is always only one active instance of the Cloud Director database per Cloud Director Server Group, with database failover being managed through the Cloud Director Appliance Management Interface, as shown in Figure 3.12.

vCloud Director Appliance Management

Navigation
vCD Database Availability

vCD Database Availability

Name	Role	Status	Following
bos1-vcd-static-211-122	standby	running	bos1-vcd-static-211-120
	PROMOTE		
bos1-vcd-static-211-120	primary	* running	
bos1-vcd-static-211-121	standby	running	bos1-vcd-static-211-120
	PROMOTE		

FIGURE 3.12 Cloud Director Management Portal

Database utilization, size, and storage I/O requirements should be monitored, and more resources should be provisioned with the growth of the Cloud Director Server Group.

Cassandra Database for Storing Historic Metric Data: VMware Cloud Director can optionally collect a defined set of metrics, which can be used by customers to view real-time and historical performance information about workloads running in their Org. VDCs for up to two weeks. The data for the historical performance metrics is stored in a Cassandra database cluster.

For historic metric data, Cassandra provides the open-source database that is required for the backing store for a scalable, high-performance solution for collecting time series data, such as virtual machine metrics. For a Cloud Provider Platform design that requires this functionality, cloud providers must install and configure a Cassandra cluster, and use the cell-management-tool to connect Cloud Director to the cluster instance. As illustrated in Figure 3.13, retrieval of real-time metrics does not require the Cassandra database, only the retrieval historic data.

This solution for providing tenant-facing metrics on running workloads is separate from that provided by the vRealize Operations Management Pack for

Cloud Director and Tenant App, which provides a far deeper insight into customer workloads and performance metrics. The use of the Cassandra database for storing historic metric data can be considered an entry-level service, which provides basic performance insight into the consumption behavior of customer workloads, specifically relating to the consumption metrics shown in Table 3.1.

TABLE 3.1 Cassandra Metric Data

Metric Name	Type	Unit	Description
<code>cpu.usage.average</code>	Rate	Percent	Host view of average actively used CPU as a percentage of total available
<code>cpu.usagemhz.average</code>	Rate	Megahertz	Host view of actively used CPU as a raw measurement
<code>cpu.usage.maximum</code>	Rate	Percent	Host view of maximum actively used CPU as a percentage of total available
<code>mem.usage.average</code>	Absolute	Percent	Usage as percentage of total configured or available memory
<code>disk.provisioned.latest</code>	Absolute	Kilobytes	Storage space potentially used
<code>disk.used.latest</code>	Absolute	Kilobytes	Storage space actually used
<code>disk.read.average</code>	Rate	Kilobytes per second	Read rate aggregated across all datastores
<code>disk.write.average</code>	Rate	Kilobytes per second	Write rate aggregated across all datastores

Figure 3.13 illustrates the recommended metric database design. Multiple Cassandra nodes (a minimum of 3) are deployed in the same network for high availability.

As highlighted above, the cloud provider can also implement a more advanced tenant-facing performance monitoring solution for their customers by using the vRealize Operations Tenant App for Cloud Director, which provides a tenant administrator visibility into their Cloud Director environment. This solution is described in more detail in Chapter 8.

Cloud Director Transfer Storage: When multiple Cloud Director cells are used, a shared temporary storage repository for uploads, downloads, and cloning operations across different vCenter Servers has to be available to all Cloud

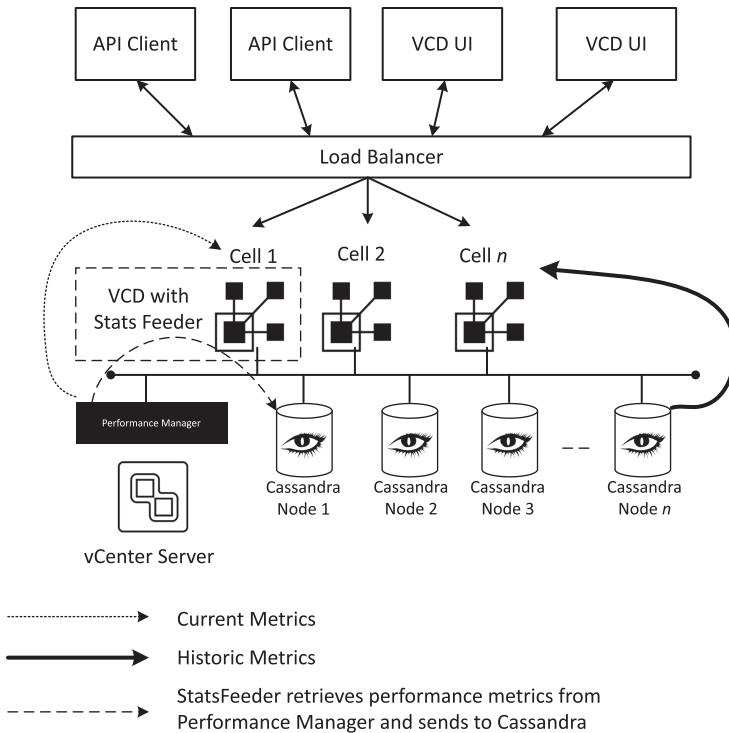


FIGURE 3.13 Virtual Machine Metric Database Design

Director cells, as illustrated in Figure 3.10. Once the transfer operation is completed the data is deleted from the transfer storage. The size of the NFS share is dictated by the number of concurrent uploads/downloads, and their average size. For example, a minimum of 500GB should be considered as the starting point with the cloud provider considering customer requirements for simultaneous user uploads, for purposes of size forecasting. In a production environment, the transfer storage should also be offered using a highly available NFS solution.

User Interface: Cloud Director offers a cloud delivery model that helps cloud service providers delegate most of the day-to-day IT operations to its consumers. The aim of the HTML5 user interface and API is to give customers more flexibility and control to operate their cloud environments.

Cloud Director provides two portals: a Tenant Portal and Service Provider Admin Portal. These can be accessed using standard URLs:

Tenant Portal: <https://cloud.cloudprovider.com/tenant/ACME>

Service Provider Admin Portal: <https://cloudprovider.com/provider>

Cloud providers can use the service provider admin portal to manage and maintain cloud resources that are made available to customers. This includes, but is not limited to, the management of Provider VDCs, external networks,

network pools, Organizations, and Org. VDCs. The cloud provider can also choose to configure the firewall to restrict access to the */provider* URL space, without impacting tenants, in order to restrict access to the provider admin portal via the public internet.

While the customer can go directly to their specific tenant URL for login, shown above, Cloud Director also offers a universal tenant login page (*/login*), shown below. This greatly simplifies the login experience for non-IT users and also simplifies global load balancing and DNS configuration for multi-regional implementations.

<https://cloudprovider.com/login>

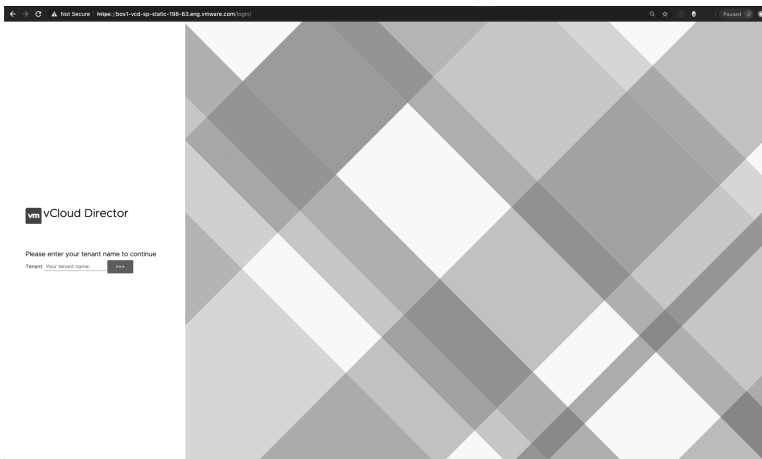


FIGURE 3.14 Default Organization Login Page

It goes without saying that the core purpose of the user interface is to provide a simple and meaningful way in which both cloud consumers and cloud providers can manage and operate the platform. A key benefit of the Cloud Director user interface is extensibility. Through the extensible HTML5 user interface, vRealize Orchestrator service delivery, multi-site capabilities, integrated tenant-facing vRealize Operations dashboards, Native NSX services, and third-party extensions all come together to make it the most flexible and comprehensive cloud management platform on the market. These, alongside a whole range of additional new features, give VMware cloud providers all the tools required to deliver end-to-end cloud services.

In addition, the enhanced vRealize Orchestrator integration further extends the options to build simple custom service integration, publishing those services to consumers, and monetizing as appropriate—in essence, further enhancing the customization capabilities of VCD to provide XaaS.

Securing Cloud Director on the Internet: A Cloud Director solution will consist of either multiple VCD appliances or, as illustrated in Figure 3.15, Linux

deployed instances within a server group. These two deployment mechanisms cannot co-exist within the same server group as the Cloud Director Linux installed software requires an external database, as opposed to the embedded PostgreSQL database provided with the appliance. As we have seen, all Cloud Director cells within the server group connect to a common database and NFS transfer server storage.

As with any public-facing application, securing the Cloud Director cloud management web user interface and API from external attacks is a key design factor for cloud providers. As with securing any application, a tiered and multi-technology approach is typically put in place, this might include, but is not limited to, IPS/IDS appliances, Web Application Firewalls, the NSX Distributed Firewall, and secured internal VCD cells, for internal access to the platform.

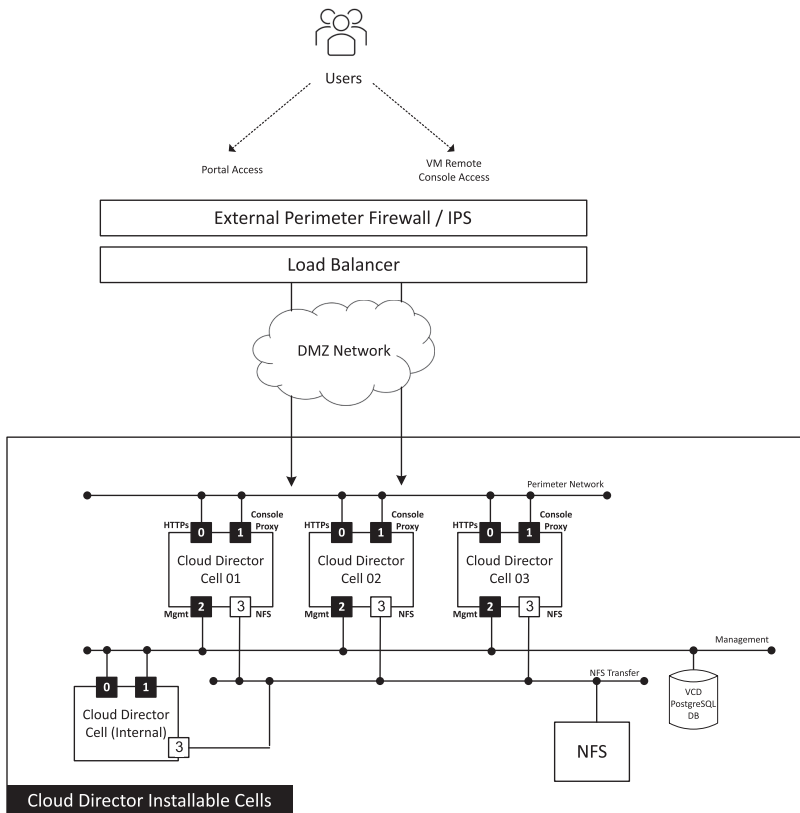


FIGURE 3.15 Securing Cloud Director

Deploying specific VCD cells on a private internal network can not only deliver a simple operational mechanism for allowing cloud provider administrators access to the environment securely over a private network, it can also help provide access to the cloud platform, should the public-facing cells be

subject to a DDOS attack. In addition, public-facing cells can be configured at the Layer 7 firewall level (as well as API) not to allow access to the */Provider* user interface. Making the internal cells the only mechanism for accessing the internal operational portal.

Strict process management and regularly audited RBAC is also key for ensuring a robust and secure platform. Cloud providers can also choose to leverage the NSX Distributed Firewall to segment and control traffic flows within the perimeter networks, preventing an attacker from hopping from one workload to another within the DMZ.

Finally, it is also possible to force the external database to only accept TLS connections, importing the database certificate into the Cloud Director cells, to ensure that communication between the VCD cells is always encrypted and that the database is unable to communicate over a plain text connection.

Cloud Director Remote Console: VMware Cloud Director provides a virtual machine console feature to allow access, from within the tenant and provider portals, to open and present the console of the VM in question. Using this feature, known as the VMware Remote Console (VMRC), the user or the administrator of the service is able to manage and control the virtual machine guest operating system on day 0 as well as troubleshoot remote connectivity, should it be required.

The Cloud Director HTML5 tenant and provider portals provides support for the VMRC from modern browsers and also provides support for the Standalone VMRC, which includes more features, such as broader keyboard input types and the management of OVA/OVF images. The use of the Standalone Remote Console significantly improves access to the virtual machine's console and gives greater flexibility over accessing customer workloads.

The VMRC communicates with ESXi servers using a custom protocol, but the Remote Console Proxy on the Cloud Director cells allows for the entire remote console communication to be tunneled over the HTTPS protocol.

As illustrated in Figure 3.16, the Remote Console Proxy runs as a process on the Cloud Director cells and communicates to the vCenter server on port 443 and to the ESXi hosts on ports 902 and 903. The VMware Remote HTML5 Console or Standalone VMRC securely communicates with the Remote Console Proxy on port 443 only.

The Remote Console only operates on powered on virtual machines, which is enforced by Cloud Director. However, the Remote Console does allow the opening of a console of a powered off VM but will by default automatically power it on.

As illustrated in Figure 3.16 in order to load balance Remote Console Traffic between Cloud Director cells and provide an additional level of availability, a load balancer can be used for console sessions in a similar way it is used to distribute user interface and API traffic.

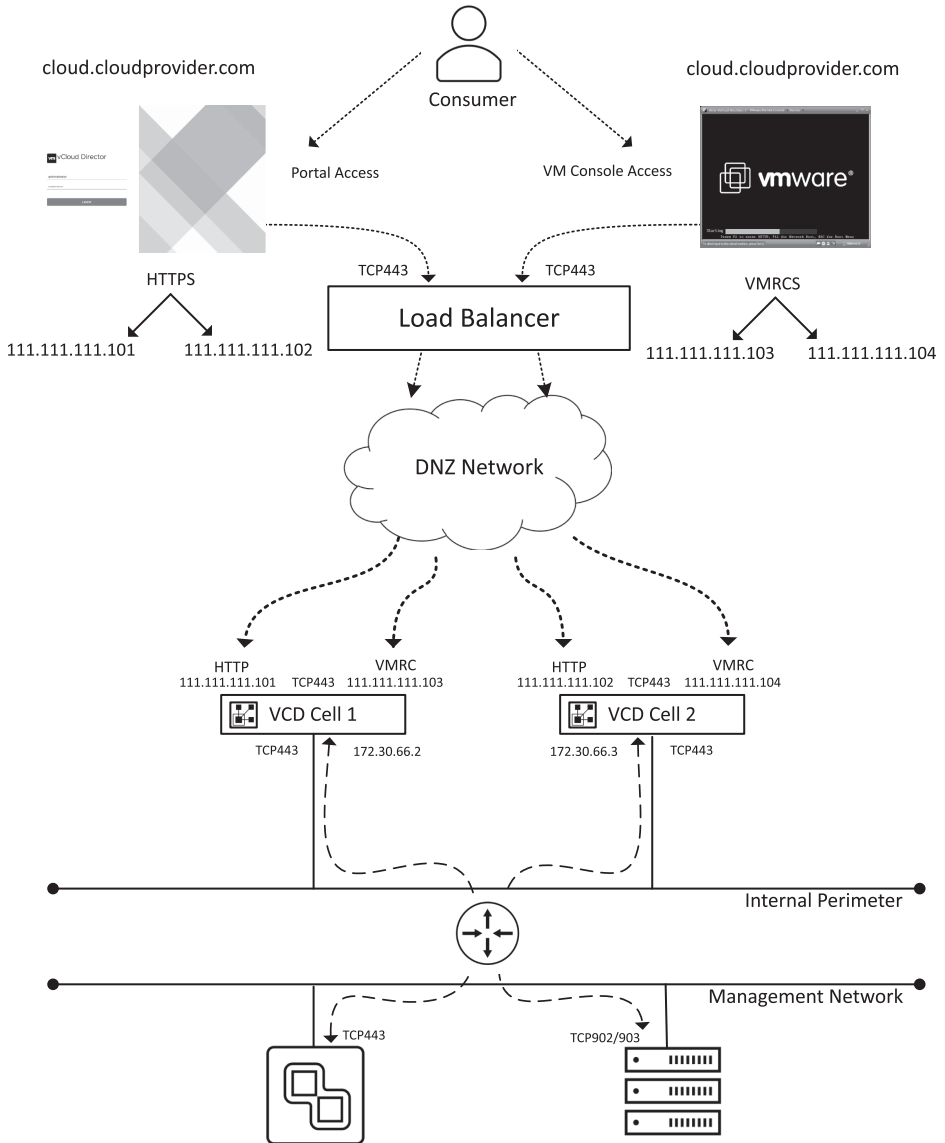


FIGURE 3.16 Remote Console

Cloud API: Cloud Director has an extensive set of RESTful APIs that can be used to perform operations on the cloud programmatically. This Cloud Director API provides an interface for managing resources and is the cornerstone to federation and ecosystem support. All federated tools talk to the Cloud Director environment through the API. For that reason, it is important that a Cloud Provider Platform environment expose the Cloud API to end users.

Cloud Director actually provides two different API surfaces, the legacy vCloud API and the new OpenAPI. Cloud Director's OpenAPI is a new Cloud Director API that is defined using the OpenAPI standards. This RESTful API borrows some elements of the legacy vCloud API and also establishes new patterns. All new features added to Cloud Director consume the new OpenAPI standard, such as user interface branding, CPOM, and NSX Data Center, which also provides a Swagger user interface for intuitive guidance.

Both of these APIs are accessible to REST Clients over HTTPS with each object and component in Cloud Director having a corresponding unique ID, by which it can be referenced. The state and properties of each component is stored in the form of XML Elements.

The Cloud Director API's can be used to run any operation on the Cloud Director Server Group, without having to go through the user interface and can be plugged into automation frameworks that execute operations on the cloud for either consumer or provider. For instance, when a new tenant is registered, a RESTful API call to create the Organization in VCD can be automatically plugged into an invoice or billing system. The Cloud Director REST API queries can also be used to generate infrastructure lists, such as listing all the Organizations belonging to a Provider VDC, or listing all Catalogs in an Organization, to further simplify the onboarding of the new customer.

The VMware Cloud Director API also provides support for developers who are building interactive clients of VCD using a RESTful application development style. For instance, the vRealize Operations Tenant App and Cloud Director servers can communicate over HTTPS, exchanging representations of vCloud objects within the Cloud Provider Platform to generate tenant-facing dashboards in the user interface.

Figure 3.17 shows some of the connections of the Cloud Provider Platform. It shows all the API interactions between the various components.

Developers can use almost any language to automate VMware Cloud Director over the API. However, there are a few ready-made solutions that make the Cloud API easier to consume for developers, such as the Python SDK (`pyvcloud`). The open-source Python SDK is available to work with Cloud Director's RESTful API's and understand its Reference Schema, such as how to parse REST calls to create, retrieve, update, and delete operations. Other tools for developers include the `vcd-cli`, which is a command-line interface for Cloud Director and the Cloud Director Terraform provider.

For security purposes, a Web Application Firewall can be used to filter API access based on the scope provided to tenants over the public internet. However, the Cloud Provider Platform components would typically get unfiltered access.

Tenant Constructs: While most of the cloud provider constructs listed previously are abstracted away from consumers by Cloud Director, they are all key to virtual data center management. However, in addition to the cloud provider constructs, there are also a number of construct types that are

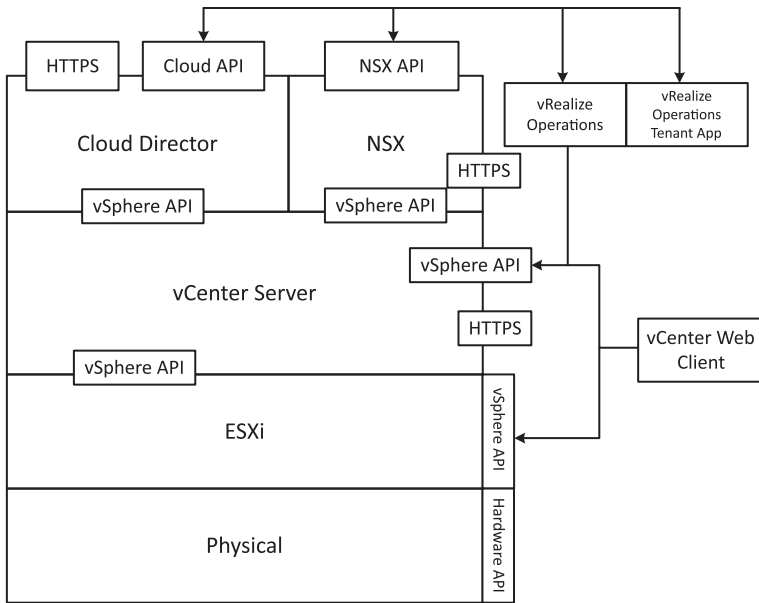


FIGURE 3.17 API Component Interactions

either managed by the tenant or in some cases managed *for* the tenant by the cloud provider. The following list of entities are, in most cases, managed by the tenant admin and users who have responsibility of management under the providers delivery model and service-level agreement.

Organization Virtual Data Center (Org. VDC): An Organization Virtual Data Center (Org. VDC) is a sub-grouping or pool of compute and storage resources allocated from a Provider Virtual Data center (Provider VDC) and assigned to a single organization. A virtual data center is a deployed environment where applications can be instantiated, deployed, and powered on.

Allocation Models: The allocation of resources to customer's Org. VDC's from the Provider VDC is governed by an 'Allocation Model.' Cloud Director has four different types of model by which it allocates resources to an Org. VDC. An Org. VDC essentially maps to a resource pool within vSphere, the first three models shown in Table 3.2 have policies that vary in terms of reservations set on these underlying resource pools in vSphere and define the upper limit of resources that an Org. VDC can consume. When a workload is created by a customer in their Org. VDC its resource allocation policies are derived from the corresponding resource pool the workload is placed in. Outlined below are the four different allocation models available for Cloud Providers to define in Cloud Director.

Allocation models are defined by limit and resource allocation; therefore, they typically dictate the basis by which a cloud provider can create service definitions for its customers, with a tenant being able to choose how they

TABLE 3.2 Cloud Director Allocation Model Resource Pooling

Allocation Models	Resource Pool Setting	Virtual Machine Setting
Allocation Pool	A % of Resource are Guaranteed, and a Max Resource Limit is set on the Resource Pool	Resource Guarantee and Limit is inherited from Resource Pool
Pay-As-You-Go	No Resource Guarantee or Limit set at the Reservation Pool	Resource Limits set at the workload level
Reservation Pool	Resource Guaranteed and Resource Limits are equal, all the resources are dedicated	No resource setting is defined on the workload; however, a user can change limits and reservation per workload.
Flex	Cloud administrators can create Org. VDC with a combination of flex allocation and VDC compute policies. Cloud administrators can control CPU and RAM consumption at both the VDC and the individual virtual machine levels. The flex allocation model supports all allocation configurations that are available in the other allocation models.	

consume resources based on their needs. For instance, a tenant who needs a fixed set of resources can choose an Org. VDC that has a predefined limit set, or the “Allocation Pool” model. A tenant might also choose a “Pay-As-You-Go” model, when they do not have empirical data on how much resource they will need to consume.

The Flex allocation model can be configured to behave equivalently to any of the legacy models: Allocation, Reservation, or Pay-as-you-Go.

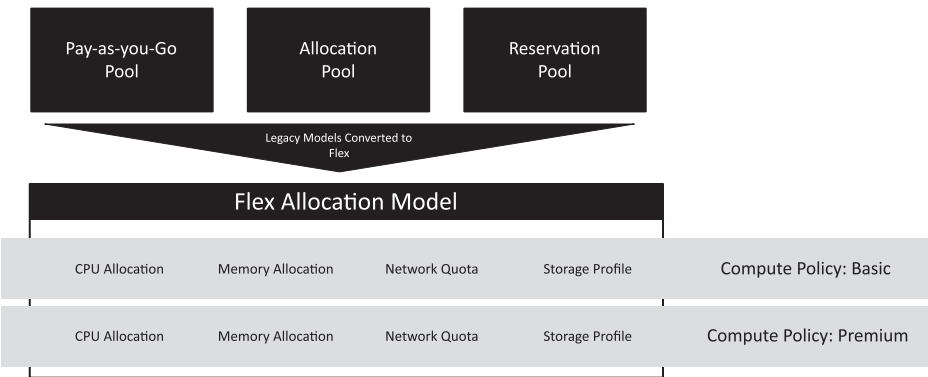


FIGURE 3.18 Flex Allocation Model

With the Flex allocation model CPU and RAM consumption can be defined at both at Org. VDC and individual VM levels. The Flex model supports all allocation configuration of the other allocation models and can be combined with *Compute Policies* to control the resource consumption of individual VMs and create tiered consumption profiles within the same Org. VDC.

The different allocation models allow cloud providers to develop consumption models based on their tenant's requirement, while ensuring that they don't need to keep separate provider resources to cater to various business needs. This helps cloud providers to manage resources across their data centers with a far greater level of efficiency, resulting in reduced costs.

Compute Policies: Cloud Director offers the concept of Compute Policies, which can help cloud providers manage resource allocation across the Cloud Provider Platform and also help optimize licensing costs by providing a placement engine for workloads.

There are two types of compute policies provided by Cloud Director, the *Provider VDC Compute Policy*, and the *VDC Compute Policy*.

The provider scoped policy defines VM-host affinity and anti-affinity rules that can be used to enforce placement of a customer's workload. The typical use case for this is license optimization, but it could also be utilized by a machine learning application, and its placement on a GPU-enabled host, or possibly for disaster avoidance purposes in a stretched vSAN cluster design. The customer has no visibility or control over Provider VDC Compute Policies as they are scoped, as the name implies, within the Provider VDC. Cloud Director administrators expose virtual machine groups and logical VM groups to tenants, which can then be referenced and control the physical compute resource allocation for the tenant's workload.

VDC Compute Policy: Placement: The VDC Compute Policy controls the compute characteristics of a workload within the tenant Org. VDC and is therefore controlled by the customer. For placement, the customer can refer to the Provider VDC Compute Policy, configured inside the VDC Compute Policy. This allows the cloud provider to deliver specific services on dedicated 'island clusters.'

The concept of island clusters is straightforward. An island cluster hosts workloads with special license requirements, or other unique characteristics, and is sometimes also referred to as an 'application cluster.' Some software vendors apply special licensing policies on their applications, middleware, or databases that are not conducive to virtualization or shared public cloud environments, especially where vSphere DRS is employed, and the application could potentially touch a high number of physical CPUs. The island cluster concept is a common approach of dealing with this challenge.

Use cases for island clusters might include:

- Running Oracle databases, middleware, or applications on dedicated clusters will not only ensure that you are able to consolidate more and more Oracle virtual machines on a small cluster of ESXi hosts, but also reduce

licensing costs by limiting the number of physical CPUs that require licensing (if this is your Oracle licensing model).

- Cloud providers also use island clusters of operating systems such as Windows or Red Hat Enterprise Linux. This helps reduce costs on data center Windows OS socket licenses, which is typically the most cost-efficient way of licensing large numbers of Windows virtual machines running on a host. Another side benefit of this approach is that it helps ESXi to take advantage of the memory management technique of Transparent Page Sharing (TPS), if enabled. This is more efficient because there are higher chances that when you are running many of the same operating system virtual machines, duplicate memory pages will be spawned by these VMs in physical memory, making your ESXi servers more efficient.
- GPU-enabled clusters might also be a great use case for an island cluster, delivering specific capabilities to applications that are designed to use the high computational capabilities general-purpose GPUs offer.

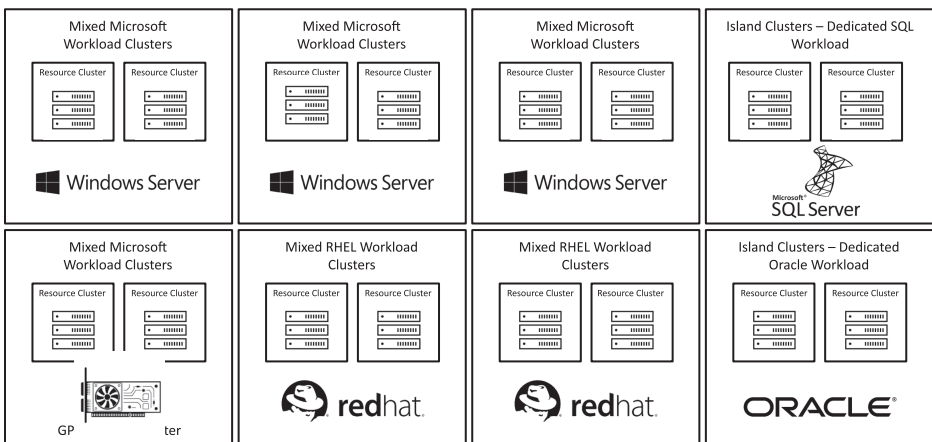


FIGURE 3.19 Island Clusters

With the use of compute placement policies, cloud providers can back this strategy with automation, guaranteeing placement and ensuring specific operating system templates or workloads are only deployed on designated resources. This is illustrated in Figure 3.20, where a Placement Policy is applied to a VM template to ensure that all VMs instantiated from the template are placed on hosts within the policy.

There are a number of benefits to cloud providers that are realized through this approach to defined workload placement, where appropriate. Not only does it ensure that workloads run only on properly licensed hosts, it can also be used to separate workloads into different hardware fault domains, in the case of vSAN stretched cluster backed Resource Groups. In addition, this approach can allow cloud providers to offer multiple tiers of performance

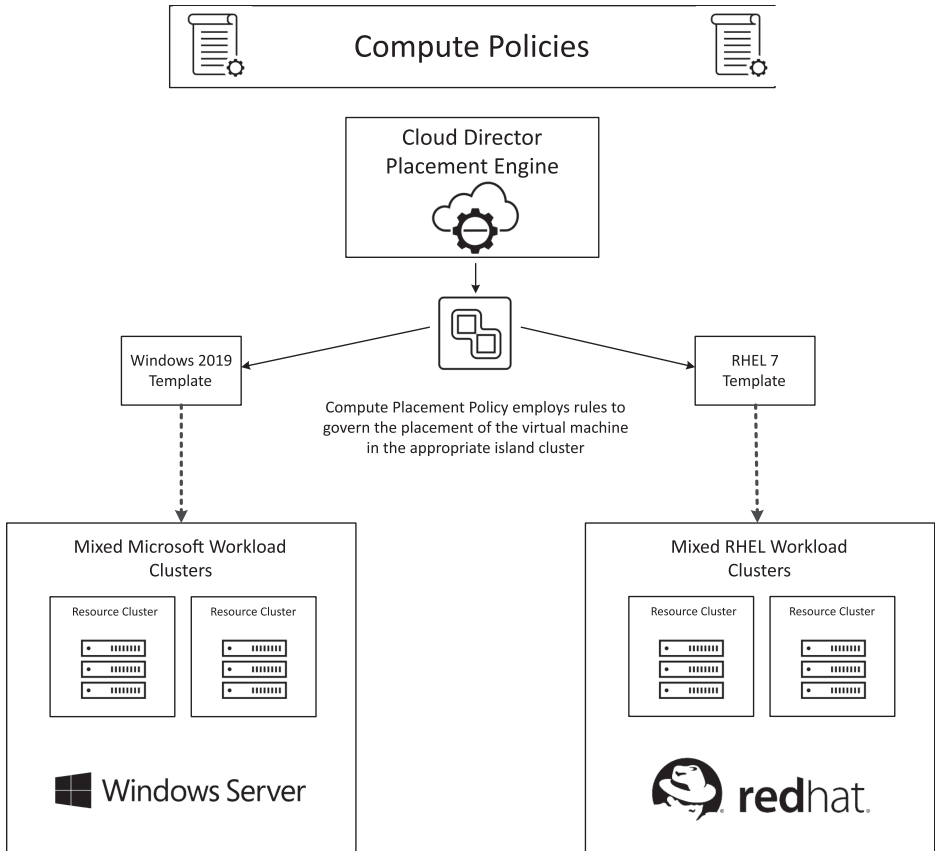


FIGURE 3.20 Island Cluster Provisioning Workflow

tied to different hardware, such as GPUs, for self-service consumption of specific services.

VCD Compute Policy: Sizing: The second Compute Policy use case is the Sizing Policy. The sizing policy allows cloud providers to assign a memory and CPU consumption profile to individual workloads and allow tenants to choose from a sizing policy, per VM, to suit the workloads consumption needs, effectively the concept of T-Shirt sizing. This is achieved by using the *VDC Compute Policy* to control the physical compute resources allocated to the customer workload, and allocate resources based on the workload-specific requirements. As illustrated in Figure 3.21, tenant users can choose either the default policy or one of the *VDC Compute Policies* defined by the cloud provider when creating and managing VMs in the Org. VDC. The sizing policy can specify:

- Number of vCPUs and vCPU clock speed
- Amount of memory allocated to the virtual machine
- Memory and CPU reservation, limit, and shares

Organization vDC				
vm	Name: T1 Basic	Memory Allocation: 1024	Number of vCPUs: 1	Compute Policy: T1 Basic Description: Small-size VM policy
vm	Name: T2 Standard	Memory Allocation: 2048	Number of vCPUs: 2	Compute Policy: T2 Standard Description: Medium-sized VM
vm	Name: X1 Premium	Memory Allocation: 4096	Number of vCPUs: 4	Compute Policy: X1 Premium Description: Large-sized VM

FIGURE 3.21 Compute Policy Examples

Sizing Policies will typically be used with the *Flex VDC* allocation model, even though they do work with the other legacy allocation models addressed earlier in this chapter, Pay-as-you-Go, Reservation, and Allocation. However, if a tenant configures any sizing policy that directly conflicts with the allocation model, it will result in a warning along a request to convert the Org. VDC to one with Flex defined as the allocation model.

It is for the Cloud Director system administrators to create and manage VM sizing policies at a global level and publish them individual to one or more Org. VDCs. For cloud providers, the use of sizing policies gives them the ability to provide per virtual machine compute resource allocation within an Org. VDC, simplifying consumption management while providing the ability to mix different consumption policies in the same Org. VDC for tiering and pricing differentiation.

Cloud providers can also restrict consumption to pre-defined policies at a per Org. VDC level for fine control over customer deployment practices.

Tenant Application Catalogs and Libraries: Each Org. VDC has a tenant level store for publishing virtual machines or Application Blueprint templates. These templates can be OVF or OVA files that help tenant users deploy their virtual machine instances or applications in the cloud faster. Catalogs can be used to store additional file types such as ISO images, which can be mounted to a virtual machine. Tenant users that are authorized can upload and maintain these catalog items without the need to have the cloud provider manage the catalog. However, as we addressed previously, cloud providers can also publish images to public catalogs that are visible to all tenants of the cloud in order to manage those centrally.

Workload Types Virtual Machine Verses vApp Blueprint: Cloud Director can deploy and manage multiple workloads types, including virtual machines, vApp Blueprints, virtual appliances, or Kubernetes Clusters. When it comes to supported Guest Operating Systems and Virtual Hardware Versions, Cloud Director supports all guest operating systems and virtual hardware versions supported by the ESXi hosts on the underlying platform which back each resource pool. This means that Cloud Director

supports the widest range of guest operating systems of any public cloud platform, including Microsoft Windows guest operating systems, Apple Releases, and Linux distributions.

A vApp Blueprint is a preconfigured virtual machine that packages applications with their required operating systems with pre-defined parameters that define operational details, effectively an application encapsulated within a vApp pool. The vApp allows multiple separate virtual machines to work together in a stack as an application by defining a number of specific things about the appliance, such as performance and resource pools, IP address allocation policies, or firewall requirements. You can also deploy new vApps based on existing ones running on the platform. VMware vApps operate on the Open Virtualization Format (OVF) standard and are exported in OVF format.

A virtual appliance is a generic term for an application delivered as a prebuilt unit. More specifically, a virtual appliance is a virtual machine image file consisting of a pre-configured operating system environment and a single application. The purpose of a virtual appliance is to simplify delivery and operation of an application. To this end, only necessary operating system components are included. Deploying an application as a virtual appliance can eliminate problems with installation and configuration, such as software or driver compatibility issues.

A Kubernetes Cluster is an open-source container-orchestration system that aims to provide a platform for automating deployment, scaling, and operations of application containers across clusters of hosts. More and more cloud service providers are offering a Kubernetes-based platform or infrastructure as a service to tenants on which Kubernetes can be deployed as a platform-providing service. This service delivery model through the VMware Cloud Director Container Service Extension will be addressed in more detail in Chapter 8.

Tenant Networking: Each Org. VDC exposes to cloud consumers multiple network features and functions through the tight integration Cloud Director provides with VMware NSX Data Center. Therefore, the following advanced network services can be exposed to consumers, with no development, integration challenges, or custom engineering.

- **Dynamic routing (DLR):** Cloud Director provides the Distributed Logical Router feature to manage routing needs between two or more Org. VDC Networks.
- **Edge Services Gateway (ESG):** The NSX Edge Services Gateway offers the capability of dynamic routing protocols (BGP and OSPF).
- **Distributed Firewall (DFW):** Performs the firewall rule processing directly within the ESXi kernel, which results in firewall rules on workload inside an Org. VDC. The distributed firewall rules define the ways that traffic can flow between workloads on an Org. VDC network.
- **Security Groups:** Cloud Director provides Security Groups that define security policies dynamically. A customer admin can create policies to match

individual workloads via Security Groups, DFW policies can be applied to these Security Groups.

- **Tenant layer 2 (L2) VPN access:** Org. VDC to an on-premises network with a tunnel that creates a Layer 2 broadcast domain between the two sites and makes the two networks look seamless across the cloud when putting workloads on both sites on the same IP subnet.
- **Load Balancing:** The NSX Edge load balancer helps evenly distribute incoming traffic to an Org. VDC pool of servers. Pools support common application types and can be based on HTTP, HTTPS, TCP, or UDP. Features included are:
 - TCP, HTTP, HTTPS with stateful high availability
 - Multiple VIP addresses, each with separate server pool and configurations
 - Multiple load balancing algorithms and session persistence methods
 - Configurable health checks
 - Application rules
 - SSL termination with certificate management, SSL pass-through, and SSL initiation
 - IPv6 support
 - Support for UDP applications
 - L7 manipulation, including URL block, URL rewrite, and content rewrite

NSX Advanced Load Balancing (Avi Networks) can also be consumed for more complex or advanced use cases.

As a result of this close network integration, Cloud Provider Platform consumers can leverage a service with full NSX support and capabilities that can be exposed and consumed in a multi-tenant, self-service fashion. This integration also gives cloud providers granular role-based access control (RBAC) over these capabilities, which enables them to carefully govern what their tenants can and cannot do and what is visible to them in the user interface. This allows for many of the features outlined above to be monetized as advanced network service options and enabled on a feature-by-feature basis, per-organization. While some cloud providers might choose to enable all features for all customers and integrate the cost into their service offerings, others might choose to provide these as value-add services, with each feature monetized for consumption. Multiple considerations play into this decision, including a go-to-market strategy for these enhanced network services, as well as the commercial choice between VMware granular billing verses a bundled approach of licensing all workloads with these features.

PERSONAS & ROLE-BASED ACCESS CONTROL

Cloud Director, as a secure multi-tenanted cloud management platform, must maintain a wide range of roles for purposes of access management. At the top level, most importantly are those of the provider and tenant.

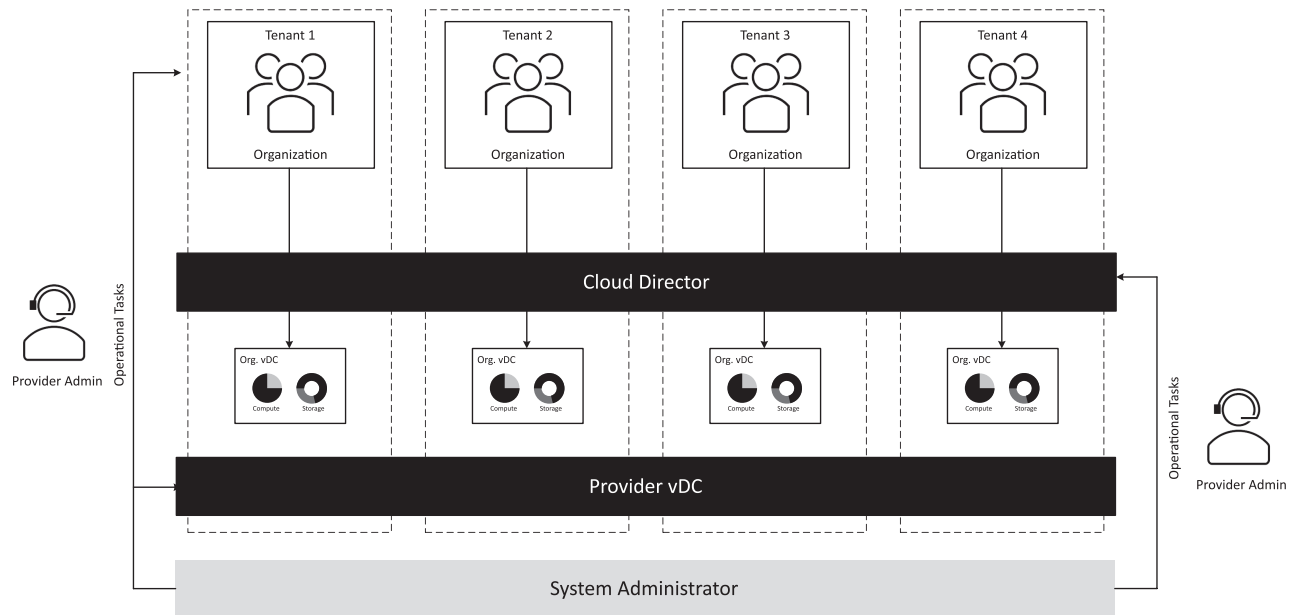


FIGURE 3.22 Cloud Director Key Personas

Provider Roles are of course defined for the cloud provider. For cloud providers a range of roles can be defined from full system administrators to provider users of custom roles, which can be defined and created to meet specific cloud operator use cases.

Provider roles will typically be responsible for:

- Infrastructure
- Platform Management
- Cloud Resource Management
- Tenant Management
- Tenant Support

By default, the cloud provider has only two roles defined, which are local to that system. These are the System Administrator and Multisite System roles. System administrators can create additional custom provider roles in the provider organization, based on the business' operating requirements:

- **System Administrator:** This role exists only in the provider organization. The System Administrator role has complete authority over the Cloud Director system. Credentials for the System Administrator role are defined during the installation and initial configuration process.
- **Multisite System:** This is a system level role used for maintaining the heartbeat process between sites in a multi-site architecture. This role has just a single right assigned, that of, `Multisite: System Operations`. This right is used to allow a remote vCloud API request that requests the status from a remote site, with which an association exists.

Cloud Director has the ability to create granular roles, using *Rights Bundles*. This works by allowing the system administrator to create rights bundles and publish them to one or more Organization on the cloud platform. This mechanism allows the System Administrator to manage the rights that are available to each Organization.

The rights bundle is a defined set of rights, which provides a flexible management hierarchy to create and publish rights that correspond to customer's service requirements. This also facilitates the functionality required to monetize specific Cloud Director services or functionality.

The rights are grouped into various categories, such as Compute, Extensibility and Networking and then subcategories, to view or manage access to the specific object to which the right relates. When creating rights bundles System Administrators can choose to select the rights globally, by subcategory, or individually.

System Administrators have full rights to edit and delete existing rights bundles or unpublish rights bundles from an organization on the Cloud Director platform. Additionally, System Administrators can publish multiple bundles to the same Organization.

It is the cloud provider System Administrators that creates the global tenant roles and publishes them to Organizations. Tenant roles can use only those rights that are published to their specific Organization and can contain a subset of the Organization rights only. These global tenant roles can then be assigned to tenant users in the Organizations to which they are published. The Organization administrators cannot edit global tenant roles. However, the organization administrators can create and edit tenant-specific roles, within their organization. The Organization administrator can then assign the new tenant-specific roles to tenant users in the organization. The tenant-specific roles will only contain a subset of the original organization rights. This ability for tenants to manage their own specific role-based access control through the user interface or API facilitates simple self-service management of permissions, without having to interact with the provider.

In addition to the ability to create and edit customer roles, tenants also have a number of Predefined Roles. The predefined roles give the organization administrator the ability to assign the default set of rights required to undertake common workflow operations in Cloud Director. Predefined roles include the Organization Administrator, Catalog Author, vApp Author, vApp User, Console Access Only, and Defer to Identity Provider, where role permissions are determined based on information received from the customer's OAuth or SAML Identity Provider. These predefined global tenant roles are published to every organization created by default.

An additional model to consider is where the physical cloud resources themselves are provided through a public cloud hyperscale platform, such as AWS, Azure or Google. In this scenario, we can add a *Cloud Operator* to the access model. In this Managed Service Provider-based solution, the Cloud Director Service, the SaaS-based offering, is likely to be chosen as the cloud management delivery mechanism, subscribed to by the cloud provider.

As illustrated in Figure 3.23 the cloud operator delivers key managed components at the vSphere layer, the Cloud Director Service, Provider VDC, tenant management and support are provided by the cloud provider, while the cloud consumer self manages their own organization, organizational resources, and applications.

CLOUD RESOURCE DESIGN

From an infrastructure platform perspective, a Cloud Provider Platform is built on a foundation of software defined data center solutions, whose components are split between management and cloud consumer resources. In building a Cloud provider Platform there is an assumption that all management components, such as vCenter Server, vRealize Operations, and NSX Manager, will run in a dedicated Management Pod.

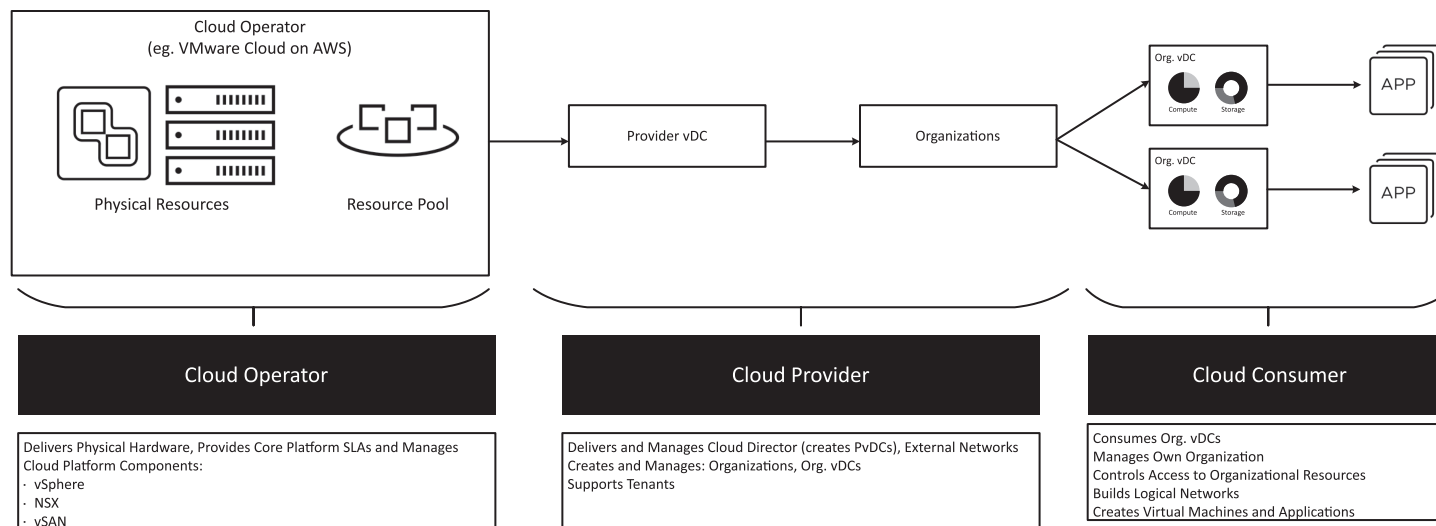


FIGURE 3.23 Public Cloud Provider Operated Model for Managed Service Providers

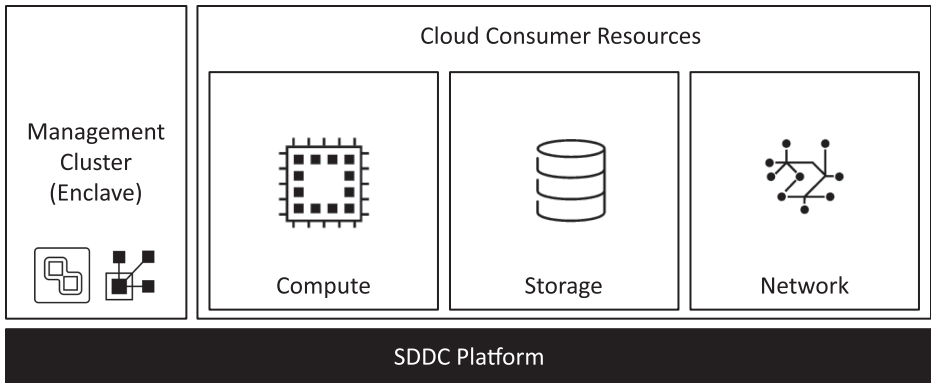


FIGURE 3.24 Cloud Provider Platform Infrastructure

Management Cluster

As illustrated in Figure 3.24, the typical approach of separating resources allocated for management functions from tenant workloads is common practice for cloud providers. Therefore, the underlying vSphere clusters will be split into two logical groups:

- A single management cluster running all components and services needed to run the Cloud Provider Platform.
- Cloud consumer resources are formed from Resource Clusters under vCenter management controlled by Cloud Director. In most cases multiple Resource Clusters are managed by a common vCenter Server, which are deployed in a scale-out model as resource consumption increases.

The primary reasons to ensure that management components are separate from the resources they are managing include minimizing overhead for cloud consumer resources (resources clusters allocated for tenant use have little overhead) and dedicating resources for the cloud to ensure resources can be consistently and transparently managed, carved up, and scaled horizontally.

The management cluster may also include virtual machines or have access to servers that provide infrastructure services such as directory (LDAP/AD), timekeeping (NTP), networking (DNS, DHCP), and security (certificates). Other management components in the management cluster will likely include the following:

- vCenter Server(s)
- Cloud Director Cell(s)
- vRealize Operations Manager
- Tenant App
- NSX Manager(s)
- vCenter Orchestrator(s)

In a federated multisite deployment, each management cluster would reside in a single physical region, consuming resources within the same physical site, ensuring a consistent level of service. Otherwise, latency issues might arise if workloads need to be moved from one site to another, over a slower or less reliable network. Considerations for connecting clouds representing different sites are addressed later in this chapter. Secondary DR sites are discussed later in the Cloud Director Availability section of Chapter 8.

The management cluster will also follow vSphere best practices to facilitate load balancing, redundancy, and high availability within the underlying vSphere Infrastructure. To facilitate VMware vSAN, a cluster of three VMware ESXi hosts will be required as a minimum (four recommended for improved vSAN high availability). While additional hosts can be added, three hosts supporting just Cloud Provider Platform management components should typically be sufficient for resource purposes. The design would typically employ a VMware HA percentage-based admission control policy in an ‘*n+1*’ fashion instead of dedicating a single host for host failures or defining the amount of host failures a cluster can tolerate. This will allow the management workloads to run evenly across the hosts in the cluster without the need to dedicate a host strictly for host failure scenarios. Additional hosts can be added to the management cluster for ‘*n+2*’ or more redundancy if this is required by the service-level agreement.

A Cloud Provider Platform design can protect the management workloads with Site Recovery Manager (SRM). moving Cloud Director cells to a secondary (DR) site should it be required and is mandated through the SLA. However, modifying the IP addresses of Cloud Director cells can result in recoverability complexities, therefore a layer 2 network extension is required.

Unlike a traditional vSphere environment, where vCenter Server is used by administrators to provision workloads, vCenter Server plays an integral role in the tenant’s self-service provisioning by handling all workload deployment requests by Cloud Director. Therefore, it is typically recommended that vCenter Servers are made highly available with the vCenter HA feature.

Other underlying platform design considerations, specific to a Cloud Provider Platform will be addressed accordingly in this document, in the following chapters.

Resource Clusters

All cloud consumer resource should be presented to the Cloud Provider Platform in separate Resource Clusters, with each Resource Cluster being a vSphere cluster of three or more hosts, as required by VMware vSAN.

This approach allows cloud consumer resources to be more dynamic in nature with a lower management overhead. When designing a Cloud Provider Platform, Resource Clusters will be represented as Provider Virtual Data Centers. The providers aim should be to make resource clusters clearly defined in nature, to reduce variability. For this reason, it makes sense that where possible the deployment and configuration should be highly standardized,

through either vSphere automation and Host Profiles or a third-party solution. VMware Cloud Foundation provides features that allows the provider to automatically deploy vSphere hosts from a predefined configuration and build workload domains that are represented as Provider VDCs to Cloud Director. Cloud Foundation will be addressed in detail in Chapter 7.

Like the management cluster, compute resources for the Resource Clusters should follow vSphere best practices where possible, with Resource Clusters typically having Distributed Resource Scheduling (DRS) enabled for CPU, Memory and Network resource balancing across the cluster, improving the performance of customer workloads and vSphere HA to automatically restart the tenants applications in the event of a host failure.

The cloud consumer resources, outlined above will be carved up by Cloud Director using the provider and tenant constructs and definitions addressed earlier in this Chapter. A high-level view of the Cloud Provider Platform logical architecture, with vSphere resource separation, is depicted as follows. The compute, storage, and networking resources design of these components will be drilled into in Chapters 4, 5, and 6.

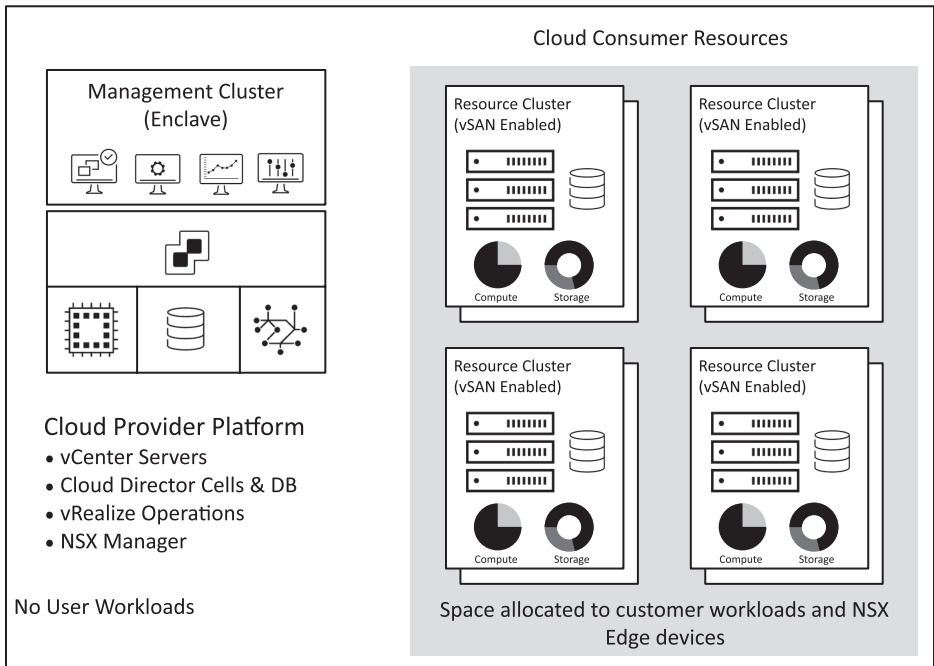


FIGURE 3.25 Cloud Provider Platform Resource Logical Infrastructure

Scalability

No software or hardware solution is without a scalability limit. Therefore, any cloud platform that must scale across hundreds or thousands of tenants

and tens or even hundreds of thousands of workloads must take a scale-out building block approach to architecture.

Pod Architecture: A common approach to creating building block infrastructure architecture is referred to as pods. Pods can include different combinations of servers, storage equipment, and network equipment, and can be set up with varying levels of hardware redundancy and varying quality of components. Pods are connected to a common network core that distributes data between them. The pod is not defined by any hard-physical properties, as it is a standard unit of connected elements within the SDDC fabric.

The cloud provider has total autonomy when it comes to pod design. For instance, a partner could choose to provide a dedicated pod to internal customers, potentially providing different services and SLAs to those internal tenants. Alternatively, a partner could provision a dedicated pod to a large enterprise customer, who had committed to migrating large numbers of workloads and has requirements beyond that of a typical client. Another option is to provide pods which meet specific regulatory standards, such as a PCI DSS pod, to meet the requirements of the card payment industry, or a pod specifically designed to meet the security requirements and service-level agreements of government tenants, similar to the AWS GovCloud in the United States. Other pods within the same architecture, could simply service multi-tenant resources carved up among hundreds of tenants.

For this example, we will address two pod types. The following pods would be deployed per region, in a multi-data center solution, which is a requirement for this example:

- **Primary Pod:** Includes the cloud management cluster, associated platform tools and portal presentation in each region
- **Expansion Pod:** Compute SDDC Resource Clusters and Shared Edge Clusters for tenant workloads plus limited cloud management components required specifically for that expansion pod

The pod is a logical boundary of functionality for the Cloud Provider Platform and encompasses a fixed number of Resource Clusters. A pod will usually span multiple racks, to allow for the distribution of multiple Resource Clusters across the rack infrastructure, as illustrated in Figure 3.26.

One of the guiding principles for such deployments is that VLANs are not spanned beyond a single pod by the network virtualization layer. Although this VLAN restriction appears to be a simple requirement, it has widespread impact on how a physical switching infrastructure can be built and on how it scales. For instance:

1. VLANs carried throughout the Fabric of a single pod
2. Layer 2 application scope is limited to a single pod, and is the failure domain

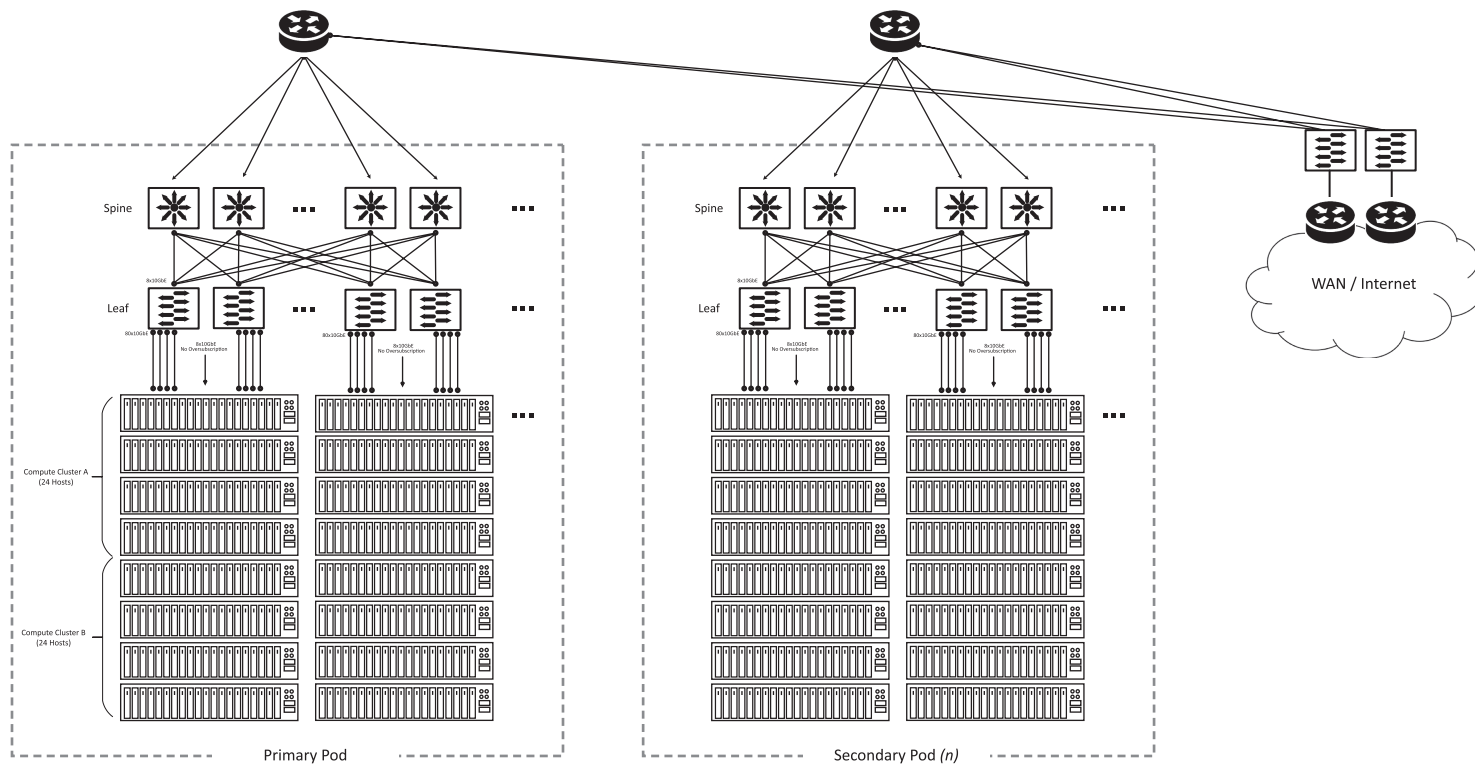


FIGURE 3.26 Pod Architecture Resource Cluster Design

- 3. Multiple aggregation modules, to limit the Layer 2 domain size
- 4. Inter pod traffic is Layer 3 routed
- 5. Extending VXLAN inter pods is the same as Layer 3 fabric

Pod Sizing: As indicated above, all software and hardware solutions have scalability limits. Whether this is tenant workload scalability, number of networks, number of petabytes of storage or physical floor space in the data center, the recommendation is that the pod sizing be dictated by these specific boundaries, and those that define the limitations in a specific environment.

Primary Pod: The Primary Pod is the first (initial) pod that will be deployed and will run the management tools and components that provide the primary tooling enclave in each region, which also manages the Expansion Pod(s) within that geographic region. This management enclave will host vCenter Server(s), vRealize Operations, vRealize Log Insight, Cloud Director, NSX Manager, and other shared management components, which are required at a regional level. In addition, the Primary Pod also runs the same management tools that will be required in each Expansion Pod, such as additional vCenter Servers and NSX Managers.

The management tools will typically be provided only in the Primary Pod, per geographic region, as illustrated in Figure 3.27.

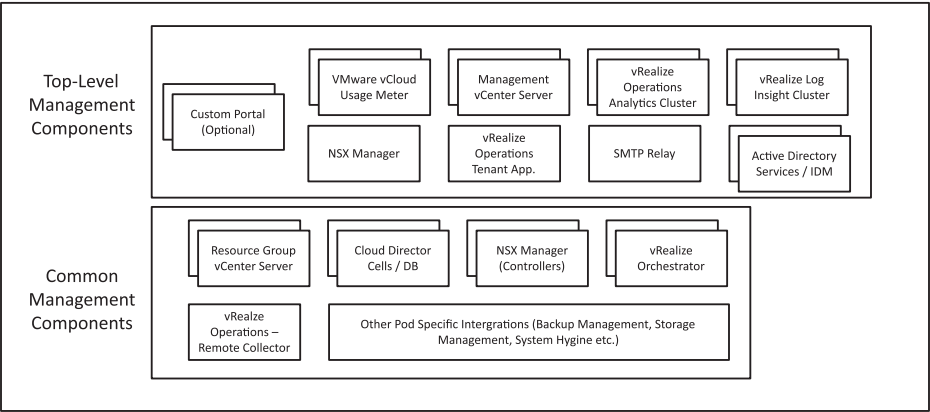


FIGURE 3.27 Primary Pod Architecture

The Primary Pod is the *top* layer of the stack within each region and is where the service consumption occurs, through the Cloud Director user portal, as well as other published services, such as the Cloud API. This layer calls for resources in one of the underlining SDDC Pods, and then orchestrates the actions at the lower layers to achieve the request. Management components in the Primary Pod will not have tenant-specific addressing or addressing that overlaps that of tenants.

As the Primary Pod hosts critical management infrastructure, it will not only be protected by implementing an extra level of local hardware redundancy, but also either be protected via Site Recovery Manager cross-site protection, or the components distributed across data centers within the same region, with high-speed connectivity, to make the management solution resilient to data center outages, as illustrated in Figure 3.28.

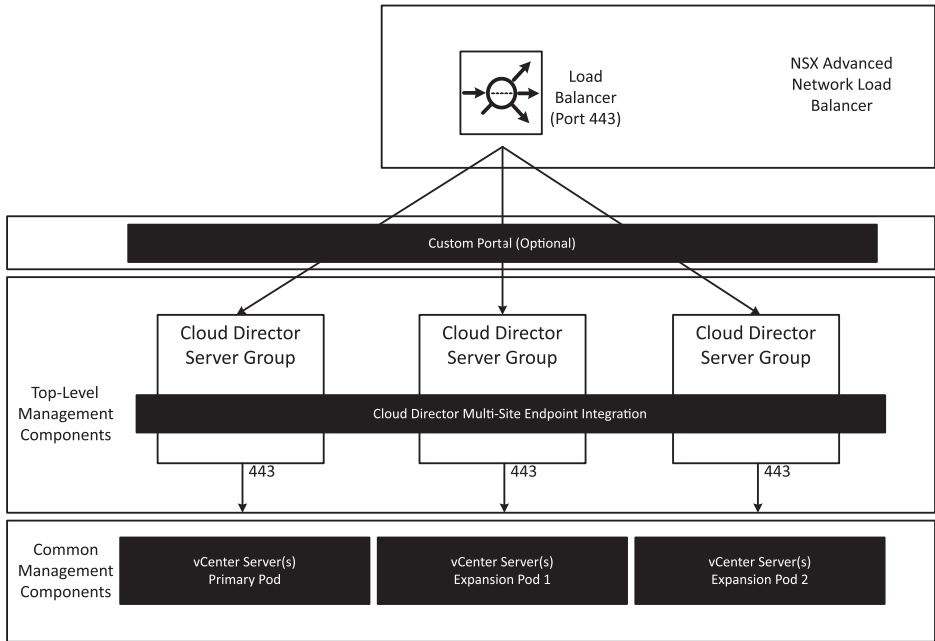


FIGURE 3.28 Distributed Architecture

To facilitate sessions across a distributed Cloud Director cell architecture, session affinity, or sticky sessions on the load balancer should be enabled to allow the same VCD cell to serve all the requests after a session is established by a user with that cell. The session affinity should be enabled on the load balancer to direct subsequent web requests from the same unique session to the same Cloud Director cell, in the load-balancer pool.

As illustrated in Figure 3.25 and Figure 3.27, the Primary Pod will also provide tenant-specific resources, as per the Expansion Pod(s), for the initial block of customers. In effect, the only difference between the Primary Pod and Expansion Pods is the larger management footprint required for the top-level management tools, employed one time only at a regional level.

Expansion Pods: The Expansion Pod provides a shared Edge cluster and compute SDDC infrastructure for Resource Clusters for tenant workloads. The Expansion Pod(s) also requires vCenter Servers, vRealize Orchestrator

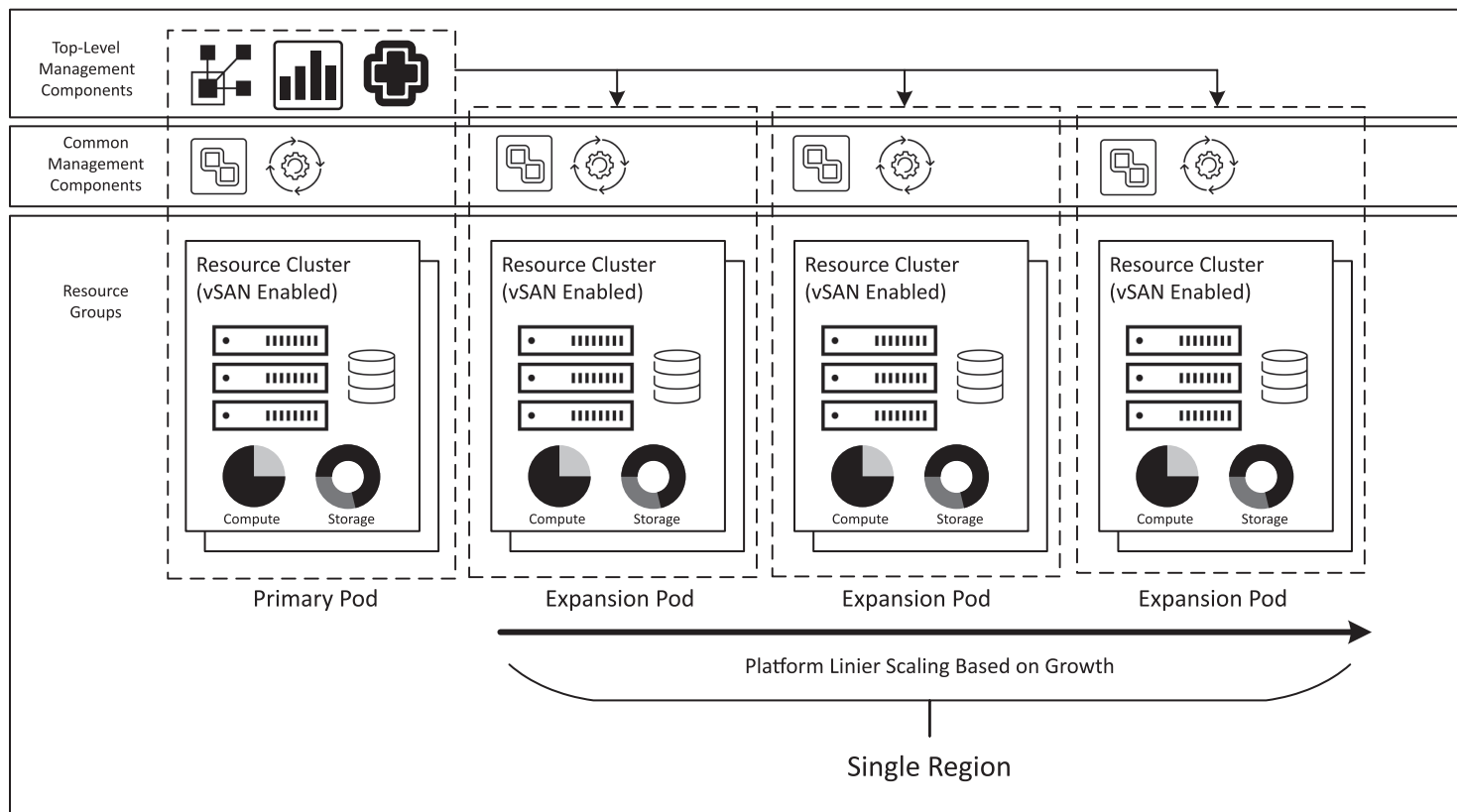


FIGURE 3.29 High-Level Pod Management Architecture

and other services required to enable the resources within that expansion pod and the Resource Clusters contained for tenant workloads. These management components will typically run on the management cluster in the Primary Pod, to give a central point of management and operations.

As the Cloud Provider Platform grows, additional expansion pods can be added to support the increase in tenant workloads within each region, or a mix of different types of workloads for different types of services or service-level agreements.

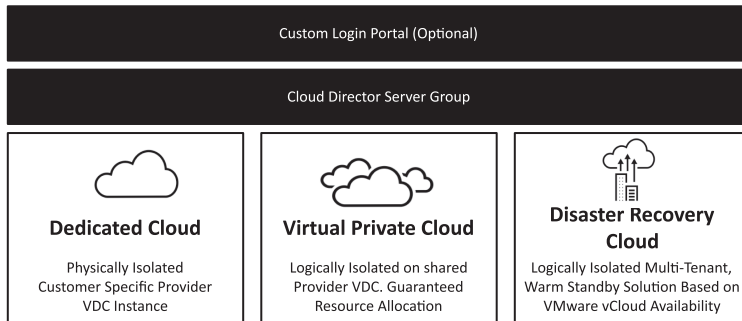


FIGURE 3.30 Flexible Pod Architecture

As highlighted previously, and illustrated in Figure 3.30, it is common for different types of expansion pods to support different provider use cases, SLAs and customer requirements.

Pod Evolution and Addressing Technology Changes: It is inevitable that over-time, typically a period of between 18 and 24 months, that the standardized technology employed in the pod architecture will required to be updated, modified or changed all together. This is likely to occur either as the natural result of technology evolution, such as CPU advancements or as a result of business or commercial drivers, such as vendor changes. This is standard practice across cloud providers, and it is important to cloud customers that their platform of choice evolves alongside trends in the industry. As a result, it is likely that a cloud provider will need to develop a plan around this evolution, to establish the criteria or drivers required to warrant a change in pod architecture. For these reasons it is recommended that cloud providers should begin considering the architecture for the evolution of pods as part of the planning for a new Cloud Provider Platform. This will ensure a seamless integration at the appropriate time.

There is no standard approach to pod design; each cloud provider will typically take their own approach based on commercial or technological drivers. However, it should be noted that initial limitations will certainly evolve over time and be mitigated in the products at some future date. Factors such as VMware vSAN sizing limits, network throughput, physical space, or power and cooling in the data center are all highly likely to evolve over time.

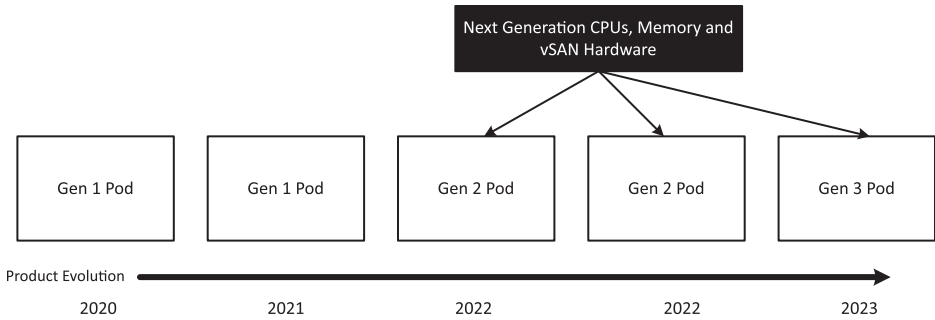


FIGURE 3.31 Building-Block Pod Architecture Evolution

Finally, as highlighted above, it is important that the architecture and scaling of the pod building-blocks is defined prior to the implementation, so that growth, resource management and cost models be established. While these design guidelines can be changed, when an adequate business case is proposed, for the overall architecture to be successful, a consistent building-block methodology is required. This approach will provide a scalable, repeatable and global architecture, which will allow for infinite growth for cloud providers and enable a simplified and common approach across regional deployments.

Hyperscale Growth: When building a platform that can support true hyperscale growth additional building constructs are required, in order to maintain a standardized approach to both physical and logical infrastructure. For instance, key design characteristics of a data center region might include the scale of the region, in terms of number of pods, should be unlimited, but might include:

- 7000 Tenants (700 per pod (with 1 Primary Pod (including management) and 9 Expansion Pods))
- 1000 hosts
- 120,000 VM/Container workloads

In theory, the scale-out architecture described in this Chapter provides limitless scalability and should be considered as part of a design and long-term strategy, along with the evolution of pods outlined previously. When planning for this level of scalability of the Cloud Provider Platform, it is important to consider the additional components or objects that are included in the solution. The first step is reviewing the respective maximum guides for Cloud Director, vSphere and NSX. While these documents reference maximum configurations for different components and features, it is important to take into account the collective maximums and remember that consuming all maximums in a through a single product or feature could limit the scalability of another product or feature.



Operational Support might be dedicated to a region or be shared across multiple regions

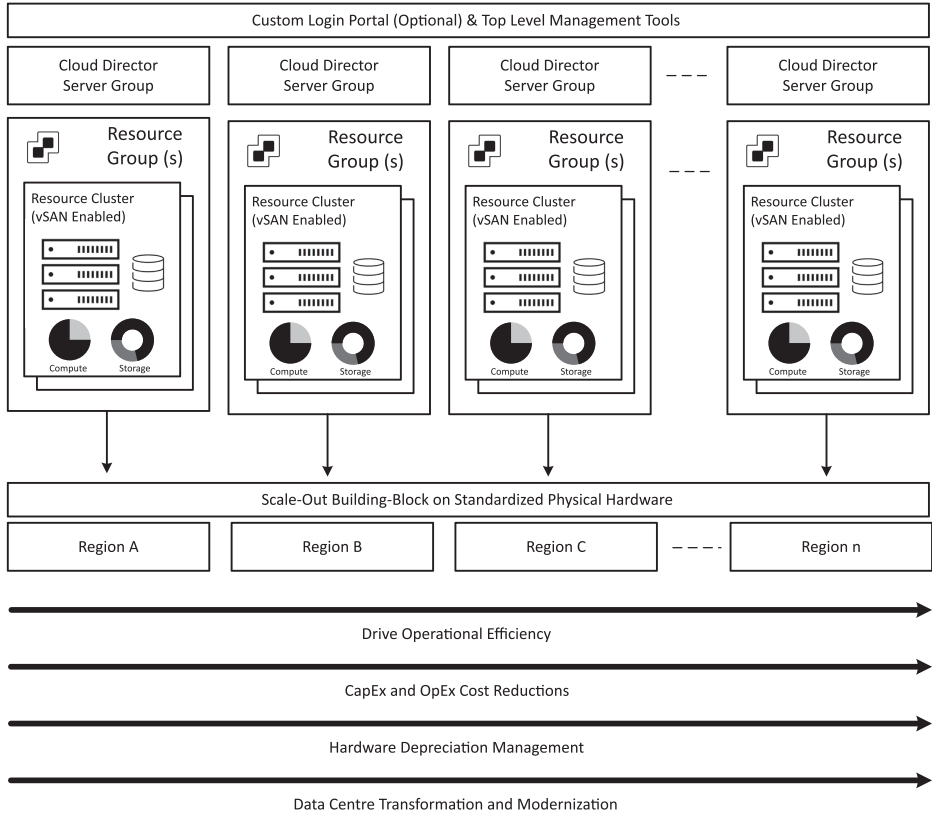


FIGURE 3.32 Hyperscale Regional Architecture

For instance, if Cloud Director supports up to 20 vCenter Servers per instance and vSphere has a maximum of 25,000 powered on VMs, if an implementation requires more than 25,000 VMs, it would be necessary to either limit the Cloud Director to one instance of vCenter Server or to split the tenant workloads across several vCenters. The design factors associated with scale must ensure that the necessary configuration maximums are evaluated for all key components to determine the scalability impact on the overall Cloud Provider Platform deployment. For cloud providers, this must be addressed alongside other key strategies as part of the lifecycle of the platform including, driving operational efficiency, reducing CapEx and OpEx, managing hardware depreciation and transforming data center infrastructure.

For more details on configuration maximums, please consult the latest VMware Product Configuration Maximums Guides.

Dedicated Hosted Cloud

There are multiple reasons why when a potential customer is searching for a new cloud service provider, they might stipulate a requirement for dedicated hardware resources. Often the customer is requiring an additional level of ‘air-gap’ separation to address security concerns, compliance certification requirements or perhaps they have concerns over workload performance and wish to ensure, as far as possible, that application performance is consistent and predictable.

The Cloud Provider Platform and more specifically, Cloud Director, allows service providers to build services such as this to meet these specific customer requirements, while still maintaining the common operating model across the platform and therefore not increasing the operating cost.

To meet this requirement, two different options exist for the provider. Depending on the specific customer requirements, in some scenarios, the cloud provider could implement managed vSphere environments for customers, using the Cloud Director Central Point of Management (CPOM) feature, and allow the customer to manage a dedicated vSphere environment through vCenter, accessed as a service through the common Cloud Director portal. The CPOM feature will be addressed in detail in Chapter 8.

The second option is to leverage Cloud Director’s IaaS engine and to provide dedicated infrastructure with a *Reservation Pool Allocation* for that customer. The key benefit of this model is, by using Cloud Director as the management platform, cloud providers gain centralized management of resources and catalogs while maintaining a physically separate compute resource pool for the customer who requires that ‘air gapped’ level of isolation, while also providing a common consumption experience for the customer across the platform. In addition, external networks can be dedicated for inbound and outbound network traffic with all network communication inside the tenant’s Org. VDC being isolated via the tunnels provided by the NSX overlay, GENEVE.

From a storage perspective, in a solution that consumes VMware vSAN storage, each Provider VDC delivers dedicated and fully isolated storage, making the storage layer fully aligned with compute in delivering the customer’s isolation requirements, with Storage Profiles based on the storage policies configured in the underlying vCenter. Where the cloud provider is implementing a traditional array based VMFS storage model, cloud providers can present dedicated storage LUNs and different tiers of storage service to the customer’s dedicated environment.

Using the Reservation Pool Allocation model, the cloud provider can deliver per tenant separation of compute resources, as illustrated in Figure 3.33. The Reservation Pool Allocation Model provides full commitment of allocated physical resources that are provided to the customer’s Organization Virtual Data Center.

As illustrated in Figure 3.33, each tenant is assigned to a single Provider VDC, which in turn is mapped to a single vSphere cluster, which is dedicated

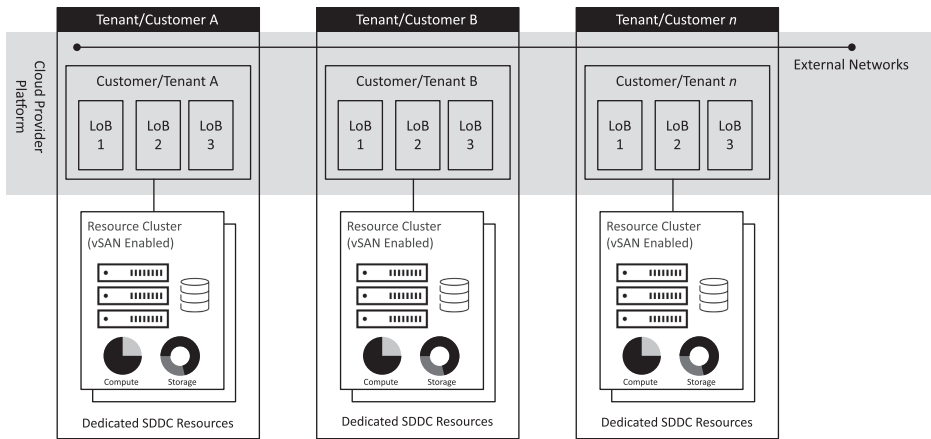


FIGURE 3.33 Dedicated Infrastructure with Reservation Pool Allocation

to that customer. This means that both the compute and storage resources of the cluster are entirely dedicated to a single tenant.

While here we have highlighted the Reservation Pool Allocation Model, the Pay-as-you-Go Allocation model is also an option to achieve this dedicated resource customer requirement. Either option can be used to provide dedicated compute and storage resources for a single tenant, while also providing a common management model for the cloud provider, for the published catalogs and other related components.

A cloud provider would typically choose to use the Reservation Pool Model when the provider will be maintaining a consistent host count for that customer. The Pay-as-you-Go allocation model is more suitable to reduce the cloud provider's overhead when the customer is likely to request additional resources for a fixed period of time, effectively bursting out and scaling back of the dedicated compute resources.

One drawback of this architecture is that due to the one-to-one mapping between Provider VDC and Org. VDC, if the customer requirement includes multiple tiers of compute, the only option is to provide physically separate compute resources for each tier, for each customer. This then brings into question the customer's willingness to pay for that, which even with cloud provider economies of scale, could prove to be prohibitively expensive.

For cloud providers offering this model, even though hardware is dedicated, it is highly beneficial to utilize a common solution across the platform and provide customers with a standardized model for scaling. This not only allows both the provider and customer to forecast costs, but also simplifies operations for the cloud provider.

Illustrated in Figure 3.34 is a model a cloud provider could also easily support with granular role-based access control, which allows a single Organization to share dedicated resources across multiple lines of business, either at the Org. VDC level or by providing physical separation at the Provider VDC level.

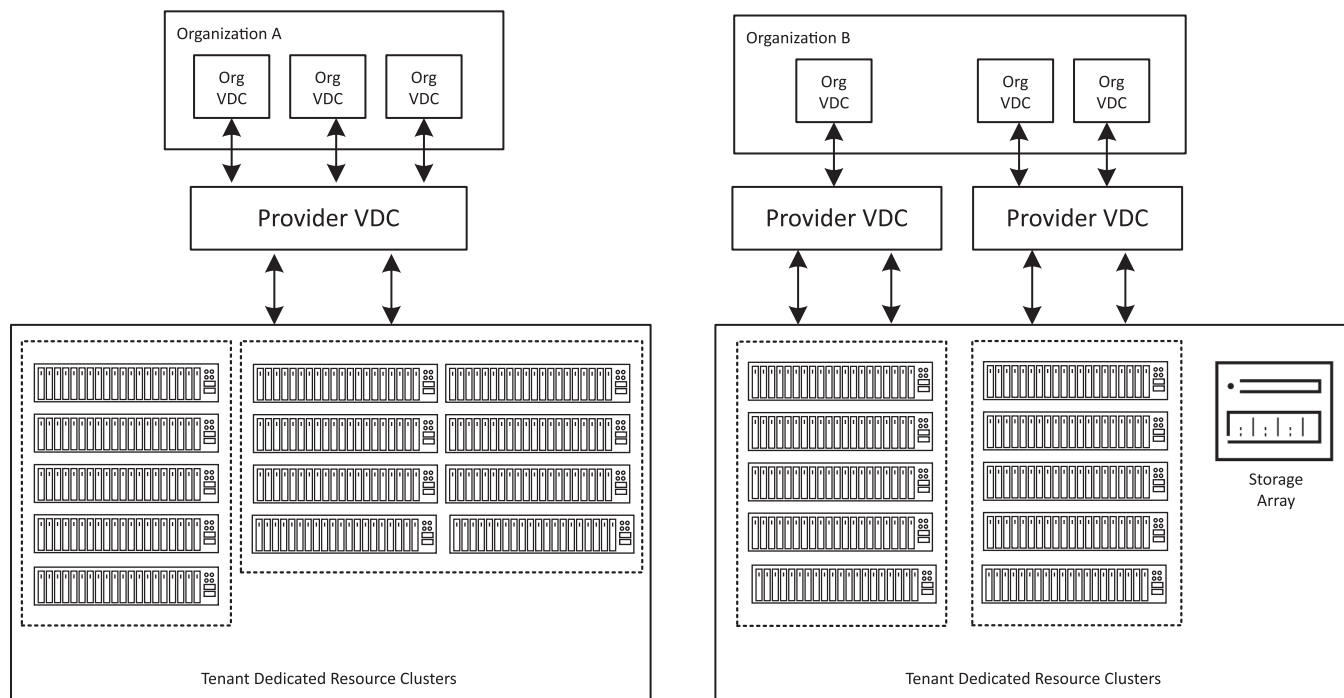


FIGURE 3.34 Dedicated Resources for Different Lines of Business

For cloud providers the key service design factor relates to scalability and the number of supported Provider VDCs per Cloud Director instance. At the time of writing, Cloud Director supports up to 100 Provider VDCs. For a large provider who offers this service to their customers, that could be considered a design constraint. However, by delivering a Pod architecture, such as that outlined earlier in this chapter, a scale-out model for Cloud Director instances can be implemented.

MULTI-SITE/MULTI-GEO CLOUDS

A multi-site architecture can mean a lot of different things to different cloud service providers. The Cloud Provider Platform is designed to support multi-site deployments in a reliable and interconnected way. When it comes to multi-Availability Zone, Multi-Region and Global deployments, a number of different design factors and cloud provider requirements must be considered.

Some providers would like to have a single user interface that encompasses all their sites and availability zones. Other providers don't mind having multiple interfaces, but they would like a common set of services available in each location. These scenarios are addressed in this section.

Cloud Director provides native multi-site capabilities to the cloud management platform in a way that allows cloud providers to avoid the trade-off between customer satisfaction and platform complexity. The HTML5-based user interface is designed to take advantage of the multi-site capability, as it is designed to interact with the Cloud API. This provides a loose coupling between the user interface and the core Cloud Director product, which is leveraged to deliver its multi-site capabilities.

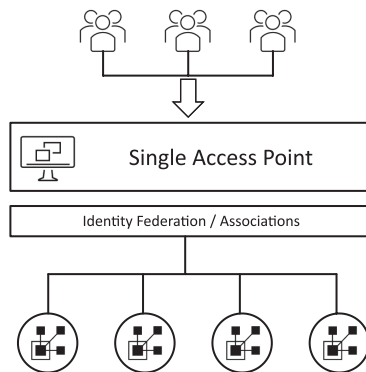


FIGURE 3.35 Multi-Site, Federated Cloud Director Architecture

The multi-site feature is key to allowing cloud providers to realize the following benefits.

- One single access point (URL) for all Cloud Director instances, regardless of location

- Identity federation for cross-site access of resources across multiple data centers
- Increasing revenue by enabling the cloud provider to offer availability zones across multiple Cloud Director instances

As stated above, multi-site can mean multiple different types of solution, depending on goals and requirements. The Cloud Provider Platform gives cloud providers the necessary flexibility to deploy multiple geographically distributed Cloud Director instances and to manage and monitor those installations as single entities, while giving customers the ability to manage resources across multiple sites or geographical regions through a single pane of glass.

For the purposes of planning a multisite platform, four basic options will be addressed. While these descriptions propose a dual site solution, except for the stretched vSAN Resource Cluster architecture, these solution designs can span multiple physical locations, either geographically dispersed regions or multiple availability zones within a single region.

As illustrated in Figure 3.36, the four base solutions that can typically be considered include:

- Multi-site architecture with discrete management
- Multi-site architecture with stretched resource clusters (vSAN Stretched Cluster(s))
- Multi-site architecture with central management stack
- Multi-site architecture with federated customer associated sites

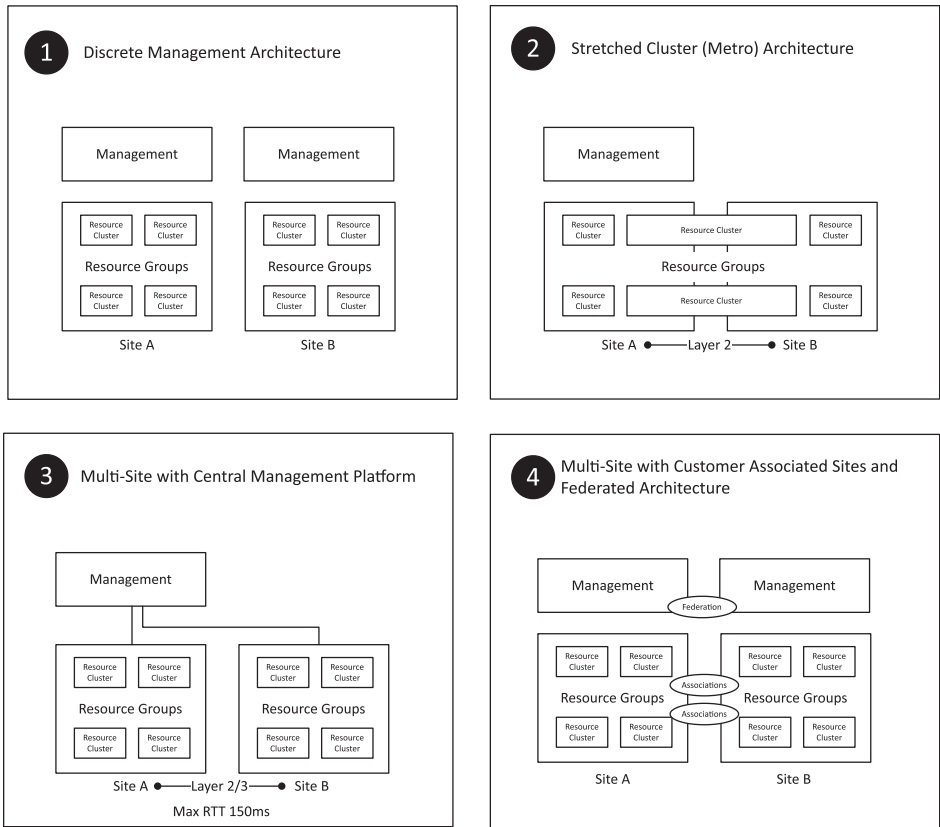
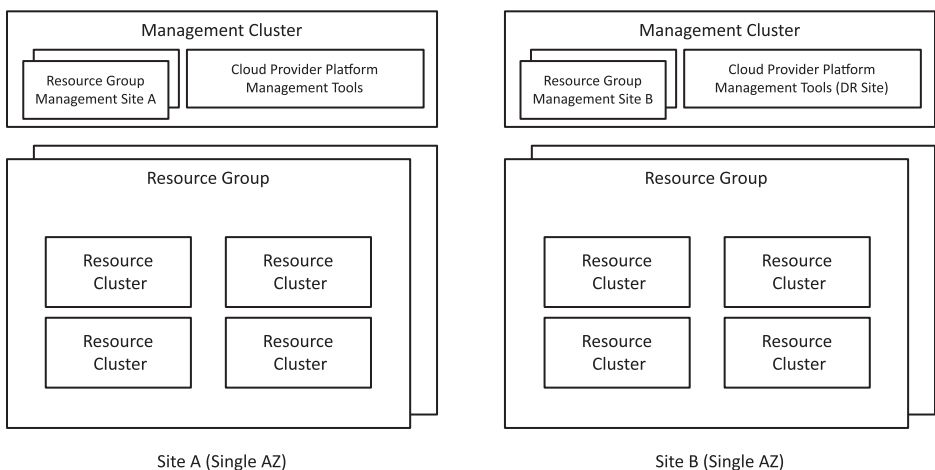
The following is a high-level overview of each of the above options alongside some key design factors, which are typically the requirements that drive the most appropriate solution for the cloud provider.

Multi-site Architecture with Discrete Management

Option 1 is for cloud providers who wish to maintain separated platforms on each site, in order to manage resources entirely independently. This is referred to as *Discrete Administration*. Some cloud providers may choose to keep things separate, although as you would expect, this model does not create a unified multi-site cloud and does not allow for native multi-site networking or migration of workloads, but ensures distinct management of each environment, even though processes and procedures might remain the same.

Multi-site Architecture with Stretched Resource Clusters (vSAN Stretched Cluster(s))

Option 2 stretches Resource Clusters across two data center locations in the same region, typically less than 100km with a Round Trip Time (RTT) of less than 5ms. In this architecture vSAN's synchronous write operations are used across data centers resulting in a Recovery Point Objective (RPO) of zero and

**FIGURE 3.36** Cloud Provider Platform Multi-Site Architectural Options**FIGURE 3.37** Cloud Provider Platform Discrete Administration

a minimal Recovery Time Objective (RTO). In a failure scenario, workloads are brought back online leveraging vSphere HA technology, with the workloads and applications being restarted in the surviving data center. However, this requires a fifty percent resource reservation to be held on both sides of the stretched resource cluster, making it a premium service, to meet specific zero RPO customer requirements.

With this solution cloud providers can choose to stretch some or all resource clusters across sites and to deploy in combination with Option 3. It might also be an option to stretch the management cluster between sites, offering the same level of availability and quick recovery of management components in a data center failure scenario.

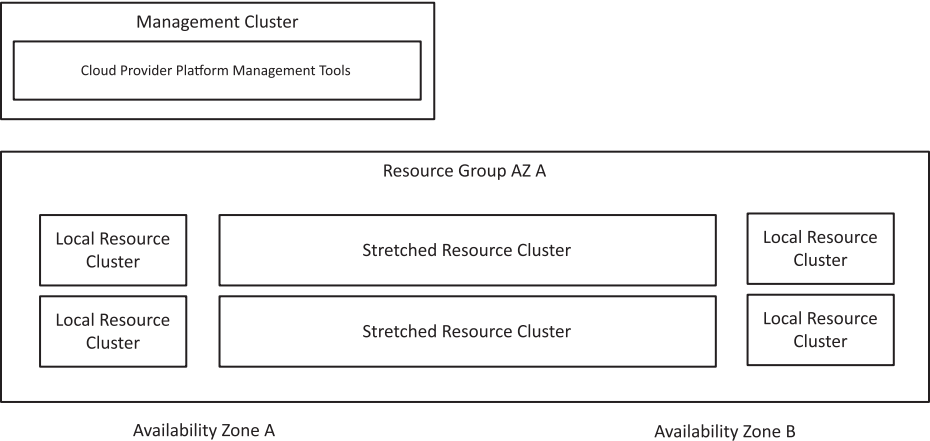


FIGURE 3.38 Multi Data Center Stretched Resource Cluster Architecture

Multi-site Architecture with Central Management Stack

Option 3 provides central management on a single site and resource clusters distributed across multiple sites. Effectively, a single Cloud Director Server Group manages Provider VDCs on both its local site and remote site(s). This architecture also allows you to stretch Organization networks across both sites and achieve Layer 2 adjacency for customer workloads. However, latency is a key design factor. While Cloud Director supports latencies of up to 150ms RTT between the Cloud Director instance and vCenter Server, the user experience degrades linearly as latency increases. For this reason, this architecture should typically only be an option when latency is consistently no greater than 40ms RTT (approximately 3300 km) between sites.

As illustrated in Figure 3.39, the management cluster can be protected to the secondary site with VMware Site Recovery Manager, or a third-party solution.

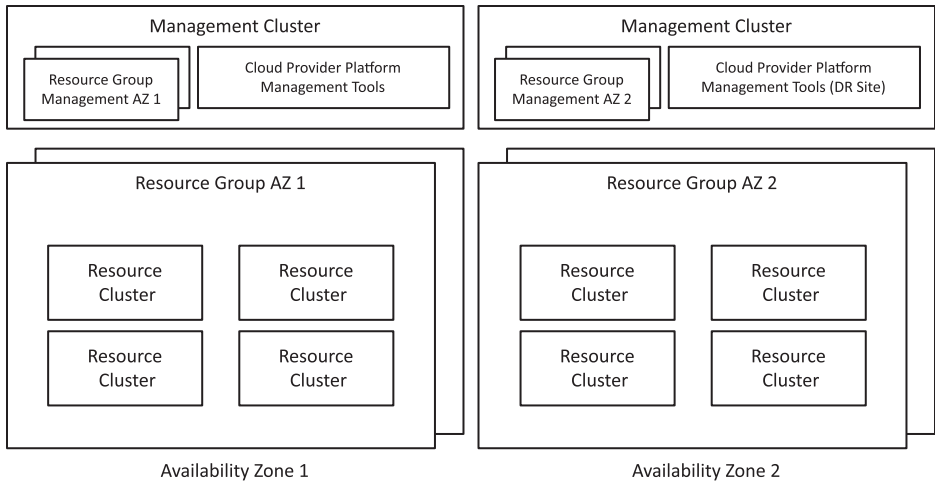


FIGURE 3.39 Multi-site Distributed Resource Groups

Multi-site Architecture with Customer Associated Sites

The final option, first introduced with Cloud Director 9, provides Organization administrators a way to associate Organizations across multiple geographically distributed Cloud Director instances and to manage and monitor their resources through a single management entry point, the Cloud Director Tenant Portal.

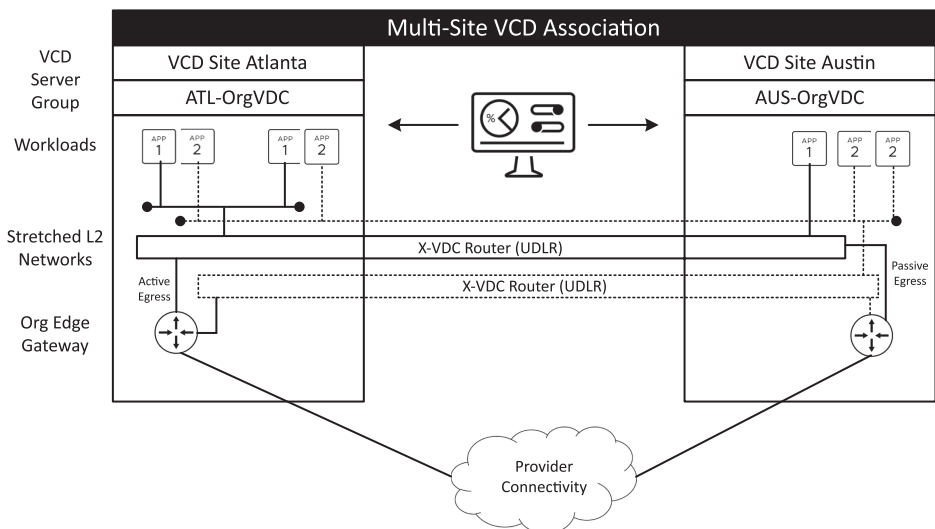


FIGURE 3.40 Multi-site with Customer Associated Sites

This option relies on the concept of associations, which allows a customer to *associate* multiple Cloud Director Organizations to enable simplified administration of those resources, through a single pane of glass. As a result, when a customer Organization is a member of an association, users who are part of that Organization in Cloud Director can use the tenant portal to access resources at any member site. However, each member organization and its assets remain local to the site where they reside.

Sites and Site Associations: The concept of associations is made up of a number of different constructs:

- **Site Object:** A site object represents the local Cloud Director server group. A cloud provider's System Administrator, who has authority across multiple Cloud Director server groups can configure those server groups as an association of Cloud Director sites. This is the first step that must be completed prior to Organization Administrators being able to associate sites.
- **Associations of Organizations:** Associations of customer Organizations can take place only after the cloud provider site association has taken place. A customer Organization Administrator at any site can initiate the process of associating their Organizations across multiple sites.

To create an association between two organizations (for instance, Org-A and Org-B), the customer must be logged in as an Organization Administrator for both Organizations to retrieve its local association data and submit that data to the other organization. The process of associating two Organizations requires two complementary pairing operations. The first operation pairs Org-1 at Site-A with Org-2 at Site-B. The second operation is to pair Org-2 at Site-B with Org-1 at Site-A. Only when both pairings are complete is the association between Organizations finished, as illustrated in Figure 3.41.

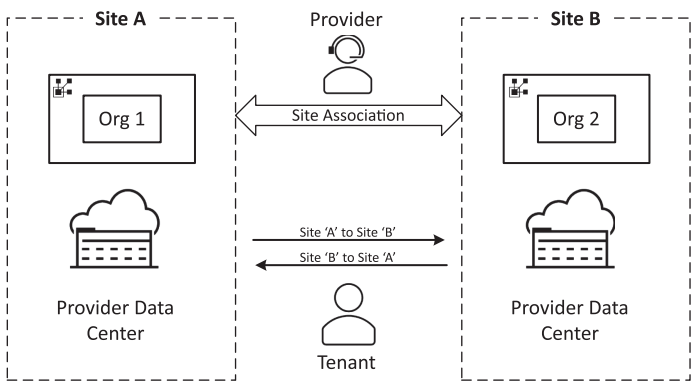


FIGURE 3.41 Organization Association Sequence

After a site association has been created, the system periodically retrieves the status of the remote site and updates that status in the local Cloud Director

database accordingly. This heartbeat process runs with the identity of a specific multisite system user, a local Cloud Director user account created in the System Organization during the Cloud Director installation. Even though this account is a member of the System Organization, it does not have system administrator rights. In fact, it has only a single right, `Multisite: System Operations`. This single right gives a specific permission to make a vCloud API request to retrieve the status of the remote member of a site association.

Organization User and Group Identities: In order to initiate the association of sites and the association of Organizations to take place, the same identity provider (IDP) must be agreed upon. User and group management for all identities across organizations, for all Organizations in the association, must be managed through this common IDP. With the exception of the System Organization, which must use the Cloud Director integrated IDP.

Each customer administrator can configure their IDP to generate user or group access tokens that are valid at either all member sites or are valid for only a subset of member sites. While user and group identities must all be in the same member organizations, user and group rights are constrained by the roles those users and groups are assigned in each member Organization. The assignment of a role to a user or group is local to that member organization, as are any custom roles that are created. The Organization Administrator controls access by Organization users and groups to assets at each member site. Association members are free to choose an IDP solution that best meets their requirements.

Also, while a site or Organization can form an unlimited number of associations with a peer, each association includes exactly two members, illustrated in Figure 3.41. Each site or organization requires its own private key. In effect, association members establish a trust relationship by exchanging public keys that are used to verify signed requests from one member to the other.

As illustrated in Figure 3.42, each site in an association is defined by a Cloud Director server group. In this architecture, each Organization in an association occupies a single site. However, this does not necessarily need to be the case. This same mechanism could also be used to scale multiple Cloud Director instances within a single site. For instance, either to provide a demarcation line across services offered, such as PCI-DSS or HIPAA, or this mechanism could be equally employed to scale out server groups at a single site, to facilitate hyperscale growth.

Federation: High-level Design: As outlined above, federating multiple Cloud Director Server groups is required for identity management across a multisite architecture, where associations are being used by customers to provide the delivery of an integrated multisite service. The identity provider (IDP) is a service that manages all user and group identities (other than local accounts). Cloud Director server groups that use a common identity provider are by result, federated.

A customer organization can define any identity provider that meets their requirements, even one that it shares with other cloud applications or

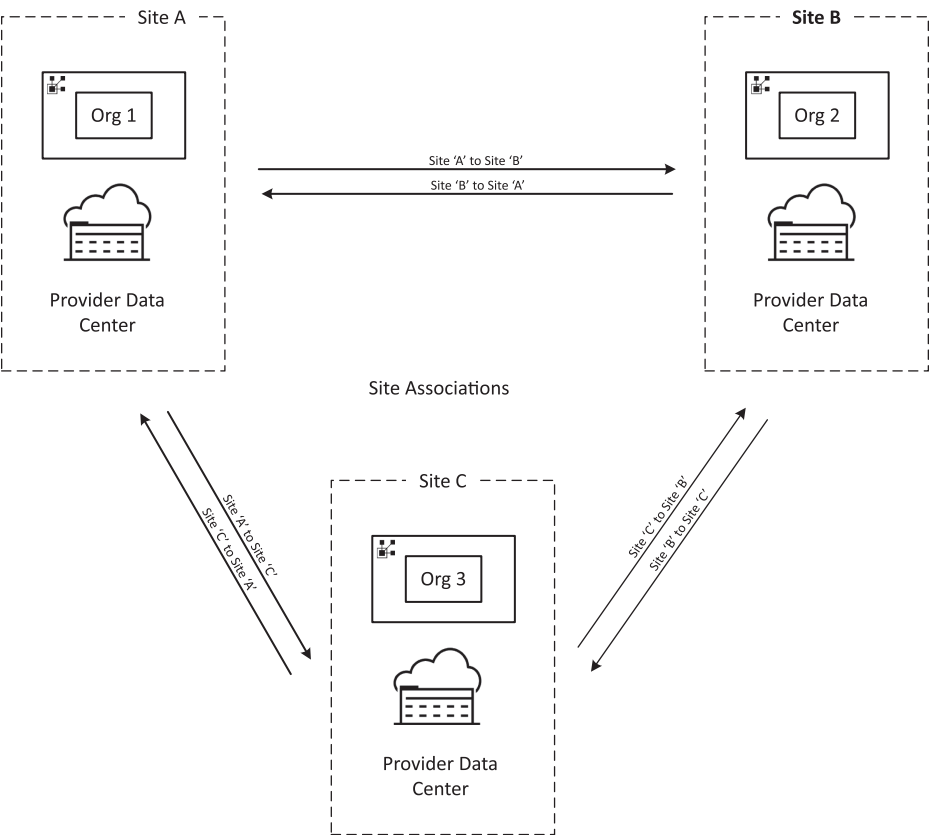


FIGURE 3.42 Association Mesh Between Multiple Member Sites

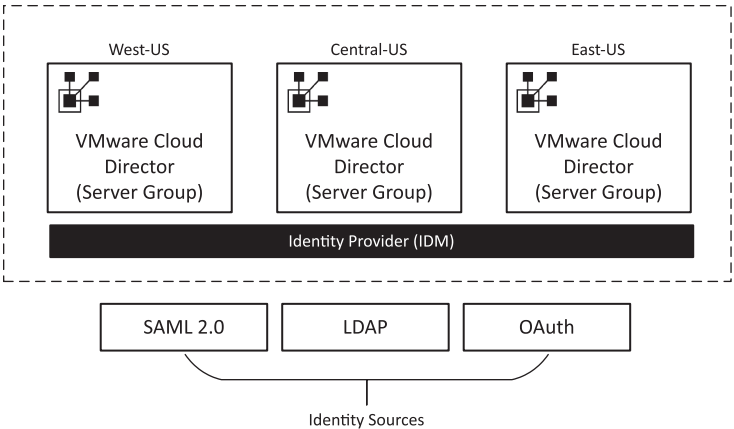


FIGURE 3.43 Cloud Director Federation

enterprises applications. This can enable a customer to provide access to multiple unrelated services, including the Cloud Provider Platform through a single set of credentials, which is often referred to as single sign-on. Cloud Director supports OAuth, SAML and Integrated IDPs. The integrated identity provider is a Cloud Director service that authenticates users who are created locally or imported from LDAP. All organizations that participate in an LDAP-based federated identity scheme must include an `OrgLdapSettings` element that specifies shared configuration parameters.

With an IDP configured, the customer's organization users authenticate to obtain a token, which they can then use to log in to the local and associated organizations, giving them the ability to manage multiple Cloud Director endpoints through a single user experience.

Load-balancer Requirements: The effective implementation of this type of multisite deployment requires the configuration of a load balancer that can distribute arriving requests to a single endpoint, such as `https://cloud-provider.com` to the endpoints for each member of the site association. For instance, `https://us-east.cloudprovider.com`, `https://us-central.cloudprovider.com`, and `https://us-west.cloudprovider.com`.

In addition, unless a site has only a single cell, the cloud provider must also configure the load balancer to distribute incoming requests across all of its cells, so that a request to `https://us-east.cloud.example.com` can be handled by `https://cell1.us-east.cloudprovider.com`, `https://cell2.us-west.cloud-provider.com`, and so forth.

More advanced load balancing options can also be considered, such as source address with regional weighting or per tenant global selection, with VMware NSX Advanced Load Balancer (Formerly Avi Networks), although these concepts are beyond the scope of this book.

Multisite Networking: Support for Cross-VDC Networking inside of Cloud Director builds on the model set out through Cross-vCenter NSX for vSphere. Therefore, Cloud Director requires a Cross-vCenter NSX configuration at the vSphere layer in order to configure tenant networking over the VXLAN overlay.

As illustrated in Figure 3.45 NSX for vSphere capabilities, including the Universal Transport Zone, Universal Logical Switch and Universal Logical router are integrated and surfaced up to the Cloud Director consumption layer. This results in the ability to stretch Layer 2 networks across Org. VDCs in different vCenters and Provider VDCs, perform static and provide DHCP IP address management for cross Org. VDC networks.

As we have addressed, there are multiple options for designing the Cloud Provider Platform across multiple data center sites, whether these have metro distances, and can be treated like availability zones, or span multiple continents in regions around the globe. This architectural flexibility greatly simplifies the user experience and allows customers to distribute applications and resources across data centers easily to improve application redundancy and provide integrated disaster recovery. In addition, allowing customers to

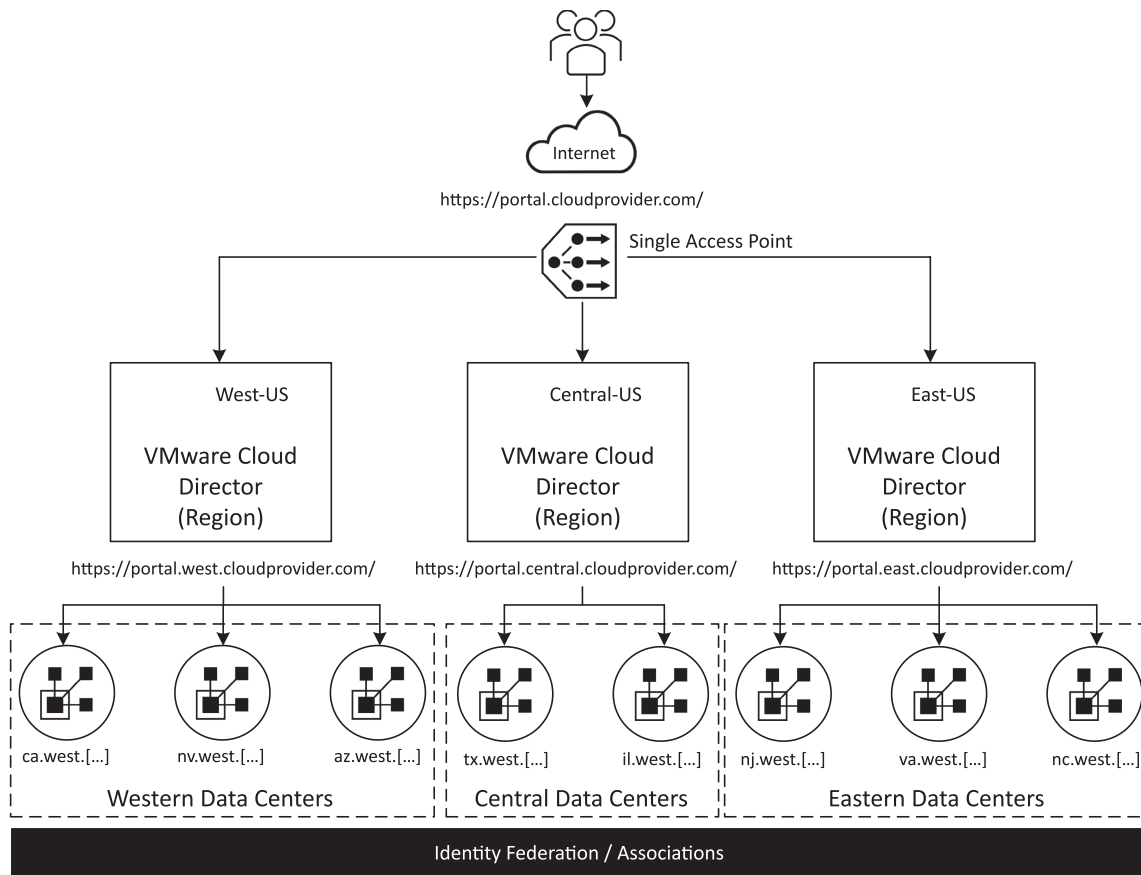
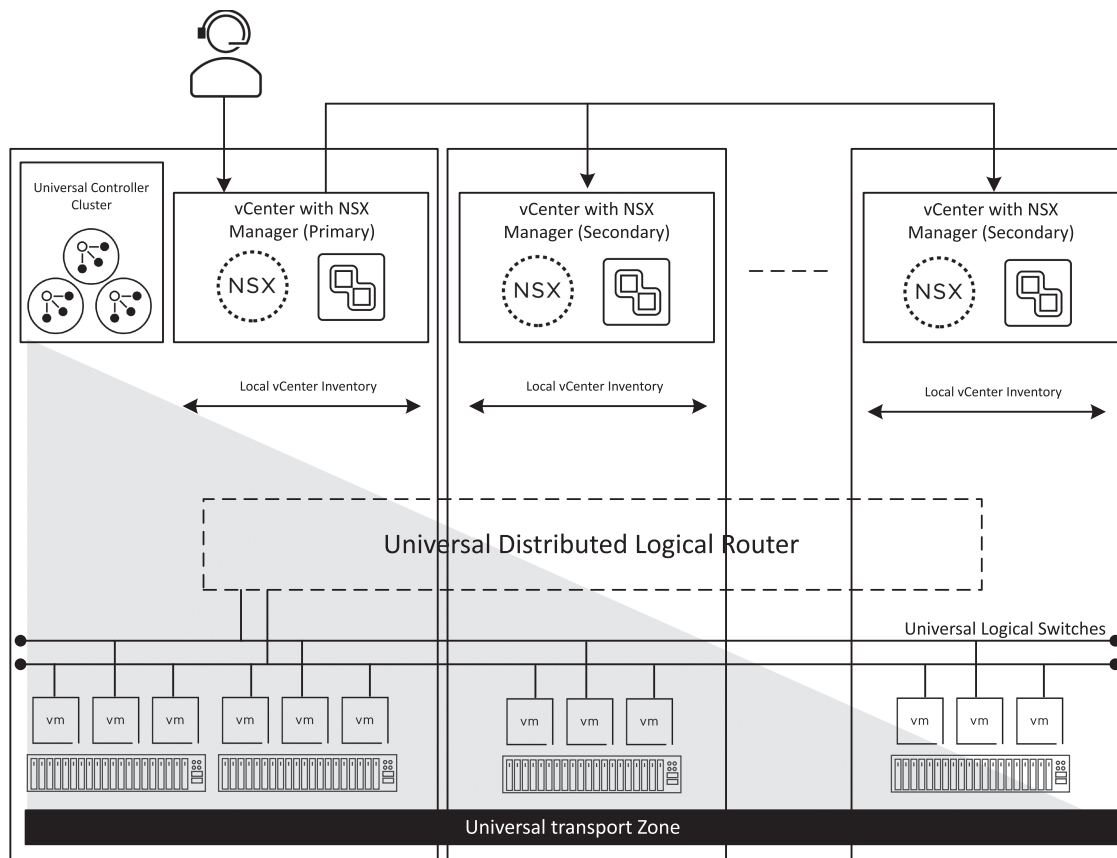


FIGURE 3.44 Multisite Cloud Provider Platform Load Balancing

**FIGURE 3.45** Cross-Site NSX for vSphere and vCenter Server

stretch networks across data centers greatly simplifies data center consolidation projects and migration, without having to re-IP applications.

CLOUD DIRECTOR CUSTOM BRANDING

Cloud providers are able to align the Cloud Director user interface with their corporate branding standards to create a fully custom experience for consumers. A number of aspects of the user interface components can be themed including, Cloud Provider Name, Portal Color (in RGB HEX format e.g., #EE00EE), Color Theme (default or dark), Logo (png file of max size 512 KB) and Browser Icon. These customizations are not limited to the tenant portal but can also be applied to the provider admin portal.

All theming and user interface customization are required to be carried out via the Cloud API through PUT requests. For instance, customizing the Portal Name, Portal Color and Portal Theme is illustrated in Figure 3.46.

```

PUT /cloudapi/branding

Headers:
Content-Type: application/json
x-vccloud-authorization: ...

Payload:
{
  "portalName": "Cloud Provider Name",
  "portalColor": "#323843",
  "selectedTheme": {
    "themeType": "BUILT_IN",
    "name": "Default"
  },
  "customLinks": [
    {
      "name": "Help",
      "menuItemType": "link",
      "url": "https://cloud.example.com/help"
    },
    {
      "menuItemType": "separator"
    },
    {
      "name": "API Documentation",
      "menuItemType": "link",
      "url": "https://code.vmware.com/apis/722/vccloud-director"
    }
  ]
}

```

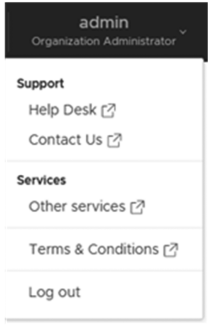


FIGURE 3.46 Customizing User Interface Code Sample & User Interface Custom Links

In addition to themes, as illustrated in the above code sample, cloud providers can also modify and add custom links to the two dropdown menus in the upper right of the tenant portal, without the need for custom extensions.

The Logo and Icon can also be customized through the use of the following PUT request to the Cloud Director CloudAPI endpoint respectively:

```
PUT https://cloud.example.com/cloudapi/branding/logo
```

```
PUT https://cloud.example.com/cloudapi/branding/icon
```

For cloud providers who have more advanced branding and theming requirements, VMware provides a toolkit to allow full customization of the look and feel of the portal using CSS themes within the Clarity framework. The toolkit itself is part of the VMware vcd-ext-sdk repository on github, available in the /ui/theme-generator folder.

<https://github.com/vmware/vcd-ext-sdk/tree/master/ui/theme-generator>

CUSTOM PORTAL (OPTIONAL) DESIGN FACTORS

Many of VMware's cloud provider partners provide a custom user interface to their customers in order to offer a wide range of services and solutions. This ranges from providing a light-weight shim layer, to simply pass-through user credentials to Cloud Director, vCenter Server or other tools, to a fully integrated experience, transferring the majority, if not all, of the platform functionality and business logic to the cloud providers custom user interface.

The core value for a cloud provider is delivering to consumers a differentiated customized offering through a custom portal, where they can not only keep their business branding and look and feel, but also potentially expose additional services.

Some cloud providers offer a basic login portal to allow a level of flexibility around presenting future services, which allow them to further differentiate their offerings. Even with the recent releases of Cloud Director, with its new flexible HTML5 user interface and login experience, that also provides a flexible and extensible solution to cloud providers and ISVs to build and customize integrations, some cloud providers continue to maintain and develop their own custom login portal for one clear and easily justifiable reason, differentiation!

However, there is nothing that matters more to cloud providers than the Cloud APIs, in other words a programmable surface. Therefore, this must mean that any custom login portal cannot inhibit access to the Cloud Director API. This means ensuring consistent access to services can be maintained alongside any custom login user interface, to ensure the delivery of services for users who wish to access the VCD environment programmatically through the API.

As illustrated in Figure 3.48, the additional cloud services required to access Cloud Director through the API surface might represent a whole range of additional value-added services that the cloud provider is offering. With these services, the same principles apply with regards carrying out functions, which can be delivered via custom orchestration, to similar backend systems or totally different components that are available within the cloud providers infrastructure.

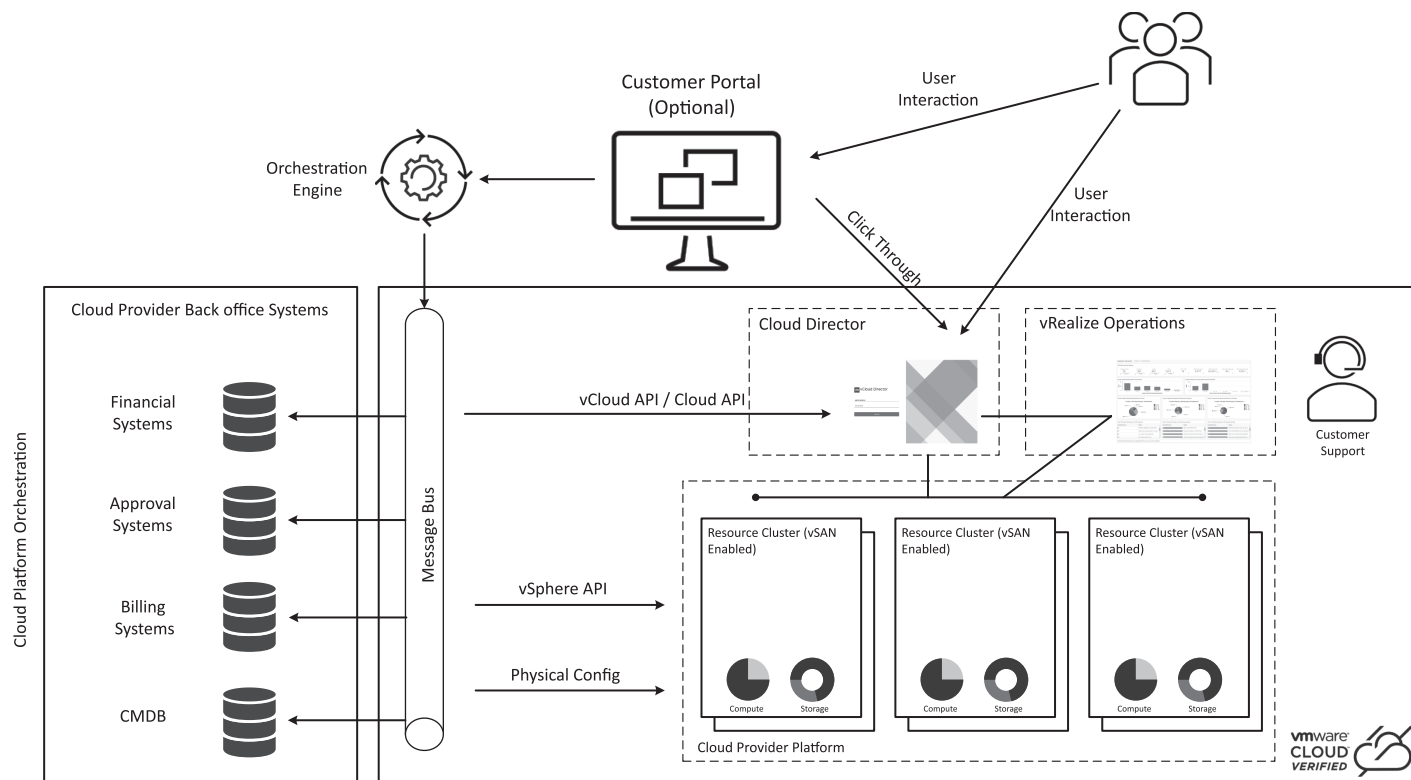


FIGURE 3.47 Custom Login Portal User Experience and Journey

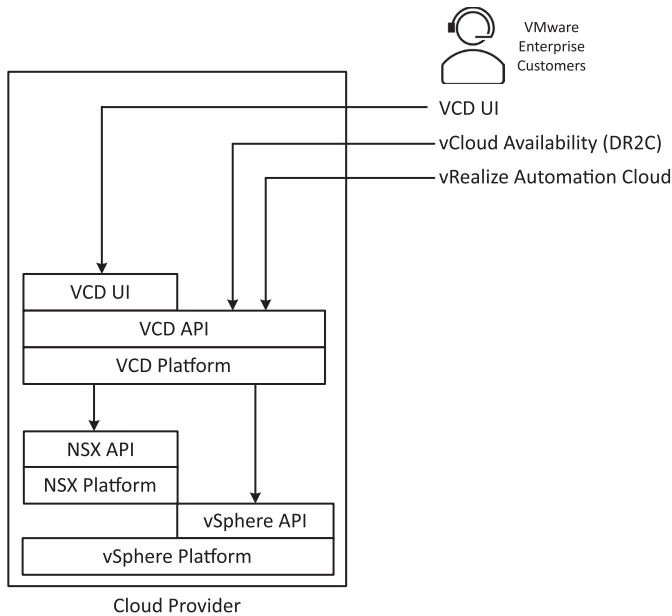


FIGURE 3.48 Access to Cloud Director API

The vCloud APIs mean that, as illustrated above, cloud providers can reach out to a large number of VMware customers, allowing them to connect to a *compatible infrastructure* using a range of tools that are easily accessible to them. These are just a few examples which demonstrate how the enterprise consumers can leverage a VCD-based public cloud. VMware, and the partner ecosystem as a whole, are launching a number of new tools that interact with the vCloud APIs natively. VMware HCX and other VMware Cloud Services such as VMware Cloud on AWS or vRealize Automation Cloud are good examples of tools consuming these programmable interfaces.

Therefore, one of the key principles which must be adhered to by cloud providers when designing a custom portal is that any services that go beyond the scope of Cloud Directors functionality, and will be exposed to end consumers, is how the VCD APIs should be exposed to ensure all required services can be delivered. This approach will typically mean that the custom portal cannot inhibit access to the VCD API or out of the box user interface, ensuring consistent access to VCD services can be maintained alongside any custom portal to deliver services for users who wish to access the VCD environment either via its user interface or programmatically through the API.

While one of the clear purposes behind implementing a custom portal is to deliver more advanced functionalities than are available through Cloud Director's portal, the design should not get in the way of the solution as a whole. Moreover, by the cloud provider building logic above the Cloud APIs,

this might obscure API access, which could as a result make the solution inconsistent.

Therefore, the cloud provider must design a solution that can meet the requirement of maintaining a consistent solution and ensuring that events that occur at the Cloud Director level are synchronized accordingly within the cloud providers backend systems. For instance, if an end consumer deploys a new virtual machine, the cloud provider would require a notification event, which might initiate logic to update a CMDB with the pertinent data. If this event cannot be captured above the VCD API in the custom portal, then how is the cloud provider tracking the activity? And while here we are discussing the update of a cloud providers CMDB, the same backend logic might be required for any system, such as ticketing or inventory systems.

For these reasons, developing the custom logic above the vCloud APIs is not typically a good approach. If the cloud provider wishes to offer VMware compatible services and solutions, the architecture must make the vCloud APIs as widely accessible as possible, and not be obscured by scripts or workflows.

Investing the engineering effort to build a custom portal that gives the cloud provider total flowability over the services they offer, how they present them to their consumers and a clear advantage in being able to provide differentiated and bespoke solutions that can separate themselves from their competitors, is a significant engineering and commercial undertaking.

As you might expect, there is no one common approach. Some cloud providers do not provide any custom user experience in front of Cloud Director. Some provide a basic landing page from which Cloud Director, and potentially other services can be launched, typically with some level of single sign-on. Other cloud providers deliver a rich, fully functioned replacement for the VCD user interface with additional custom PaaS and SaaS services already integrated.

Illustrated in Figure 3.49 is just one example of how a cloud provider can develop a custom portal solution to deliver multiple VMware-based cloud services. In this platform offering, the cloud provider is able to offer premium dedicated vCenter Services via VCF and the Cloud Director Central Point of Management (CPOM) feature, a lower cost and easier mechanism to manage dedicated vCenter platforms through Cloud Director. Also, this architecture is offering VMware Integrated OpenStack (VIO) and VMware Cloud Provider Hub solutions, delivering a single integrated platform, giving end customers a wide range of choice through self-service provisioning, and single pane of glass management, all with a common network backbone.

Figure 3.49 attempts to layer the various components that might be required to meet a cloud provider's requirements. For a comprehensive solution, such as the one illustrated, the cloud provider would need to build their own portal solution, or reuse and extend an existing portal that they want to complement with VMware Cloud Director-based services.

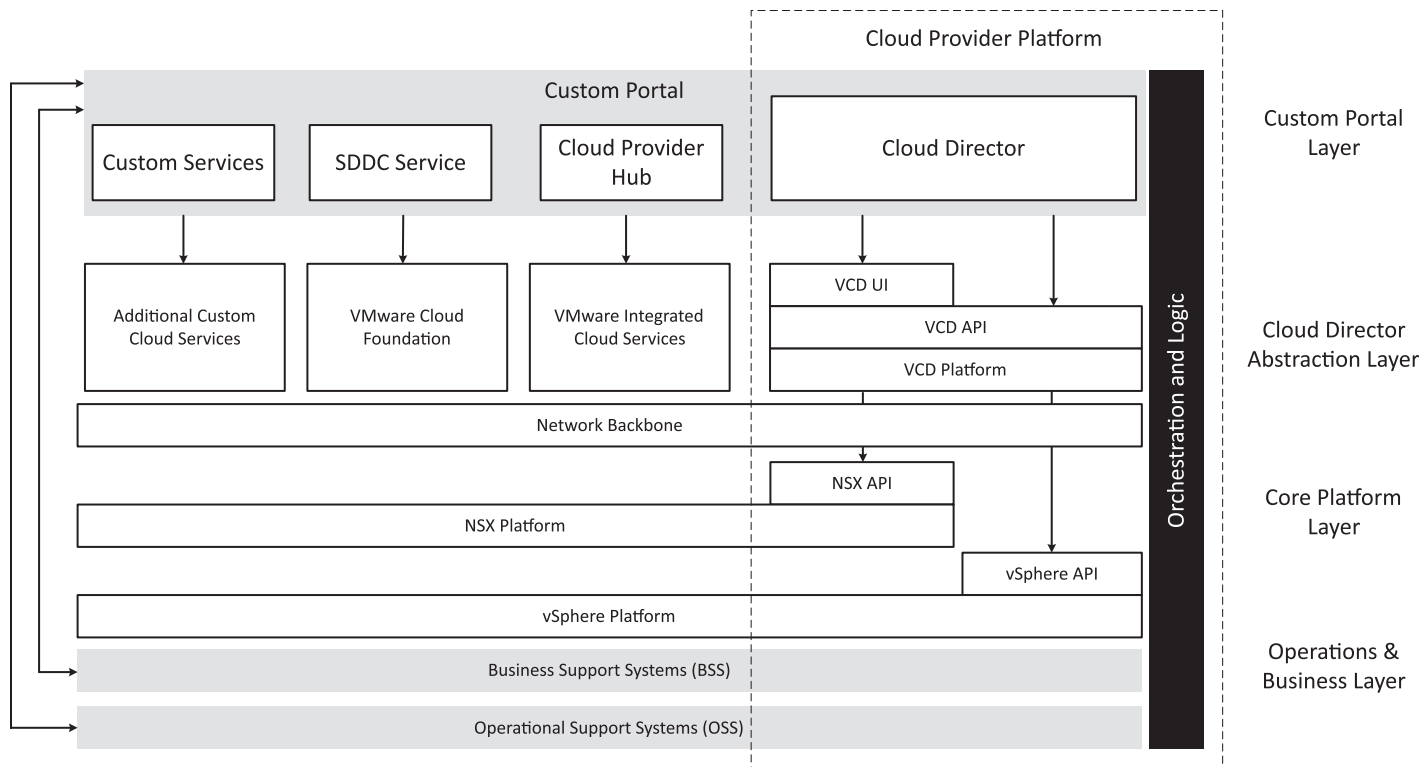


FIGURE 3.49 Custom Portal Architecture Example

However, despite the variety of solutions and service offerings available, the core value that the cloud provider is trying to deliver to consumers is that they want to have a differentiated and customized offering, which delivers choice to the customer, via a custom portal, where they can not only keep their business branding, look and feel, but also expose additional services. The next priority is typically the integration with their back-office systems, such as billing, asset management, and configuration management databases. The common approach to address updating these backend database systems is through orchestration tools, as well as potentially serverless *Function*-based capabilities. This topic will be addressed along with Cloud Automation in Chapter 7.

As we have addressed, it is strongly recommended that the native Cloud Director APIs and user interface are exposed in an unencumbered way to allow cloud consumers to interact with the core services offered by Cloud Director. This ensures the provider is not limiting their offering to potential consumers and are providing a platform that can be consumed by the maximum available market. In addition, the opportunity around differentiation via a custom portal is there for the cloud provider to take to market and is only limited by their own imagination or the cost of development.

Finally, this is of course just a very high-level overview of what is a very complex subject, with a wide range of technical design factors and business-level decisions that need to be addressed and fed into any proposed solution.

Chapter 4

Designing a VMware *Cloud Verified* Compute Platform

CLOUD VERIFIED OVERVIEW

For most cloud providers hybrid capability is a key strategy and includes disaster recovery, management of third-party clouds, and management of replication from a customer's infrastructure off-premises to the cloud provider's data center. To achieve these goals, the provider cannot afford to write a full orchestrator, automation, and customer portal from scratch. In working with VMware, the cloud provider can leverage the best practices and off-the-shelf tools, such as VMware Cloud Director, which offers the core framework for the Cloud Provider Platform, to avoid large development efforts and delays that would be involved in creating something from scratch.

As outlined in Chapter 2, the *Cloud Verified* badge signals to customers and prospective customers that the foundational cloud technologies and services being implemented by that partner are based on the most comprehensive VMware-based cloud platform possible. The *Cloud Verified* badge is built on the stability, integration and interoperability of the core compute, storage and network building blocks, which gives customers the freedom and flexibility to run their applications where they want, in the cloud, in their own data center or a combination of the two. The VMware *Cloud Verified* mark signifies to customers that this cloud service provider, distinguished by the logo, delivers the full value of VMware Cloud infrastructure including compatibility, choice and control.

The Cloud Provider Platform is a reference design consisting of a pre-validated, pre-tested solution of VMware infrastructure to address cloud provider deployments and is the foundation of the VMware *Cloud Verified* program. This validated design guidance is inspired by best practices outlined

in VMware's Validated Designs (VVD) but is designed for cloud providers specifically.

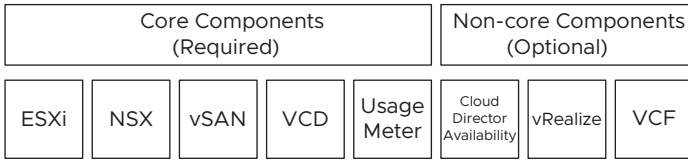


FIGURE 4.1 Cloud Provider Platform – Core and Non-core Components

The VMware Validated Design for Cloud Providers is a pre-validated, pre-tested solution of VMware infrastructure to address cloud provider deployments, which is focused on interoperability, performance and multi-regional scale. The design factors addressed in chapters 5, 6 and 7 is provided in addition to the guidance given in the validated design documentation.

The core Cloud Provider Platform is designed to deliver on-demand access and control of network bandwidth, servers, storage, and security while maximizing asset utilization. Specifically, the underlying VMware software-defined solution allows cloud service providers to integrate all the key functionality that enterprise business customers are demanding, including:

- A self-service portal and API for end user and administrative provisioning (VMware Cloud Director)
- A library of available compute services (Private, Public, Marketplace and Bitnami Catalog)
- Rapid, precise, and automated service provisioning (VCD IaaS Engine and vRealize Orchestrator Extensions)
- Multi-tenant monitoring, reporting, and billing integration capabilities (vRealize Operations and Tenant App)

The Cloud Provider Platform's software-defined architecture also helps cloud service providers offer highly customized services that meet the complex demands of their business customers. These features help the cloud provider's business customers maintain control over infrastructure and management, while realizing the benefits of service provider economies of scale through flexible and commercial delivery models.

The focus of this book is to assist the cloud provider community in creating solutions that will drive a scalable compute foundation in the infrastructure, and to deliver value-added services that create new revenue streams. Through the Cloud Provider Platform, end customers can benefit from their providers' ability to meet agreed upon service-level agreements and from the rapid provisioning of new services in anticipation of ongoing and changing business requirements.

The cloud provider must be able to position the various components within an applicable service, to identify their suitability to a consumer's application or requirements. Whether a consumer uses a virtual data center, hosted private cloud, or a blend of both private and public service offerings, these should form part of an integrated solution, regardless of the underlying infrastructure. To achieve this, the provider must standardize across the infrastructure platform and layer on clear, repeatable, and supportable orchestration and automation to manage these on a per-resource, per-consumer, and per-platform level.

COMPUTE PLATFORM RESOURCES

The software-defined compute resource for the Cloud Provider Platform is delivered by the VMware ESXi hypervisor, which enables service providers to build an enterprise-grade, scalable, multi-tenant platform for complete compute service lifecycle management.

The hypervisor replaces the traditional operating system, such as Microsoft Windows or Red Hat Linux, and provides the ability to create a number of virtual servers on the same hardware. This technique provides multiple benefits, one of which is isolation of the virtual server from the underlying hardware, allowing hardware resources to be utilized more fully, and providing the mechanism for dense server consolidation that is the basis for enabling cloud-based computing.

Employing this VMware-based software-defined computing layer gives the cloud provider the ability to seamlessly deliver highly scalable, on-demand infrastructure services to customers, while reducing power, saving space, maintaining reliability, and reducing the overall cost to serve.

While a virtual machine is a logical entity, to its operating system and end user, it appears to be a physical host with its own CPU, memory, network controller, and disks. All virtual machines running on a host share the same underlying physical hardware, but each consumes its share in an isolated manner. From the hypervisor's perspective, each virtual machine is a discrete set of files that include a virtual machine configuration file, data files, and so on.

Scalability and Designing Physical Resources

With any virtual infrastructure that is required to scale extensively across hundreds or even thousands of hosts, provide petabytes of storage, and support large complex networks, extensibility is a key factor. For a successful cloud provider deployment, scaling a large physical platform while maintaining control, compliance, and security is critical. Taking a predefined building block approach to this type of architecture, from day one, is paramount in planning for scalability.

In addition, the configuration and assembly process for each system should be standardized, with all components installed identically. Standardizing the

configuration of the physical components is critical in providing a manageable and supportable infrastructure at scale by eliminating variability, which in turn, reduces the amount of operational effort involved with patch management and helps provide a flexible building block solution that meets a service provider's requirement for elasticity.

While the configuration and scaling are likely to be hardware vendor dependent, this model should form part of the Cloud Provider Platform design, which will be linked to commercial growth estimates of the services offered.

For instance, the example in the following figure represents one possible building block scenario where the cloud provider for the compute clusters is employing a hyper-converged model, with VMware vSAN ready (HCI) compute nodes deployed.

From a physical platform perspective, each HCI Cloud Director Resource Group is made up of 96 converged ESXi compute hosts configured as four 24-node VMware vSphere clusters split equally across server cabinets. In addition, a four-node vSphere management cluster, distributed across hosts, as illustrated in Figure 4.3. Each HCI rack also houses two 48-port 25-GbE *leaf* switches and two 1-GbE IPMI management switches for out-of-band connectivity. Each of these clusters is designed to provide a fault domain for compute, network, and storage.

The number of HCI Clusters in this example can be scaled out accordingly, depending on hardware, software, and power limitations. The vSphere components in each HCI Cluster are managed by a single VMware vCenter Server instance.

In addition to the HCI Clusters for customer workloads, this example architecture requires a management solution to house the Cloud Provider Platform and VMware NSX Edge components required to support all the customer workload clusters within the cloud service providers data center. In this example architecture, the Cloud Provider Platform management cluster and NSX Edge Cluster are also hosted on vSAN ready nodes.

Figure 4.4 illustrates the Logical Data Center Layout of this scaled-out approach to building highly scalable infrastructure.

However, there is no single solution to scaling the compute, storage, and network resources available from each component layer of this building block architecture. During the cloud provider's design phase, several factors will play an important role in designing the building blocks including, but are not limited to:

- Expectations regarding services and provider growth
- Hardware availability and lead times
- Physical hardware scalability limitations (such as with system management tools)
- Capital expenditure and hardware depreciation considerations
- Data center power, space, and cooling limitations

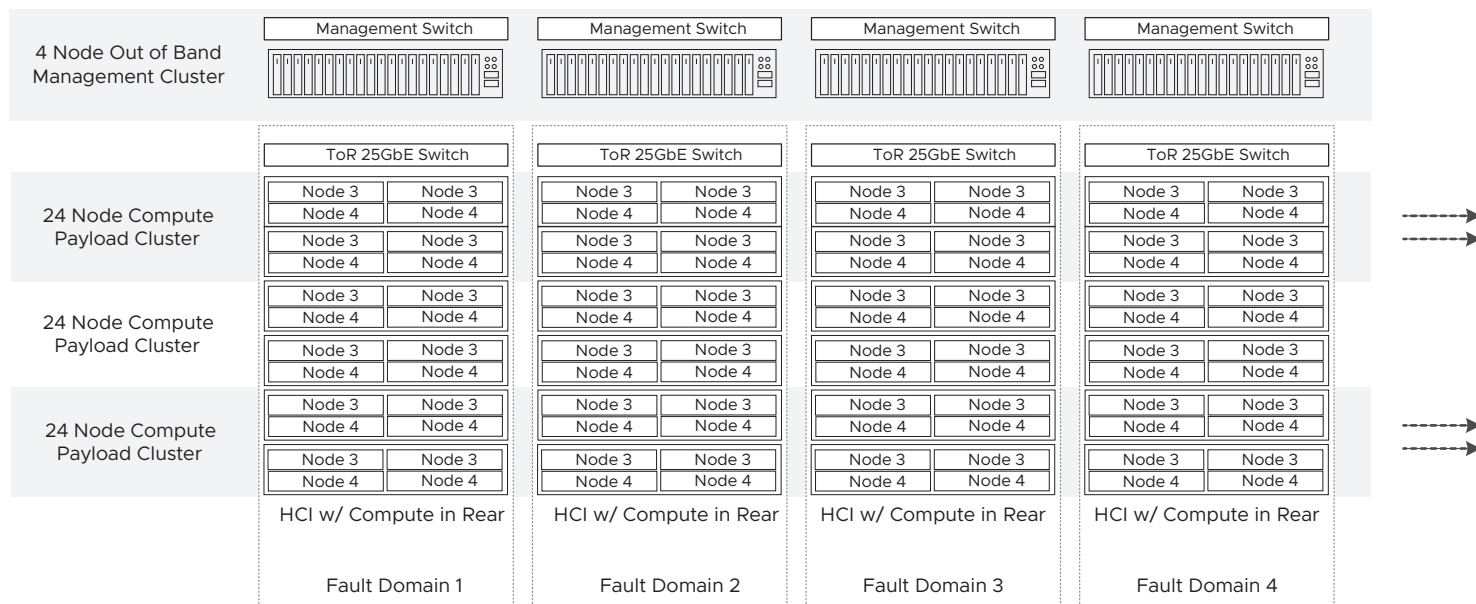


FIGURE 4.2 Sample HCI Cluster Architecture

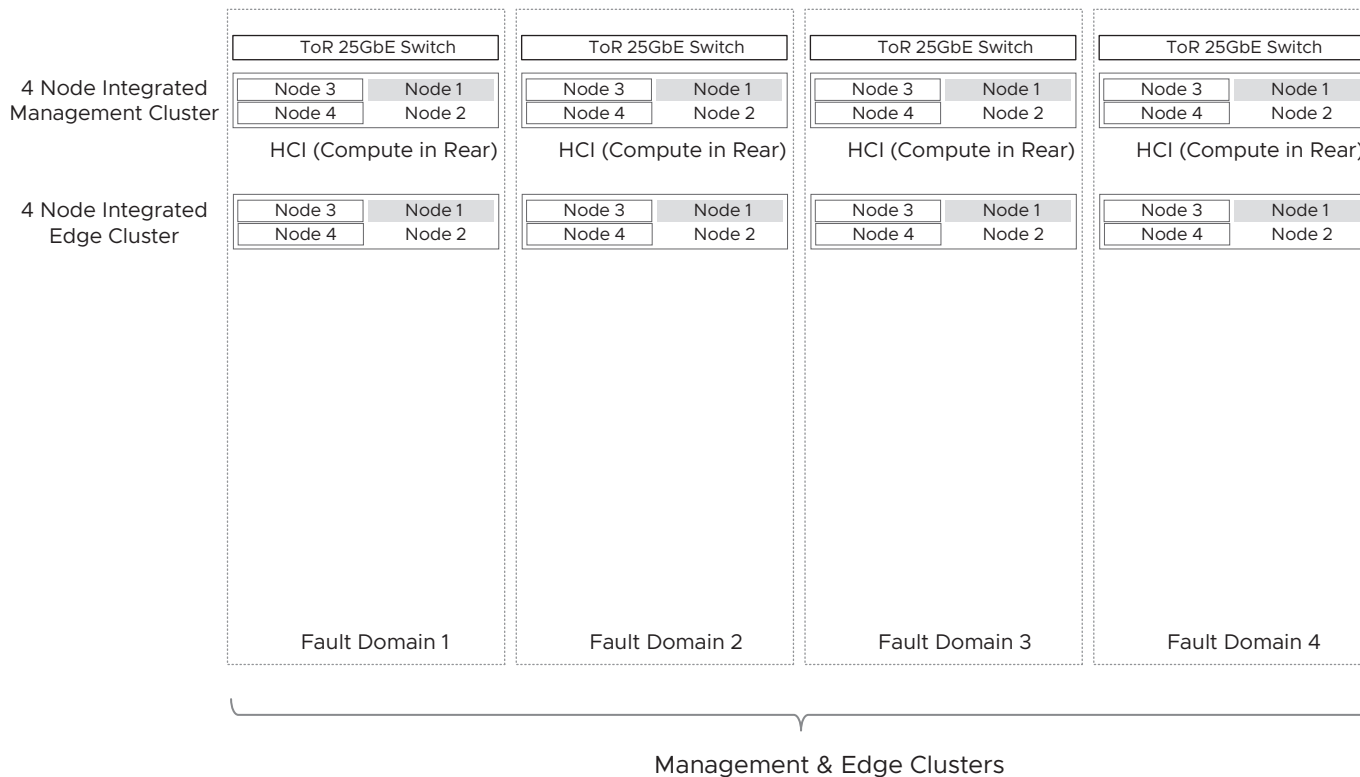


FIGURE 4.3 Sample Management Cluster Architecture

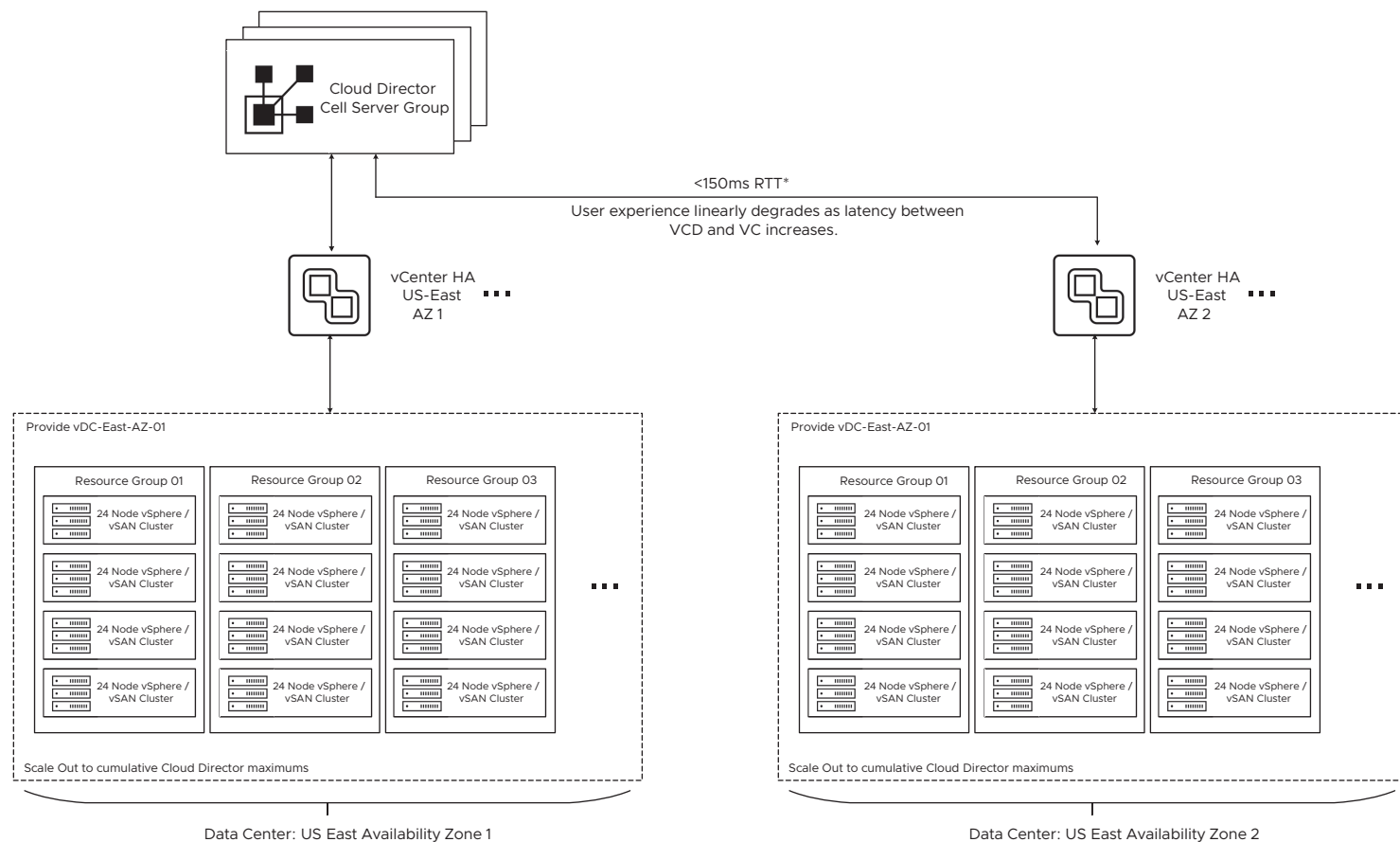


FIGURE 4.4 Planning for a Highly Scalable Deployment

Infrastructure Protection & Availability

There are two key considerations for protecting the building blocks that make up the physical infrastructure—local availability, referring to the single physical data center, and cross-site availability, referring to the business continuity processes that are undertaken when either a partial or complete data center incident takes place on the first site.

Local availability is achievable through employing a number of technologies and strategies to prevent service disruption when either a single component or multiple components fail. The following tactics can together form a strategy to mitigate these types of component failures:

- Redundant hardware to eliminate single points of failure.
- Employing vSphere clustering and HA technologies to return services to an online state in case of host outage.
- Use of snapshots to quickly return to a previous healthy state in response to application failure or corruption resulting from operating system of application updates.
- Securing data both onsite and off-site to prevent data loss through traditional backup and replication mechanisms.

In addition, in order to maintain enterprise level service-level agreements and provide consistently repeatable scalable infrastructure, other data center design factors should also be considered. These include, but are not limited to:

- Power Distribution (the use of uninterruptable power supplies (UPS))
- Cable Management (cable management best practices for easy access to device components during maintenance windows and ensure proper cooling efficiency)
- Environmental Considerations (to achieve maximum efficiency and optimization in this solution, an appropriate cooling design, power distribution, and grounding configuration must be employed)
- Eliminating Single Points of Failure (high availability at the core of every design decision, therefore, eliminating single points of failures)

Comparing Blade Servers with Rackmount Servers

With the rapid adoption by cloud providers of VMware vSAN hyper-converged-based storage, a new dimension has been added to the blade versus rackmount debate. Due to their small form factor, blade systems do not typically provide adequate local disk space to deliver enough capacity, and as such, make vSAN with blade solutions an unviable option. For this reason, purchasing a new blade system for cloud platform projects typically removes the option of either deploying vSAN as part of the initial implementation, or

later repurposing hardware to support vSAN distributed datastores as part of, or as a complete, storage solution.

Outside of VMware vSAN, the long-standing discussion regarding blade servers compared with rackmount server systems for use with cloud platforms has no definitive correct answer because every use case is different, and no one solution has superiority over the other.

The likelihood is that each service provider has their own preference when it comes to their cloud platform compute hardware. However, it is important to be able to work through the advantages and drawbacks of each option when it comes to server form factor.

Blades have the density advantage, and with rack space often provided at a significant cost, higher density is a good thing. Centralized management also provides a major advantage. A single interface where you can see and work with all the blades, creating server profiles, adding networks, adding SAN connectivity, configuring VLANs, and managing power provides a serious advantage over rackmount servers. The ability to perform all these tasks, and many more, on hundreds or thousands of blades from a single pane of glass has many operational advantages.

On the downside, blades can be expensive and require special power connections. You are also locked into a single vendor when you buy a chassis or enclosure. The initial setup of blade systems can be time consuming and require additional skills and experience and some blade systems add another layer of network/storage complexity to the overall design. Their setup is often a manual process although once done, adding additional blades to an existing chassis is often extremely simple, through cloning the configuration for an existing system.

Though blade system density is higher, there is often a maximum of two enclosures per cabinet because of the weight of a full enclosure. Two full enclosures can weigh over 900 pounds, so it is necessary to spread the load across additional floor space as well as spread the power consumption across multiple 32-amp connectors.

In addition, if you consider the commodity-based hardware concept of the software-defined data center, the vendor tie-ins associated with blade systems are not typically conducive to this model.

Rackmount servers are independent, which means that you can insert them into any 19-inch server cabinet without the need for special wiring, power, or enclosure. Rackmount systems are often more affordable than their blade counterparts and typically come equipped with KVM access. Older rack systems can also generate much more heat and pull a great deal of power compared to modern blade systems.

There is no single correct answer or a single ultimate architecture for a Cloud Provider Platform solution. In reality, no cloud provider data center has just one compute form factor in it. Most have every kind of system imaginable, with systems from every vendor, often dating back decades.

Whichever compute architecture type you choose, the aim is to create a building block approach for the platform that will enable both horizontal and vertical scaling of resources. Standardizing the configuration and using a consistent hardware platform helps to provide a manageable and supportable infrastructure by eliminating variability. Maintaining this consistent and standardized building block approach to compute resources:

- Simplifies capacity planning and scaling-out
- Simplifies automated installation, patching, and configuration of hosts
- Simplifies troubleshooting, fault finding, and configuration management
- Preserves VMware vSphere vMotion compatibility across the entire platform
- Minimizes unused resources in VMware vSphere High Availability (vSphere HA) configurations

Compute Host Sizing – Scale-Up Compared with Scale-Out

When it comes to sizing hosts to meet the service providers requirements for CPU, memory, and network throughput, there are two basic strategies:

- The “scale-up” approach (fewer hosts, but larger in capacity)
- The “scale-out” approach (greater in number, but lower in capacity)

The scalability of the hypervisor platform will determine the level of consolidation that can be achieved, as well as the underlying design of the compute layer. When examining host hardware choices, there are several key characteristics that influence the design:

- The number of physical CPUs (cores) and the virtual CPU-to-physical CPU ratio, which will impact the number of tenant virtual machine workloads.
- The amount of supported memory.
- The expandability, including the number and type of I/O cards for network and storage connectivity—the key factor in the amount of bandwidth available to a host.
- The number and size of storage devices, storage groups and storage controllers are particularly relevant in a hyper-converged or vSAN architecture.
- The virtual machine consolidation ratio will be the definitive limit for host scaling.

Typically, this hardware choice is transparent to the consumers. From the perspective of the tenant’s virtual machines, the host must provide enough virtual CPUs and virtual memory to facilitate the virtualization of the guest operating system and running applications.

If the design includes the use of the Sphere HA mechanism (which will almost certainly be the case), be sure to weigh the hosts’ capabilities against

the clusters' capabilities. For instance, because a limitation of 64 hosts per cluster exists in vSphere 6,7 the amount of resources available will be capped, although as discussed further in the cluster design section of this chapter, this is not a factor that will typically limit any design. The reason for this is that it is considered a good practice that standard compute configurations for the host design include the same blade or rackmount server type within each cluster, and each cluster should, if possible, span across chassis' or server cabinets to maximize availability. Therefore, very large clusters might not offer the levels of availability that a more determined building block approach to cluster design will ensure.

Determining Host CPU and Memory Requirements

There are two considerations for compute sizing—processing requirements and memory requirements. On a dynamic Cloud Provider Platform, designing for empirical data with regard to CPU and memory requirements is unlikely to be possible. Instead, sizing will typically be based on the anticipated workloads that will run on the infrastructure.

Typically, CPU and memory requirements are defined during the requirements analysis and specific metrics are based on anticipated workloads and expected growth for the lifecycle of the platform. From this information, the architect can determine the aggregate CPU and memory requirements. When designing an environment, look at current requirements and design a solution that will allow the environment to grow without re-designing the platform every time you need to add capacity.

The processing capabilities of each compute node can be determined by multiplying the number of cores times the speed of the processors. For sizing purposes, we typically want to plan for no more than 80 percent utilization of the processing capabilities.

To fully utilize the processing capabilities that the compute nodes offer, systems must be configured with sufficient memory. In recent years, memory costs have been reduced and the memory density of hardware has increased. With these new extended memory capabilities of the servers, in terms of multi-terabytes of memory, the balancing point between CPU and memory capacity might now favor denser multi-core CPUs.

After the total memory requirements, established during the planning process, have been calculated, divide by the memory per host to determine the number of compute nodes that are required. For this calculation, you can usually ignore the overhead of the hypervisor, which is minimal when you consider the available memory and the efficiency of consolidation.

The final predicted values should not only take into account the aggregate CPU and memory requirements, but also the service provider's desired maximum utilization thresholds, anticipated growth factors (as expressed by business stakeholders), and any expected benefit from virtualization technologies, such as Transparent Page File Sharing (TPS), if enabled.

Also, consider that if TPS is enabled as part of the platform design, an important design note is that estimated savings from memory sharing should intentionally be kept low where the guest operating systems will be 64-bit and as such, large memory pages will be used.

Finally, when sizing the platform, it is important to document all assumptions about how you calculated the service provider's requirements. Because these are estimations, in most cases, based on theoretical workloads, it is important that the business stakeholders accept the methodology that has been employed and understand the confidence level in the information that was used in assessing the compute platform requirements.

Determining an Appropriate Virtual CPU-to-Physical CPU Ratio

In a virtual machine, processors are referred to as virtual CPUs or vCPUs. When the vSphere administrator adds vCPUs to a virtual machine, each of those vCPUs is assigned to a physical CPU (pCPU), although the actual pCPU might not always be the same. There must always be enough pCPUs available to support the number of vCPUs assigned to a single virtual machine or the virtual machine will not boot.

However, one of the major advantages of vSphere virtualization is the ability to oversubscribe, so there is of course no 1:1 ratio between the number of vCPUs that can be assigned to virtual machines and the number of physical CPUs in the host. For every workload beyond a 1:1 vCPU to pCPU ratio to get processor time, the vSphere hypervisor must invoke processor scheduling to distribute processor time to virtual machines that need it. The higher the ratio becomes, the higher the performance impact will be, because you must account for the length of time that a virtual machine must wait for physical processors to become available. The metric that is by far the most useful when looking at CPU oversubscription, and when determining how long virtual machines must wait for processor time, is CPU Ready.

The vCPU-to-pCPU ratio you should aim to achieve in your design depends upon the application you are virtualizing. In the absence of any empirical data, which is generally the case on a heterogeneous cloud platform, it is a good practice, through the use of templates and blueprints, to encourage your service consumers to start with a single vCPU and scale out when it is necessary. While multiple vCPUs are great for workloads that support parallelization, this is counterproductive in the case for applications that do not have built in multi-threaded structures. Therefore, while a virtual machine with four vCPUs will require the hypervisor to wait for 4 pCPUs to become available, on a particularly busy ESXi host with other virtual machines, this could take significantly longer than if the VM in question only had a single vCPU. This performance impact is further extended as the vSphere ESXi scheduling mechanism prefers to use the same vCPU-to-pCPU mapping to boost performance through CPU caching on the socket.

Cloud providers should, where possible, try to educate their consumers on provisioning virtual machines with the correct resources, rather than indiscriminately following physical server specifications that software vendors often refer to in their documentation. If they have no explicit requirements, advise the consumer to start with a single vCPU if possible, and then scale up once they have the metric information on which to base an informed decision.

To ensure that the vCPU-to-pCPU ratio is optimized and you are able to take full advantage of the benefits of over provisioning, in an ideal world you would first engage in dialog with the customer and application owners to understand the workload prior to allocating virtual machine resources. However, in the world of shared platform and multi-tenant public cloud computing, where this is unlikely to be the case, and the application workload will be unknown, it is critical to not overprovision virtual CPUs, and scale out only when it becomes necessary. Employing VMware vRealize Operations, with the Management Pack for VMware Cloud Director, as a monitoring platform that can trend historical performance data and identify virtual machines with complex or mixed workloads is highly beneficial and its capacity planning functionality assists in determining when to add pCPUs. Again, CPU Ready Time is the key metric to consider as well as CPU utilization. Correlating this with memory and network statistics, as well as storage I/O and virtual disk I/O metrics, enables the cloud provider to proactively avoid any bottlenecks and correctly size and scale the Cloud Provider Platform to avoid performance problems or overprovisioning.

The actual achievable ratio in a specific environment depends on several factors:

- The vSphere version—The vSphere CPU scheduler is always being improved. The newer the version of vSphere, the more consolidation that should be possible.
- Processor age—Newer processors are much more robust than older ones. With newer processors, service providers should be able to achieve higher vCPU:pCPU ratios.
- Workload type—Different kinds of workloads on the host platform will result in different ratios.
- As a guideline, the following vCPU:pCPU ratios can be considered a good starting point for a design:
 - 1:1 to 3:1 is not typically an issue
 - With 3:1 to 5:1, you might begin to see performance degradation
 - 6:1 or greater is often going to cause a significant problem for VM performance

The use of vCPU-to-pCPU ratios can also form part of a tiered service offering. For instance, offering lower CPU consolidation ratios on higher tiers of service or the reverse. Use Figure 4.5 as a starting point to calculate consolidation ratios in a Cloud Provider Platform design but remember for every

single rule there are exceptions and calculating specific requirements for the customer is key to a successful architecture. This figure is for initial guidance only.

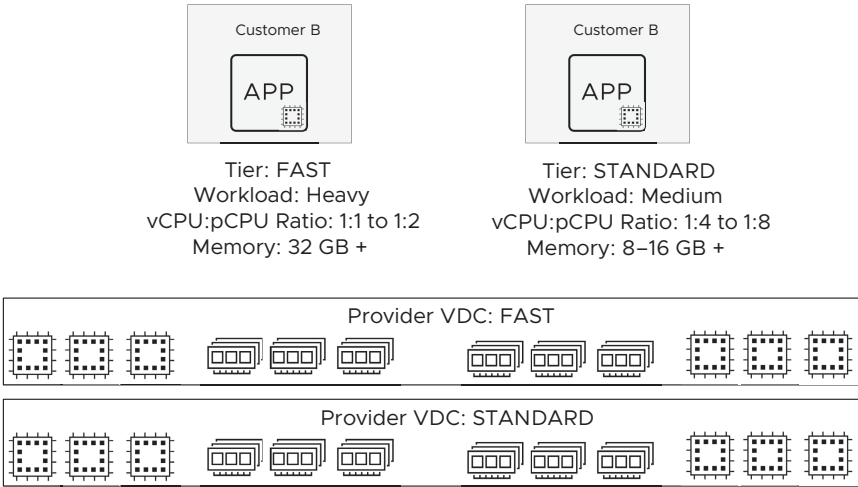


FIGURE 4.5 Virtual CPU-to-Physical CPU Ratio

As a general guideline, you should attempt to keep the CPU Ready metric at 5 percent or below. For most types of platforms, this is considered a good practice.

Performance Tuning with NUMA

When considering virtual machines *Compute Policy* ‘T-Shirt’ sizing templates, keep non-uniform memory access (NUMA) recommendations in mind to optimize platform performance.

Most modern servers have CPUs with directly attached memory. The CPU scheduler is aware of this physical architecture when it is available in the hardware, and targets processes to run on a CPU with fast access to the local memory, as shown in Figure 4.6. Because the process of accessing memory on a different CPU is not as efficient, the scheduler tries to keep processes on the same physical CPU to take advantage of the CPU cache and local memory (unless you manually override this functionality). Some NUMA systems provide the ability in the BIOS to disable NUMA by enabling node interleaving. Typically, you will get optimum performance by disabling node interleaving (i.e., leaving NUMA enabled). When configuring the compute node system BIOS, maintaining consistent settings that conform to the manufacturer’s recommendations is the starting point for any design. It is also noteworthy that the default hardware BIOS settings provided out-of-the-box on servers might

not always be the optimal choice to maximize performance when the ESXi hypervisor is configured.

VMware recommends not assigning more vCPUs to a virtual machine than a physical CPU has cores. This recommendation should be considered when designing the Cloud Director *Compute Policies* for service catalogs and virtual machine ‘T-shirt’ sizing. Employing this recommendation as part of your service design means that the scheduler will not split vCPUs across multiple CPUs logical cores.

Likewise, in the service design, do not assign more memory to a virtual machine than is available to a single NUMA node. If necessary, check the server configuration to see how much memory each CPU can directly access. When under CPU contention, the scheduler might move vCPUs to other NUMA nodes, which will have a temporary performance impact.

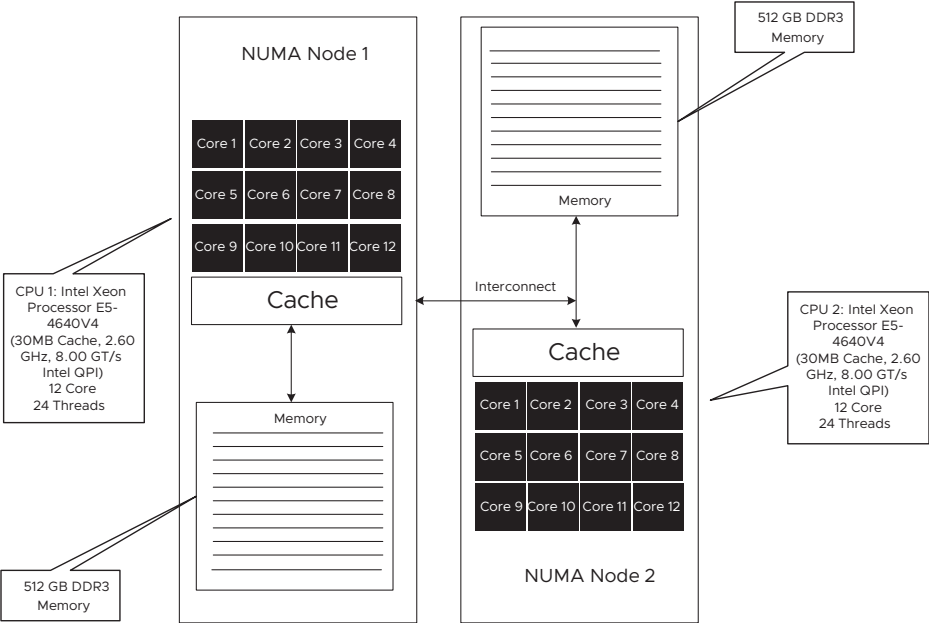


FIGURE 4.6 NUMA Architecture

For the Cloud Provider Platform architecture, ensure to examine the service design for high performance, oversized virtual machines and make the recommendation to stakeholders to stay within the size of a single physical NUMA node for vRAM and vCPU for *Compute Policy*-based ‘T-Shirt’ sizing of templates. This way local memory with the highest speed access is employed, with CPU and memory within the virtual machine being maintained from a single NUMA node, where possible.

vNUMA: If a virtual machine has more than eight vCPUs, virtual non-uniform memory access (vNUMA) is automatically enabled (although this behavior can be modified if required). Virtual NUMA is especially useful with large, high-performance virtual machines or multiple virtual machines within a vApp or application blueprint. With vNUMA awareness, the virtual machine’s memory and processing power is allocated based on the underlying NUMA topology, as outlined previously, even when it spans more than one physical NUMA node.

As the ESXi hypervisor creates a NUMA topology that closely mirrors the underlying physical server topology, it allows the guest operating systems to intelligently access memory and processors in the single NUMA node. A virtual machine’s vNUMA topology will mimic the topology of the host on which it starts. This topology does not adjust if the virtual machine migrates to a different host, which is one of the reasons why using consistent hardware building blocks is the recommended approach for cluster design.

Capacity Management for Cloud Service Providers

Achieving true elasticity is a challenge that every service provider faces. Allowing tenants to scale-up and scale-down their infrastructure, based on

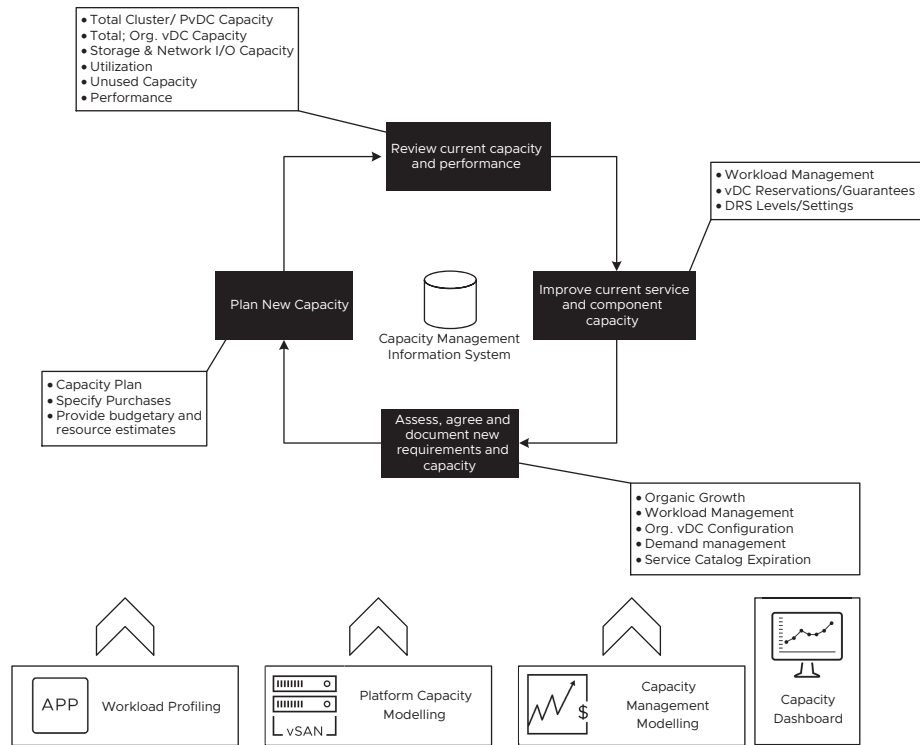


FIGURE 4.7 Capacity Management Planning

demand requirements, takes carefully considered capacity planning, hardware contingency, and customer trend metrics. The ability to forecast demand and either allocate or reserve resources for future needs is a key requirement for the Cloud Provider Platform.

The first crucial factor in capacity management is to get the architecture correct—the wrong design makes management and operations unnecessarily complex.

It is also important that the cloud provider is constantly aware of their usable capacity and lead-time for expansion of compute, storage, and network resources. The final step for capacity management is configuring VMware vRealize Operations, with appropriate management packs, to address and alert, based on the following core questions:

- What is our typical virtual machine profile?
- How many more virtual machines can we support based on actual workload as opposed to theoretical workload?
- When do we need to add capacity for compute, storage or network?

Automating Host Deployment

Deploying ESXi hypervisor hosts into the data center can be achieved in a number of different ways, which gives us a number of deployment scenarios to consider, depending on the hardware choice made. The first key decision is whether to opt for a traditional stateful installation of the ESXi hypervisor or to go stateless. The second pivotal decision to make is the destination media, where the hypervisor binaries will be installed, assuming that they are not being stored only in memory as part of a stateless implementation. Options include:

- ESXi is installed on to local hard disks (typically configured as a RAID 1 mirror)
- Removal media, such as SD or a USB device
- Remote installation such as a SAN LUN (in a boot from SAN scenario)
- vSphere Auto Deploy—a network boot mechanism which enables quick provisioning and configuration of ESXi hosts

The physical location where logs are written depends on the installation device used during the ESXi deployment. When the ESXi installation device is an SD card, USB key, or a remote boot LUN from a SAN device, a local scratch partition is not created on the installation media automatically during the deployment. Regardless of its size, ESXi 6.7 always sees this type of installation as remote, and as such, logs are stored in RAM disk (a disk drive that is made up of a block of volatile memory) and therefore lost when the host is rebooted.

The reason for this is that USB/SD devices are sensitive to high I/O volumes, so the installer will not place the scratch partition on this type of device. During installation, the ESXi installer first scans for a local 4-GB VFAT partition. If it is unable to find one, it will then scan for a local VMware vSphere VMFS volume to use to create a scratch directory. If no local VFAT partition or VMFS volume is found, the last resort is to put the scratch partition in `/tmp/scratch` (a scratch location on the local RAM disk—a block of volatile memory).

Because log messages that are stored on RAM disk are not retained after a reboot, troubleshooting information contained within the logs and core files is lost. In addition, if a persistent scratch location on the host is not configured properly, you might experience intermittent issues due to lack of space for temporary files, and the log files will not be updated. This can be problematic in low memory hosts but is not a critical issue for ESXi operation.

If the installation device is considered local during deployment, the ESXi host is not typically required to be manually configured with a scratch partition. The ESXi installer automatically creates a 4-GB Fat16 partition on the target device during the installation process, if there is sufficient space to do so. If persistent scratch space is configured, most of the logs are located on the scratch volume, and the `/var/log/` directory contains symlinks (symbolic links—a type of file that contains a reference to another file in the form of an absolute or relative path) to the persistent storage location.

Where the number of servers being deployed best suits a “light touch” approach, a scripted installation might offer several advantages. Employing the use of a kickstart script for the mass deployment of hosts or ongoing deployment of hosts, not only provides a means to fully automate the operating system deployment, but also:

- Provides a repeatable and consistent process for ESXi deployment
- Provides a standardized build, reducing operational overhead
- Speeds up the deployment process
- Provides additional post-deployment configuration options

The ESXi answer file (`ks.cfg`) mimics the syntax found in Red Hat Linux kickstart script and can provide a scalable and consistent approach to host deployment by being practically ‘hands-free’ after the initial effort to set up, test, and customize is complete. Therefore, this type of build environment is ideally suited to large cloud service providers who are potentially deploying hundreds of servers each week.

In Chapter 7 we will address VMware Cloud Foundation (VCF) in detail. At the core of VCF is lifecycle management (LCM) which allows for a fully automated deployment, configuration and patching and upgrade of the entire SDDC stack. VMware Cloud Foundation is more and more becoming the preferred model for cloud providers to operate their core platform. We will take a closer look in Chapter 7.

Planning a Scalable vSphere Host Clusters Architecture

vSphere ESXi hosts and their resources are typically pooled together into clusters, which in turn will be consumed by VMware Cloud Director as Provider Virtual Data Center resources. These vSphere clusters contain the CPU, memory, network, and storage resources available for allocation to virtual machines. Clusters can scale up to a maximum of 64 nodes in vSphere 6.7 and can support thousands of virtual machines.

When designing clusters, there are two basic strategies:

- ‘Scale-out clusters,’ which have fewer ESXi hosts, but the architecture will end up with a larger overall number of clusters configured.
- ‘Scale-up clusters,’ which each have more ESXi hosts, but a fewer number of overall clusters.

Therefore, the design decision is whether to have many smaller clusters or fewer larger clusters.

As with ‘scale-out’ or ‘scale-up’ at the host level, each approach to cluster design has a number of advantages and drawbacks, but the decision will be driven by a number of potential factors as described in Table 4.1.

TABLE 4.1 Advantages and Drawbacks of vSphere Cluster Design Options

Advantages	Drawbacks
Scale-out clusters (fewer numbers of hosts per cluster)	
You can be less concerned about staying within the vSphere HA cluster virtual machines per host maximum.	Depending on the admission control policy, more resources might be reserved for failover, reducing the amount of resources available to host virtual machines. Reduced VMware vSphere Distributed Resource Scheduler™ (DRS) migration options.
Scale-up clusters (larger numbers of hosts per cluster)	
Fewer resources might be reserved for failover, increasing the amount of resources available for virtual machines. If DRS is enabled, more hosts offer greater migration choices and more opportunities to achieve a better workload balance across the cluster.	There is DRS algorithm overhead on vCenter Server due to the larger number of calculations required. Depending on the storage employed, presenting LUNs to large numbers of hosts could affect performance when locks are placed on datastores.

Other general factors for a VMware Cloud Verified platform, including the following, might also play a part in the design decision:

- Application and OS Licensing (consider Oracle or Red Hat)
- Redundancy ($n+1$, $n+2$, $n+3$ and $n+4$ models)

- CPU architectures (Intel/AMD)
- Network connectivity
- Total numbers of virtual machines

The vSphere cluster is typically the boundary of shareable resources. Therefore, when planning the design of clusters, consider the following key design factors:

- Capacity planning—It may be simpler to plan for growth with a small number of large clusters. However, limitations on the number of hosts per cluster, and therefore the number of virtual machines per cluster, might warrant a scale-out approach to cluster design.
- Hardware cost—Because each cluster requires a defined amount of spare resources to accommodate failures, depending on the scale of the environment, having a large number of smaller clusters will result in a higher hardware cost to virtual machine ratio.
- Security—Isolating tenants or tenant applications into dedicated clusters is one way to segment workloads and control access through role-based access control (RBAC).
- Performance—Separating tenant's workloads or specific tenant applications into dedicated 'island' clusters provides a mechanism to ensure resources are consistently available for customers.

The number of hosts in a cluster will affect consolidation ratios. For example, you have eight compute nodes, and you have to decide whether to deploy two 4-node clusters or one 8-node cluster. You would have to reserve one ESXi host in each of the two 4-node clusters for HA failover to achieve $n+1$. To achieve the same level of availability in the 8-node cluster, you would still only need to reserve a single node, which will provide the design with one extra ESXi host for the running of virtual machine workloads.

In a production environment, you should always consider at least one host as a failover minimum per 8 to 10 ESXi servers to achieve $n+1$. Therefore, in a 16-node cluster, do not stay with only one host for failover. Aim to increase this number to two. The reason for this is that you need to ensure that you cover the risk of dual failure as much as possible by providing an additional node for this failover scenario, but also provide the ability to carry out maintenance on a host without a single host failure affecting the customer's resources. The vSphere HA calculations should not be overlooked, and are detailed further in the section, Planning for Server Failure.

The minimum size of a cluster is two nodes for vSphere HA to protect workloads in case one host stops functioning. However, in most use cases, a 3-node cluster is far more appropriate because you have the option of running maintenance tasks on an ESXi server without having to disable HA.

3-Nodes is also the minimum host requirement to enable VMware vSAN storage on those hosts.

When hosting tenant mission-critical applications, which are of the utmost importance and must have consistent performance at all times, you might need to place them in their own dedicated cluster. Whether you place multiple different applications in the same cluster or only place, one application per cluster (the concept of an “island cluster”) will depend on the critical nature and resource requirements of the application, and possibly other factors such as “air gap” isolation.

It is considered a VMware best practice to not run mixed-host clusters operating on different versions of ESXi code. However, typically during upgrade or patching activity there is likely to be a period of mixed-mode operation.

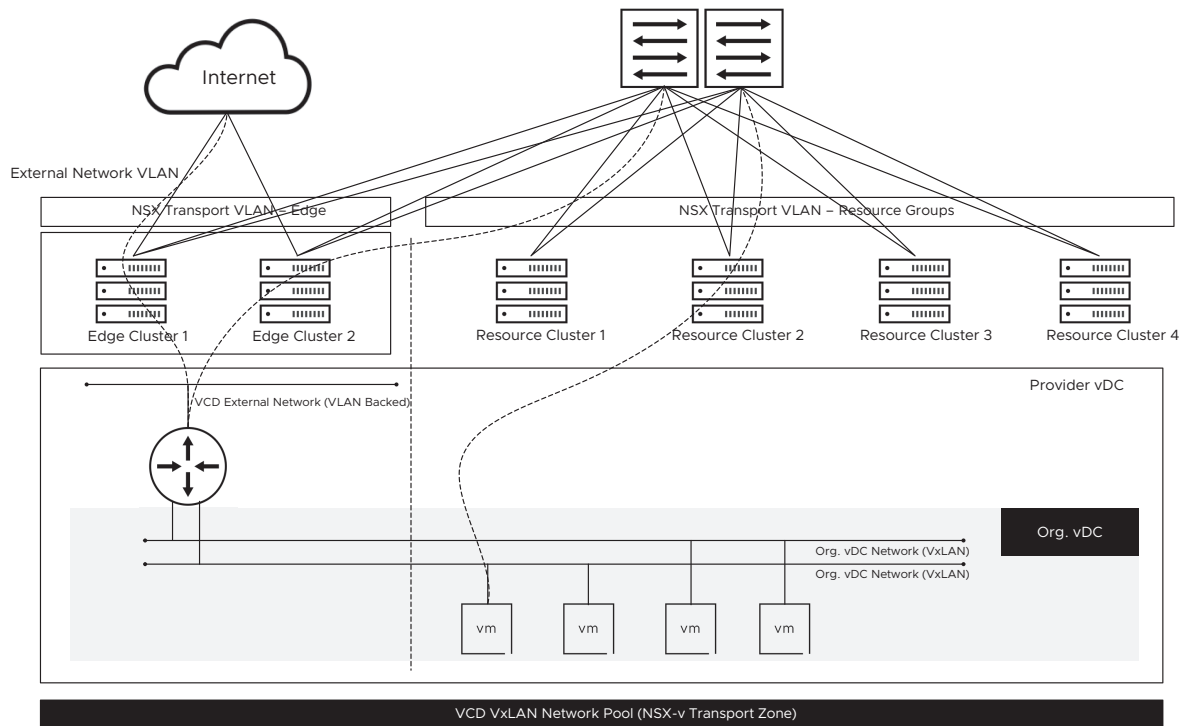
Cloud Platform Edge Cluster

VMware Cloud Director allows a cloud provider to define specific primary, and if necessary secondary, NSX Edge clusters. The purpose of the Cloud Provider Platform edge cluster is to place all tenant edge devices that are provisioned as part of the onboarding and ongoing tenant provisioning process onto a dedicated and segmented cluster, providing North-South routing for the workloads in the tenant compute clusters. No compute workloads are placed by Cloud Director into the edge clusters.

As illustrated in Figure 4.8, where this type of architecture is defined, tenant traffic traverses from a North/South perspective through a specific route through the transport VXLAN (with NSX for vSphere) to its defined tenant Edge device, residing on the Edge cluster. From there, the Edge device will egress out to the northbound network via its defined gateway.

A dedicated edge cluster hosting all VMware NSX Edge devices can act as a demarcation line to the Internet, or to corporate VLANs, so that the network administrator can provide management in a more secure and centralized way. For instance, this might mean that external connectivity needs to be configured only on the edge cluster hosts. Therefore, the use of Edge clusters in a cloud providers architecture can simplify physical network design with clear egress for north and south traffic separation, minimizing broadcast domains and streamline resource accounting by separating edge appliance consumption from customer workloads.

However, in smaller environments, deploying NSX Edge gateways appliances, also known as Edge Services Gateways (ESG) into a common resource pool owned by the customer’s Org. VDC is a fully acceptable approach. However, as the NSX Edge gateways provide essential network functions, such as NAT, Load Balancing, VPN Services, and Edge Firewall Services, in a multi-tenanted cloud provider platform architecture, the ability to deliver these services in a dedicated edge cluster enhances security capabilities, network hygiene and with primary and secondly edge clusters deployed, availability.



- Traffic egresses out of Compute Cluster to switching fabric
- Destination is Tenant Edge which resides on Edge Cluster
- Edge egresses it northbound to External Networks – which in turn will exit out of NSX to internet

FIGURE 4.8 Cloud Platform Edge Cluster Design and Traffic Flow

A cloud provider administrator must create a primary edge cluster and specify the resource pool via API:

```
POST /cloudapi/1.0.0/edgeClusters
GET /cloudapi/1.0.0/edgeClusters
```

The cloud admin can then create a network profile for each Org. VDC, and assign the edge cluster by updating the vDC's network profile:

```
PUT /cloudapi/1.0.0/orgVdcs/{orgVdcId}/networkProfile
GET /cloudapi/1.0.0/orgVdcs/{orgVdcId}/networkProfile
```

Different Org. VDCs can share the same edge clusters, or have their own dedicated ones.

Illustrated in Figure 4.9, the JSON body is formulated on each cluster under the OpenAPI `/cloudapi/1.0.0/edgeCluster`. Also illustrated, each tenant Org. VDC is defined with a `/networkProfile` available for configuration.

A few key design factors associated with Edge Cluster services include:

- At time of writing this feature is only available with NSX for vSphere and must be configured via the API.
- This architecture is intended for public-facing Org. VDC Edges only. Other internal Edge devices, such as vApp Edges (NSX for vSphere) and Universal Distributed Logical Routers (UDLR), continue to traverse traffic through standard placement.

The use of Edge clusters is a key requirement for large-scale deployments of the Cloud Provider Platform as it allows the cloud provider to optimize traffic flows through a secure architecture, while maintaining high levels of availability for network services. In addition, this provides a simplified physical network design with clear egress routes for north-south traffic. Finally, for customers who want to leverage a fully air-gapped solution, their Org. VDCs do not have to share the same Edge cluster but can be defined to route traffic through an isolated solution specific to that tenant.

Planning for Server Failure with Resilience

An important consideration for cluster design is planning for server failure, or planned maintenance. The ability for a multi-tenant infrastructure to adapt to unforeseen events is imperative. The multitenant nature of the Cloud Provider Platform reinforces the importance of both data protection and high availability to provide business continuity for each tenant. The underlying vSphere architecture offers several technologies to enable an otherwise commodity infrastructure to cope with unplanned circumstances at the platform layer,

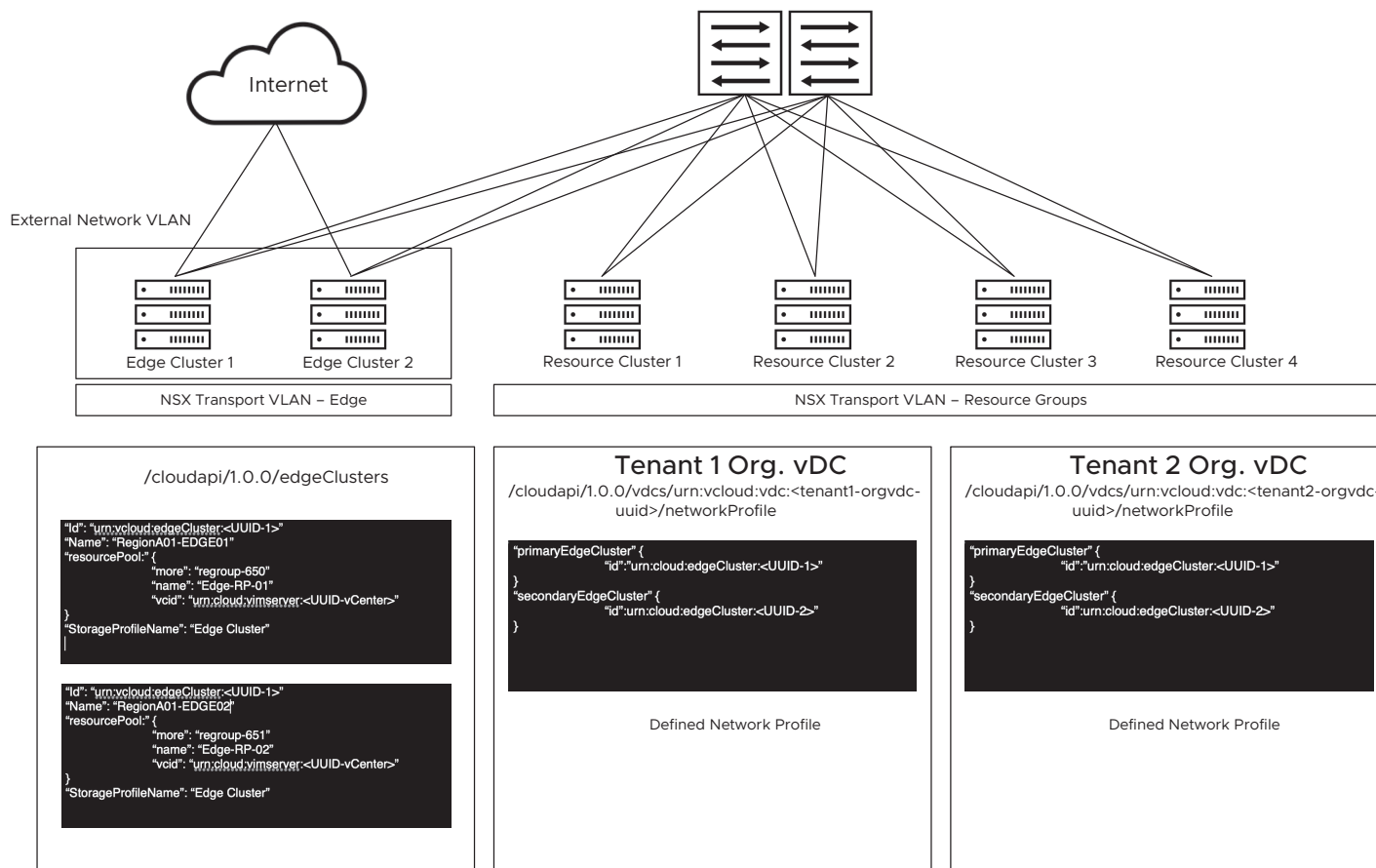


FIGURE 4.9 JSON Network Profile Definitions

meaning applications don't need to be rewritten for resilience purposes, as is the case with other public cloud environments. The result is a resilient solution that enables uninterrupted operations in the face of the inevitable real-world problems encountered in the data center.

vSphere clusters are typically used to provide increased availability and resource load balancing and are a key component of a Cloud Provider Platform design. One part of the cluster design, discussed previously, is the proportion of the cluster's total capacity that is reserved for a failure event or reserved to undertake operational maintenance tasks. Determining the number of compute nodes in each cluster and the total capacity you want to reserve (which makes it unavailable for tenant workloads) will provide the percentage of capacity that must be allocated for failure scenarios.

Where possible, using multiple server cabinets for the ESXi hosts, and distributing the vSphere clusters across the cabinets to minimize the impact of a single component failure. While a blade chassis or physical server rack is not typically considered a single point of failure, additional protection can provide higher levels of mitigation against major component outage.

The next consideration beyond chassis failure is mitigating against rack failure or even data center zone failure. The level of attention given to these requirements for availability and reducing the possibility of single points of failure should be stipulated in the design requirements.

vSphere High Availability: vSphere High Availability (vSphere HA) operates at the cluster level, and when enabled, provides the capability to monitor vSphere hosts for failures and automatically restart one or more virtual machines if deemed necessary. At a cluster level, vSphere HA monitors all hosts in the cluster through both network and datastore heartbeats. In the event that the network heartbeat between a slave node and master node in a cluster is lost, the master node attempts to use the datastore heartbeat to verify that the slave node is still operational. If the datastore heartbeat has stopped as well, the slave node is determined to have failed and the master node begins restarting the appropriate virtual machines on other nodes in the cluster. This is the reason for the requirement for reserved unused capacity within the vSphere cluster. The Cloud Provider Platform uses this vSphere core functionality as the foundation for service availability to their enterprise business tenants. Understanding the vSphere HA mechanism is key to a good, resilient and functioning design.

In a vSAN-based cluster, this mechanism works slightly differently. On a vSAN platform, datastore heartbeats become irrelevant, and the vSphere HA agent uses the vSAN network to communicate instead of the management network. However, the management gateway is still used by the host to detect if it has become isolated.

Designing a scalable and reliable vSphere HA environment for the Cloud Provider Platform does not differ significantly from designing the same mechanism for the enterprise business customer. However, a cloud provider's

platform is heterogeneous by its very nature due to its lack of control over the applications provisioned by the tenants.

Each vSphere cluster within the Cloud Provider Platform is typically configured for vSphere HA to automatically recover virtual machines should any ESXi host fail, or even if there is a specific virtual machine failure. An exception to this rule might exist where a lower cost service offering is being provided to the tenants that does not include a strict availability clause in its service-level agreement.

vSphere 6.7 supports cluster sizes of up to 64 nodes. This brings with it considerations for designing vSphere HA for such a large failure domain. For instance, with 64 nodes in a cluster, having such a large failure domain should mean you can increase your consolidation ratio by reducing the number of reserved hosts for admission control. However, you must ensure the design allows for sufficient host failures to guarantee SLAs can be met in a multiple host failure scenario, and during host maintenance and patching.

Resource Balancing and Transparent Maintenance

vSphere Distributed Resource Scheduler (DRS) dynamically balances compute capacity across a collection of hardware resources aggregated into logical resource pools based on CPU, memory, and network load status, and is required for Cloud Director environments. vSphere DRS continuously monitors utilization across resource pools, or the root cluster resource pool, to intelligently allocate available resources among virtual machines based on a predefined set of rules that reflect business or application requirements. These vSphere vMotion migrations of virtual machines can be performed automatically or manually based on the defined parameters.

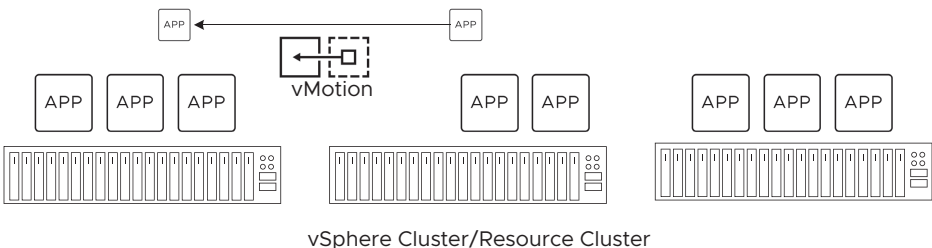


FIGURE 4.10 vSphere Distributed Resource Scheduler (DRS)

An additional benefit of DRS is realized when there is an operational requirement to place a host in maintenance mode. For instance, when servicing is required or to install additional memory on a host, all virtual machines that are running on the targeted host will be automatically migrated to other hosts within the cluster (assuming the DRS policies are configured appropriately).

to allow automatic migration). This provides a significant additional benefit gained from the vSphere HA admission control policy reserving spare capacity for both planned and unplanned outages. This can prove invaluable when performing rolling hardware maintenance or orchestrated patching through vSphere Update Manager. These types of actions can be carried out without any need to disrupt tenant services and are completely transparent to the consumers, providing non-stop IT services, maintenance without downtime, and greatly improved availability.

DRS Automation: vSphere DRS employs vSphere vMotion to migrate virtual machines in either a fully automated, partially automated, or manual manner, depending on the parameters defined.

The screenshot shows the 'Edit Cluster Settings' dialog for cluster 'RAX01'. The 'vSphere DRS' toggle is turned on. The 'Automation' tab is selected, showing the 'Automation Level' set to 'Fully Automated'. A description states: 'DRS automatically places virtual machines onto hosts at VM power-on, and virtual machines are automatically migrated from one host to another to optimize resource utilization.' The 'Migration Threshold' is set to 'Conservative' on a scale from 'Conservative' to 'Aggressive', with a description: 'DRS provides recommendations when workloads are moderately imbalanced. This threshold is suggested for environments with stable workloads. (Default)'. The 'Predictive DRS' checkbox is checked and labeled 'Enable'. The 'Virtual Machine Automation' checkbox is unchecked and labeled 'Enable'. 'CANCEL' and 'OK' buttons are at the bottom right.

Automation Level	Migration Threshold	Predictive DRS	Virtual Machine Automation
Fully Automated	Conservative	Enable	Enable

FIGURE 4.11 vSphere DRS Configuration Options

As shown in Figure 4.11, there are three different automation levels. When Fully Automated is selected, a migration threshold must be configured.

As you would expect, the conservative setting leads to fewer migrations whereas the aggressive configuration will lead to more frequent virtual machine migrations.

For a cloud provider environment, in which clusters are usually heterogeneous in nature, it is typically desirable to set the vSphere DRS automation level to Fully Automatic and the migration threshold to the default moderate (Level 3) configuration to avoid automatic vSphere vMotion actions that might have either only a minimal or short-term performance benefit.

TABLE 4.2 DRS Automation Levels

Setting	Initial VM Placement	Load Balancing
Manual	Recommendation is displayed to administrator	DRS makes a recommendation but will not migrate VMs without validation from the administrator.
Partially Automatic	Automatic placement	While the initial placement is completed automatically by DRS, the migration of powered on virtual machines will only be performed after the administrator validates the recommendation in vCenter.
Automatic	Automatic placement	Fully automated migrates powered on virtual machines automatically. The level of aggressiveness employed by DRS for this automatic migration is based on a threshold corresponding to five different recommended levels from conservative (5 stars) to aggressive (1 star).

TABLE 4.3 Migration Threshold Options

Level	Stars	Description
Level 1	5	Conservative. Migrations will only take place if rules are not being respected or if a host is placed into maintenance mode.
Level 2	4	A migration will only take place if level 1 is met or if a migration will bring about significant improvements in performance.
Level 3	3	A migration will only take place if the first two levels are met or if a migration brings about a good amount of improvements to virtual machine performance.
Level 4	2	A migration will only take place if the first three levels are met or a migration brings about moderate improvements to virtual machine performance.
Level 5	1	Aggressive. Migration will occur only if all recommendations from Level 1 to 4 are met or if the migration will bring about minor improvements to virtual machine performance.

For a cloud provider’s implementation, typically DRS is enabled to automate workload balancing. DRS will always benefit the overall infrastructure improving performance, scalability, and manageability while providing transparent maintenance to the consumers. A Cloud Provider Platform should typically configure all DRS cluster parameters consistently across all clusters in all data centers. This helps to limit variability and simplify operational management.

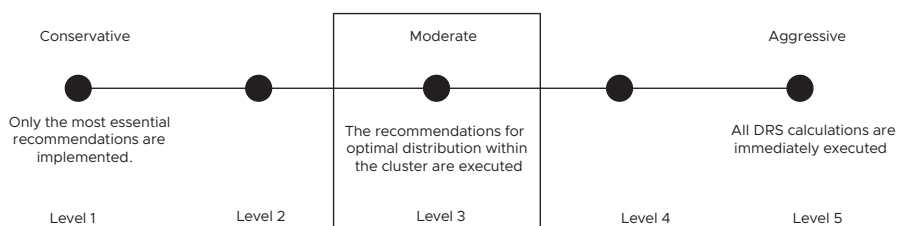


FIGURE 4.12 Migration Threshold Slider – Recommendation for the Cloud Provider Platform

Enhanced vMotion Compatibility

Enhanced vMotion Compatibility (EVC) addresses the challenge that each generation of processor hardware presents through improvement with new functionalities. EVC makes sure that all host servers, where slight variations in processor generation exist, offer the same CPU instruction sets to virtual machines, safeguarding vSphere vMotion compatibility among the mixed processors.

EVC employs a baseline across the CPU instruction sets and this baseline enables a set of functionalities supported by all processors found in the cluster, defining a common level for all processors.

The CPU is the most restrictive component when it comes to vSphere vMotion. Enabling EVC permits the masking of some differences, offering more compatibility between servers with different generations of processors. When employed, EVC does not hinder performance and does not affect the number of processor cores or the size of the CPU cache. The only possibility of degraded performance is where certain processor instructions, such as SSE 4.2, have been masked and therefore not used.

The Cloud Provider Platform design decision on whether or not to enable EVC will depend on a number of questions that you will need to address:

- Do you intend to mix hardware in the cluster?
- Do you think there is a high possibility of adding newer hardware to your cluster?
- If cross vCenter vSphere vMotion or long-distance vSphere vMotion or HCX migration forms part of a design or service offering, do you know what hardware types and processor generations exist at the other target data centers?

If the answer to any of these questions is yes, the general advice is to enable EVC on all clusters. Doing so makes adding hosts easier and saves having to split the cluster in the future if you have the foresight to enable it from the outset.

- **Passive Node:** Runs as the passive instance of vCenter Server and constantly receives state updates from the Active node in synchronous mode. The Passive Node Takes the role of Active node after a failure event.
- **Witness Node:** This is a lightweight virtual appliance that serves as a quorum node and does not take either an Active or Passive role.

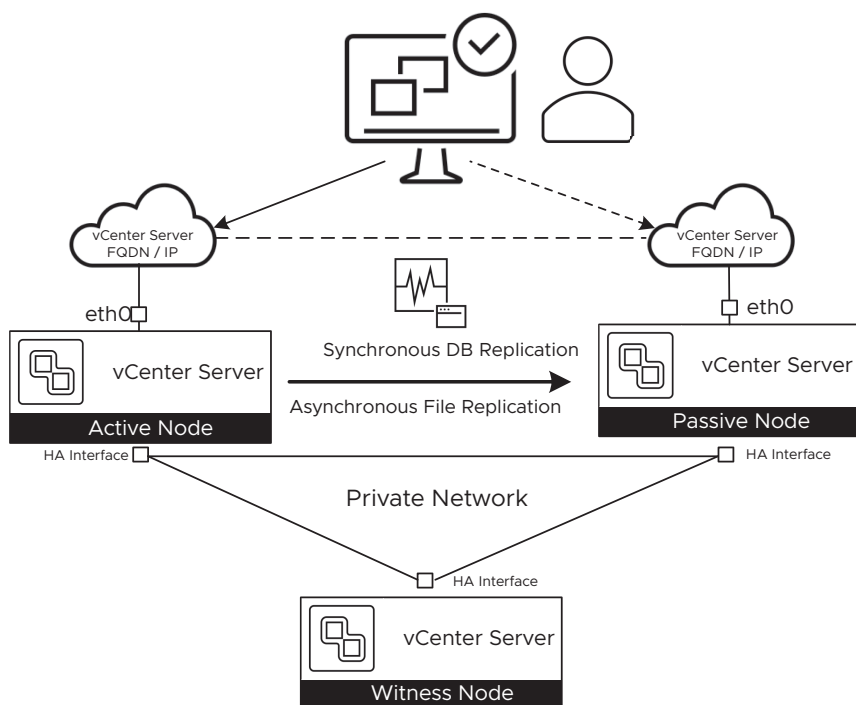


FIGURE 4.14 vCenter High Availability

At least three hosts are required in the management cluster to complete the deployment of this solution for vCenter High Availability, as well as good network connectivity for Active and Passive nodes to guarantee zero RPO. In addition, a dedicated network separated from the management network must be configured. In the event of the Active vCenter Server failing, this architecture behaves as follows:

- **Active node fails:** Until Passive and Witness node can communicate with each other, the Passive node will take over the Active role and starts serving client requests.
- **Passive node fails:** If Active and Witness node can communicate with each other, the Active node will continue to server client requests.
- **Witness node fails:** As long as the Active and Passive node can communicate with each other, the Active node will continue to serve client requests.

- **Isolated node behavior:** If a node gets isolated from the cluster, it is automatically taken out of the cluster and all services are stopped. For example, if an Active node is isolated, all services are stopped to ensure that the Passive node can take over as long as it is connected to the Witness node.

Cloud Platform (vSphere) Patch Management

A Cloud Platform patch management automation tool is essential for the ongoing security and stability of the Cloud Provider Platform. One tool that is often implemented as part of a vSphere platform to keep the ESXi hosts, virtual machine hardware, and VMware Tools up to date is vSphere Update Manager.

One of the key benefits of VMware Cloud Foundation is its orchestrated approach to Lifecycle Management (LCM), which allows for a fully automated deployment, configuration, patching, and upgrades. VMware Cloud Foundation will be addressed in more detail in Chapter 7.

However, if VCF is not in use, vSphere Update Manager can also perform orchestrated vSphere upgrades. An orchestrated upgrade allows you to upgrade the objects in your vSphere inventory in a two-step process: host upgrades, followed by virtual machine upgrades. You can configure the process at the cluster level for higher automation, or at the individual host or virtual machine level for more granular control. You can upgrade clusters without powering virtual machines off, as long as vSphere DRS is available to the cluster. To perform an orchestrated upgrade, first remediate a cluster against a host upgrade baseline, and then remediate the same cluster against a virtual machine upgrade baseline group that contains the virtual machine hardware upgrade and VMware Tools upgrade to match host baselines.

Chapter 5

Designing a VMware *Cloud Verified* Storage Platform

The Cloud Provider Platform requires that their storage infrastructure not only be reliable and scalable, but also cost effective. VMware vSAN transforms standard x86 servers with directly attached storage into a simple Hyper-Converged Infrastructure (HCI) solution, which aligns with this Cloud Provider Platform requirement, to help eliminate high storage costs, management complexity, and performance limitations.

The Cloud Provider Platform aims to improve cloud providers storage infrastructure and services by simplifying operations and ensuring continuous uptime with VMware vSAN. The software-defined storage platform offered by vSAN is critical to delivering hybrid cloud service offerings under the VMware *Cloud Verified* program and provides an infrastructure that can be consumed to deliver a building block hyper-converged storage solution.

There are several approaches for architecting a VMware software-defined storage solution and the cloud providers approach will depend on the use case and the technology deployed. However, the end goal is always the same. This *VMware Cloud Architecture Toolkit 5.0* publication describes the Cloud Provider Platform solution for supporting a software-defined storage solution with VMware vSAN. However, the Cloud Provider Platform also supports a mixed storage model that extends beyond vSAN and allows pods of any kind to be deployed with any supported storage solution or a mix of storage offerings to meet varying client and application needs.

VMware Cloud Director also supports storage profiles as part of its software-defined storage management layer. This allows cloud providers to leverage traditional storage such as SAN or NAS-based solutions and software-defined storage like vVols and vSAN by defining multiple data-stores that provide varying levels of service and capability, as illustrated in Figure 5.1. Consuming vSAN alongside other storage types is fully supported

in the Cloud Provider Platform model, and is commonly seen among VMware most strategic cloud providers.

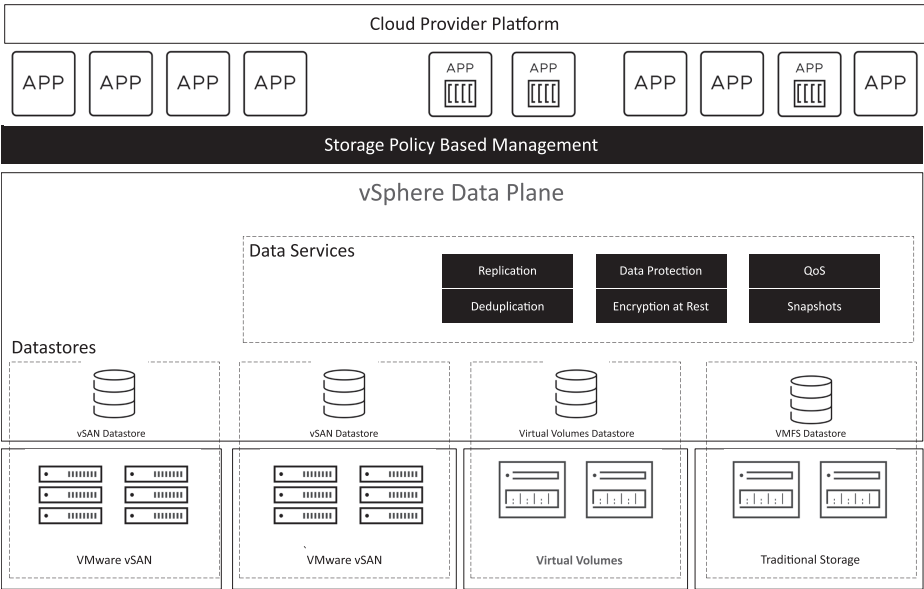


FIGURE 5.1 Flexible Mixed Storage Layer

VMWARE VSAN OVERVIEW

VMware vSAN is designed to provide storage services and service-level agreement automation through a software layer on the hosts that integrates with and abstracts the underlying hardware. There is no need to repurpose LUNs or volumes. Virtual machine workloads might change over time, and the underlying storage can be adapted to the workload at any time. This means with vSAN, workload storage requirements can be dynamically instantiated based on requirements and not the underlying platform.

vSAN is a software-defined storage solution that is fully integrated with vSphere and aggregates locally attached disks in a vSphere cluster to create a storage solution that is provisioned from VMware vCenter during virtual machine provisioning operations. It is an example of a hypervisor-converged platform—that is, a solution in which storage and compute for virtual machines are combined into a single device, with storage being provided within the hypervisor itself, as opposed to through a storage virtual machine running alongside other virtual machines.

Under the hood, vSAN is a proprietary object-based storage system designed to provide virtual machine-centric storage services and capabilities

through a Storage Policy-Based Management (SPBM) mechanism, designed to simplify virtual machine storage placement decisions. Object-based storage is considered a leading technology for hybrid cloud deployments because many of its most prominent features, such as massive scalability, geographic independence, and multi-tenancy, have proven ideal for cloud storage.

There are many different workload profiles in the data center, but from the perspective of storage, workloads have traditionally been separated into two profile types: file-based and block-based. While these technologies have had a long track record of success in the enterprise space, their track record has been somewhat limited in the cloud space and overtaken by the HCI model and in many cases, object-based storage.

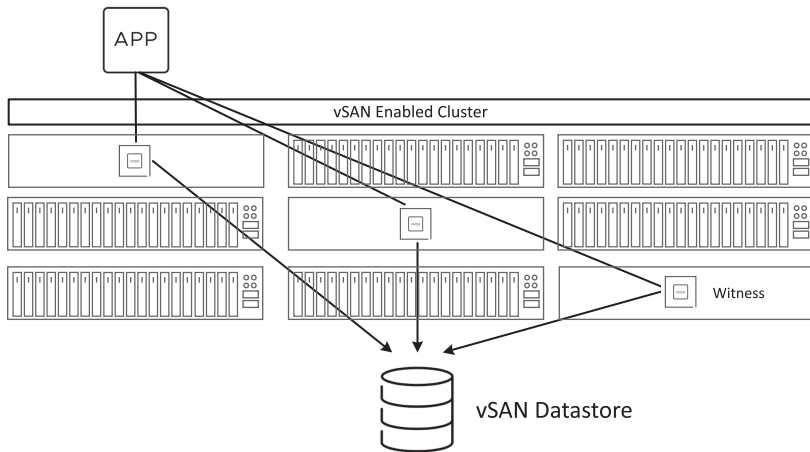


FIGURE 5.2 vSAN Object-Based Storage System

VMware vSAN has been tightly integrated with VMware vSphere to deliver a solution that offers simplicity, reliability, and performance for nearly any cloud provider use case. vSAN is fully integrated with core vSphere features, such as vSphere High Availability (HA), vSphere Distributed Resource Scheduler (DRS), and vSphere vMotion.

The goal of vSAN is to provide both high availability and scale-out storage functionality. It also can be considered in the context of Quality of Service (QoS) because virtual machine storage policies can be created to define the levels of performance and availability required on a per-virtual machine basis.

vSAN is easy to implement with automated configuration and includes proactive tests to help verify functionality and performance. vSAN is optimized for all-flash storage with efficient nearline deduplication, compression, and erasure coding capabilities that lower TCO while delivering cloud provider-required performance.

New and improved features such as the performance and health services make it easier than ever to verify vSAN configurations and closely monitor key metrics such as IOPs, throughput, and latency at the cluster, host, virtual machine, and virtual disk levels. Quality of service can be managed by using IOPs limits on a per-virtual machine and per-virtual disk basis. VMware vSAN also provides the Cloud Provider Platform with:

- A converged platform (compute + storage using commodity hardware)
- A simple to manage storage platform, as with CPU and memory
- The ability to scale on demand, in lockstep with application needs, but also in a flexible way
- Per-VM automated SLA management

In addition, one of the key features of vSAN is that the product runs as a kernel module in vSphere, rather than running as a virtual appliance in a cluster, as many competitor solutions do. This allows for greater performance and better interaction with the vSphere hypervisor.

VMware vSAN uses a number of terms that are good to define before proceeding:

- Each vSAN host contributing storage does so in the form of a ‘disk group.’
 - A disk group is a combination of exactly one cache-tier SSD and one or more capacity-tier drives.
 - The cache tier is used for write-buffering and/or read caching.
 - The capacity tier is used for long-term persistence and data at rest.
- vSAN object—Every element of the VM (home directory, VMDKs, snapshots, swap files) is an object in vSAN, not a file.
- vSAN Component—Objects are distributed constructs made of components.
 - Components are the data and witnesses that allow the object to be assembled and deliver data to the end user (the VM).
- Witness—A witness is a metadata-only component. It exists to maintain quorum and ensure integrity during split-brain scenarios.
- Storage Policy-Based Management (SPBM)—virtual machine storage policies are solutions designed to simplify virtual machine storage placement decisions for vSphere administrators.

Being natively integrated with vSphere allows vSAN to be configured with just a few mouse clicks. Because disks internal to the vSphere hosts are used to create a vSAN datastore, there is no dependency on external shared storage. In addition, virtual machines can be assigned specific storage policies based on the needs of the applications. These workloads can then benefit from this dependable shared storage with predictable performance and availability characteristics.

VMWARE VSAN ARCHITECTURE OVERVIEW

VMware vSAN clusters are made up of physical hosts that contain either a combination of mechanical disks and flash devices (hybrid configuration) or all flash devices (all-flash configuration) that contribute cache and capacity to the vSAN distributed datastore.

For most Cloud Provider Platform use cases, and for purposes of this content, the all flash configuration provides optimum flexibility and predictable capacity verses performance cost modeling. For this reason, the primary focus for this book will be the all flash configuration. Therefore, unless specifically called out, all references to vSAN configuration and design will relate to this model.

In an all-flash deployment, the flash device in the cache tier is used for write-cache only (no read cache) because read performance from the capacity flash devices is more than sufficient for even the most demanding application. In an all-flash configuration, two different grades of flash devices are commonly used, a lower capacity, higher endurance device for the cache layer, and a more cost effective, higher capacity, lower endurance devices for the capacity layer. Writes are performed at the cache layer and then de-staged to the capacity layer, only as needed. This helps extend the usable life of the lower endurance flash devices in the capacity layer.

VMware vSAN is part of the core SDDC platform and requires vCenter Server for configuration. vSAN cannot be managed through VMware Cloud Director, therefore, vSAN is configured and monitored from the vSphere Client.

Space Efficiency: Multiple space efficiency technologies, optimized for all-flash configurations, to minimize storage capacity consumption while ensuring performance and availability are supported by vSAN. This feature includes deduplication, compression, and RAID5/6 erasure coding to reduce capacity consumption, for a lower total cost of ownership, while still ensuring the same levels of availability and performance.

Enabling deduplication and compression can reduce the amount of storage consumed by as much as seven times. However, the actual reduction value will vary, because this depends primarily on the types of data present, number of duplicate blocks, how much these data types can be compressed, and distribution of these unique blocks. For example, video files do not compress well, while documents and spreadsheets typically yield more favorable results. The following figure shows deduplication and compression efficiency being enabled in an all-flash vSAN environment.

Deduplication and compression is a cluster-wide setting that is disabled by default and can be enabled using a simple drop-down menu, illustrated here. Note that a rolling reformat of every disk group on every host in the vSAN cluster is required, which can take a considerable amount of time. However, this process does not incur virtual machine downtime but is most

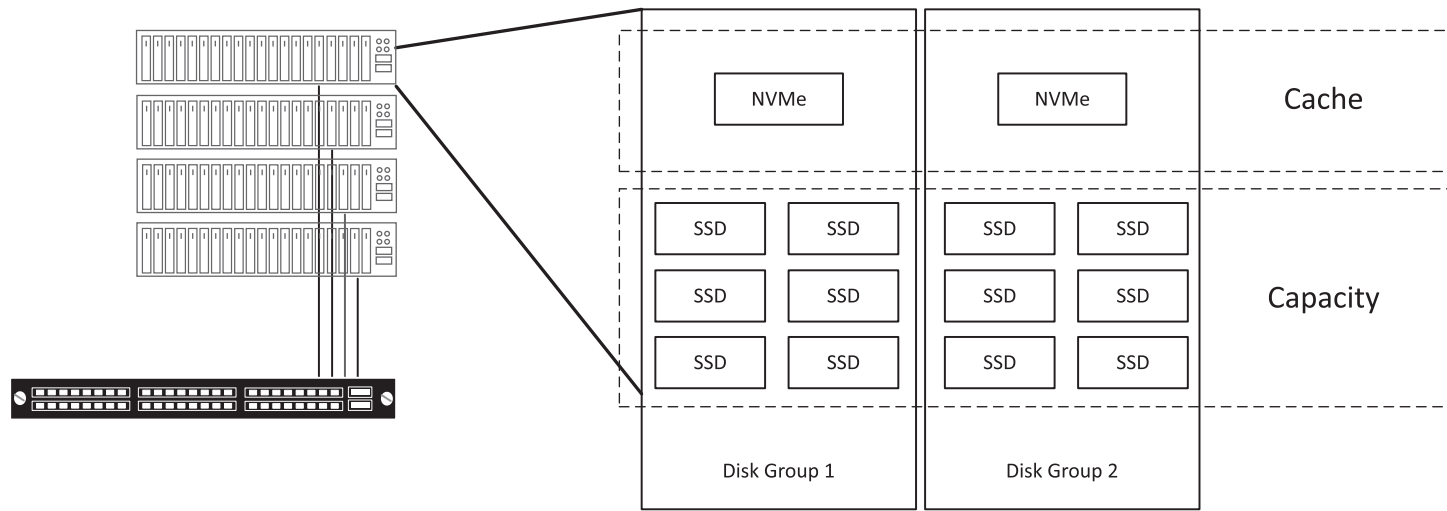
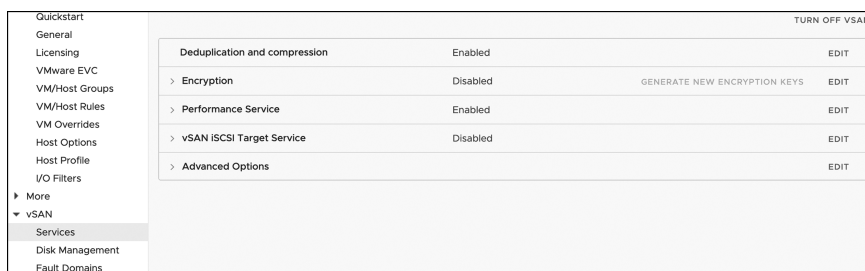
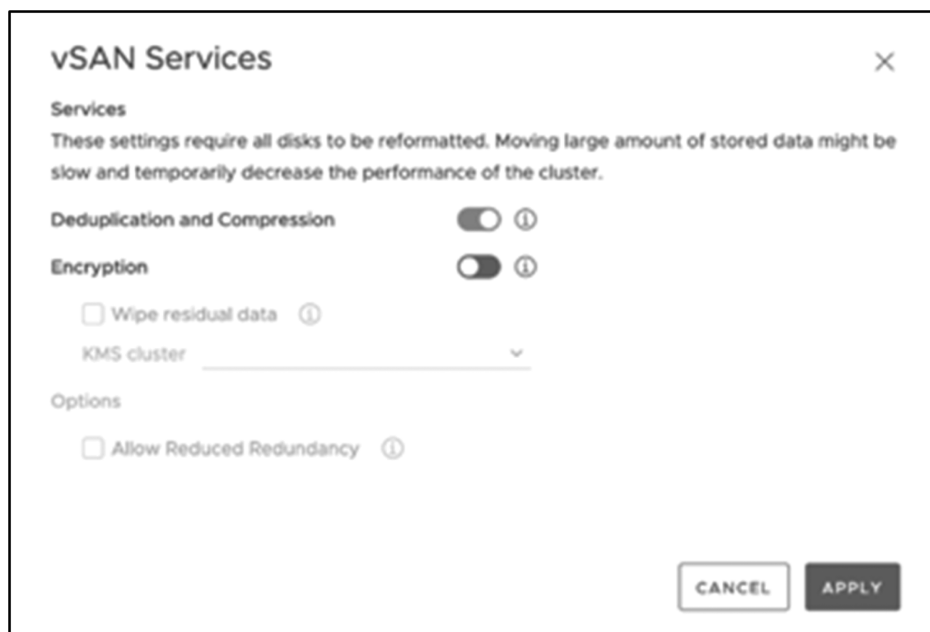


FIGURE 5.3 vSAN All-Flash Architecture

**FIGURE 5.4** Enabling vSAN

effectively a parameter that should be configured at time of deployment. In addition, deduplication and compression are enabled as a single parameter. It is not possible to enable deduplication or compression individually.

**FIGURE 5.5** Enabling Deduplication and Compression

Deduplication occurs when data is destaged nearline from the cache tier to the capacity tier of an all-flash disk group. The deduplication algorithm utilizes a 4K fixed-block size and is performed within each disk group independently. In other words, redundant copies of a block within the same disk group are reduced to one copy, but redundant blocks across multiple disk groups are not deduplicated. The deduplication at the disk group level by vSAN using a 4K block size helps provide a good balance between space efficiency and performance.

The compression algorithm is applied after deduplication has occurred, just before the data is written to the capacity tier flash devices. Considering the additional compute resource and allocation map overhead of compression, vSAN will only store compressed data if a unique 4K block can be reduced to 2K or less. Under all other circumstances, the block is written uncompressed.

There is a storage policy implication to be aware of when deduplication and compression are enabled. This is especially true when upgrading from a previous version of vSAN, where one or more storage policies contain an object space reservation rule with a value other than 0 or 100 percent. When deduplication and compression are enabled, object space reservation rules must be set to 0 or 100 percent. Values from 1 to 99 percent are not supported when deduplication and compression are enabled. An object that is assigned a storage policy containing an object space reservation rule of 100 percent will be analyzed for deduplication, but no space savings will be realized because capacity for the entire object is reserved. Before upgrading vSAN, all policies containing an explicit object space reservation rule should be configured to 0 or 100 percent.

Note: *The implicit default value for this rule is 0 percent, so there is no need to adjust a policy that does not have the object space reservation explicitly defined.*

Naturally, the processes of deduplication and compression on any storage platform require additional CPU cycles and can potentially impact performance in terms of latency and maximum IOPS. vSAN is no exception. However, considering deduplication and compression are only supported in an all-flash vSAN configuration, these effects will be negligible in the majority of use cases due to the high levels of performance available from modern enterprise flash devices.

Host and vSphere Cluster Failure Tolerance: The vSAN Storage Policy-Based Management (SPBM) rule, called the fault tolerance method, allows virtualization administrators to choose which method of fault tolerance to employ, mirroring or erasure coding. While mirroring techniques excel in workloads where performance is the most important factor, they are expensive in terms of capacity required. RAID-5/6 (erasure coding) data layout can be configured to help provide the same levels of availability, while consuming less capacity than RAID-1 (mirroring). Use of erasure coding reduces capacity consumption by as much as 50 percent compared with mirroring at the same fault tolerance level. This method of fault tolerance does require additional write overhead in comparison to mirroring as a result of data placement and parity.

RAID-5/6 (erasure coding) is configured as a storage policy rule and can be applied to individual virtual disks or an entire virtual machine. Note that the failure tolerance method in the rule set must be set to RAID5/6 (erasure coding).

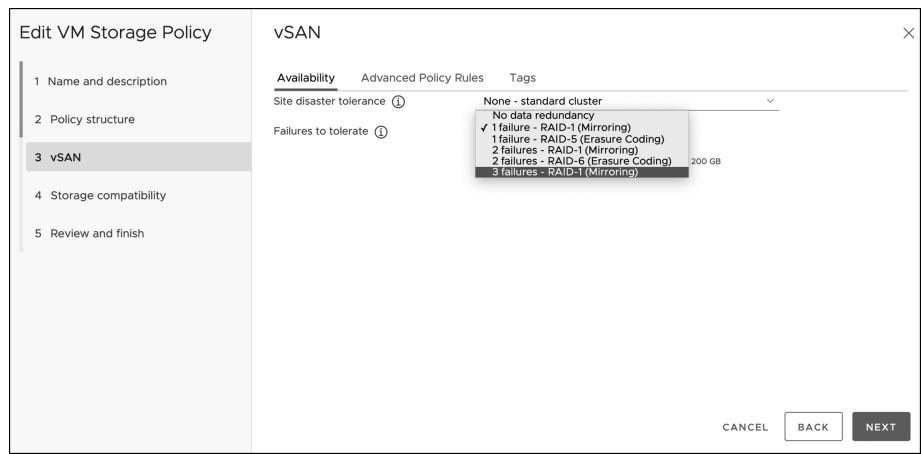


FIGURE 5.6 Enabling Storage Policy-Based Management Fault Tolerance Methods

RAID-1 (mirroring) in vSAN employs a $2n+1$ host or fault domain algorithm, where n is the number of failures to tolerate. RAID-5/6 (erasure coding) in vSAN employs a $3+1$ or $4+2$ host or fault domain requirement, depending on 1 or 2 failures to tolerate respectively. RAID-5/6 (erasure coding) does not support three failures-to-tolerate. Table 5.1 details the host and capacity requirements.

TABLE 5.1 vSAN Data Protection Space Consumption

	Tolerated Failures	RAID-1		RAID-5/6		Erasure Coding Space Savings vs Mirroring
		Minimum Hosts Required	Total Capacity Requirement*	Minimum Hosts Required	Total Capacity Requirement*	
FTT=0	0	3	1x	n/a	n/a	n/a
FTT=1	1	3	2x	4	1.33x	33% less
FTT=2	2	5	3x	6	1.5x	50% less
FTT=3	3	7	4x	n/a	n/a	n/a

*Without Deduplication/Compression taken into account

Erasure coding can provide significant capacity savings over mirroring, but it is important to consider that erasure coding incurs additional overhead, which is common among any storage platform for this type of data striping.

Because erasure coding is only supported in all-flash vSAN configurations, effects to latency and IOPS are negligible in most use cases due to the inherent performance of flash devices.

Data Integrity (Software Checksum): Software checksum enables service providers to detect the corruptions that could be caused by hardware/software components, including memory, drives, and so on during the read or write operations. In case of drives, there are two basic kinds of corruption. The first are latent sector errors, which are typically the result of a physical disk drive malfunction. The second type are silent corruption errors, which can happen without warning (these are typically called silent data corruption). Undetected or completely silent errors can lead to lost or inaccurate data and significant downtime. There is no other effective means of detection for these errors without an end-to-end integrity checking mechanism.

During the read/write operations, vSAN checks for the validity of the data based on the checksum. If the data is not valid, vSAN will take the necessary steps to either correct the data or report it to the user to take action. These actions could be:

- To retrieve a new copy of the data from other replica of the information, stored within the RAID1, RAID5/6 constructs. This is referred to as recoverable data.
- If there is no valid copy of the data found, an error is returned. These are referred to as non-recoverable errors.

In the case of data errors, issues are reported in the vSphere Web Client user interface and in log files. These include impacted blocks and their associated VMs, allowing the administrator to see the following:

- List of the VMs/blocks that are hit by non-recoverable errors within the vSphere Web Client user interface.
- Historical/trending errors on each drive from the vSphere Web Client user interface.

The data integrity feature uses a CRC32 algorithm, which also supports CPU offload to reduce overhead. In addition, there are two levels of scrubbing employed:

- Component-level scrubbing—every block of each component is checked. If there is a checksum mismatch, the scrubber tries to repair the block by reading other components.
- Object-level scrubbing—for every block of the object, data of each mirror (or the parity blocks in RAID-5/6) is read and checked. For inconsistent data, all data in the affected stripe is marked as bad.

The repair can either happen during normal I/O operations by the DOM owner of that object, or by the scrubber, although the repair mechanism for a mirrored or RAID-5/6 operation is different. If checksum verification fails, the scrubber or DOM owner will read the other copy of the data (or other data in the same stripe in case of RAID-5/6) and rebuild the correct data by writing it out to the bad location.

This end-to-end checksum of the data aims to prevent data integrity issues, which could be caused by silent disk errors, with the checksum being calculated and stored on the write path, or with silent corruptions being detected when reading the data through checksum data.

When checksum verification fails, vSAN will automatically read a different copy of the data (or other data in the same stripe in the case of RAID-5/6), then rebuild the corrected data, writing it out to the bad location, based on a 4KB block size.

Data Integrity (Software Checksum) is a cluster wide setting, which by default is switched on. However, this value can be disabled on a per-object basis, using storage policies.

Rack Awareness (Fault Domains): vSAN fault domains provide the ability to tolerate rack failures and power failures in addition to disk, network, and host failures.

The idea behind fault domains is to tolerate groups of hosts (chassis or racks) failing without requiring additional data copies. Their implementation allows vSAN to save replica copies of the virtual machine data in different domains, such as different racks of compute. An example of using fault domains is shown in Figure 5.7.

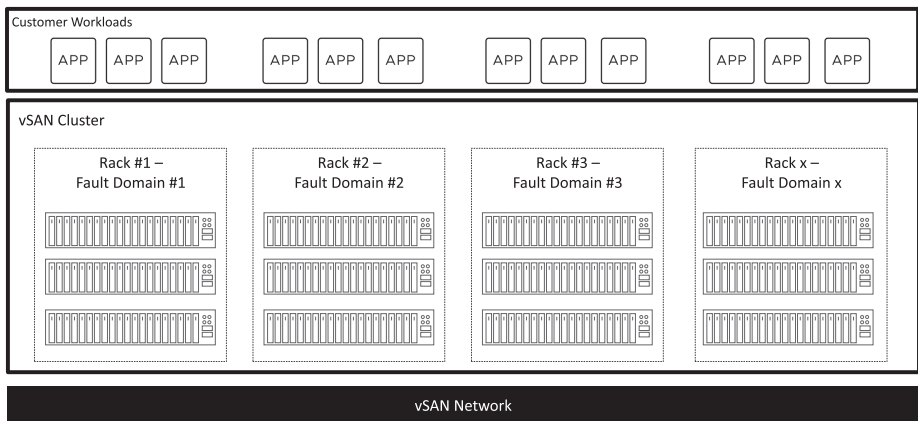


FIGURE 5.7 Rack Awareness (Fault Domains) Example

When working with fault domains, to tolerate n number of failures, rather than needing $2n+1$ hosts you require $2n+1$ fault domains instead.

This allows for replica data to be spread out across the domains rather than hosts and increases the ability to handle a greater than single host failure.

A minimum of three fault domains are required, however, VMware recommends having four or more, for redundancy purposes. In addition, to have a minimum level of redundancy, VMware also recommends having a minimum of six hosts (two per fault domain).

Fault domains are used to further protect the environment from failure, assuming there are enough hosts to properly support the configuration.

Swap Efficiency: Virtual swap files are created when virtual machines are powered on. In cases where physical host memory is exhausted, the virtual swap file is used in place of physical memory. Virtual swap files are sized according to the allocated memory minus reserved memory. For instance, a virtual machine with 4 GB of RAM allocated, with a 2GB memory reservation, will create a 2-GB virtual swap file. In large cloud provider deployments of thousands of virtual machines, this additional capacity could be substantial and require planning.

To help mitigate this, for additional space savings the default behavior of the vSAN swap object is set to thin (`SwapThickProvisionedDisabled`) and matches the storage policy attributes assigned to the workload. This setting creates virtual swap files as a sparse object on the vSAN datastore, which will only consume capacity on vSAN as they are accessed. The result can be significantly less space consumed, providing virtual machines do not experience memory over commitment, requiring use of the virtual swap file. Depending on the workload types, this has the potential to significantly decrease the effective storage capacity used in the Cloud Provider Platform's vSAN environment.

Quality of Service: The quality of service feature that limits the number IOPS an object can consume is provided as a storage policy in vSAN. In underutilized configurations, limits may or may not be necessary, because objects likely have sufficient resources to effectively meet the needs of their workload. While it is entirely desirable to have more than enough resources, it does not come without cost. Efficiently sized configurations are typically a good mix of cost and available resources. The metrics of appropriate resources for workloads can change over time, especially as utilization grows, or as workloads are added over the lifecycle of a platform.

There are several situations where it might be advantageous to limit the IOPS of one or more virtual machines. The term noisy neighbor is often used to describe when a workload monopolizes available I/O or other resources, which negatively affects other workloads on the same platform. In environments where there is a mix of both low and high utilization, it could be detrimental for a virtual machine with low utilization during normal operations to change its pattern and start to consume massive resources, starving others

of enough to operate properly. Particularly at a larger scale, the impact of this situation might affect multiple workloads, tenants, or customers.

A good example of a noisy neighbor situation includes end-of-month reporting. Consider those virtual machines generating reports residing on the same four-node Hybrid vSAN cluster as other Tier 1 applications that have stringent service-level agreement requirements. What would happen if one or more large reports were generated, consuming a larger percentage of IOPS than were available for the minimum service-level requirements of the Tier 1 applications? It is quite possible the Tier 1 applications would not be able to satisfy the requirements of the service-level agreement. Implementing IOPS limits on the reporting solution VMs could prevent the situation where the reporting solution starves the Tier 1 applications for IO.

As quality of service is applied to vSAN objects through a storage policy, it can be applied to individual disks or the entire virtual machine without interruption.

The screenshot shows the 'Edit VM Storage Policy' window for a vSAN storage policy. The left sidebar lists five steps: 1 Name and description, 2 Policy structure, 3 vSAN (selected), 4 Storage compatibility, and 5 Review and finish. The main area is titled 'vSAN' and has three tabs: Availability, Advanced Policy Rules (active), and Tags. Under 'Advanced Policy Rules', the following settings are visible:

- Number of disk stripes per object:** 2 (with a dropdown arrow)
- IOPS limit for object:** 500 (with a range selector icon)
- Object space reservation:** Thin provisioning (with a dropdown arrow)
- Flash read cache reservation (%):** 0 (with a range selector icon)
- Disable object checksum:** A toggle switch that is currently turned off.
- Force provisioning:** A toggle switch that is currently turned off.

At the bottom right, there are three buttons: CANCEL, BACK, and NEXT. Small informational text below the 'Object space reservation' and 'Flash read cache reservation' settings states: 'Initially reserved storage space for 100 GB VM disk would be 0 B' and 'Reserved cache space for 100GB VM disk would be 0 B' respectively.

FIGURE 5.8 vSAN QoS Enforced by SPBM

Quality of service for vSAN is normalized to a 32-KB block size, and treats reads the same as writes. An example with an IOPS limit of 500 (regardless of block size up to 32 KB) will result in 500 IOPS, while a block size of 64 KB would result in 250 IOPS. It is important to consider the workload profile when configuring IOPS limits.

Data Locality: In computer science, data locality, also known as locality of reference, is the behavior of computer programs according to which a workload accesses a set of data entities or storage locations within some period with a predictable access pattern.

There are two main types of data locality:

- Temporal locality—The probability that if some data (or a storage location) is accessed at one point in time, it will be accessed again soon afterwards.
- Spatial locality—The probability of accessing some data (or a storage location) soon after some nearby data (or a storage location) on the same medium has been accessed. Sequential locality is a special case of spatial locality, where data (or storage locations) are accessed linearly and according to their physical locations.

Data locality is particularly relevant when designing storage caches. For example, NVMe flash devices offer impressive performance improvements over SSD, at a cost, so efficient use of these resources becomes an important design factor.

Like any storage system, vSAN makes use of data locality. vSAN uses a combination of algorithms that take advantage of both temporal and spatial locality of reference to populate the flash-based read caches across a cluster and provide high performance from available flash resources.

Examples include:

- Every time application data is read by a virtual machine, vSAN saves a copy of the data in the Read Cache portion of the flash device associated with the disk group where the copy of the data resides. Temporal locality implies that there is high probability that said data will be accessed again before long. In addition, vSAN predictively caches disk blocks in the vicinity of the accessed data (in 1MB chunks at a time) to take advantage of spatial locality as well.
- vSAN uses an adaptive replacement algorithm to evict data from the Read Cache when it is deemed unlikely that the data will be accessed again soon and uses the space for new data more likely to be accessed repeatedly.
- vSAN makes replicas of storage objects across multiple servers for protection purposes. Reads are distributed across the replicas of an object for better load balancing. However, a certain range of logical addresses of an object is always read from the same replica. This approach has two important benefits:
 - Increases the chances that the data accessed is already in the read Cache.
 - A data block is never cached in more than one flash device.

vSAN Stretched Clusters: vSAN Stretched Clusters introduce the ability to create a cluster between two or more geographically separated sites, synchronously writing data in both locations.

The stretched cluster feature introduces enterprise-level availability where an entire site failure can be tolerated, with no data loss and near zero

downtime. In addition, this feature enables disaster and downtime avoidance by allowing the proactive migration of virtual machines from one site to another in order to avoid an impending outage, or for planned maintenance purposes.

vSphere vMotion and vSphere DRS are natively integrated into a stretched cluster architecture and can help load balance workloads between two active sites, while vSAN ensures read efficiency by making sure that virtual machines use the local site copy exclusively for read operations.

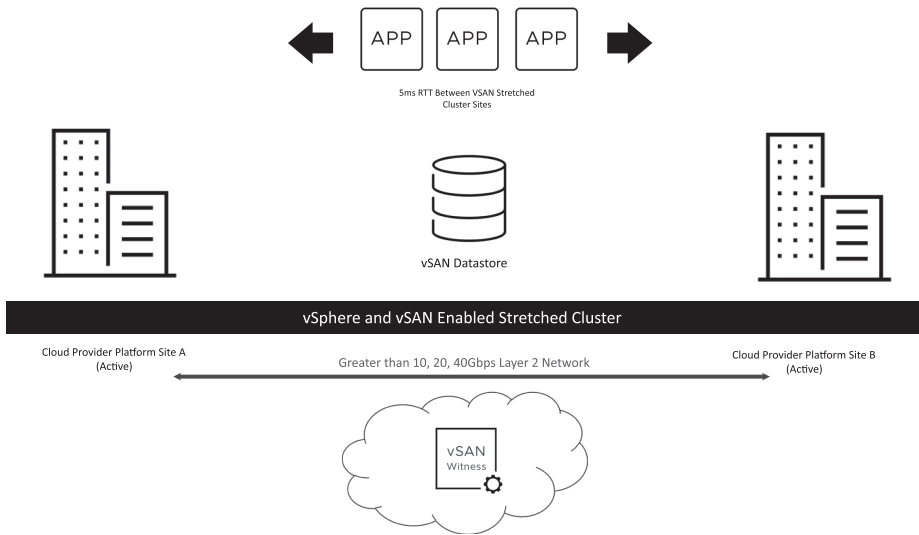


FIGURE 5.9 vSAN Stretched Cluster Architecture

Because of the tight integration with the VMware advanced features like vSphere vMotion, vSphere DRS and vSphere HA, the stretched cluster solution works seamlessly with little additional operational overhead, and expands the knowledge of vSphere administrators, who are already familiar with how to operate and run everything around vSphere.

vSAN removes all the complexity around storage management traditionally seen in a multisite environment and also simplifies storage tasks to a few clicks in the vSphere Client.

Data Replication: vSAN can also utilize VMware vSphere Replication for its replication data services. vSphere Replication with vSAN supports Recovery Point Objectives (RPOs) from 5 minutes to 24 hours. VMware Cloud Director Availability can be used as part of the Cloud Provider Platform solution to deliver a fully orchestrated disaster recovery solution for tenant workloads.

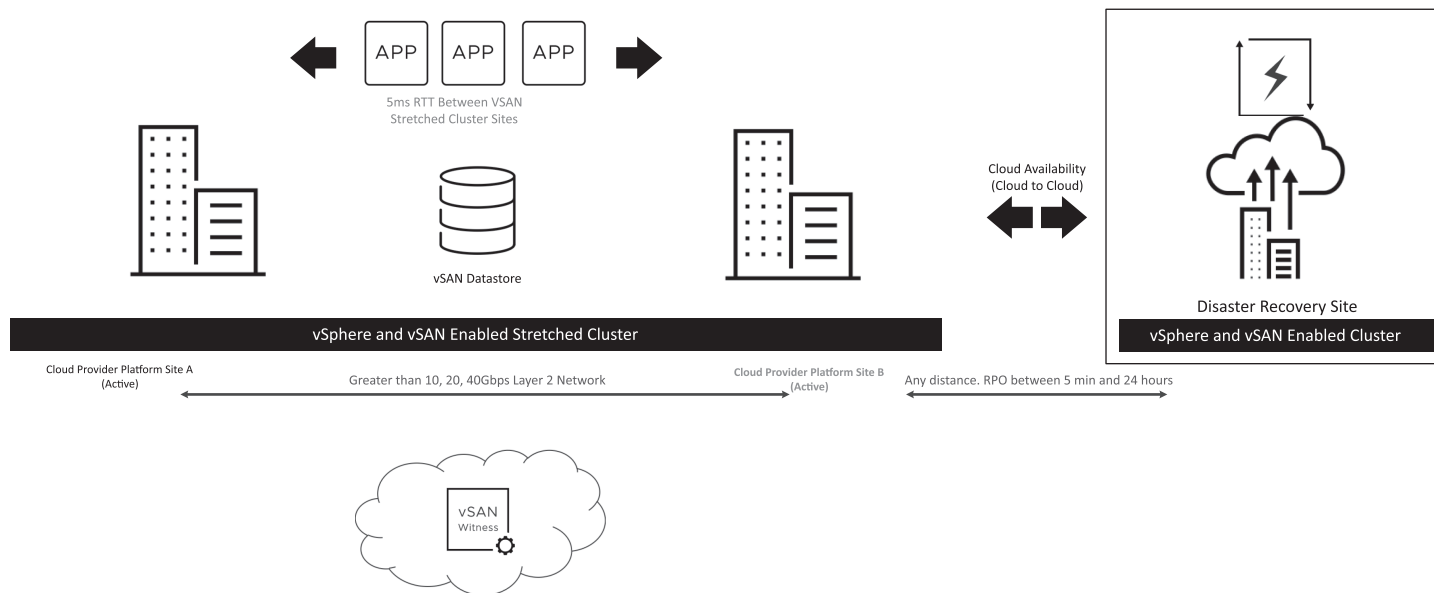


FIGURE 5.10 Three-Site Disaster Recovery Architecture

vSAN Encryption: Data at rest encryption is built into vSAN, and is fully compliant with all major security standards. vSAN builds data-at-rest encryption into the vSAN kernel and enables it at the cluster level. vSAN Encryption encrypts all objects in the vSAN data store using a two-level encryption method: It uses a key encryption key (KEK) to encrypt a data encryption key (DEK). The DEK is a randomly generated key that encrypts data on each disk. Each vSAN host stores the encrypted DEKs but does not store the KEK on disk. If the host requires the KEK, it requests it from the third-party Key Management Service (KMS) server.

As illustrated in Figure 5.11, vSAN Encryption occurs when the vCenter Server requests an AES-256 KEK from the KMS. vCenter Server only stores the KEK's ID, not the key itself. The ESXi host then encrypts disk data with the industry standard AES-256 XTS mode. Each disk has a different randomly generated DEK. Each ESXi host then uses the KEK to encrypt its DEKs and stores the encrypted DEKs on disk. As mentioned above, the host does not store the KEK on disk. If a host reboots, it requests the KEK with the corresponding ID from the KMS. The host can then decrypt its DEKs as needed.

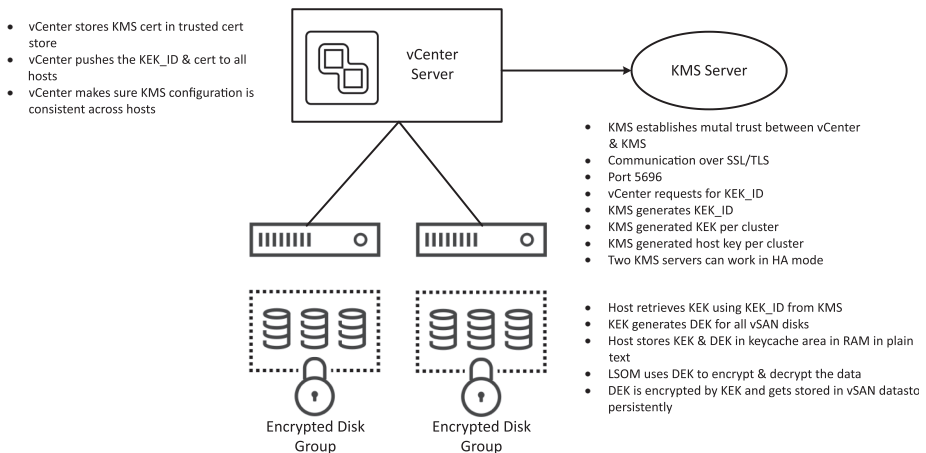


FIGURE 5.11 vSAN Encryption Framework

vSAN Encryption integrates with all KMIP compliant key management technologies, including SafeNet, Hytrust, Thales, Vormetric, etc., and importantly, has no vCenter dependency. However, encryption needs CPU resources, and can introduce some level of CPU overhead. The amount will be dependent on the type and volume of I/O activity.

As with other cluster-wide data services (deduplication and compression), encryption should be enabled before running production workloads. However, it is possible to enable encryption at any time assuming there is adequate capacity in the cluster for a rolling upgrade. When encryption is

enabled, vSAN evacuates each disk group, one at a time, and upgrades/encrypts the disk group. As you can imagine, this is easier and faster if there is little or no existing data in the cluster.

vSAN encrypts the data when it is at rest in the cache and capacity tiers, not when it is in flight up the vSphere stack (to workloads) or across the network. Therefore, vSAN Encryption works fine with third-party backup solutions. Data is also decrypted when it leaves the cache tier and then re-encrypted when it is destaged and written to the capacity tier. This allows vSAN Encryption to work seamlessly with vSAN Deduplication and Compression (Deduplication and compression are applied when data is de-staged from the cache tier to the capacity tier).

Application Support: As vSAN is increasingly utilized for business-critical applications, it has been important to provide additional certification, support, and validation specific to several key workloads. Examples include support for Microsoft Windows Server Failover Clustering, as well as, support for Oracle Real Application Clusters (RAC) and SAP applications including HANA and S4.

Business critical applications, and other application types have increased scalability with vSAN, without the sizing challenges often associated with legacy storage platforms. As a result, vSAN built on vSphere can easily be sized for small workloads, medium or large workloads, or anything in between.

VSAN DESIGN OVERVIEW

Designing storage resources for a cloud model differs from the traditional approach used in modeling storage for non-cloud onsite enterprise environments. Historical data is usually the best guide for designing and determining the size and scale of a vSAN infrastructure. For example, if the average virtual machine requirements, in terms of storage, memory, IOPS, and compute can be correlated, this information can be used to efficiently design the Cloud Provider Platform vSAN infrastructure.

Application data is also required in the design. Information, such as whether the application is cache-friendly, requires file services, and has availability requirements, is critical. After this information has been calculated, it must be correlated and aligned to future growth needs and the financial requirements of the provider with respect to operational and capital expenditure, as well as profitability goals. The objective is to design the most efficient architecture that will meet current and future demand.

vSAN Cluster Design: Generally, for a good design, VMware highly recommends that each node contained within a vSphere cluster has the exact same configuration. When designing a vSAN cluster, similarly configured and sized hosts should be used. This allows for an evenly balanced vSAN cluster

configuration. Otherwise, the performance and capacity expectations will be compromised.

Most cloud providers have a preferred server vendor and, more than likely, they have a preconfigured and validated server (Ready Node) architecture for vSAN. If not, providers can also build their own vSAN nodes using the information provided by the compatibility guide. From an operational standpoint, the Ready Nodes have been the preferred choice by cloud providers for vSAN infrastructure.

A Cloud Director Management Cluster typically has a different profile type than a Resource Cluster. Using vSAN for the management cluster provides an entire integrated stack that is automated by vCenter. The following best practices should be taken into consideration when designing a management cluster backed by vSAN:

- Use a minimum of four nodes, for best availability and space efficiency
- Use a balanced cluster with identical host configurations
- Do not use a stateless boot image. Use an SD card/USB/SATADOM (Serial ATA Disk on Module)

In addition, as a good practice, you should enable vSphere HA) on the vSAN cluster to protect it from host and virtual machine failure.

A single vSAN cluster can contain up to 64 nodes. As a starting point, it is a commonly defined configuration for cloud providers to implement 20 to 24 node clusters, with each node configured with two disk groups to provide flexibility of growth. This approach also provides flexibility over Fault Domain design, for improved data availability and protection.

Determine the number of hosts for a cluster based on the following:

- Amount of available space on the vSAN datastore
- Number of available failures you wish to tolerate in the cluster
- The number of hosts is generally a balance between hardware and space:
 - More ESXi hosts and/or disk groups means higher hardware costs
 - Fewer ESXi hosts and/or disk groups means resource availability could suffer

For instance, if the vSAN cluster has only three ESXi hosts, only a single failure is supported. If a higher level of availability is required, additional hosts are also required. For this reason, VMware recommends that the minimum production configuration should include four hosts per cluster.

Host Memory Requirements: VMware vSAN memory requirements are defined based on the number of disk groups and disks that the ESXi hypervisor manages. To support the maximum number of 5 disk groups in a host, 32 GB of RAM is required. A best practice is to assign a consistent amount of RAM to each host in the cluster. If there are separate host configurations with

disparate amounts of RAM in the hosts, it is recommended to create multiple vSAN clusters. Other considerations when calculating memory requirements for hosts participating in a vSAN cluster include:

- vSAN scales back on its memory usage when hosts have less than 32 GB of memory.
- vSAN consumes additional memory when the number of nodes in the cluster is greater than 32.
- vSAN features such as Deduplication and Compression, and Encryption consume additional memory resources.

Host CPU Overhead: vSAN does not have specific CPU requirements because it introduces less than 10 percent CPU overhead. This might impact available resources in high consolidation ratio workloads and CPU-intensive applications. Therefore, when designing and sizing the environment, take the projected resource requirements for the environment into account so that performance does not suffer due to a lack of resources or the overcommitment of available resources.

Disk Group Design: Disk groups can be thought of as storage ‘containers’ on vSAN hosts. They contain a maximum of one flash cache device and up to seven capacity devices. Either mechanical disks (hybrid configuration) or flash devices (all-flash configuration) are used as capacity devices. Each disk group assigns a cache device to provide the cache for a given capacity device. The recommendation is to have at least a 10-20 percent cache-to-capacity ratio. This provides a degree of control over performance as the cache-to-capacity ratio is based on disk group configuration. This also needs to be taken into account when planning future growth. For instance, you want to ensure that the flash layer devices are large enough to scale the capacity layer for growth. Otherwise, you will not be able to maintain the minimum flash-to-capacity ratio. Depending on the use case, it might be necessary to design with additional cache upfront to allow for future growth of the capacity layer.

To rebuild components after a failure, the design should be sized so that there is a free host’s worth of capacity to tolerate each failure. There should be at least one full host’s worth of capacity free for maintenance. The number of failures to tolerate will determine whether there is a requirement for additional host capacity. For example, to rebuild components after one failure (FTT=1), there must be one full host’s worth of capacity available. To rebuild components after a second failure (FTT=2), there must be two full hosts’ worth of capacity free.

When evaluating hardware for cluster nodes for a Cloud Provider Platform environment, the hardware should be identical, with special attention given to the storage I/O controllers. Queue depth should be as large as possible. At a minimum, the queue depth should be able to accommodate the throughput of current and future devices. Equally important, verify that the storage

I/O controller supports pass-through mode. RAID 0 is not recommended in a Cloud Provider Platform environment due to the increased maintenance of setting up and replacing disks.

Key design decisions must be made about the number of disk groups and the cache ratio. Consider that vSAN:

- Supports up to one flash device for cache and a maximum of seven flash devices for capacity per disk group in an all-flash configuration.
- Supports up to a maximum of five disk groups per host.
- Allow 30 percent slack space when designing capacity.
- vSAN begins automatic rebalancing when a disk reaches the 80 percent of full threshold.
- Target configurations should be approximately 10 percent of the 80 percent threshold.
- Multiple disk groups typically provide better performance and smaller fault domains but might sometimes come at a cost and consume additional disk slots.

The more disks that are configured per disk group, the more cache is needed, and the more capacity that will be available for virtual machines. However, this leads to additional costs due to the disk group limits. Multiple disk groups require one flash device per group for cache and at least one flash device for capacity.

Disk group sizing is also important to consider when designing the total datastore volume. Include the following data points when deciding on the number of disk groups per host:

- Available space on the vSAN datastore.
- Number of failures you want to tolerate in the cluster.
- The optimal number of disk groups is a balance between hardware and space requirements for the vSAN datastore. More disk groups increase space and provide higher availability. However, adding disk groups can be cost-prohibitive.
- The total amount of space in the configuration can be quite significant, based on the number of disks configured in the hosts. Configure the disk groups based on:
 - The end size of the datastore.
 - The tolerance for failure required for the design (both failures-to-tolerate and fault domains).

Sizing the vSAN volume can be approached in several ways, depending on the factors that are most important to the environment. All different policy settings can impact the available space in the cluster. Also, the resultant sizing can mean additional hardware is required or that disk groups must be configured in a specific way.

VMware recommends using the vSAN Sizing tool to determine the correct number of hosts and sizing for the cluster. For proper recoverability, use eight or more hosts in the cluster to support the configured fault domains. The end datastore size and number of hosts is dependent on the end workloads being hosted, and it should be sized accordingly. Therefore, before selecting a drive size, consider disk groups, sizing, and expected future growth.

The Impact on Sizing of the Number of Failures-to-Tolerate Policy: The number of failures-to-tolerate policy setting is one of the core availability mechanisms with vSAN. This policy controls the number of replicas (mirrors) of a virtual machine component, although the policy can be applied to all of a virtual machine's disks, or individual VMDKs. This policy is important when planning and sizing storage capacity because it directly relates to the consumption of capacity that the virtual machine will have on the storage.

For every n failures-to-tolerate, $n + 1$ copies of the data are needed, and $2n + 1$ hosts contributing storage are required. For instance, if the Number of Failures-to-Tolerate capability is set to 1, the virtual machine or disk will have two replica mirrored copies of its components created across the cluster. If the number is set to 2, three mirror copies are created, and so on. If a failure occurs, the mirrors take over.

As you can see, this will lead to greater storage utilization than is actually configured because of the replicas created, which has an impact on I/O amplification in the subsystem. Higher PFTT policies or R5/6 require more I/O work, but they also consume additional components. This setting should be configured based on the availability requirements of the virtual machine or disk.

These requirements are defined in a storage policy and applied to the appropriate workloads. The default is to have one failure to tolerate unless the policy is changed to be a different value. The maximum number of failures to tolerate is three.

The recommendation is to keep the failures to tolerate setting at the default value of 1, unless there is specific need to configure a separate policy that provides a higher level of fault tolerance. If this is the case, this policy should be used sparingly with higher priority virtual machines.

Disk Controllers

The I/O controllers are just as important to a vSAN configuration as the selection of disk drives with performance being highly dependent on the choice of I/O controller. Each vSphere host that contributes storage to the vSAN cluster requires a disk controller. This can be a SAS or SATA host bus adapter (HBA) or a RAID controller. However, the RAID controller must function in one of two modes:

- Pass-through mode (preferred)
- RAID 0 mode

Pass-through mode, commonly referred to as JBOD or HBA mode, is the preferred configuration for vSAN because it enables vSAN to manage the RAID configuration settings for storage policy attributes based on availability and performance requirements that are defined on a virtual machine. These are the only two modes that are supported, with many storage adapters supporting both modes.

Consider the following when selecting a storage adapter:

- What modes does it support? (RAID 0, pass-through, or both.)
 - In RAID 0 mode, SSD performance should be reviewed.
 - RAID 0 mode also comes with operational overhead that can impact performance.
- Storage controller interface speed.
- Number of devices supported for the controller.
- Number of controllers to be used. Multiple controllers can reduce the failure domain and increase speed, but they also increase the cost.
- Controller queue depth is important for performance. Ideally, you should select a queue depth of 256 or higher. This is a significant determining factor for the performance of vSAN.

Regardless of the choice made, vSAN requires complete control of the drives. Performance between pass-through and RAID 0 modes is generally very similar for most interfaces.

When utilizing RAID 0 mode, the storage controller cache should be disabled so that it does not conflict with the SSD drive caches that are controlled by vSAN. The storage controller cache is configurable on some, but not all, storage controllers.

When the storage controller cache cannot be completely disabled in a RAID 0 configuration, VMware recommends that you configure the storage controller cache for 100 percent read cache, effectively disabling the write cache.

The main consideration when utilizing RAID 0 mode for storage controllers within vSAN is the impact on the operational model. RAID 0 mode controllers will typically require interaction with the storage controller software to manage the addition and removal of drives, because each drive is configured as a separate array, individually presenting each disk to vSAN, rather than simply providing a JBOD solution. vSAN performance and reliability can be impacted if this is not configured correctly.

The recommended configuration for the Cloud Provider Platform is to use I/O controllers suited for the design characteristics of the environment, including:

- Model of the SSDs
- Queue depth of the controller
- Number of disks (and corresponding disk groups) being configured

Flash-Based Devices: The vSAN all-flash architecture is based on a two-tier model for performance and capacity. The cache is set to 100 percent writes because read performance is not a factor.

For endurance of the SSD used, standard industry write metrics are the primary measurements used to gauge the reliability of the drive. The measurement VMware uses has been updated to the industry standard, Terabytes Written (TBW) over the vendor’s warranty. Previously the specification used was Drive Writes per Day (DWPD).

When designing a Cloud Provider Platform, it must be understood that a direct correlation exists between the SSD performance class and the level of vSAN performance. In general, the highest-performing hardware supports optimal performance of the solution. Cost, therefore, is the determining factor and might make a lower class of hardware more attractive, even though the performance or size might not be ideal.

The general guidance is to select an SSD/NVMe size that is, at a minimum, 10 percent of the anticipated size of the consumed capacity storage, before the Number of Failures to Tolerate capability is considered. For instance, the SSD/NVMe should be at least 100 GB if usage is estimated to be 1 TB of capacity storage, consumed in a 2TB disk group.

TABLE 5.2 General Caching Guidelines

Read/Write Profile	Workload Type	AF-8 80K IOPS	AF-6 50K IOPS	AF-4 25K IOPS
70/30 Read/Write; Random	Read Intensive, Standard Workloads	800 GB	400 GB	200 GB
>30% Write; Random	Medium Writes, Mixed Workloads	1.2 TB	800 GB	400 GB
100% Write; Sequential	Heavy Writes, Sequential Workloads	1.6 TB	1.2 TB	600 GB

NOTE: Flash-based caching devices do not contribute to the overall size of the distributed vSAN shared datastore. Because they are utilized for write caching, they count only toward the capacity of the vSAN write cache tier.

Network Requirements: VMware vSAN uses the network to transport all information, including communication between the cluster nodes and virtual machine I/O operations. The required network transport is accomplished by

a specially created VMkernel port group that must be configured on all hosts in the cluster.

When designing the network, the architect must consider how much replication and communication traffic is running between hosts, because with vSAN, the amount of traffic directly correlates to the number of virtual machines that are running in the cluster and how write-intensive the I/O is for the running applications.

As with other communications, such as vSphere vMotion, vSAN traffic should be isolated on its own Layer 2 network segment. You can do this with dedicated switches or ports, or by using VLANs, as shown in Figure 5.12.

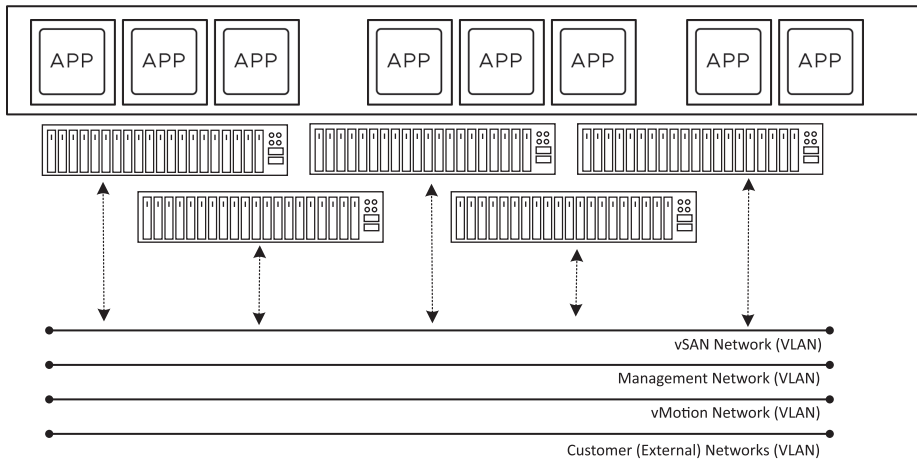


FIGURE 5.12 Network Diagram

Network Interface Cards (NIC): The all-flash architectures are only supported with a minimum 10Gb Ethernet capable network adapter. For redundancy and high availability, a team of network adapters can be configured on a per-host basis. The teaming of network adapters for link aggregation to improve performance is not supported. However, the teaming of network adapters to provide increased availability will be addressed in further detail later in this section. VMware considers this to be a best practice, but it is not necessary in building a fully functional vSAN cluster.

Virtual Switches: vSAN is supported on all vSphere and NSX switch types, including the N-VDS (NSX vSphere Distributed Switch).

The benefit of using the N-VDS switch is that they allow Network I/O Control to be used, which allows for prioritization of bandwidth when there is contention in an environment. In addition, N-VDS switch using Network I/O Control is an attractive option for environments that employ a limited number

of 25-GbE network ports, because it allows the interfaces to be shared, while it prioritizes bandwidth in contention scenarios.

VMkernel Network Port Group: As highlighted previously, the VMkernel interface is used for host intra-cluster communications as well as for read and write operations whenever a vSphere host in the cluster is the owner of a particular virtual machine, but the actual data blocks making up that virtual machine's objects are located on a remote host in the cluster. In this case, I/O must traverse the network configured between the hosts in the cluster. If this interface is created on a N-VDS, the Network I/O Control feature can be used to set shares or reservations for the vSAN traffic.

On each vSphere host, a VMkernel port for vSAN communication must be created and used for synchronization and replication activities.

Jumbo Frames: vSAN supports using jumbo frames for vSAN network transmissions. The environment is supported fully whether or not jumbo frames are employed. The performance gains are often not significant enough to justify the underlying configuration necessary to enable jumbo frames properly on the network.

For this reason, VMware only recommends using jumbo frames for vSAN traffic if the physical environment is already configured to support them, and they are part of the existing design, or if the underlying configuration does not create a significant amount of added complexity to the environment.

VLANs: VMware recommends segregating vSAN traffic on its own VLAN. When multiple vSAN clusters are used, each cluster should use a dedicated VLAN or segment for their traffic. This will prevent interference between clusters and aid in troubleshooting cluster configuration. VMware strongly recommends that a production design always employ separate VLANs for vSAN traffic.

Network QoS and Network I/O Control: Quality of Service (QoS) can be implemented using Network I/O Control for a vSAN environment. This will allow a dedicated amount of the network bandwidth to be allocated to vSAN traffic. When using Network I/O Control, no other traffic will impact the vSAN network, and likewise, through the use of shares, other traffic types can also be protected. This is useful in a shared environment but does require that a vSphere distributed virtual switch or N-VDS be used.

In Network I/O Control, reservation and shares for the vSAN outgoing traffic should be configured as follows:

- Set a reservation so that Network I/O Control guarantees that minimum bandwidth is available on the physical adapter for vSAN.
- Set shares so that when the physical adapter assigned for vSAN becomes saturated, certain bandwidth is available to vSAN to prevent vSAN from

consuming the entire capacity of the physical adapter during rebuild and synchronization operations. For example, the physical adapter might become saturated when another physical adapter in the team fails, and all traffic in the port group is transferred to the other adapters in the team.

In addition to this, priority tagging can be set. Priority tagging is a mechanism to indicate to the connected external network devices that vSAN traffic has higher QoS demands. You can assign vSAN traffic to a certain class and accordingly mark the traffic with a Class of Service (CoS) value from 0 to 7 by using the traffic filtering and marking policy of the vSphere Distributed Switch/N-VDS. The lower the CoS value is, the higher the priority of the vSAN data.

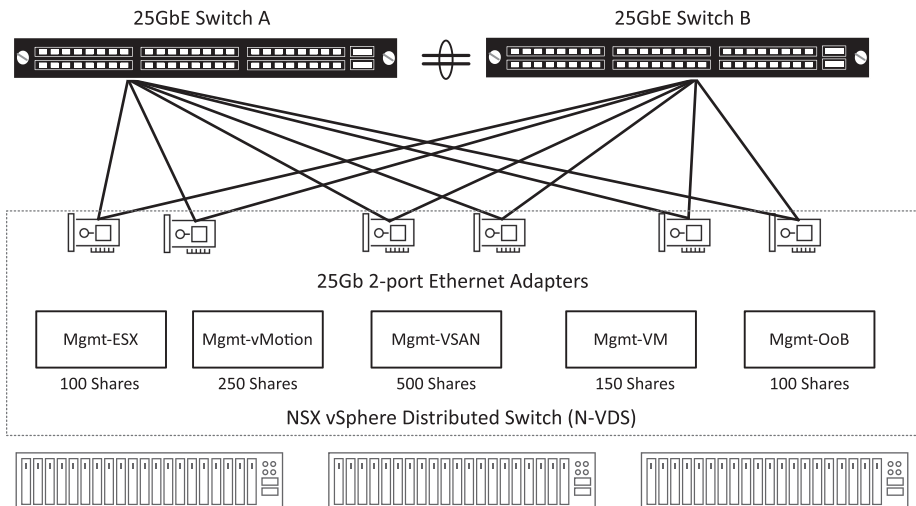


FIGURE 5.13 Network I/O Control

For most designs, make use of Network I/O Control and QoS when the NIC is shared between traffic types, such as if a 25-GbE network card team is used for all traffic on a host system.

vSAN Policy Design: After vSAN is enabled and configured, storage policies that define the virtual machine storage characteristics must be created. These characteristics allow configuration of different levels of service to be provided to a virtual machine. If no specific policy is applied, a default policy

tolerating a single failure and having a single disk stripe is applied. Creating multiple policies with vSAN allows for policies to be applied on a per virtual machine basis and changed as needed.

As discussed earlier, vSAN is an object storage technology. Each virtual machine deployed on vSAN comprises of a set of objects. VMDKs, snapshots, VM swap space, and the VM home namespace are all object types. Each of these objects is comprised of a set of components that are determined by capabilities configured in the VM storage policy. For instance, if a virtual machine is deployed with a policy to tolerate one failure, objects will be made up of two replica components. If the policy contains a stripe width, the associated object will be striped across multiple devices in the capacity layer. Each stripe is a component of the object.

As you plan the design with respect to component maximums, be aware that vSAN might decide that an object needs to be striped across multiple disks, even with the default policy of one stripe in place. Normally, this is the result of an administrator requesting that a VMDK be created that is too large to fit on a single physical drive. As previously highlighted, the maximum component size is 255 GB. To elaborate, objects that are greater than 255 GB in size will automatically be divided into multiple components. For example, if a customer deploys a 2-TB VMDK, it will result in eight or more components being created in the same RAID-0 stripe configuration.

vSAN guarantees the storage policies that have been configured. If a policy cannot be met, virtual machine provisioning fails, unless force provisioning has been set. Almost any combination of the possible attribute capabilities can be set in a policy. Some additional benefits of vSAN policy application include the ability to configure policies at any time and to switch policies ‘on-the-fly’ for virtual machines.

Finally, all virtual machines created on vSAN will be thin-provisioned, so a design must deal with capacity management accordingly to avoid overcommitting resources. An understanding of the policies that can be set allows the administrator to avoid configuring policies that do not make sense.

vSAN Policy Capabilities: vSAN allows you to set several different policy attributes in a storage policy. These attributes can be used alone or combined to provide different service levels.

Before making design decisions, understand the policies and the objects to which they can be applied. The policy options are listed in Table 5.3.

vSAN Default Policy: By default, policies are configured based on the application requirements. However, they are applied differently depending on the object. Table 5.4 lists the default policy options applied to the different objects in vSAN.

TABLE 5.3 vSAN Policy Options

Capability	Use Case	Value	Comments
Number of disk stripes per object	Performance	Default 1 Maximum 12	This is a standard RAID 0 stripe configuration used to increase performance for a virtual machine disk. This setting defines the number of disks on which each replica of a storage object is striped. If the value is higher than 1, increased performance can result. However, an increase in system resource usage might also result.
Flash read cache reservation (%)	Performance	Default 0 Maximum 100%	Flash capacity reserved as read cache for the storage is a percentage of the logical object size that will be reserved for that object. Use this setting only for workloads that must have read performance issues addressed. The downside is that other objects cannot use a reserved cache. VMware recommends not using these reservations unless it is necessary because unreserved flash is shared fairly among all objects.
Number of failures-to-tolerate	Redundancy	Default 1 Maximum 3 Note: Maximum value is 1 if disk size is greater than 16 TB	Defines the number of host, disk, or network failures a storage object can tolerate. The higher the value, the more failures can be tolerated. When the fault tolerance method is mirroring, for n failures tolerated, $n+1$ copies of the disk are created, and $2n+1$ hosts or fault domains contributing storage are required. The higher n value indicates that more replicas of virtual machines are made, which can consume more disk space than expected. When the fault tolerance method is erasure coding, to tolerate 1 failure, 4 hosts (or fault domains contributing storage) are required, and to tolerate 2 failures, 6 hosts or fault domains are required.

(continued)

TABLE 5.3 vSAN Policy Options (*Continued*)

Capability	Use Case	Value	Comments
Failure tolerance method	Performance or Capacity	Default: RAID 1 (Mirroring)	Defines the method used to tolerate failures. RAID 1 achieves failure tolerance using mirrors, which provides better performance. RAID 5/6 achieves failure tolerance using parity blocks, which provides better space efficiency. RAID 5/6 is only available on all-flash vSAN clusters, and when the number of failures-to-tolerate (FTT) is set to 1 or 2. A value of 1 FTT implies a RAID 5 configuration, and a value of 2 FTT implies a RAID 6 configuration.
IOPS limit for Object	Performance	Default: 0	Defines IOPS limit for a disk IOPS is calculated as the number of I/Os using a weighted size. By default, the system uses a base size of 32 KB. Thus, by default, a 64 KB I/O will represent 2 I/O. When calculating IOPS, read and write are regarded as equivalent and cache hit ratio or sequence is not taken into account. If IOPS of a disk exceeds the limit, I/O will be throttled. If the limit is set to 0, it means no limit is being applied.
Disable Object Checksum	Override Policy	Default: No	vSAN uses end-to-end checksum to ensure the integrity of data by confirming that each copy of a file is exactly the same as the source file. The system checks the validity of the data during read/write operations, and if an error is detected, vSAN repairs the data or reports the error. If a checksum mismatch is detected, vSAN automatically repairs the data by overwriting the incorrect data with the correct data. Checksum calculation and error-correction are performed as background operations. This setting determines whether checksums will be calculated for the data being written to the volume or not.

Force provisioning	Override policy	Default: No	Force provisioning allows for provisioning to occur even if the policy configured cannot be satisfied by the currently available cluster resources. This is useful if there is a planned expansion of the vSAN cluster, and provisioning of VMs must continue, because vSAN will automatically try to bring the object into compliance as resources become available.
Object space reservation (%)	Thick provisioning	Default 0 Maximum 100%	The percentage of the storage objects that will be thick provisioned upon VM creation. The remainder of the storage objects will be thin provisioned. This is useful if a predictable amount of storage will always be filled by objects, cutting back on repeatable disk growth operations for all but new or non-predictable storage use.

TABLE 5.4 vSAN Default System Policy

Object	Policy	Comments
Virtual Machine Namespace	1 Failures-to-Tolerate	Configurable. However, changes are not recommended.
Swap	1 Failures-to-Tolerate	Configurable. However, changes are not recommended.
Virtual Disks	User-Configured Storage Policy	Can be any storage policy configured on the system.
Virtual Disk Snapshots	Uses Virtual Disk Policy	Same as Virtual Disk policy by default. Changing is not recommended.

If there is no user-configured policy, the default system policy of one failure-to-tolerate and one disk stripe is used for virtual disks and virtual disk snapshots.

Policy defaults for the VM namespace and swap are set statically and are not configurable, so that there is appropriate protection for these critical virtual machine components. Policies must be configured based on the application’s requirements. This feature gives vSAN its power, because it can adjust how a disk performs on demand based on the configured policies.

Application Requirements: Policy design starts with an assessment of business needs and application requirements, because policies allow any configuration to become as customized as needed. Using general policies is advisable if no specific use cases exist, or if a proof of concept (PoC) is under development.

Start by assessing the following different application requirements:

- I/O performance and profile of your workloads on a per-virtual-disk basis
- Working sets of your workloads
- Hot-add of additional cache requires repopulation of cache
- Specific application best practice (such as block size)

Aim to design policies for availability and performance in a conservative manner, so that space consumed, and recoverability properties are balanced. In many cases, the default system policy is adequate, and no additional policies are required, unless there are specific requirements for performance or availability.

vSAN and Recoverability: vSAN enables providers to protect consumer data, even in the event of disk, host, network, or rack failures, with built-in distributed RAID and cache mirroring. Resiliency is enhanced by defining availability on a per-VM basis. The number of hosts, network, disk, or rack failures to

tolerate in a vSAN cluster can be customized based on customer, application, or business requirements.

vSAN supports clustering technologies from both Oracle and Microsoft. With Oracle Real Application Cluster (RAC), customers can run multiple Oracle RDBM instances, accessing the vSAN datastore to deliver better performance, scalability, and resilience. Additionally, Windows Server Failover Clustering (WSFC) can provide protection against application and service failures.

Technically, there is a minimum requirement of three ESXi hosts in a vSAN cluster. Although VMware fully supports three-node configurations, there are limitations when compared with configurations of four or more nodes. In particular, in the event of a failure, there is no way for vSAN to rebuild the components on another host in the cluster to tolerate another failure. There is also an operational limitation because vSAN will not have the ability to migrate all data from a node during maintenance.

Business continuity and disaster recovery are paramount for ensuring that consumers' business critical environment, data, and online presence are available with minimal downtime. vSAN recoverability can be addressed through a range of solutions, from near-line capabilities such as snapshots, offline capabilities provided by vSphere storage integrated backup solutions (using APIs for data protection), or inter-data center solutions, such as a stretched cluster.

Understanding Failure Impact: vSAN was built with a native ability to handle failures. It can handle transient failures, such as a network misconfiguration, or a full device failure, such as an SSD. vSAN classifies failures in two distinct ways:

- Absent failures
- Degraded failures

The way that vSAN handles a failure depends on its type. The following sections describe these two types of failures.

Absent Failures: Absent failure events are recognized in vSAN whenever I/O failures are detected with any of the following hardware devices:

- Physical network switches
- Network interfaces (NICs)
- ESXi hosts or fault domains

When an absent failure occurs, all data is not immediately resynchronized to another host in the environment. The resynchronization operation is in fact not initiated for 60 minutes from the detection and acknowledgement of the failure event.

The reason for this is that these types of failures can often be temporary. For instance, the network could be down, or a host might be rebooting for regular maintenance. Therefore, the system waits for these components to come back online before committing to a resynchronization of all the data on the node.

While the default delay is 60 minutes, this is an adjustable system parameter that can be increased or decreased based on business requirements. For example, you might want to adjust the parameter if a host takes longer than the set delay to reboot, or if the host will not be available during a maintenance window that is a longer period of time than the set delay.

Degraded Failures: Degraded failure events are recognized in vSAN whenever I/O failures are detected with storage devices or storage controllers. Unlike absent failure events, degraded failure events immediately activate the resynchronization of the data among the other hosts in the cluster. This function is not configurable because when one of the listed devices fail, the failure is not likely to be temporary. Therefore, there is no benefit to waiting for the device to come back online.

When this type of event is detected, the resynchronization operation is performed for objects that are no longer in compliance with their policies due to the failure. It is an expensive operation because it creates new replicas of the data that existed on the failed device or component, using the remaining replica for the component as the source. This data exists elsewhere in the cluster, on other hosts or on other mechanical disks of the same host, based on the configured failures-to-tolerate (FTT). The replicas of individual objects are not all created in the same place. Rather, the replicas are distributed around the rest of the cluster wherever there is spare capacity. Thus, the entire cluster is used as a hot spare.

Regardless of when it is activated, the resynchronization operation contends with the I/O of virtual machines and the resource available in the cluster. The operation could have a detrimental impact on the overall capabilities of the vSAN cluster if not sized and designed correctly. From a performance perspective, the resynchronization operation essentially limits the IOPS available to virtual machines in the cluster because of the operations being performed to recover the affected objects and components.

From a data availability perspective, whenever the resynchronization operation is unable to be completed, due to an inadequate amount of collective spare capacity in the cluster, data accessibility could be at risk. For instance, consider a scenario with a three-node cluster configured with the default availability policy setting of FTT=1. When a host failure occurs, the remaining two hosts are excluded from the resynchronization operation because they are already hosting objects and components for the affected virtual machines. In this scenario, the resynchronization operation waits until the failed node is brought back online, or a new one is added to the cluster, before resuming the resynchronization operation, and restoring the compliance value for data availability.

Summary of Key vSAN Design Factors

There are a number of considerations that are needed, prior to configuring a cluster to participate in vSAN on the Cloud Provider Platform. Table 5.5 shows the minimum key requirements to implement a vSAN-enabled cluster.

TABLE 5.5 Key Virtual SAB Design Factors

Design Factor	Detail
Controller Queue Depth	Controller queue depth impacts the rebuild/resync times. A low controller queue depth may impact the availability of production virtual machines during rebuild/resync. A minimum queue depth of 256 is required for vSAN.
Number of disk groups	The number of disk groups impacts fault isolation as well as rebuild/resync times. Fault isolation: Configuring more than 1 disk group allows better tolerance against failures, because data is spread across more disk groups. Rebuild/resync times: Configuring more than 1 disk group allows faster rebuilds/resyncs.
Number of drives in a disk group	The number of SSDs in a disk group has an impact on the performance of vSAN more capacity drives improves performance when there are more virtual machines and provides better handling of rebuild/resync activities.
Class of SSDs	The class of SSD you choose has a direct impact on the performance of your overall system.
Balanced compared with unbalanced cluster	An unbalanced cluster can impact vSAN performance, as well as the rebuild/resync times. A balanced cluster delivers more predictable performance, even during hardware failures. In addition, performance impact during resync/rebuild is minimal when the cluster is balanced.
10-Gb Ethernet	Ethernet speed has an impact on the vSAN performance. 10 GbE or greater is required.

vSAN Stretched Cluster Deployments

Stretched storage with vSAN allows you to split the vSAN cluster across two sites, so that if a site fails, providers are able to seamlessly fail over to the other site without any loss of data. vSAN in a stretched storage deployment accomplishes this by synchronously mirroring data across the two sites. The failover is initiated by a witness VM that resides in a central location, accessible from both sites. This is a specific implementation for environments where disaster avoidance and no unscheduled downtime are key requirements.

A vSAN stretched cluster is a specific deployment where the provider sets up a vSAN cluster with two disparate Active/Active sites with an identical number of ESXi hosts distributed evenly between the two sites. The witness host resides at a third site and the sites are connected by way of a high-bandwidth, low-latency link. The third site, hosting the witness host, is connected to both of the Active/Active data sites. The sites can be a combination of cloud provider, customer, and third-party data centers.

In a vSAN stretched cluster implementation, each site is configured as a vSAN fault domain, and there is only one witness host in any configuration. Each site can be considered a fault domain and a maximum of three sites (two data, one witness) is supported. For deployments that manage multiple stretched clusters, each cluster must have its own unique witness host.

A virtual machine deployed on a vSAN stretched cluster will have one copy of its data on site A, a second copy of its data on site B, while any witness components are placed on the witness host in site C. This configuration is achieved through fault domains and affinity rules. In the event of a complete site failure, there will be a full copy of the virtual machine data available, as well as greater than 50 percent of the components. This will allow the virtual machine to remain available on the vSAN datastore. If the virtual machine needs to be restarted on the other data site, vSphere HA will accommodate this task.

Geographic distance is, in theory, not a concern when designing a vSAN stretched cluster. The key requirement is the latency between the respective sites. VMware requires a maximum latency of no greater than 5 ms RTT (Round-Trip Time) between data sites and no more than 200 ms RTT between data sites and the witness host. As long as the latency requirements are met, there is no restriction on geographic distance.

As highlighted in Figure 5.14, a vSAN stretched cluster requires three disparate sites, and each site must communicate on the management, vSAN, VM, and vSphere vMotion networks. To minimize uncertainty with the implementation, VMware recommends that providers implement a stretched Layer 2 network between the data sites, and a Layer 3 configuration between the data sites and the witness site.

In the example illustrated in Figure 5.14, the virtual witness is connected over Layer 3 with static routes. The witness is deployed on a physical ESXi host with two preconfigured networks for the management and vSAN networks, respectively. The data sites are connected by way of stretched Layer 2 which backs the management, vSAN, VM, and vSphere vMotion networks. All hosts in the cluster must be able to successfully communicate, and to facilitate this communication, static routes must be configured (per host) between the data hosts in Site A and B, and the witness host in Site C, for vSAN traffic to flow between the data sites and witness site.

A stretched cluster architecture can be offered as part of the Cloud Provider Platform, often for a premium service, offering higher levels of availability to customers who demand continuous uptime for their business-critical applications.

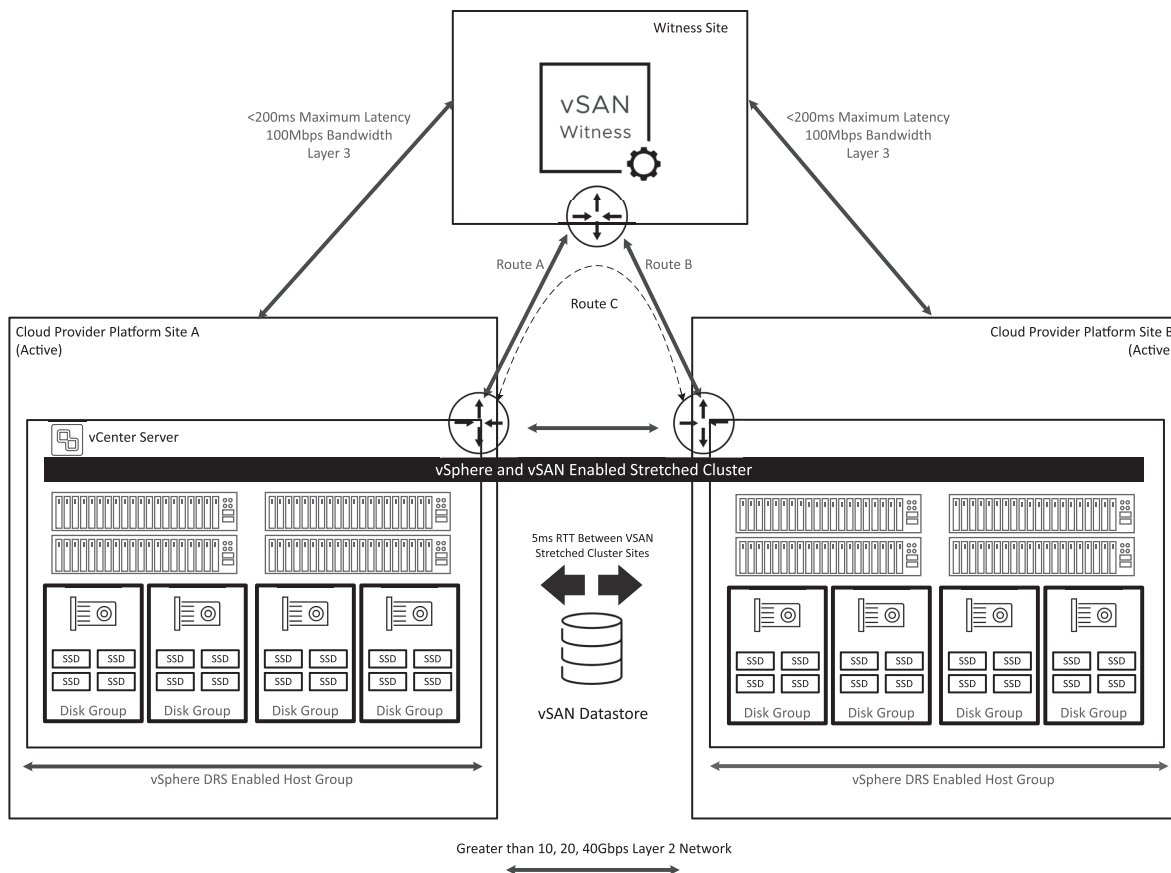


FIGURE 5.14 Stretched Cluster Architecture

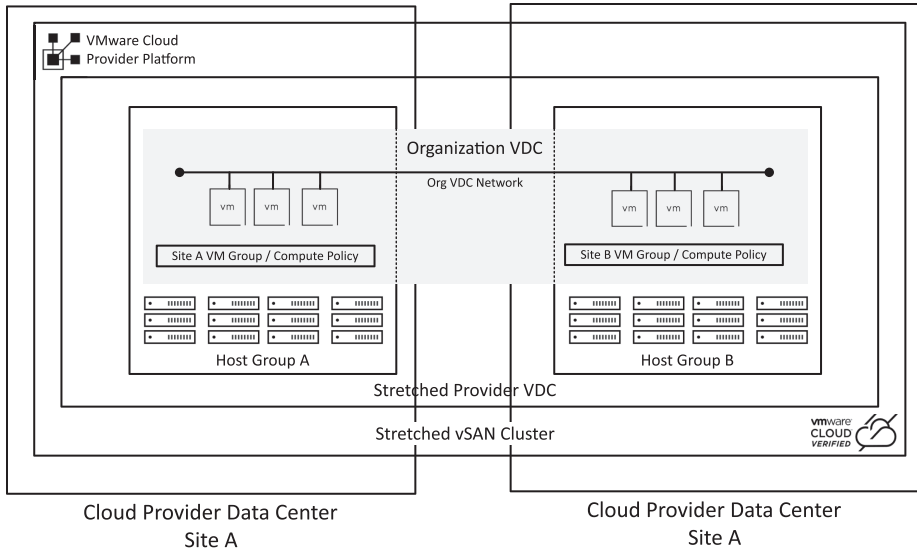


FIGURE 5.15 Cloud Director Provider VDC over a vSAN Stretched Cluster

Even though the stretched architecture is entirely transparent to Cloud Director itself, with a well thought out design, alongside the use of the VCD workload placement engine, stretched solutions such as these can offer a cloud provider a key differentiator against the competitive offerings from hyperscale cloud services and deliver levels of availability and disaster avoidance that can meet the requirements of the most demanding customers.

Ultimately, the success and design considerations of a specific vSAN stretched cluster implementation depend upon many factors ranging from choice of topology to the physical capabilities of a provider's networking infrastructure.

VSAN STORAGE AND CLOUD DIRECTOR

Storage is of course a key resource, which Cloud Director abstracts to deliver infrastructure within the Virtual Data Center (VDC), and therefore configuring vSAN storage resources to support a Cloud Provider Platform requires the consideration of a range of provider use cases and workload types. In addition, as HCI storage with vSAN is probably one of the most significant areas of change for many cloud providers it therefore will be one of the topics of most interest.

As we have seen in Chapter 3, VMware Cloud Director creates an abstraction layer on top of vSphere which pools data center resources, such as compute, memory, networks and storage, along with their associated policies,

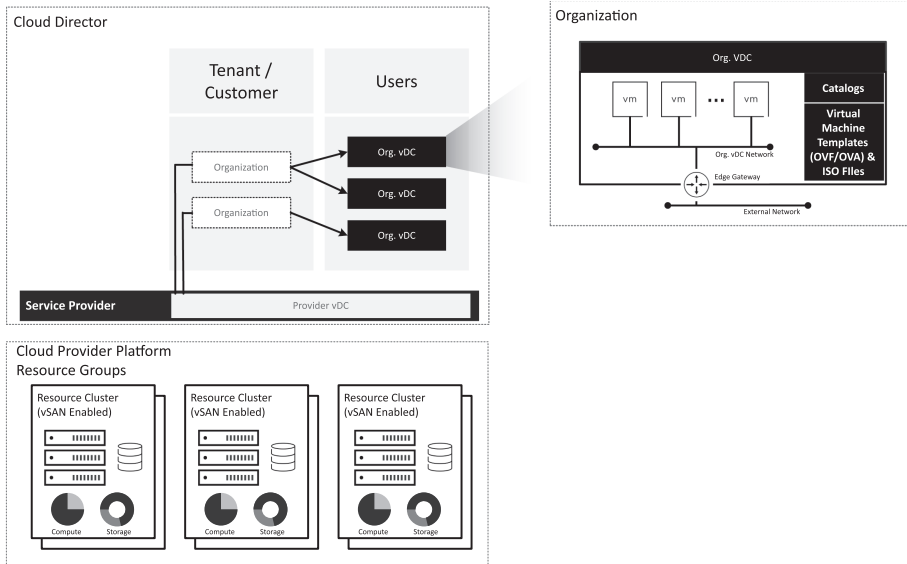


FIGURE 5.16 Cloud Director Platform

into Virtual Data Centers (VDCs). In a Cloud Director environment all physical hardware resources are presented firstly to VMware vCenter Server, and subsequently carved up into logical data centers, referred to as Provider Virtual Data Centers.

The customer's organization, defined within Cloud Director, draws on the Provider VDC resources, through the pooled resources assigned to their respective Org. VDCs. Each Organization VDC can only consume resources from a single Provider Virtual Data Center, although a Provider VDC can provide resources to multiple Org. VDCs. The consumer's deployed workloads reside within the tenants Org. VDC.

By employing these constructs, Cloud Director allows cloud providers to create logical data centers, composed of compute, network and storage resources, mapped from the underlying physical components. This architecture is illustrated in Figure 5.16.

Provider VDCs play an important part in employing VMware vSAN as part of a Cloud Provider Platform and pose the following design factors for consideration:

- A Cloud Director Organization may utilize multiple Org. VDCs
- An Org. VDC can only consume resources from a single Provider VDC
- Provider VDCs can provide resources to multiple Org. VDCs, as illustrated in Figure 5.16

As outlined above, Cloud Director customer's vApps and virtual machines represent the tenants deployed workloads, the resource requirements for

which can vary significantly. Some will require more compute resources, while others, may require more disk intensive operations to be processed. Cloud Director, by design, allows Org. VDCs to be categorized based on workload requirements, or tier of performance required.

In order to address these varying workload requirements, Provider VDCs must typically be configured to offer multiple service levels within the Cloud Provider Platform environment. These are sometimes categorized as either precious metals, such as Gold, Silver and Bronze, or as various tiers of service, which may include Tier 0, Tier 1, Tier 2 or Tier 3.

However, many cloud providers also avoid these types of labels, after all, cloud customers may not like the idea of placing workloads on lower tier or lower values of storage. Therefore, in the following examples I will categorize storage as either Fast, Standard or Capacity. The following table provides an example of tiered Provider VDCs.

TABLE 5.6 Example of resource tiered Provider VDCs within Cloud Director

PROVIDER VDC Tier of Service	Host CPU	Host Memory	Storage Configurations
Fast (Tier 0)	24 Core CPUs	1024GB	vSAN All-NVMe-Flash Configuration
Standard (Tier 1)	12 Core CPUs	512GB	vSAN NVMe Cache/SAS SSD Capacity Drives
Capacity (Tier 2)	12 Core CPUs	512GB	External Storage Array (Multiple Storage Profiles)

Accurately capturing tenant’s application performance requirements, which must be satisfied by each level of service, remains a challenge for cloud providers. While guest operating systems typically have well understood requirements, it is more difficult to quantify application performance requirements, particularly when it comes to storage. In addition, workload storage requirements can be measured in various different ways. Some cloud customers will define their requirements into Read/Write IOPS, maximum latency requirements and sequential or random I/O profiles, whereas other customers define their storage requirements from an application perspective, and request resources to meet specific application requirements.

Either way, capturing tenant’s workload performance requirements can be challenging for cloud providers, who might have little or no control over the workloads being deployed, and while applications and operating systems typically have well-defined requirements for CPU and memory consumption, it is often far more difficult to quantify the performance characteristics of storage resources. For instance, a customer may simply request storage to support:

TABLE 5.7 Capturing Tenant Storage Workload Requirements

	Storage Profile Requirement 1	Storage Profile Requirement 2	Storage Profile Requirement 3
I/O Profile Requirement	1000 heavy Microsoft SQL 2019 users, with 25 databases.	Oracle Application (1250 IOPS) with a 60:40 Read/Write.	Data Warehouse Application, (600 IOPS) with a 100% storage read profile.
Capacity Requirement	25TB capacity requirement.	3TB capacity requirement.	16TB capacity requirement.

It's important to define both performance and capacity requirements from cloud consumers, as often application storage profiles are not simply about performance or capacity, but a balance between the two, which can be the storage design challenge within a Cloud Director environment.

Providing storage resources which can meet all of the various performance requirements set out by tenants, while also providing the ability to scale and manage changes or spikes in demand for resources are key design requirements for cloud provider partners, when implementing the Cloud Provider Platform with vSAN storage resources. The key aim for any service provider is to allocate storage as efficiently as possible across the platform, so that resources can be utilized by Org. VDCs with a minimum of waste and optimum performance, delivering operational stability and accurate billing. This optimized configuration and allocation of storage across Provider VDCs and consumption by Org. VDCs, is key to achieving an efficient tiered storage model.

vSAN Storage Resource Isolation

VMware vSAN offers a storage layer for multitenant environments and is fully capable of offering isolation of shared storage resources to tenants. In this shared storage model, cloud providers can allow tenants to place workloads across shared storage resources. The exact location of the tenant's data on the vSAN-enabled cluster is unable to be specified with this model, as vSAN provides common storage resources within the Provider VDC only.

As a result of this architecture, in any multi-tenanted environment, where various different business entities are storing data across a common vSAN-enabled cluster, tenant data is highly likely to be shared across all disk groups, with no way of creating tenant isolation. However, this does not differ significantly from the storage model employed with array-based shared storage systems, as illustrated in Figure 5.17.

In the past it has not been uncommon for cloud providers to deliver *dedicated* storage to Cloud Director customers by provisioning LUNs on a customer by customer basis. However, even in this model, it is highly likely that

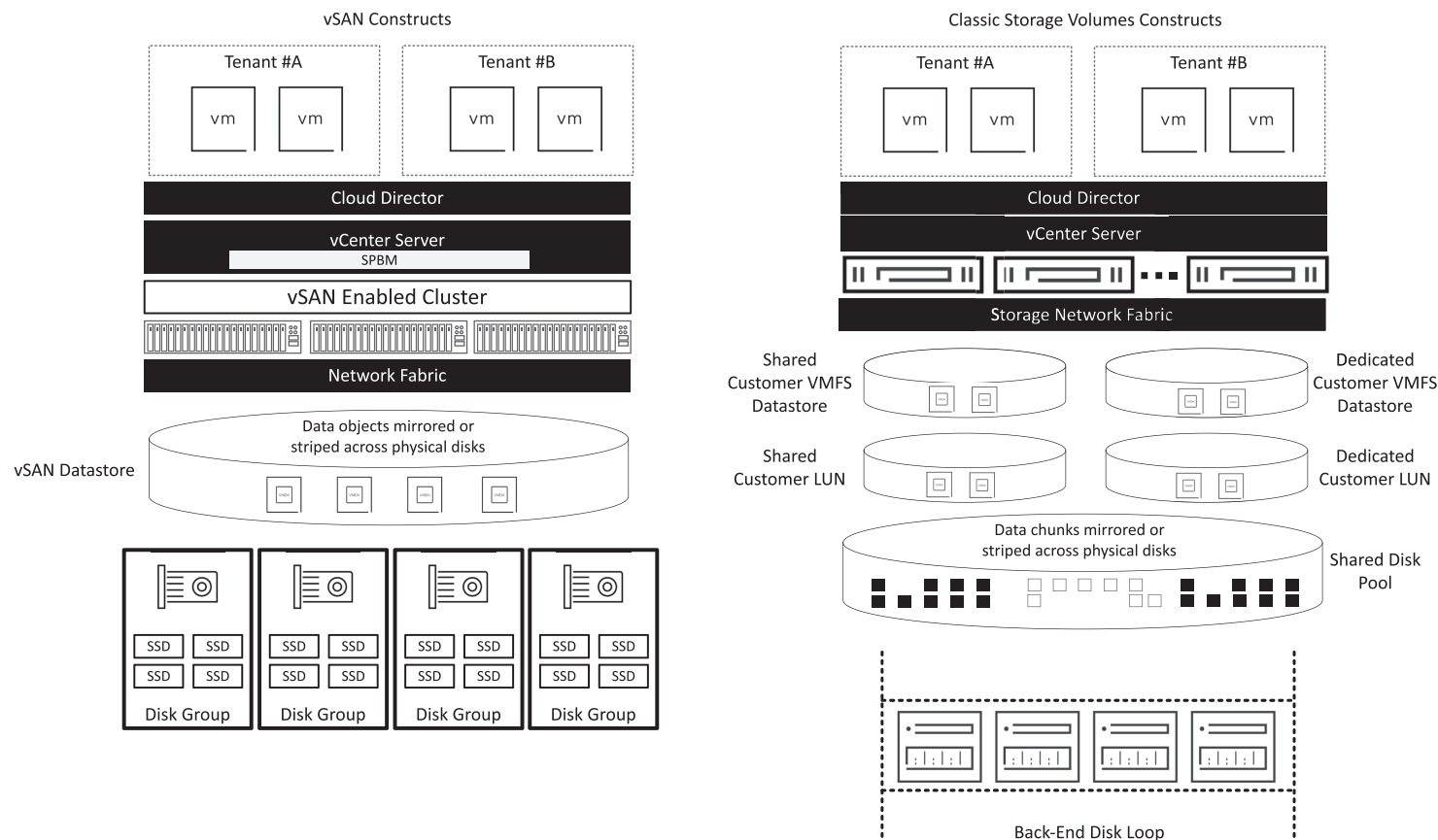


FIGURE 5.17 vSAN Datastore Abstraction verses Storage Array Volumes

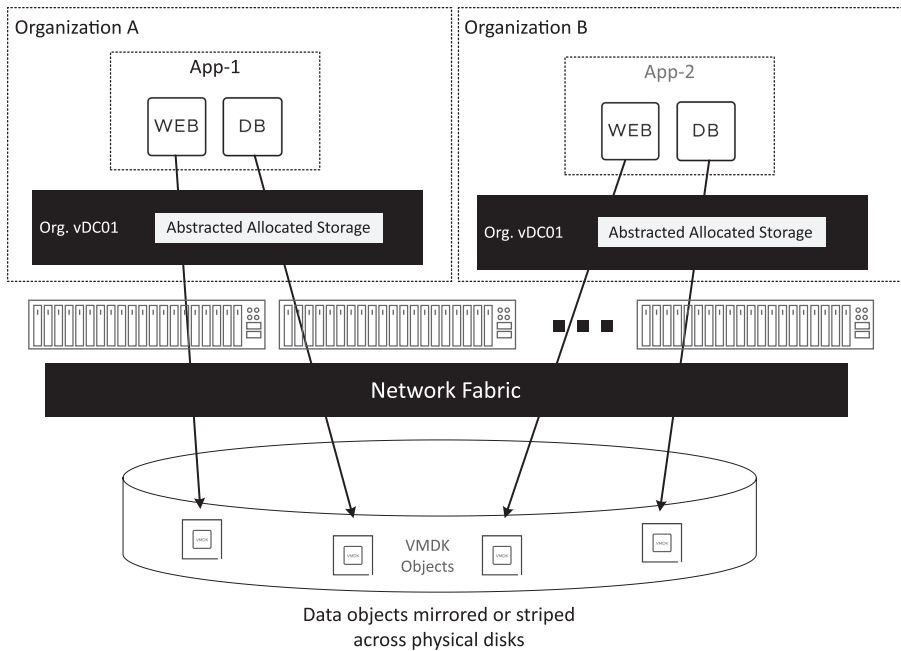


FIGURE 5.18 Cloud Director Storage Isolation

the actual tenant data is striped or mirrored across a common set of physical disks, as illustrated in Figure 5.17.

However, in a Cloud Director environment, this is not typically of concern. The reason for this is while a vSphere administrator can see the physical storage systems, including a vSAN datastore, these datastores are created through vCenter Server, which requires specific rights, which are not required by a cloud customer, consuming resources through Cloud Director.

Therefore, even if the Cloud Director cell is configured with an user with rights to create new datastores, that privilege is not exposed to the provider system administrator or Organization users. Instead, system administrators enable datastores, disable datastores and assign datastores to Organizations.

In addition, Organization administrators and users, who create and upload vApps, simply store the VM VMDKs on the vSAN datastores assigned to the Org. VDC, without any visibility of the backend disk system providing the resource.

For this reason, virtual machines can never see any storage outside of their VMDKs, unless they have network connectivity to other storage systems (such as NFS shares or CIFS), where a cloud provider gives access to external storage to workload VMs, Containers or vApps as a network service, but this must be separate from the vSAN datastore assigned to the vSphere hosts backing the Cloud Director platform.

Likewise, cloud users only see the storage policies backed by one or more of the vSAN datastores assigned to them, but even that view is limited to the Cloud Director abstraction layer and are therefore unable to browse

the datastore. They only see what’s published in catalogs or that exist within workloads they manage.

Therefore, for the reasons outlined above, cloud providers can only achieve true storage resource isolation when employing vSAN in a Cloud Director environment, by assigning dedicated Provider VDC resources per-tenant, which is not typically an efficient approach to resource management, and for most customer’s, unnecessary.

However, this approach does enable cloud providers to eliminate any risk of resource contention across tenant applications, while simultaneously ensuring the secure separation of user data. This approach might also provide advantages to cloud providers where performance and capacity requirements from the tenant are dynamic and encompass a wide range of application types.

Cloud Director Storage Design with vSAN

vSAN storage policies focus on availability, with some slight exceptions such as IO limits, read cache and striping, depending on the working dataset, read

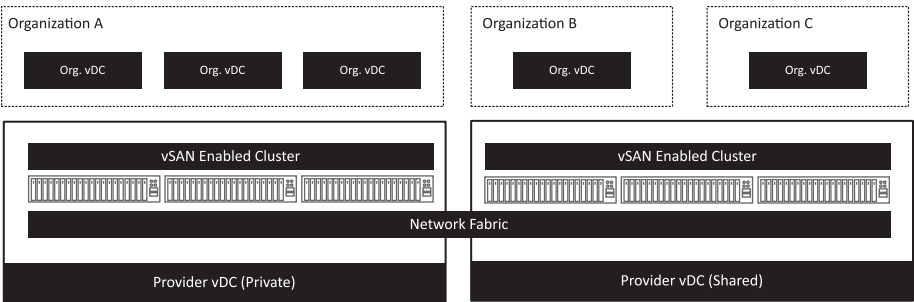


FIGURE 5.19 vSAN Storage Resource Isolation

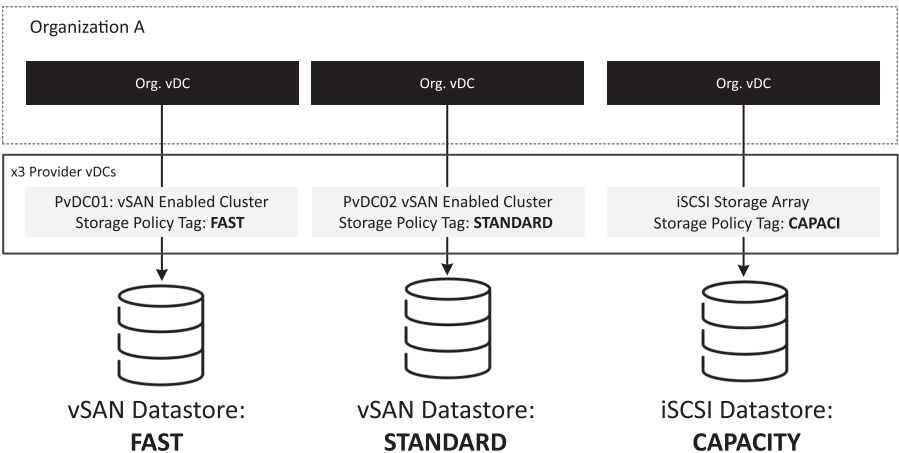


FIGURE 5.20 Model 1: Multiple Provider VDC Storage Tiers, Supporting Multiple Org. VDCs

profile, configuration and design. Therefore, in order to provide multiple tiers of storage performance, multiple resource clusters, each with different defined capabilities are required to be configured, each containing a unique vSAN configuration, with the resources being consumed by Organizational VDCs. In this paper we will address two potential design models which can be considered when employing vSAN as the primary storage platform for a Cloud Provider Platform environment:

1. Multiple Tiered Resource Clusters as dedicated Provider VDCs (1:1 Resource Cluster/Provider VDC)
2. Single Provider VDC, hosting multiple tiered resource clusters

Model 1: Multiple Tiered Resource Clusters as Dedicated Provider VDCs:

As outlined in the illustration, in model 1 each Provider VDC (mapped 1:1 with resource cluster tier) has been configured with a vSAN datastore, which meets the specific capability requirements set out by the cloud provider's service-level agreement for that tier of service.

As illustrated in Figure 5.21, the Cloud Director design contains two Provider VDCs offering different tiers of storage. The Fast performance Tier (Tier 0), backed by an all-NVMe-Flash vSAN-enabled cluster, which has been offered up to the *Production Org.* VDC, via Provider VDC01, while Provider VDC02 offers the Standard of storage to its consumers, within the *Test & Dev Org.* VDC.

The compute consolidation ratio and vSAN capability, based on the disk group configuration and storage policy, defines how these offerings will perform for a customer. In addition, not shown in the configuration below, NIOC and QoS are being employed by the cloud provider to ensure an appropriate balance of network resources are assigned, based on tier of service.

Figure 5.21 illustrates how two different tiers of storage can be manually configured. However, the exact disk configuration will vary depending on hardware manufacturer and provider SLAs.

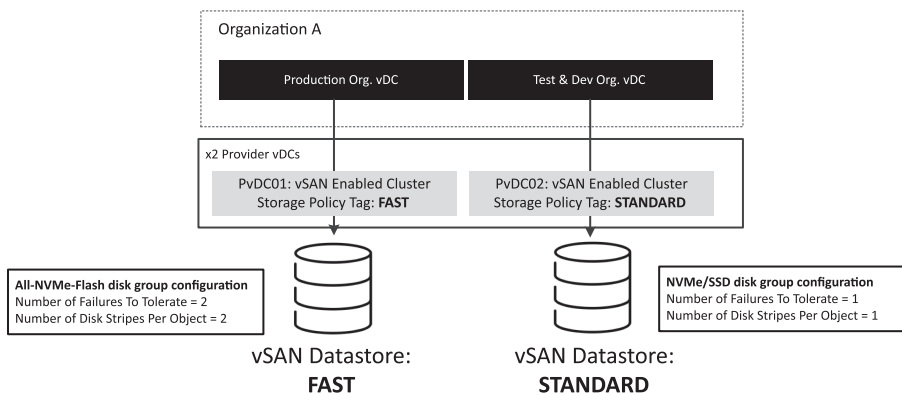


FIGURE 5.21 Manual Storage Tiering within a Single Organization

As you would expect, the cloud provider would configure individual cost models and rate factors specific to each storage tier, to accurately charge for the utilization of the storage resources by the various workloads within the customer Organizations.

By employing this design strategy, the Cloud Director administrator can take advantage of the multiple tiers of storage offered through vSAN. Different tiers of storage can also be easily mapped to cost models, which can be configured to reflect the level of storage being offered and consumed within the organization.

Typically, tiered service offerings such as these are defined by more than just storage capability. vCPU consolidation ratios, levels of guaranteed memory and network resources and backups etc. will all be employed by a service provider to define the SLAs. This will be addressed in more detail later in this paper.

Figure 5.22 shows how the solution is constructed on VMware technologies. The core vSphere platform provides the storage capability through vSAN, which in turn is abstracted via Cloud Director. The vSAN disk group configuration, along with the storage policy, is configured at the vSphere level and automatically presented through to the Cloud Director layer. These constructs define the performance, availability and capacity capabilities of the vSAN datastore, which in turn is employed to define the SLAs for this tier of the cloud offering.

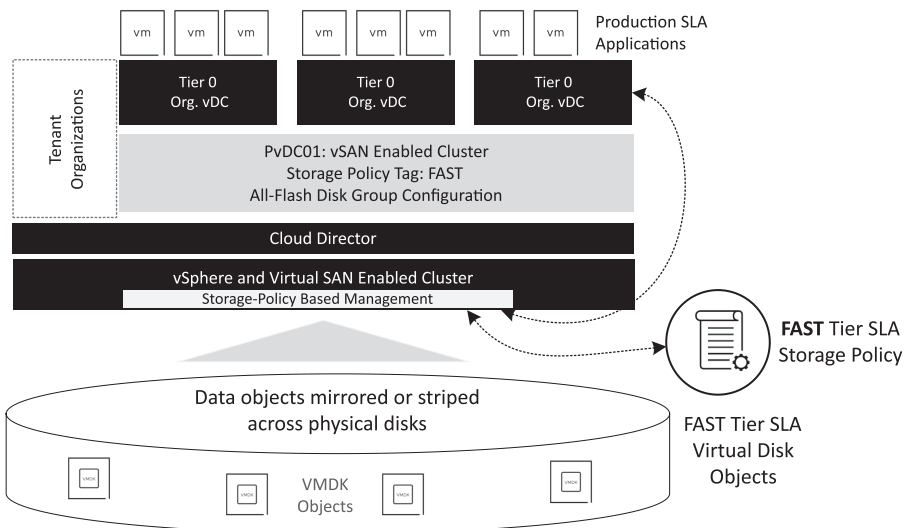


FIGURE 5.22 VMware Technology Solution Stack

As is illustrated in Figure 5.22, the vSphere resources are abstracted by Cloud Director into a Provider Virtual Data Center. These resources are then further carved up into individual Virtual Data Centers (VDC), assigned to Organizational Tenants. The overall result is that the workloads that reside

within the Organizational VDCs represent the vSAN storage capability defined by the cloud provider.

When employing this model, storage tiering, within a Provider VDC resource cluster, is not possible. Where vSAN is providing the storage resources, only a single type of storage can be configured within each vSphere cluster.

However, for a vApp author, building a typical multi-tier LAMP application often requires *Fast* storage for a database and *Standard* storage for the web components and middle or application layer, this may present a problem, which cannot be solved with this design.

The reason for this issue, while Cloud Director is able to employ Storage Policies, and storage policies can be employed to identify the performance characteristics, as defined at the vSphere layer, vSAN storage policies are predominantly focused on availability, and not performance. Therefore, this means that different performance characteristics cannot be assigned within a single resource cluster, and as discussed earlier in this paper, in this model, an Org. VDC may only map to a single Provider VDCs resources. This lack of available storage types within the Provider VDC means that a vApp author cannot deploy their multi-tier application within a single Org. VDC, where multiple different tiers of storage are required.

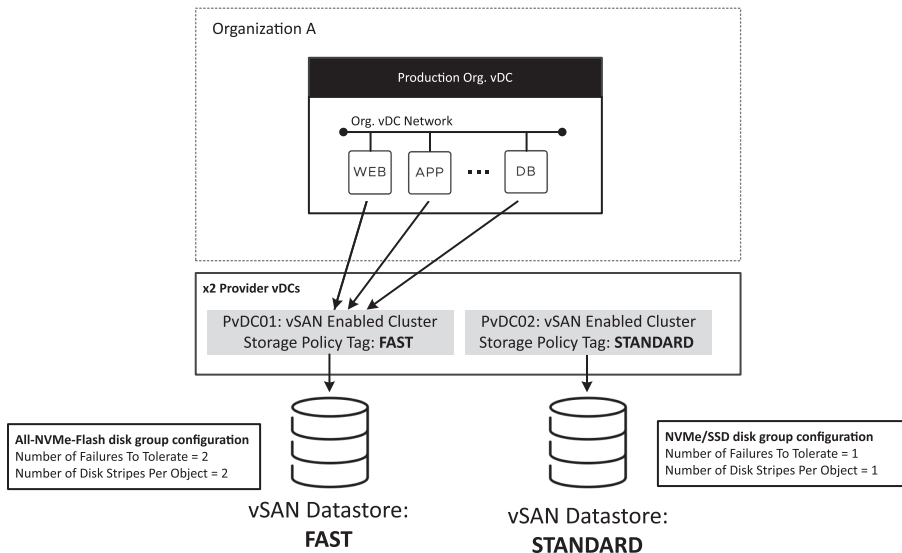


FIGURE 5.23 Multi-Tiered Application Storage Mapping

As illustrated in Figure 5.23, each of these application components might have different performance requirements for their respective workload, the Cloud Director administrator however can only map the Org. VDC to the most appropriate Provider VDC, based on their requirements.

For instance, the vApp author might want the databases to reside on the *Fast* Tier (All-NVMe-Flash) vSAN storage, while the web and application

workloads utilize the *Standard* Tier vSAN storage. However, to manually tier the storage in this way, when employing this model is not possible. Therefore, for each Org. VDC, the workloads will be restricted to the single storage type supporting that Org. VDC, via its Provider VDC, unless secondary external storage is provided to the vSAN hosts and surfaced to Cloud Director.

Model 2: Single Provider VDC Hosting Multiple Tiered Resource Clusters:

The second design model being addressed takes advantage of Cloud Directors Elastic VDC capability, the ability to span a Provider VDC across multiple vSAN-enabled Resource Clusters.

Cloud Director gives cloud providers the ability to create Elastic VDCs, which allows an Org. VDC to consume resources from multiple vSAN-enabled Resource Clusters. By having the Provider VDC abstract the resources from multiple vSAN clusters, it not only becomes simpler to grow capacity when needed, but also allows providers to design multiple storage performance and availability capabilities through storage policies and deliver them to a single Org. VDC. However, in this model the host's compute resources must remain consistent throughout the Elastic Provider VDC.

In Figure 5.24, there are three different tiers of storage to the Org. VDC, supporting Organization A. An all-NVMe-Flash-based vSAN cluster (Fast), a NVMe/SSD backed tier (Standard), and an external storage array backed Capacity tier.

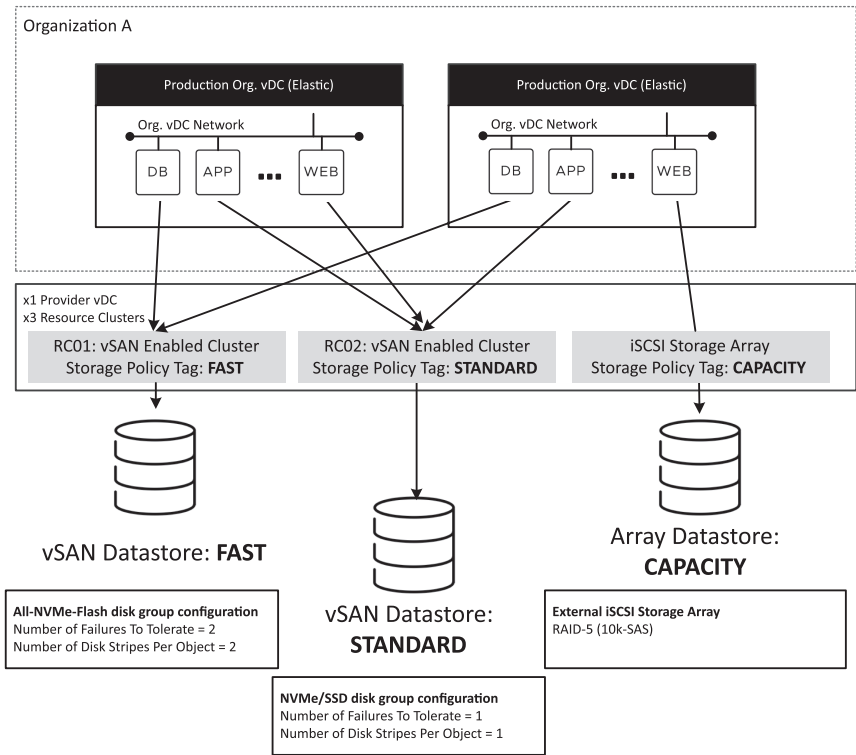


FIGURE 5.24 vCloud Elastic Provider VDC vSAN Logical Architecture

It is noteworthy however, the way in which Cloud Director system administrators extend a Provider VDC requires the creation of a Provider VDC primary vSphere resource pool, before additional cluster resources can be added, as the backing vSphere resources. The additional vSphere resource pool clusters must be added in subsequent steps. Note that a Provider VDC can only span clusters managed by the same vCenter Server under the same vCenter data center construct.

An additional option, which might be employed with both design models illustrated, is the inclusion of vSAN's Storage Policy-Based Management framework, with the single type of storage being provided by the Resource Clusters, being employed to enable multiple capabilities, within the single storage tier, based on availability. For instance, if employed, the common storage devices, across the resource clusters, can be assigned multiple storage policies to reflect availability characteristics, such as the Number of Failures to Tolerate (FTT) capability, as illustrated in Figure 5.25. These storage profiles can then be categorized as 'Enhanced Availability,' with FTT=2, 'Standard Availability' with FTT=1, and 'No Enhanced Availability,' with FTT=0, with all storage policies configured across a common shared storage class, such as an All-Flash NVMe/SDD configuration. For instance, within the Org. VDC, the vApp can be deployed on a single storage type, but by leveraging the storage policies, multiple availability capabilities can be chosen.

Once the datastore storage capabilities are configured, the storage policies must first be created in vSphere, which are then automatically presented to Cloud Director, where they can be viewed within the Provider VDC. By defining Storage Profiles through Cloud Director, SPBM can optimize the vApp provisioning process and leverage the vSphere Storage Policy Based Management API to automate storage management operations for the Software-Defined Storage infrastructure.

By default, the '(Any)' Storage Profile is created and employed by Cloud Director. However, the '(Any)' Profile should typically not be used and is provided for compatibility reasons only. If the design requires storage tiering within an Elastic VDC, as illustrated in model 2, then cloud administrators must use the storage policies leveraged from vCenter, to place a vApp on a specific storage profile within Cloud Director.

By default, a storage profile is set when the Org. VDC is created, which means that when provisioned by a vApp Author, all virtual machines, unless otherwise specified, will utilize that storage profile, and therefore be located on that tier of storage. However, by using this example architecture, where a three-tier vApp is required to be deployed, each individual virtual machine within the vApp can have its own storage profile, with the storage profiles being specified individually when the vApp is provisioned.

As a result, compliance and deployment of multi-tier applications can be achieved through SPBM. Furthermore, using SPBM in this way, surfaced through Cloud Director, allows provisioning in an on-demand basis, reducing the burden

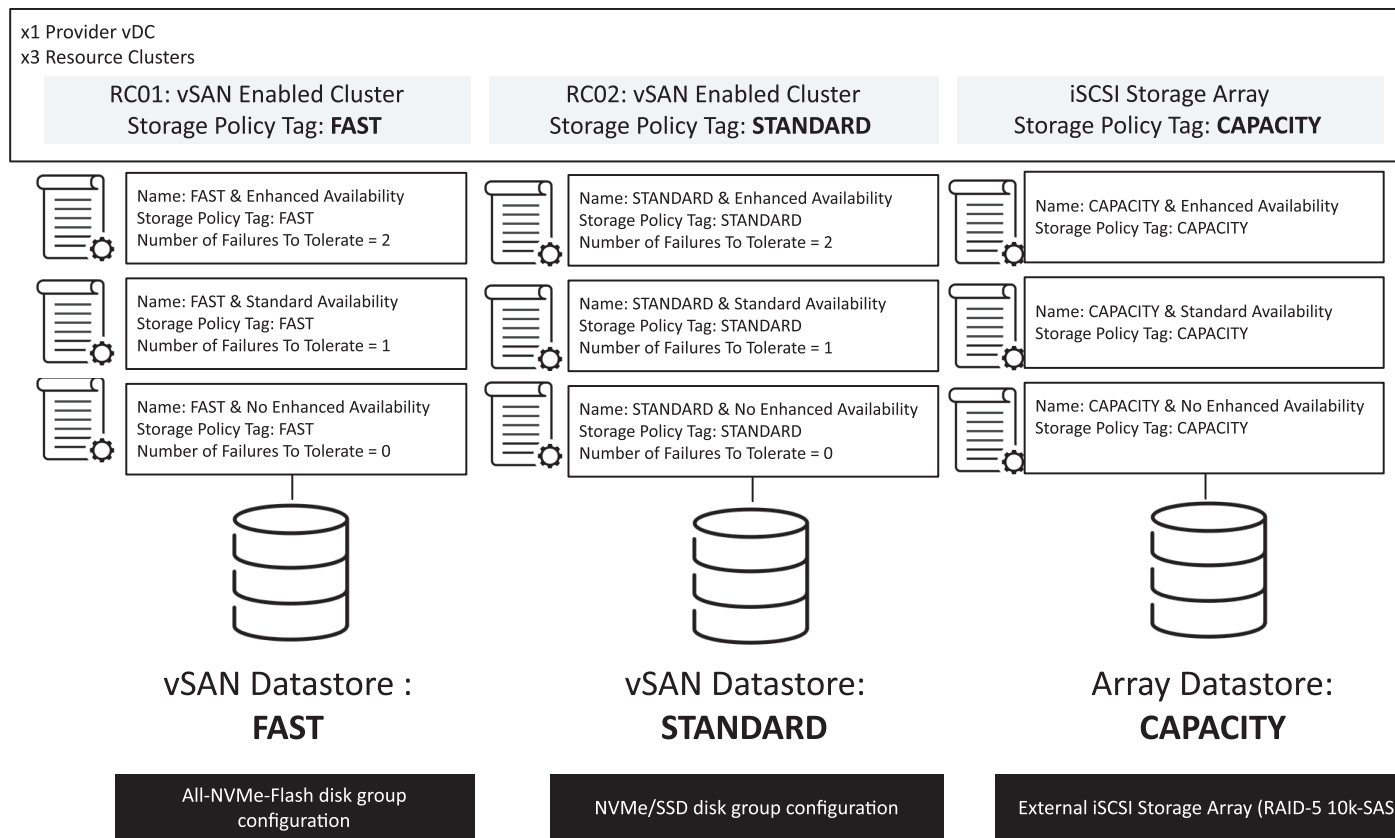


FIGURE 5.25 Provider VDC with Mixed Storage Availability Policy Capabilities

of storage administrators and empowers the cloud customers and vApp authors to take responsibility for vApps and automated provisioning tasks.

This architecture enables the vApp authors to tier the individual virtual machines in their vApps, as required by the application, in each Org. VDC. This is made possible by the tiered storage resources available within the single Elastic Provider VDC. In addition, the storage profile of a virtual machine can also be changed at any time. By changing the storage policy of the virtual machine, the virtual machine and each of its VMDK files will be migrated to a cluster and datastore which is compliant with the selected storage profile.

The Tenant App billing mechanism integrates with the storage policy features presented to Cloud Director from vCenter Server and can be employed to accurately identify and track storage policies utilized by customer virtual machines. This makes it possible to meter cloud consumers based on the virtual machine storage profile consumed, or classification of the storage on which the virtual machine resides.

The SPBM policies, which must be configured at the vSphere layer, before being imported into Cloud Director at the Provider VDC level, can be manually assigned by the vSphere/Cloud Director administrator.

In this design, the vApps and associated virtual workloads operating in the Cloud Director Provider VDC, will have their data located across multiple Resource Clusters and their respective vSAN datastores. Although vSAN, and therefore Cloud Director, offers no control over the dynamic placement of the vApp data on the underlying vSAN storage. However, vSAN will be employing its native replication technologies at the backend, to meet the appropriate availability requirement, defined in the vSphere storage policy.

Impact of vSAN on Cloud Director Compute and Network Resources: When employing vSAN as part of a vSAN design, as outlined above, it is important to consider the impact on the other resources in the Cloud Director environment, as it is of course the case that tenant organizations also require CPU and memory compute resource to operate. Organizations across a Cloud Director platform draw all of their resources from Org. VDCs, whose resources are drawn from the respective Provider VDC, with the Provider VDC resources being a collection of CPU, memory, network and storage abstracted from the underlying vSphere platform.

In a hyper-converged vSAN environment, the appropriate allocation and balance of required CPU, memory and network resources is just as important to calculate as storage. In order to implement a successful design, and to ensure the ongoing operation of the Provider VDCs across the Cloud Provider Platform, a balance approach to resources across all resource types is critical.

Just as a Provider VDC allows the cloud provider to pool infrastructure resources to create standard offerings to tenants, Provider VDCs can also be employed to group different levels of compute resource to compliment the different tiers of storage. Ensuring an appropriate balance of all three resource types is an important design calculation, which must be addressed

in order to provide optimum balance of compute, storage, and networks within the Provider VDC.

For instance, Tier 0 CPU, memory, and network resources can be combined with the fastest storage type to create a balanced Provider VDC, which can be priced to consumers at the appropriate level, which would typically be at a higher cost than Tier 1 or 2, as illustrated in Figure 5.26.

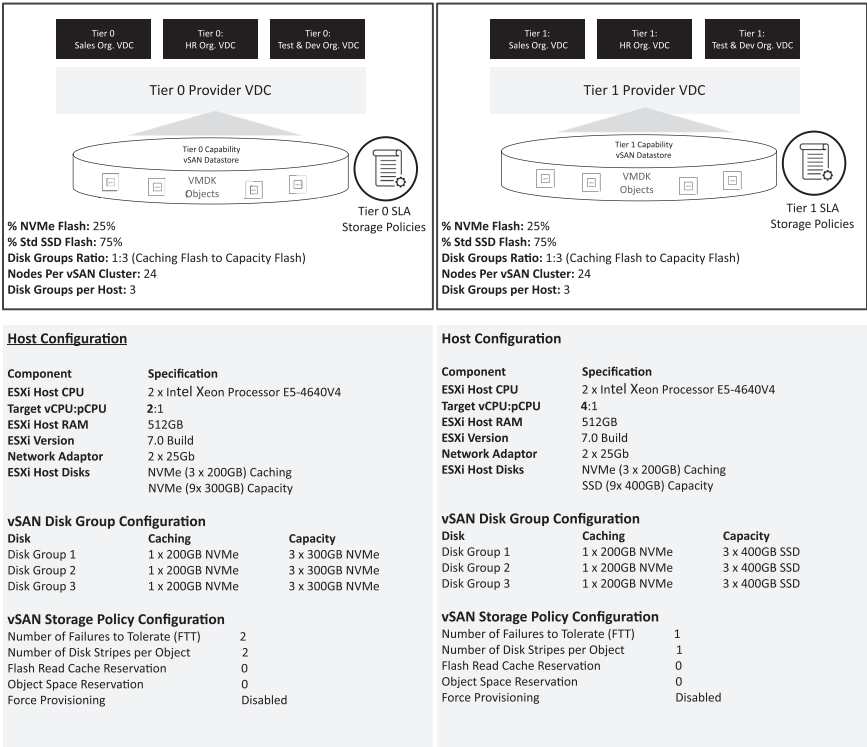


FIGURE 5.26 Balanced Provider VDC Resource Offerings

Cloud Director NFS Storage Requirement

As part of the design of a Cloud Director platform with vSAN, the architect needs to address the location of the organization catalogs, and the type and tier of storage on which it is hosted.

A catalog provides cloud consumers with a library of content, such as ISO media and vApp templates. Public catalogs allow multiple organizations to share specific catalog items with authorized users and developers have the ability to upload and publish their content to these public catalogs.

Catalog media files and other content do not require the disk performance of *active* systems and can therefore typically be located on a much lower tier of storage, with the primary design consideration being based around

capacity, and ideally thin provisioned disk. NFS storage is ideally suited to meet this requirement and the use of 3rd party independent software vendor (ISV) solutions can be employed to provide an NFS layer of storage offered up by a virtual appliance running on top of vSAN.

In general VMware does not recommend utilizing native vSAN storage for catalogs. The reason behind this recommendation is that all catalog media images, such as ISO files are uploaded by Cloud Director as file objects, into a directory structure under a single folder. Therefore, assuming the same storage policy is used for different catalogs, they will all share one VM Home Namespace object, with a maximum size of 256 GB. For this reason, either a native NFS storage array or third-party virtual storage appliances, which consume vSAN storage, and provide NFS file services, are better placed to be used to provide catalogs across clusters, providing the ability to scale beyond this 256 GB imposed limit.

In addition, another design consideration is that a Cloud Director server group, that is a deployment made up of multiple cells, requires the use of a shared volume, which is referred to as the transfer server storage. The storage employed for the transfer server storage must be available to all Cloud Director cells in the server group, and whilst any type of shared storage, such as iSCSI, is possible, because NFS storage is easier to set up and more flexible than other shared storage file system types, NFS is typically the preferred choice to meet this requirement.

When employing NFS for the transfer server storage, a number of configuration settings must be set, so that each Cloud Director cell in the Cloud Director server group can mount and use the NFS-based transfer server storage. The ability for each cell to mount the NFS-based location, and use it as the transfer server storage, is related to the user and group permissions settings that Cloud Director employs.

Another design consideration, which plays into the use of NFS, is the inclusion of maintenance storage. For instance, it is common for all local disks on a host to be allocated to vSAN disk groups and therefore the deployment of the ESXi hypervisor binaries could be onto USB or SD devices. However, if you allocate all local storage to vSAN, you will not have local disks to deploy a local datastore, which is required for persistent local logging. The reason for this is due to the I/O sensitivity of USB and SD devices, and the hypervisor behavior which means the installer will not create a scratch partition on these devices. When installing on USB or SD devices, the installer attempts to allocate a scratch region on an available local disk or datastore. If no local disk or datastore is found, `/scratch` is placed on the ramdisk. After the installation, you should reconfigure `/scratch` to use a persistent datastore location.

Therefore, an additional design option to maintain logs and vSAN trace files is to redirect them to an NFS datastore, which can be employed as a centralized maintenance location. For instance, the vSAN trace files can be redirected by employing the `esxcli vsan trace set` command: `vsantraces -> /vmfs/volumes/NFS-Extent/`

Note: *This would typically be done in addition to the configuration of an ESXi Dump Collector and Syslog Collector, to redirect ESXi memory dumps and system logs onto a centralized network server.*

For cloud providers introducing a Cloud Provider Platform they are now able to develop multiple new ways to reduce the cost of provisioning and resource consumption. Cloud Director layers on top of the vSphere infrastructure in a way that offers great agility in provisioning, isolation for multitenancy and resource metering. Employing vSAN storage on a Cloud Director public cloud platform allows the further abstraction of storage resource and further simplification of the allocation of storage, with new benefits being gained from automation and reduced operational overhead, whilst still maintaining the building blocks required to develop and manage a highly scalable Cloud Provider Platform.

Designing a Highly Scalable HCI Platform

One simple approach to a highly scalable HCI platform is the building block approach that is employed already by many cloud providers. The idea is that each cluster is a standard container of resources that is provisioned consistently to provide a simple, scalable, and building block approach to compute resource provisioning. Not only does this methodology scale consistently across data centers, but it also eliminates variability, configuration drift, and the amount of operational effort involved with patch management and day-to-day operations. This approach is the simplest and most effective way to provide a flexible building block solution that can meet the cloud provider's requirement for elasticity. This building-block approach standardizes the configuration of the HCI hosts, clusters, and even server cabinets to help provide a manageable and supportable infrastructure.

Standardizing not only the model, but also the physical and logical configuration of the HCI hosts and clusters, is critical to providing a manageable and supportable infrastructure in large-scale deployments by eliminating variability. A cloud provider might also wish to utilize vSphere host profiles to configure additional values consistently across hosts and clusters wherever applicable to ensure the cloud providers scaling and elasticity requirements for the shared multi-tenanted virtual data center can be met. For more information on this building block methodology, please revert to Chapter 3.

CONSUMING EXTERNAL ARRAY STORAGE

Having now looked at VMware vSAN and consuming that storage model through Cloud Director, we will now address other storage solutions. Today, it is commonplace to see cloud providers consume more than one storage type when implementing the Cloud Provider Platform. While vSAN can provide and meet the storage requirements of the vast majority of customer

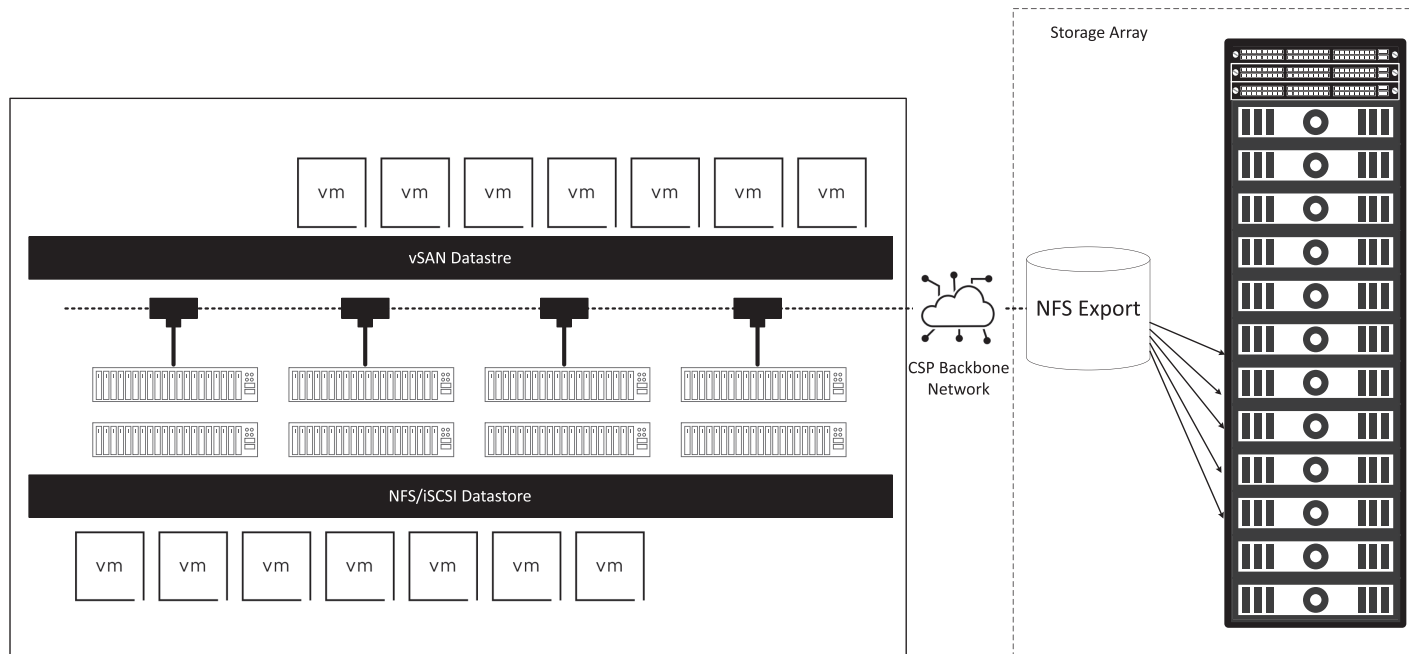


FIGURE 5.27 Cloud Provider Platform—Mixed Storage Model

workload types, adding additional storage for either high-capacity use cases, where vSAN cannot reasonably meet the required capacity through the HCI model, or high IO use cases, where dedicated consumption is required by a specific application.

In most examples seen, these secondary storage use cases consume either the NFS or iSCSI storage protocols, removing the need for a storage only Fibre Channel network to be deployed and maintained within the cloud provider's environment.

Traditional VMFS Storage Considerations for Cloud Director: By providing access to a shared storage resource, either exclusively or alongside vSAN storage in the Provider VDC, an additional level of flexibility for cloud providers to deliver services can be achieved by:

- Providing a mixed model of storage, with vSAN
- Providing all of Cloud Director's managed storage datastores
- Providing multiple datastores with a single capability and utilize datastore clusters
- Provide tiered levels of storage capabilities defined by manually created Storage Profiles

Cloud Director Storage Profiles with External Storage Devices: As we have addressed earlier in this Chapter, Cloud Director supports native Storage Policy Based Management which allows storage policies to be defined in vSphere and to be assigned to Provider VDCs, allowing for the adoption of the capabilities defined to be consumed by tenant via their Org. VDCs. Storage policies must first be defined on vCenter, which then become listed in Cloud Director.

For instance, the following example of manual tiering addresses typical tenant use cases. However, manual tiering of storage is only one option available to cloud providers. Also available from most storage vendors are auto-tiering solutions, which are fully integrated into the array software, which can reduce the operational overhead associated with the manual tiering of data.

With manual tiering, perhaps the more traditional method, the storage team has to monitor workloads and move data between the tiers periodically. This can prove to be repetitive and operationally time consuming, ranging from hours to days to complete. Automated storage tiering on the other hand addresses many of these same challenges, around data retention and access, but in a more automated way.

Datastore Clusters with External Storage Devices: It is considered a best practice to consume datastore clusters in a Cloud Provider Platform deployment. Datastore clusters allow cloud providers to pool storage resources of the same capabilities into a single managed entity. If enabled, datastore clusters use Storage DRS to balance and distribute resources within the datastore cluster. You can set resource management policies at the cluster level

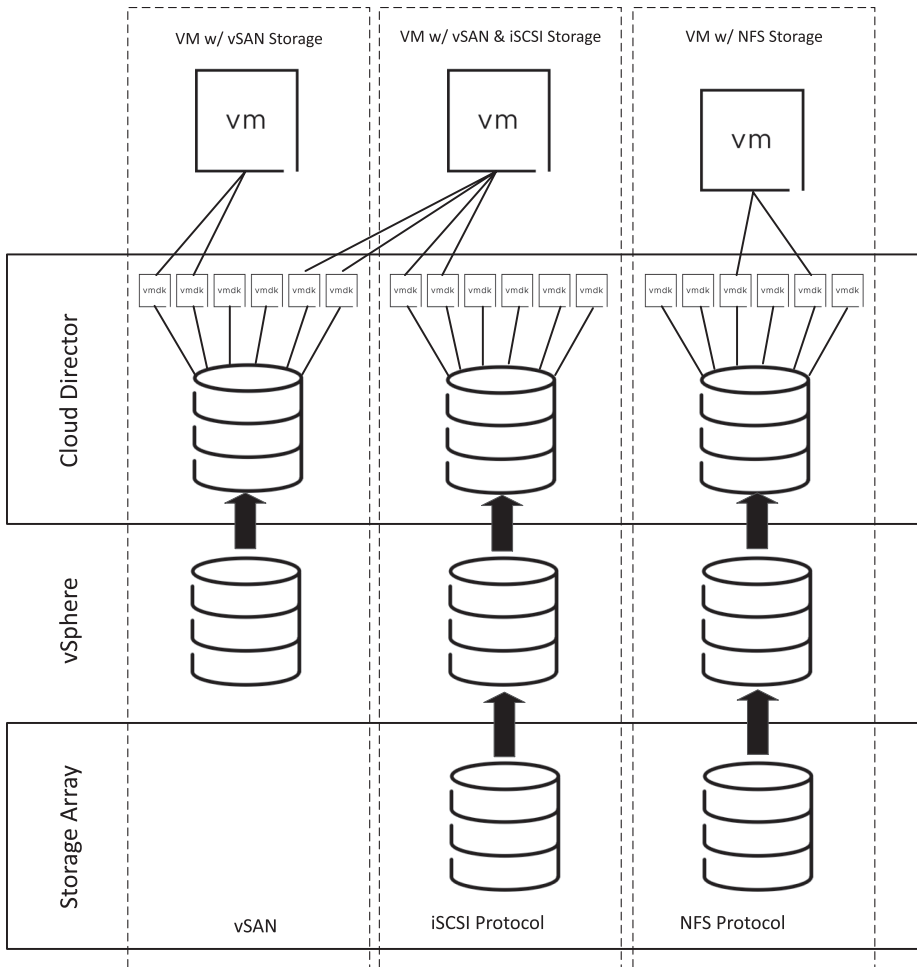


FIGURE 5.28 Logical Boundaries Separating Storage Types

which gives you features such as Space Utilization Load Balancing and I/O Latency Load Balancing.

It is considered best practice that datastores in a datastore cluster have equal performance capability, protocol and quality of service. If this is not the case, the performance of that Provider VDC storage pool will be impacted by the slower storage in the collective pool.

IOPS Management with External Storage Devices: Cloud Director can manage the storage IOPS for custom workloads, with performance being defined at virtual disk granularity.

As illustrated in Figure 5.29 cloud providers can offer different tiers of storage to customers. The use of IOPS management can help differentiate

	Service Levels		
	Specification	Attributes	Value
Primary vSAN Storage	Guaranteed Performance	Performance throughput (I/O sec.)	5,000+
		Response time (ms)	< 8ms
	Availability	Maximum unplanned downtime per year (mins)	< 26.5
Secondary Storage	Guaranteed Performance	Response Time	< 1 Second
		Throughput	<= 300Mbps
	Availability	Maximum unplanned downtime per year (mins)	< 5.25 mins
	Retention & Disposition	Application Retention Period Required	< 10 Years
		Data Shredding Compliance	Yes
	Accessibility	Read Access Frequency	< Hourly
	Data Integrity	Guarantee of Authenticity	Yes
	Offsite	Recovery Point Objective	< 1 minute
Operational Recovery (OR)	Recovery Granularity	Recovery Granularity	Complete App. Restore
	Recovery Point Objective (RPO)	Amount of Data Loss	1 hour
	Recovery Time Objective (RTO)	Time to Restore Data	< 30 minutes
	Recoverability	Ability to Recover Backed Up Data	100%
	Retention Period	Time Data is Retained	2 hours
Disaster Recovery (DR)	Recovery Point Objective (RPO)	Amount of Data Loss	0 minutes
	Recovery Time Objective (RTO)	Time to Restore Data	< 2 hours

FIGURE 5.29 Manually Tiered Storage Capabilities and Attributes Example

these tiers and enforce the virtual disk performance based on an IOPS metric. This not only eliminates noisy neighbor scenarios, but also helps address consumption and capacity management across the platform.

Cloud Director leverages vSphere capabilities to control the IOPS limit for workloads based on the defined storage policy. This is achieved through the Storage I/O Control functionality which is supported on VMFS (block) and NFS datastores. VMware vSAN has its own I/O limit mechanism. However, vSphere is unable to manage available the IOPS capacity of a datastore the same way it is able to with compute capacity. This is where Cloud Director is required.

This is achieved by the cloud provider first creating a new vSphere custom field called `iopsCapacity`. This is then used to define a Cloud Director managed datastore IOPS capability, which is done via the vCenter Managed Object Browser user interface, and is described in KB 2148300.

Considerations for Use of vSphere Storage and Arrays Features: There are a number of storage-management features in vSphere and cloud providers need to consider their impact when used with a Cloud Provider Platform deployment. There are some features that are transparent to Cloud Director, which improve overall performance of the platform, while others that are transparent can be disruptive. For instance, the use of VMware Storage vMotion can be used with good results if the Provider VDC consists of several datastores with datastore clusters enabled, where Storage DRS is used to balance the load of the workloads on those datastores. However, it could also be very disruptive if at the vSphere layer an administrator moves a virtual machine from one Provider VDC to another Provider VDC, without realizing that this would stop Cloud Director from managing that workload.

Features within the underlying storage array, such as thin provisioning or deduplication, Multi-Path Plug-ins (MPP) are all without impact, whereas storage array snapshots or array-based replication may have significant implications for a Cloud Provider Platform deployment due to the additional layer of abstraction.

In summary, when designing a storage array-based Cloud Provider Platform:

- Build Provider VDCs with large dedicated datastores.
- Define storage tiers by creating separate SPBM definitions, which are surfaced in Cloud Director.
- Use a simple naming convention for Provider VDCs and Organization VDCs that will help a storage administrator recognize that the objects are Cloud Director-based objects, to simplify operations.
- Work closely with storage administrators to ensure that infrastructure is provisioned and monitored for best performance and ensure that multipath and redundant components are configured for maximum high availability.
- Zone storage only for the hosts in the Cloud Director Resource Clusters; this will help prevent the storage vMotion of virtual machines to other clusters on different vCenter Servers that are not used by Cloud Director.

VIRTUAL VOLUMES (VVOLS)

vSphere Virtual Volumes leverages the same management and automation engine of vSAN, Storage Policy-Based Management, to automate the delivery of storage service levels with virtual machine granularity. This means that Virtual Volumes offer many of the benefits realized by the software defined model of vSAN but with an external array-based storage solution. Virtual

Volumes significantly changes the way external storage is managed. Also, it simplifies the day-to-day operations, enables granular management of the workloads by utilizing SPBM to manage capabilities, such as efficient I/O control and handling of snapshots. The aim of vVols is to increase the value of array-based storage to cloud providers through better utilization and potentially increasing performance.

Traditional storage can be complex to manage, slow to provision and rigid in deployment model. Having to create LUNs or volumes for a specific capability is inefficient and can often be a waste of capacity and increase costs. Other traditional storage challenges include:

- When a LUN or Volume is provisioned, it is set up with specific features or capabilities. Changing the capabilities of an existing LUN or volume requires migrating workloads to another datastore
- To change an existing capability requires destroying and recreating a new LUN/Volume with a different capability.
- Management requires some form of tracking utility or spreadsheets.

With vVols, there's no file system on the storage container, so resizing is a simple process. Also, you no longer need to over allocate space as it's a logical 'quota,' which can be increased as required, assuming capacity is available.

With traditional LUN-based infrastructure, the I/O and data are not understood by the storage, therefore it doesn't know what each workload is, the I/O profile or how to allocate or distribute the data. With the VASA provider being written by the storage vendor and linking vSphere to the array, both are now completely aware of all I/O and data profiles. Subsequently, each workload can be handled differently or given different capabilities, based on the storage features presented to vSphere. With multiple capabilities being available on the storage array, any number of vSphere storage policies can be created based on the customer's specific requirements, as illustrated in Figure 5.30.

Traditional storage LUNs or volumes are rigid and fixed in their capabilities. They also require a file system, such as VMFS or NFS. With vVols, there is no file system, it's inside each vVol that a file system is created, and that file depends on the OS or the vVol function.

With regards to the capabilities, generally a LUN or volume is provisioned with a specific set of capabilities, which cannot easily be changed without creating a new LUN or volume and migrating the workloads. With vVols, changing the capability is simple, with the simple change of a storage policy. This allows the capability to be modified without storage vMotion or data migration being required. Depending on the array, you can even change from spinning disk media to all-flash with a policy change. The array handles the performance and data migration, with no storage administrator action required. Another key aspect of vVols over traditional storage is that

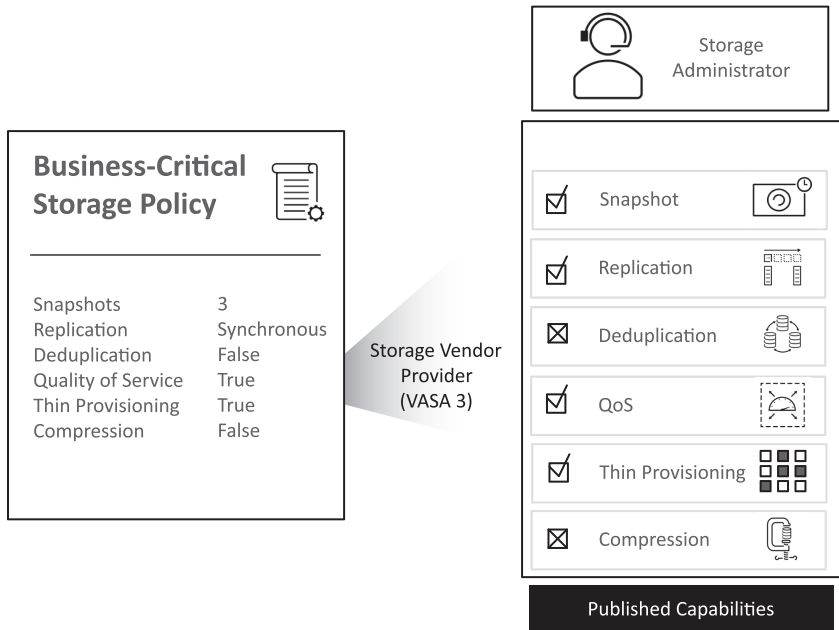


FIGURE 5.30 Virtual Volumes Storage Policy-based Management

the array and vSphere have complete insight into the data and I/O requirements. This allows the array can then manage the I/O and data accordingly.

vSphere Virtual Volumes is an industry-wide initiative supported by all the major storage vendors. This can allow cloud providers to leverage existing storage array-based investments, while transitioning to this new operational model. The vVol feature leverages a number of standard components that are important in the VVol architecture and therefore need to be addressed:

- Virtual Datastore and a Storage Container
- Protocol Endpoints (PE)
- VASA/Storage Vendor Provider
- Storage Policies
- The Virtual Volume (vVol)

Virtual Datastore and a Storage Container: Logically a vVols datastore, VMFS, or NFS datastore are the same. It's how a vVols storage container is provisioned and utilized that is the difference. Each vVol datastore corresponds 1:1 with a storage container on an array.

When a vVols datastore or storage container is created on the array, there is no physical allocation of space, it's a quota, with a limitation in size. When created, the array creates a container and limits its allocation of space to the size decided by the storage administrator. For instance, if the storage admin creates a 15TB container, no storage is provisioned, just a quota of

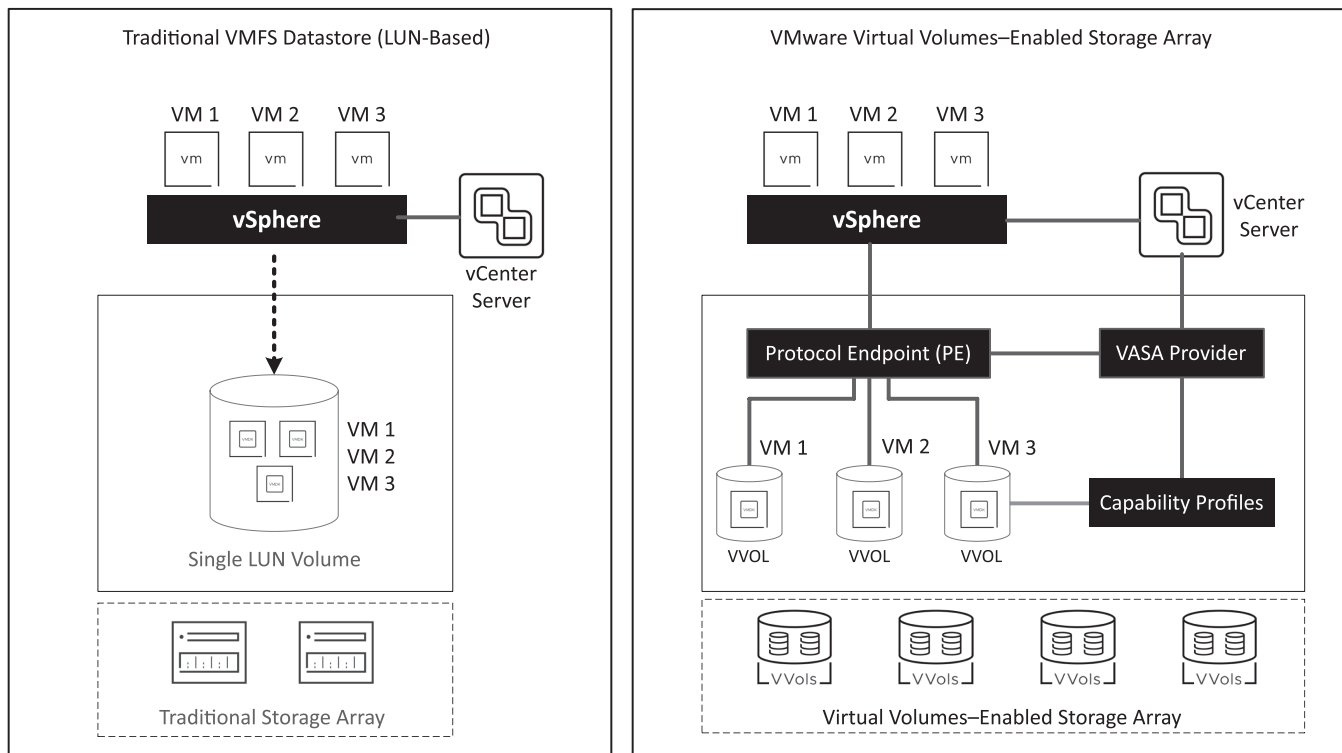


FIGURE 5.31 Traditional storage LUNs Verses Virtual Volumes

space where the content will reside. The other key difference is the capabilities, which are handled by the VASA provider and the storage policies, created from the published capabilities from the storage array. When a policy is applied, and workload created, the array creates the individual vVols for that virtual machine based on the storage policy, which grants or limits the capabilities assigned.

Protocol Endpoint (PE): The storage system manages all aspects of vVols, vSphere hosts have no direct access to the vVols on the storage side. Instead, vSphere uses a logical I/O proxy, called the Protocol Endpoint, to communicate with vVols and virtual disk files that vVols encapsulate.

The Protocol Endpoint is used by vSphere to establish a data path, on demand, from virtual machines to their respective Virtual Volumes. PEs are compatible with iSCSI, Fibre Channel, Fibre Channel over Ethernet (FCoE) and NFS protocols. Typically, when the array's storage provider is configured, the Protocol Endpoints is set up when the storage provider is registered. Each Virtual Volume is bound to a specific Protocol Endpoint. When a virtual machine on the host performs an I/O operation, the PE directs the I/O to the appropriate vVol. In most environments, the storage system will require a very small number of Protocol Endpoints, as a single Protocol Endpoint can connect to up to 16,383 vVols.

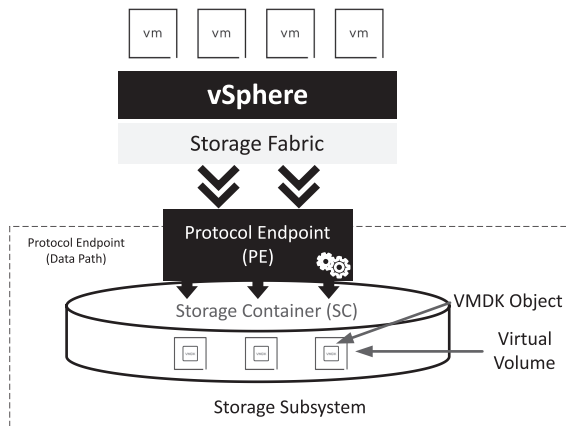


FIGURE 5.32 Protocol Endpoint (PE) Architecture

VASA Provider (VP)/Storage Provider: A VASA Provider is a storage-side software component that provides storage awareness services to the underlying vSphere environment. ESXi hosts and vCenter Server connect to the VASA provider to obtain information about available storage topology, capabilities and status. Effectively the VASA Provider creates storage aware

services and announces them to the vSphere environment through out of band communication between the vSphere hosts and the storage array.

Therefore, the VP provides the key information link between the array and vSphere. It is not a data link, however, no data I/O goes through the VASA Provider, all I/O goes through the Protocol Endpoints, as directed by the VASA provider, offloading all virtual machine related operations directly to the array.

As a result of this architecture, illustrated in Figure 5.33, the array owns the vVols and capabilities, such as snapshots, clones, replication, etc. All of which are handled by the array, instead of vSphere. This way of utilizing the array's native capabilities is much more efficient, meaning that these actions, which are handled by the array, all take place on the array itself, and not back and forth between the hosts and the array, as was the case with VMFS-based storage LUNs. For instance, capabilities, such as those listed below, can now entirely be handled on the storage array, without consuming resources from ESXi hosts:

- Clones and snapshots
- Instant deletion
- Policy change (no storage migration required to modify workload capabilities. The policy changes happen behind the scenes, giving QoS on the fly)

Each VASA Provider is written by the array vendor and can typically manage multiple arrays, of that specific array model. The VASA Provider can run anywhere, in firmware, in management code, in a virtual machine or even as a vCenter Server plugin.

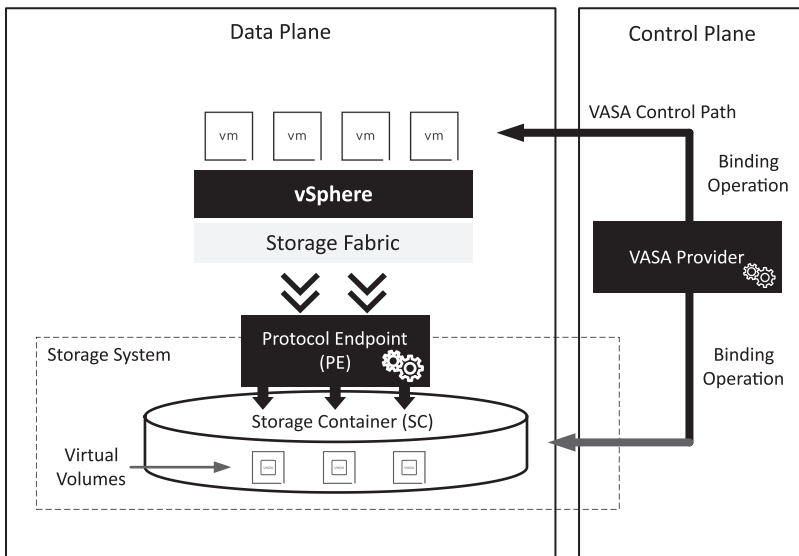


FIGURE 5.33 VASA Provider Architecture

Because the array owns the data and uses the VASA Provider to communicate with vSphere, the information is no longer just I/O and some blocks. The array knows specifically what each vVol is and how it should be managed. Because capabilities are managed via policy, creating several different datastores for each capability is no longer needed. This greatly reduces the complexity and operational management of the storage environment. Examples of array functionality or capabilities, which can be presented to vSphere through the VASA Provider include, but are not limited to:

- Thin Provisioning
- Array Snapshots
- Auto Zero Detect
- De-duplication
- Virtual Copy
- Space Reclamation (UNMAP)
- Adaptive Flash Cache
- Priority Optimization
- Drive type (Nearline, SSD, Fiber Channel, etc)
- RAID Level

Utilizing the array's native capabilities in this way is much more efficient and often times significantly faster.

Vvol Storage Policies and Storage Policy-based Management: Just like VMware vSAN, one of the key features of vVols is Storage Policy Based Management. With SPBM, different capabilities can be added, removed changed on workloads within the same vVols datastore with no storage or data migration being required. For instance, test and dev workloads can be stored in the same vVols datastore as business-critical applications but be assigned different policies. Storage Policies, including QoS capabilities, can be used to ensure the business-critical applications are always allocated array resources over test and dev workloads. Other policies can include replication, disk type, RAID or encryption. However, these are entirely dependent on what the array vendor publishes via the VASA Provider.

The vSphere administrator creates the SPBM policies based on the capabilities presented and the requirements of the cloud provider. As these are policies, and not physically allocated storage, the management of the capabilities is more granular. For instance, you can granularly apply different policies to different disks belonging to the same workload, just as you can in vSAN. An example might be an SQL server where you create a policy for the OS, database, backup disk and log disk, each with its own defined capabilities, with all of the vVols existing on the same vVol datastore.

The Virtual Volume (vVol): What actually is a Virtual Volume? When a virtual machine is created, there are a minimum of 4 vVols created. Config, Data

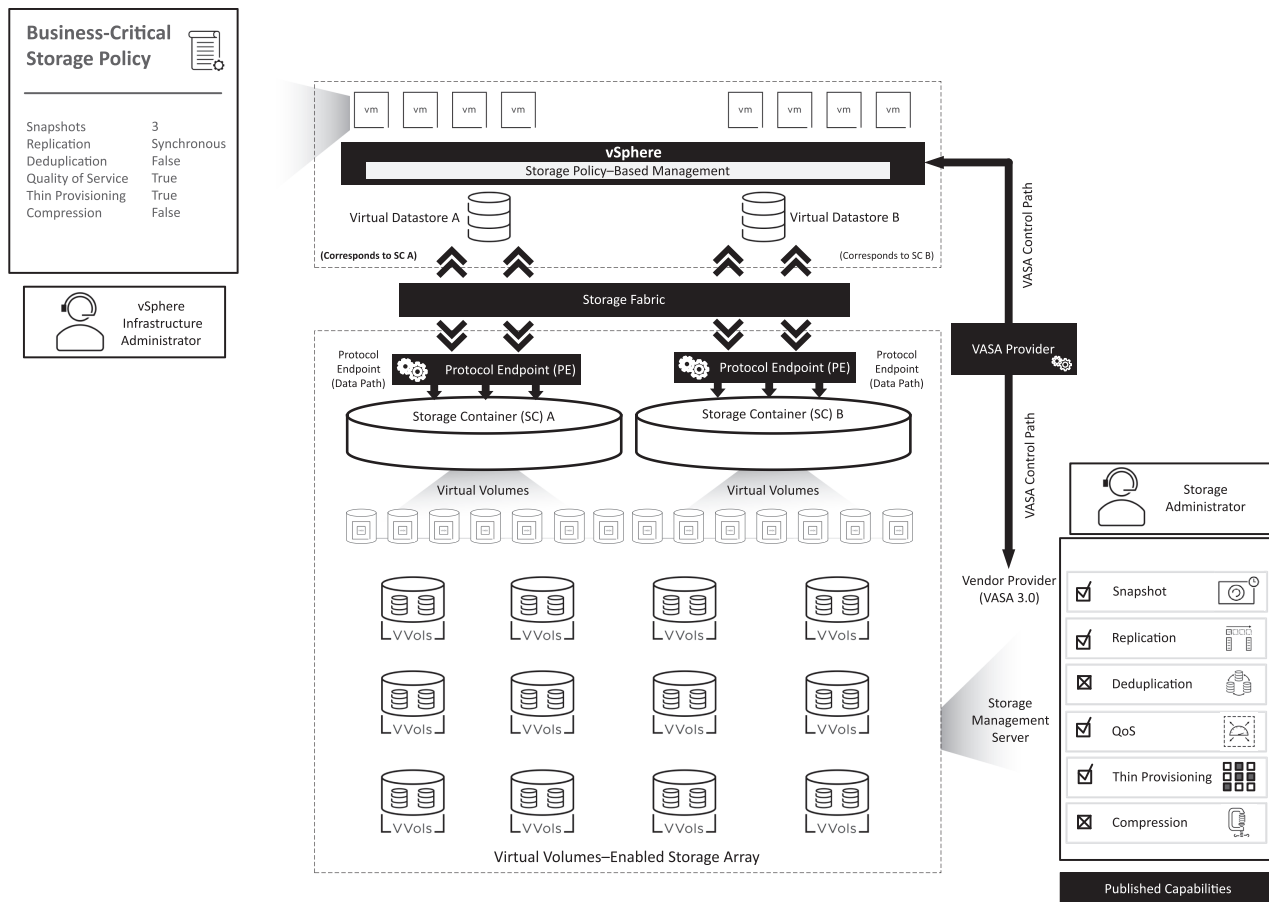


FIGURE 5.34 Storage Policy Based Management and VASA Provider Solution

(VMDK), Swap (memory) and Mem (Snapshots). If more disks are created for the VM, then another vVol is also created. There are also other vVols that can be created based on I/O filters or other vSphere solution-specific functionality. CBRC (Content Based Read Cache) disks would be one example of *Other* vVols where no distinct handling by the array is expected, so no specific vVol type is defined.

These Volumes then get a file system based on the functionality. For instance, a VMDK disk gets whatever native file system the OS uses, making it optimized for that OS.

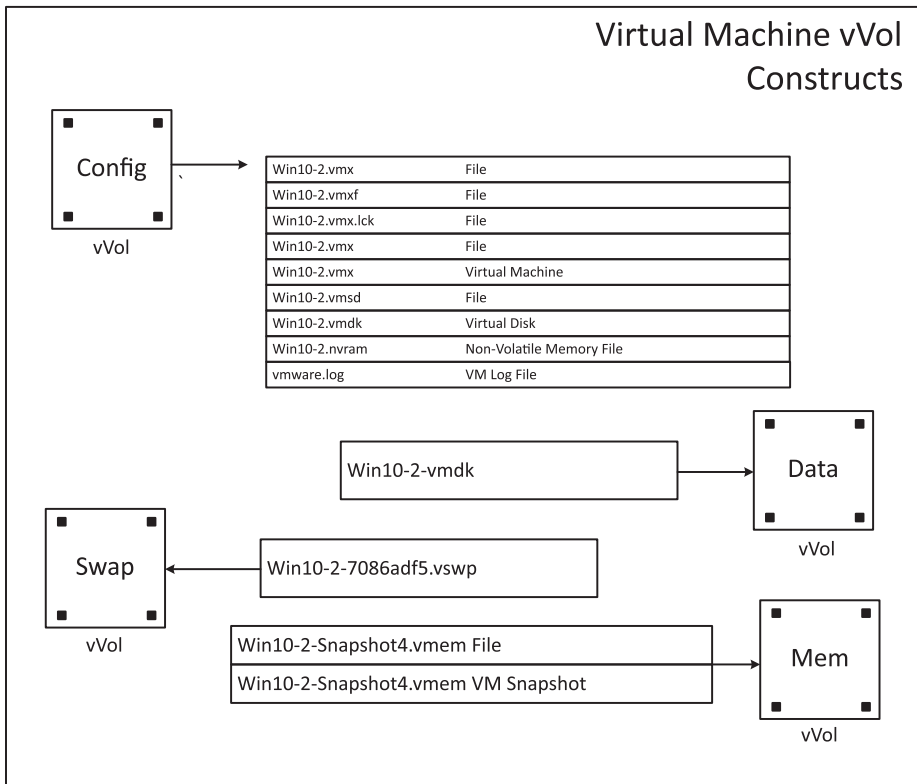


FIGURE 5.35 Virtual Volume

Another unique feature of vVols is that they are mounted or 'bound' when the virtual machine is powered on. This way there is no waste of IDs by vVols that are not being used. When the VM is created, the vVols are created, when the VM is powered on, the VASA provider tells the array which Protocol Endpoint to use and binds the vVols to that PE host where the VM resides. If the VM is migrated to another host, the vVols are unbound and rebound to a new PE on the new host. These bindings are simply data path coordinating

mechanisms that occur between VASA providers and ESXi hosts, for accessing the virtual volume.

The Virtual Volumes feature was released as part of vSphere 6.0. Since that release, momentum among cloud providers has been slowly building as more and more storage vendors developed support for the feature and introduced new capabilities. One of the key inhibitors to widespread adoption by providers had been its lack of support for array-based replication (ABR). However, this feature was added with the release of vSphere 6.5 Virtual Volumes 2.0.

The primary focus of vVols 2.0 is to extend the fine granular storage approach to provide data protection and recovery. Providing per VM/VMDK disaster recovery allows cloud providers to deploy Virtual Volumes beyond test/dev workloads to all use cases. Just like vSAN, vVols can work alongside other storage types to deliver a heterogeneous storage model, illustrated in Figure 5.1 to further extend the software-defined-storage control plan to existing SAN and NAS storage, with VM centric operations.

Chapter 6

Designing a VMware *Cloud Verified* Network Platform

For end customers achieving the required security in the cloud is key to adoption. VMware NSX enables the implementation of virtual networks on a physical network fabric. Within the Cloud Provider Platform, NSX allows the infrastructure to programmatically initialize, control, change and manage security in the cloud on-demand. By reproducing the physical network in software, NSX consolidates the entire networking stack within each tenants Organization. For instance, it offers a distributed logical architecture for Layer 2-7 services including, logical switch, router, firewall, load balancer and VPN, within the Org. VDC. These logical network services are provisioned programmatically by Cloud Director, making the whole idea of multi-tenant network support for the Cloud Provider Platform align perfectly to the use of NSX.

End-to-end security is crucial to cloud customers, but in the traditional way of working, often higher security compromises flexibility. Making changes in the data center or on the physical network takes time and is hard to automate. For cloud providers this approach is not cost effective. However, with NSX, because you do not need to make changes on the physical network and all updates and changes can be performed by the customer through Cloud Director programmatically, NSX offers significant advantages.

Like a server hypervisor allows the creation of software virtual machines, NSX enables the creation of software virtual networks. These virtual networks can be created, saved, deleted, restored, just like virtual machines but for the network through Cloud Director in an easy to consume manor. NSX combined with Cloud Director abstracts away from the tenant the complexity of building and maintaining secure physical network infrastructure. This cloud-based consumption mechanism for NSX allows the customer to

solve problems surrounding security, automation and agility in the modern hybrid cloud world by adding the following networking elements to the Cloud Provider Platform:

- **Logical switching:** Reproducing the complete L2 and L3 switching functionality in a virtual environment, decoupled from the underlying hardware.
- **NSX Gateway:** L2 gateway for seamless connection to physical workloads and legacy VLANs.
- **Logical Routing:** Routing between logical switches, providing dynamic routing within different virtual networks.
- **Logical Firewall:** Distributed firewall, kernel enabled line rate performance, virtualization and identity aware, with activity monitoring.
- **Logical Load-Balancer:** Full featured load balancer with SSL termination.
- **Logical IPSec VPN:** Site-to-Site
- **NSX API:** RESTful API for integration into the Cloud Provider Platform and integrated third-party components.

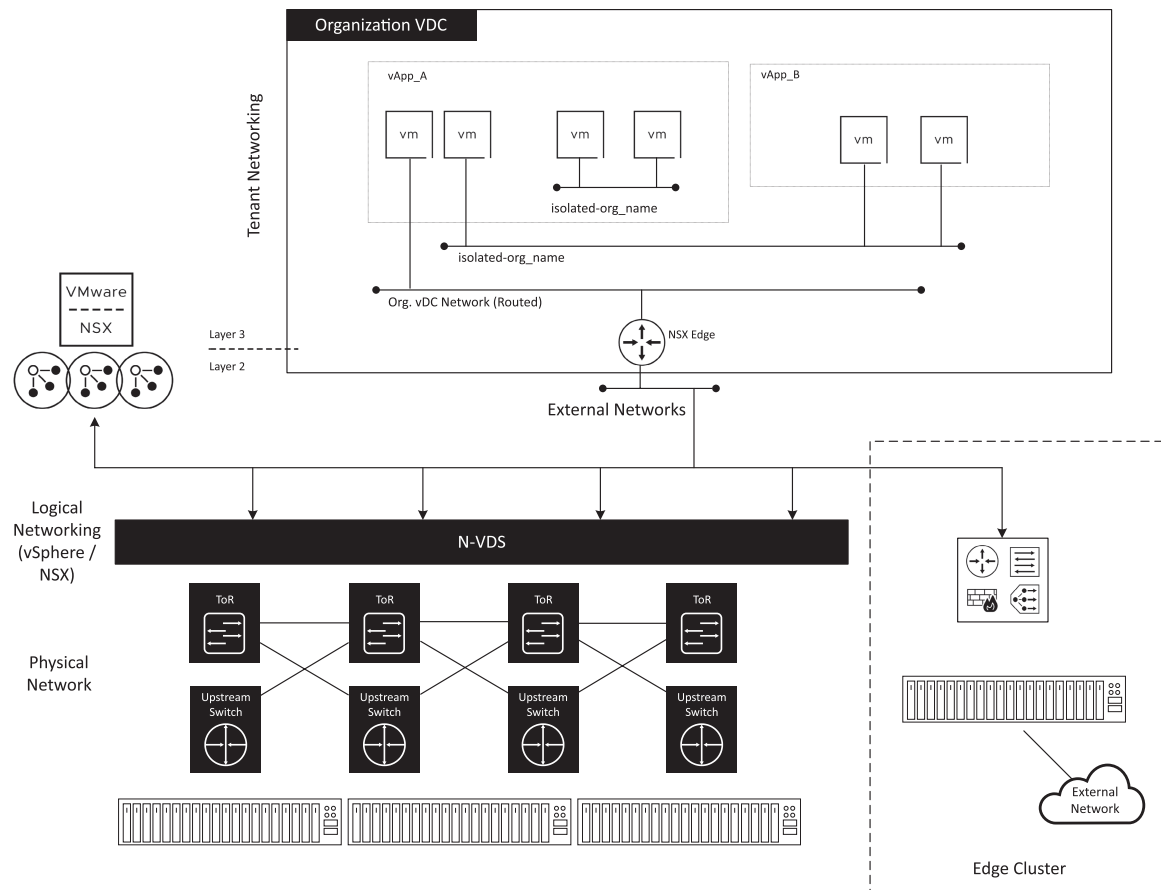
As we have seen, VMware NSX integrates into the core of the Cloud Provider Platform allowing cloud providers to treat their physical network as a pool of transport capacity that can be consumed and repurposed on demand. Through this unique integration with Cloud Director, virtual networks are programmatically created, provisioned and managed, utilizing the underlying physical network as a simple packet forwarding backplane. This unique combination of NSX deployed and completely integrated with Cloud Director and VMware vSphere delivers significant benefits to both provider and customer, including:

- Network provisioning time reduced from days to seconds
- Operational efficiency through automation
- Place and move workloads independent of physical topology
- Integrate third-party network and security solutions through standard RESTful APIs
- Non-disruptive deployment over existing physical networks or next generation topologies

VMWARE NSX FOR VSPHERE VERSES NSX-DATA CENTER

NSX for vSphere (NSX-v), as the name implies, is built around the VMware vSphere ecosystem and therefore requires vCenter Server and ESXi hosts. The VMware NSX-v solution is the older of the two technologies as it spawned from the vSphere product line and Nicira acquisition.

NSX Data Center, rebranded from NSX-T, is the next evolution of software-defined networking from VMware. The major difference between NSX

**FIGURE 6.1** NSX and Cloud Director Logical Architecture

Data Centre and NSX-v is that the newest solution is *decoupled* from VMware vSphere, which means there is no specific requirement to have a vCenter Server and ESXi hosts. From a bigger picture perspective, this allows VMware to move into multi-cloud and more cross-cloud hybrid infrastructure. NSX Data Center also has support for different hypervisors and environments, including cloud-native applications, bare metal workloads, public clouds and multi-cloud environments. In fact, NSX-T is supported on ESXi, KVM, bare-metal servers, Kubernetes, OpenShift, AWS and Azure.

For the purposes of the Cloud Provider Platform, the focus of NSX Data Center is aligned with a native VMware stack of components, although these multi-cloud and public cloud capabilities do open up a range of exciting options and opportunities for cloud providers to build services that can also coexist with hyperscale public clouds.

VMware is now moving fully to the newer platform, although NSX-v continues to be fully supported. However, in order to level set on a single technology, this book will focus predominantly on NSX Data Center (NSX-T) for the implementation of a greenfield Cloud Provider Platform, as all NSX-v functionality has been folded into the newer product. However, having said that, at the time of writing, we are still during a period of transition, and as such there are some features and functionalities that remain currently unique to one or the other NSX solution, when implemented alongside Cloud Director. This will remain to be the case until complete feature parity and the NSX Data Center APIs mature. These will be called out during this chapter as appropriate.

It is noteworthy to highlight, during this period of transition, NSX Data Center and NSX-v functionality can coexist across the same Cloud Director instance. However, keep the following design factors in mind when looking to support both NSX architectures in the same VCD instance:

- NSX Data Center and NSX-v must manage different clusters, but these can be in the same or across different vCenter Servers.
- In a brownfield deployment, retain the current NSX-v backed Provider VDCs/Org. VDCs as-is, and create new Provider VDCs/Org. VDCs for NSX Data Center gradual transition.
- NSX-v and NSX Data Center backed Org. VDCs can be connected with Layer 3 through external networks. No Layer 2 connectivity/stretched network between the different NSX architectures can be supported.

NSX DATA CENTER ARCHITECTURE OVERVIEW

A deep insight into NSX Data Center is beyond the scope of this book, however, understanding its key architectural components is vital to their integration with Cloud Director and the delivery of network services on the Cloud Provider Platform.

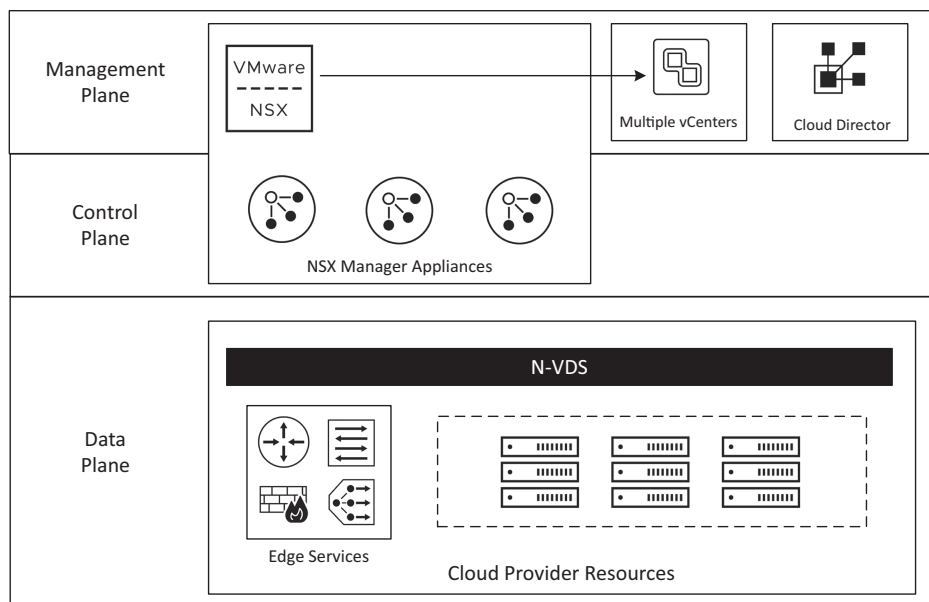


FIGURE 6.2 NSX Data Center High-Level Architecture

From a high-level architecture perspective, NSX Data Center contains a management plane, data plane, and a control plane, as illustrated in Figure 6.2.

Management Plane: The Management Plane provides the point to perform operational tasks, such as configuration and monitoring of all management, control, and data plane components. These changes can be made using either the NSX Data Center user interface or RESTful API. The NSX Manager and NSX Controller are bundled in a virtual machine called the NSX Manager Appliance. NSX Data Center relies on a cluster of three such NSX Manager Appliances for scaling out and for redundancy.

NSX-Manager is the central point of management, but it does not have a one-to-one relationship with vCenter Server, like NSX-v, it can work with multiple vCenter Servers at once. NSX-Manager can be integrated with various Cloud Management Platforms (CMP) via RESTful API. However, for the purposes of the Cloud Provider Platform, the focus is Cloud Director.

Control Plane: The Control Plane is an advanced distributed state management system that provides control plane functions for logical switching and routing, along with propagating the distributed firewall rules. It contains NSX Controllers which are provided as part of the NSX Manager Appliance Cluster. The Control plane is further segregated into two parts, the Central Control Plane (CCP) and Local Control Plane (LCP), which are beyond the scope of this book. For more information, refer to the official NSX documentation and Reference Design Guide.

Data Plane: The Data Plane performs stateless packet forwarding, following the tables populated by the control Plane and maintains statistics to report back to the NSX Manager Cluster. Many heterogeneous systems can be part of the Data Plane, such as Multiple Hypervisors, Bare-metal Servers, NSX Edge nodes, Public Clouds, and Containers.

NSX Data Center uses in-kernel modules for ESXi for constructing the Data Plane. As NSX Data Center is decoupled from vSphere, it doesn't rely on a vSphere Distributed Switch (VDS) for network connectivity. NSX Data Center introduces a specific host switch called N-VDS, NSX Managed Virtual Distributed Switch.

The VMware NSX Virtual Distributed Switch is the newest type of switch in the line of VMware virtual switches. The N-VDS is decoupled from VMware vCenter and is a host switch created on the transport nodes, which for the Cloud Provider Platform are the ESXi hypervisor hosts. The N-VDS switch allows the Data Plane to offer features such as Logical routing, Logical switching, DFW, NAT, DHCP, as well as:

- Provide different Uplink Profiles
- Offer VLAN and Overlay Logical Switches
- Create, read, update and delete operations are performed via the NSX-T Manager

Cloud Director and NSX Integration

As highlighted previously, NSX Data Center is tightly integrated with Cloud Director via its RESTful APIs. As illustrated in Figure 6.3, Cloud Director makes the required API call to the NSX Manager, which is responsible for deploying, configuring, and managing NSX Edge devices that are consumed by tenants to isolate and protect their business applications.

It is noteworthy to highlight again that Cloud Director is not surfacing 100 percent of the configuration potential NSX Data Center Edge functionalities. There are certain things that can be done natively using the NSX tools that cannot be done via VCD API calls. This is largely due to the continuing development of NSX- Data Center and the transition of VCD from consuming NSX-v to being able to consume both NSX-v and NSX-Data Center. These functional differences are more of an exception, and will be mitigated completely in future VMware Cloud Director releases. Figure 6.3 represents a simplified view of this integration.

This end-to-end integration is out-of-the-box and provides the self-service capabilities across the board for workload deployments, as well as security configurations. This doesn't mean that in a Cloud Provider Platform deployment you can't use anything else from a security perspective. NSX Data Center offers an extensible platform to insert other vendor services. Integrated software and hardware partner products can range from network gateway services, application delivery services, and network security platforms to security services from vendors.

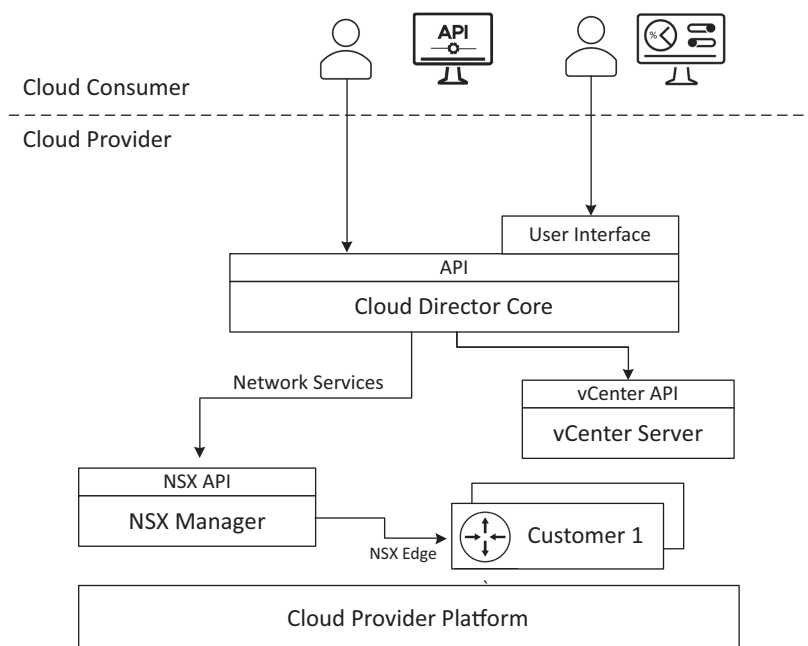


FIGURE 6.3 NSX Data Center and Cloud Director API Integration

This means that if you are required to use something else, other than out-of-the-box NSX functionality, to provide Guest Introspection, for example, you do not need lose out on the integrated experience the Cloud Provider Platform offers consumers.

CLOUD DIRECTOR NETWORKING WITH NSX DATA CENTER

For many cloud consumers, networking is likely the most complex concept of Cloud Director and can be confusing. So let's first step back and understand what we're trying to achieve.

The VMware Cloud Provider Platform is about providing cloud consumers with a level of flexibility to consume resources in a secure and simple way. Cloud Director provides both a user interface and API, which allows consumers to build networks that often mirror what is available to them on-premises as a vSphere administrator. This approach means giving users a lot of flexibility and control. However, from a cloud provider's perspective, delivering this secure isolation model in the networking subsystem is incredibly difficult. Giving customers the ability to create, in a self-service model, separate layer 2 network segments, defining custom layer 3 IP policies, configuring services such as DHCP, NAT Distributed and Edge firewalls, without causing networking conflicts across thousands of customers, is no trivial task.

For this reason, in order for customers to build networks in Cloud Director, a layered approach is required, as illustrated in Figure 6.4.

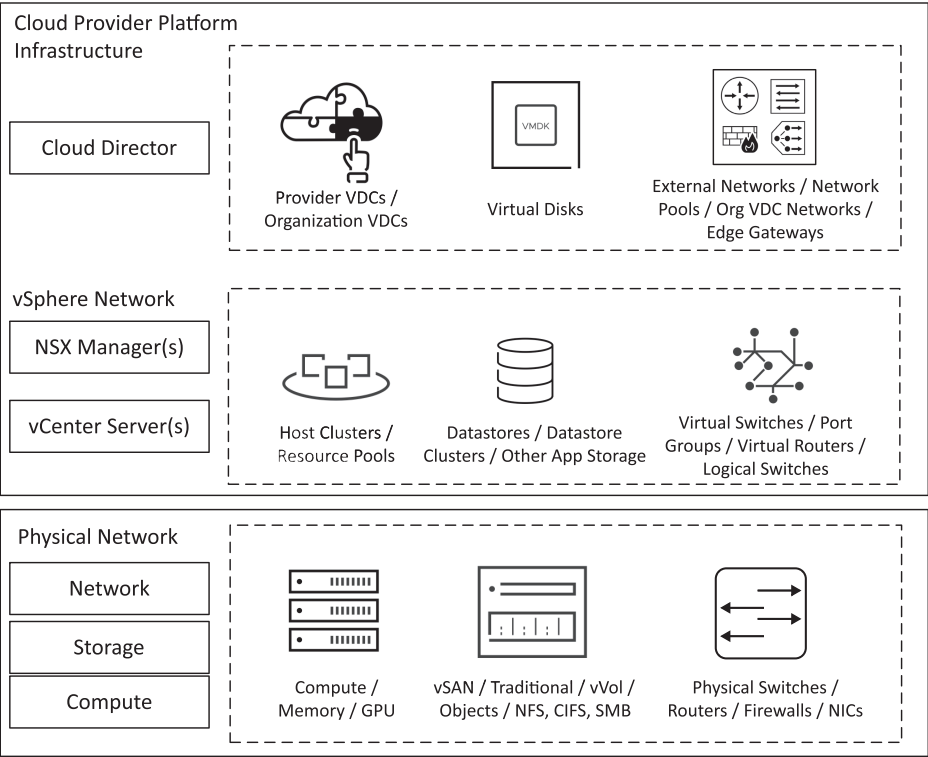


FIGURE 6.4 Cloud Director Network Layers

Here I will try to explain the three major networking levels for the Cloud Provider Platform, from the bottom up, but when it comes to providing guidance on physical network architecture a number of variables, which relate to hardware, come into play. For this reason, in this section the aim is to simply provide some high-level guidelines, which most cloud providers should consider commonly adopted good data center practices.

Physical Network Architecture (NSX Underlay)

The whole purpose of the software defined network is to abstract away the complexity of physical network transport design. However, that does not mean that a sound foundation to the platform is not required, nor does it mean that key physical network design factors do not impact the layers that reside above.

While possible to implement the Cloud Provider Platform on a Layer 2 switch fabric, for cloud provider scalability, a Layer 3 transport should be the

only option considered. This approach allows cloud providers, where necessary, to leverage existing investments in physical network infrastructure.

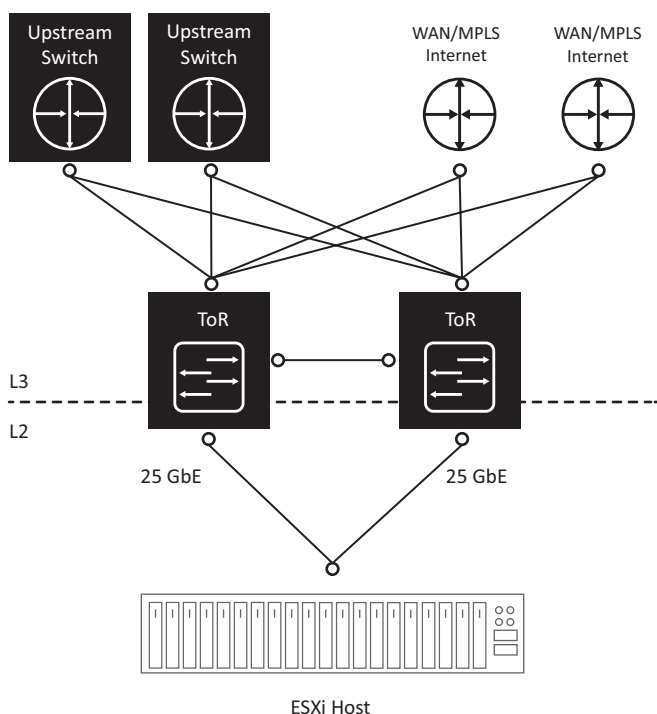


FIGURE 6.5 Host to Top of Rack (ToR) Connectivity and Layer 3 Transport

Top of Rack Connectivity: As illustrated in Figure 6.5, each ESXi host is connected across two 25 GbE ports on ToR switches. These redundant connections can use features in the N-VDS and NSX Data Center to ensure traffic is balanced and that no physical interface is flooded.

Network Transport: For the choice of switch type and specific network connectivity implemented in the data center, the following common best practices should be considered for Top of Rack (ToR) VLAN, Subnet, and access port configuration:

- Redundant physical switch architecture to ensure SLA availability requirements can be met
- Configure each ESXi switch port as an 802.1Q trunk on the access ports that connects the ToR switches to the corresponding hosts. Also consider the following additional network settings and their necessary configuration based on environment variables:

Spanning Tree Protocol (STP)/PortFast

Set MTU for all VLANs and SVIs (Management, vMotion, Geneve/VxLAN, and Storage) to jumbo frames for consistency purposes

DHCP Helper/Relay

- Use only/24 subnets to reduce confusion and mistakes when handling IPv4 subnetting
- Use the IP address .254 as the (floating) interface with .252 and .253 for Virtual Router Redundancy Protocol (VRPP) or Hot Standby Routing Protocol (HSRP)

Jumbo Frames: Network throughput can benefit from the configuration of jumbo frames for multiple traffic types, including vSAN, vMotion, NFS storage, and vSphere Replication, as used by Cloud Director Availability. In addition, using jumbo frames also improves traffic throughput of the Geneve overlay, which requires an MTU value of 1600 bytes or greater.

This is due to the increase per-frame payload from 1500 bytes to the jumbo frame setting, typically an MTU value of at least 9000 bytes. This configuration is required end-to-end, including the physical switch ports, inter-switch links (ISL), switched virtual interfaces (SVIs), vSphere Distributed Switches, vSphere Distributed Switch port groups, and N-VDS switches that support the traffic types listed above.

vSphere Network Constructs

The virtual switches which operate at the vSphere layer perform complex operations of moving data packets in and out of the Cloud Provider Platform. vSphere supports multiple switch types, either running independently or in conjunction with each other.

- vSphere Standard Switch (VSS)
- vSphere Distributed Switch (VDS)

Also, as we have seen, NSX Data Center brings with it the N-VDS, which was introduced earlier in this chapter, a virtual switch specifically designed for use with NSX Data Center, which has some noteworthy characteristics:

- Decoupled from vCenter
- Cross-platform support (KVM and Cloud Endpoints)
- Different Uplink Profiles
- VLAN and Overlay Logical Switches

These characteristics mean that NSX Data Center offers a number of advantages over NSX-v, for multi-cloud environments, in that it is decoupled from vCenter Server. Other noticeable characteristics are illustrated in Figure 6.6. These include:

- pNics are physical ports on the host
- pNics can be bundled to form a link aggregation (LAG)

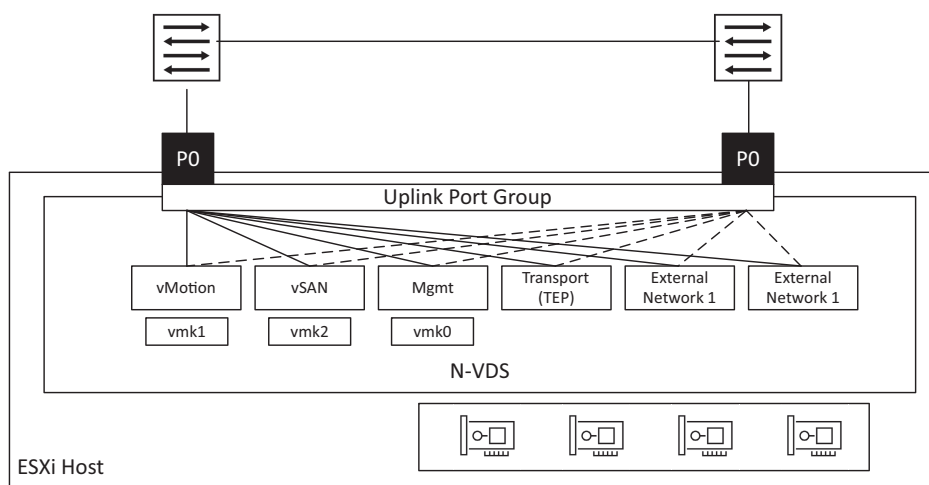


FIGURE 6.6 NSX Data Center N-VDS Architecture

- Uplinks are logical interfaces of an N-VDS
- Uplinks are assigned pnic's or LAGs

N-VDS Teaming Policies: The teaming policy defines how the N-VDS balances traffic across its uplinks and the uplink redundancy and failover model.

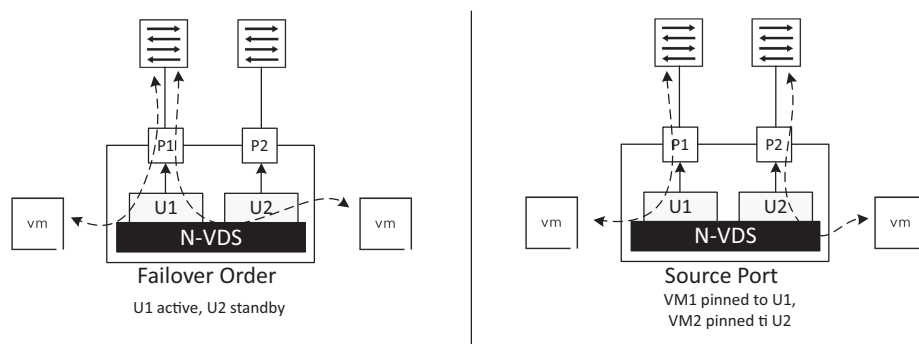


FIGURE 6.7 N-VDS Teaming Policies

Figure 6.7 illustrates two examples of how traffic from two different VMs, on the same network segment, can be distributed across uplinks. The uplinks of the N-VDS can provide any combination of single pNICs or LAGs, which has no impact on the way traffic traverses the uplinks. A LAG is only considered to be failed when all the physical members of the LAG are down. When designing the transport, the cloud provider must specify a default teaming

policy, which is applied by default to the network segments available to this transport node.

Note that a LAG uplink has its own hashing options, however, those hashing options only define how traffic is distributed across the physical members of the LAG uplink, whereas the N-VDS teaming policy define how traffic is distributed between N-VDS uplinks.

GENEVE Overlay: The N-VDS also provides the GENEVE (**Generic Network Virtualization Encapsulation**) overlay, which gives NSX Data Center the capability to create isolated, multi-tenant broadcast domains across the physical data center fabric. Within the Cloud Provider Platform, this enables end customers to create elastic, logical network segments that span physical network boundaries.

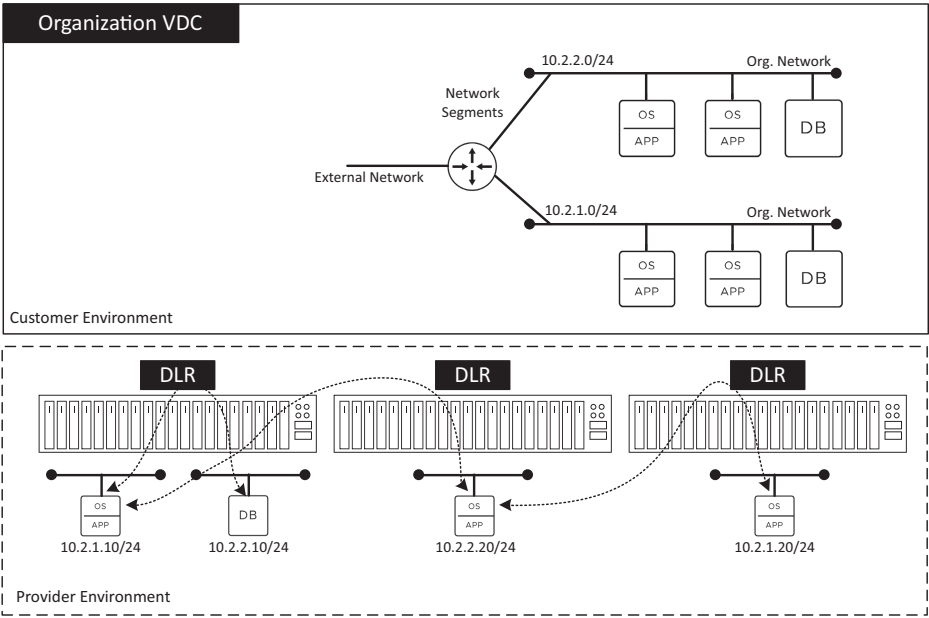


FIGURE 6.8 VMware Cloud on AWS Logical Routing

As illustrated in Figure 6.8, Logical Routing provides East-West routing between different logical switches. The GENEVE network overlay encapsulates networks from the underlying infrastructure so that intra-segment traffic is either switched locally, for workloads on the same host, or is encapsulated and sent through the underlay, for workloads on different hosts. Likewise, for inter-segment traffic, the routing will take place locally where possible, or be encapsulated and sent through the underlay network for cross-host traffic.

The key advantages of the GENEVE network overlay over VLANs is that it is a software defined approach, which can:

- Increase the logical network ID (VLANs have physical ID)
- Easy to create L2 adjacencies across IP networks with tunnels

- You can have almost a million tenants coexist on the same physical network using the GENEVE overlay; therefore, this makes it suitable for both hyper-scale multi-tenant cloud environments and container-based workloads

For more information on any aspect of NSX Data Center design and configuration, please refer to the official VMware documentation.

VMware Cloud Director Tenant Networking

Tenant networks on the Cloud Provider Platform, through Cloud Director, has three different layers of networking construct, with each layer providing a specific purpose:

- External Network
- Org. Network
- vApp Network

However, it is important to highlight immediately that a vApp network, where the network only resides with the vApp, is only applicable when designing a Cloud provider Platform with NSX-v. With that in mind, and as the focus of this book is NSX Data Center (T), we will only be addressing this layer at the most basic level, and for information only.

NSX Data Center and Cloud Director Integrated Architecture: As illustrated in Figure 6.9. The Cloud Director NSX Data Center integrated architecture has some subtle but key differences to that of NSX for vSphere. With NSX Data Center, Cloud Director still acts as a proxy for NSX Managers to maintain the tenant boundary by maintaining the networking function at the Organization level and per user level. As highlighted previously, the NSX Data Center is not dependent on vCenter Server and is therefore attached to Cloud Director independently from vCenter Server.

Tier-0 (T0) or Tier-1 (T1) Gateway offer similar functionality to NSX-v Edge Services Gateways. However, in NSX Data Center these devices are not instantiated as dedicated VMs, instead it is a cluster of *Edge Transport Nodes* where each Edge Node is provided as a VM or baremetal host.

It is also important to highlight that while T0 and T1 Gateways are similar, they are not identical, with each type having a specific purpose or set of services it can offer. As illustrated in Figure 6.9, Tier-1 Gateways are typically connected to single Tier-0 Gateway, a relationship that is automatically managed by NSX Data Center.

However, at the time of writing, T0 devices and corresponding uplinks are created in NSX Manager and then imported into Cloud Director with creation of external networks, instead of just importing an external network port group from vCenter, which is how you would do the same task with NSX for vSphere. As we will address next, this Tier-0 Gateway will provide external networking access for the tenants and should be routable from the internet.

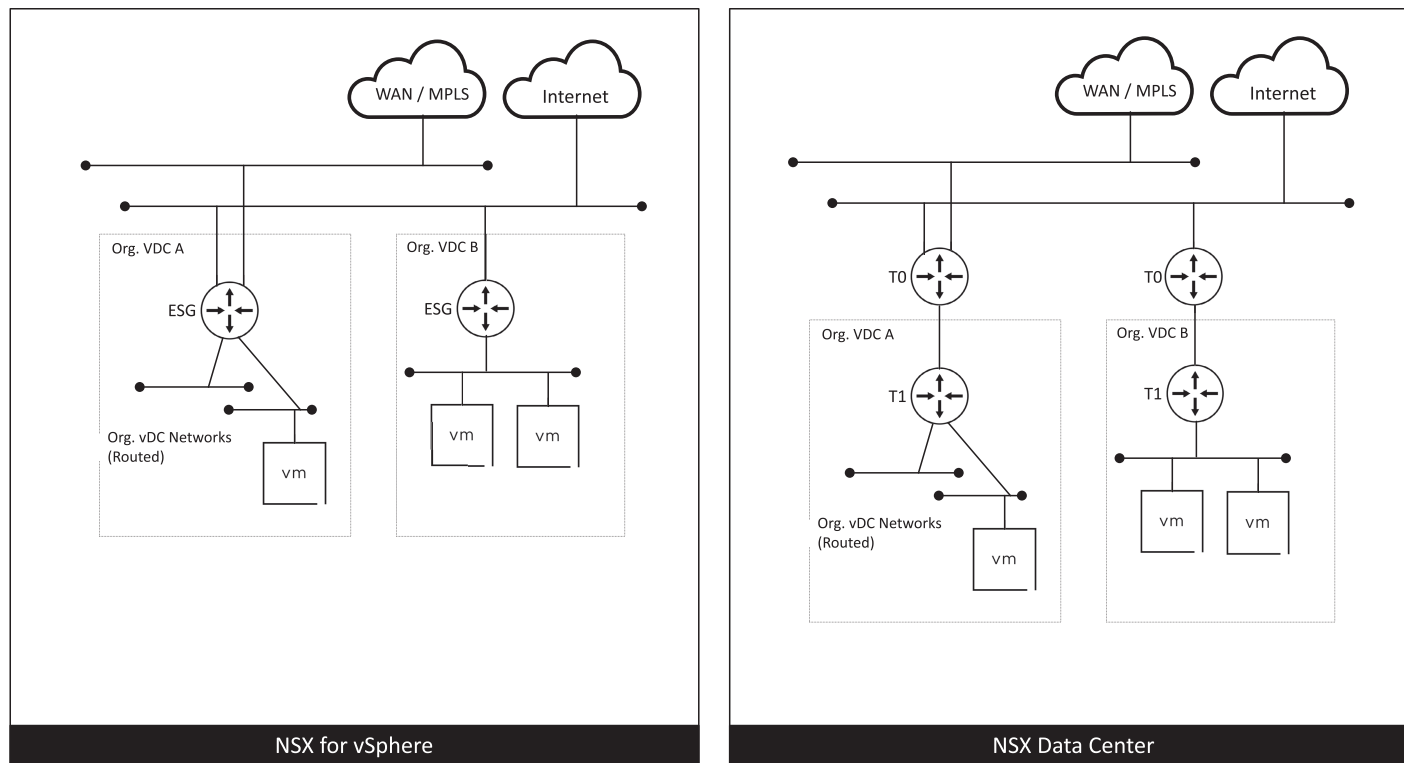


FIGURE 6.9 Cloud Director NSX Integration Comparison

In a typically Cloud Provider Platform design the cloud provider would configure one T0 Gateway in ECMP mode, providing North-South routing (Provider Edge), for each Org. VDC. Tier-0 Gateways cannot be configured with overlapping IP address space.

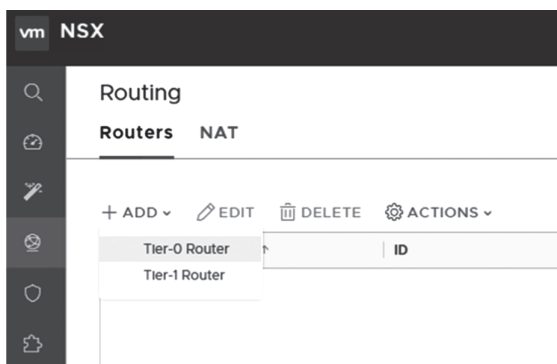


FIGURE 6.10 Creating Tier-0 devices in NSX Manager

The T1 acts as the Org. VDC Edge, providing routing, NATing, firewalling, DNS, DHCP, VPN, and load-balancing services to the workloads that reside in the Org. VDC. NSX Data Center network segments are backed at the vSphere layer by the N-VDS and operate on the GENEVE overlay. All tenant networks connect to the T1 are advertised to the T0, unless configured not to do so.

Cloud Director External Networks: External Networks, also known as *Provided or Provider Networks*, are always created and managed by the cloud provider. This network type is used to deliver external connectivity to the tenant's applications. External Networks are the doors of Cloud Director to the outside world, whether that be private connectivity over MPLS or similar or the public internet. Typically, external networks are created by mapping a Port Group coming off a Virtual Switch at the VMware vSphere layer. Outlined next are some of the typical use cases for the External Networks:

1. Public Internet Access
2. Provider supplied network endpoints such as:
 - 2.1. IP-based storage delivered by the cloud provider or an Org Admin
 - 2.2. Online or offline backup services provided by the cloud provider or an Org Admin
3. Backhauled networking services for customers such as:
 - 3.1. VPN access to a private cloud of the consumer
 - 3.2. MPLS termination to provide access to the customer's private cloud networks

Cloud Director External Networks map one-to-one with vSphere Port Groups. However, Cloud Director does not automatically create vSphere

Port Groups when you create an External Network. The vSphere Admin must create the Port Groups first, before the Cloud Director Provider Admin can map External Networks to them, although as we will address in more detail in Chapter 7, this is a key example of how automation can be created to simplify common operational tasks. Despite how it is created, External Networks require pre-configuration on vCenter Server, and for all hosts to have access to the newly created Port Group. Once this is completed, the External Network can be added to Cloud Director. Setting up an External Network requires the following Layer 3 information:

- Gateway
- Subnet mask
- DNS address
- IP address pool

As illustrated in Figure 6.11, this information is abstracted from the end user and used to provide workloads with a direct connection to the External Network. The pool of IP addresses is reserved for use only by Cloud Director.

The External Networks are the foundation of tenant networking. If the Organization network must access a corporate network or the internet, it will require a shared or dedicated External Network to be configured.

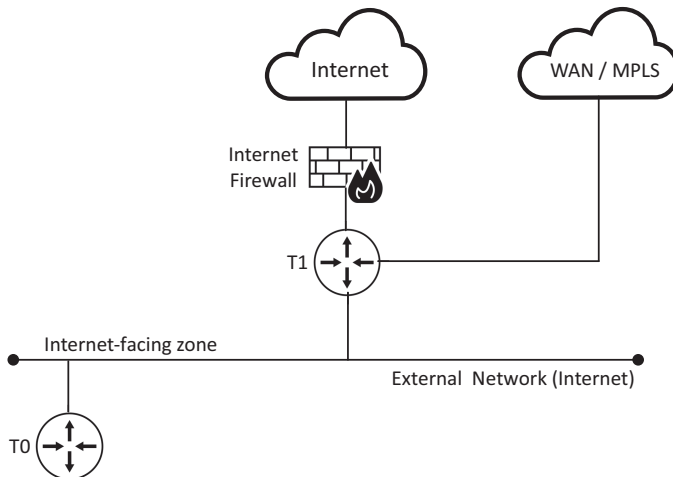
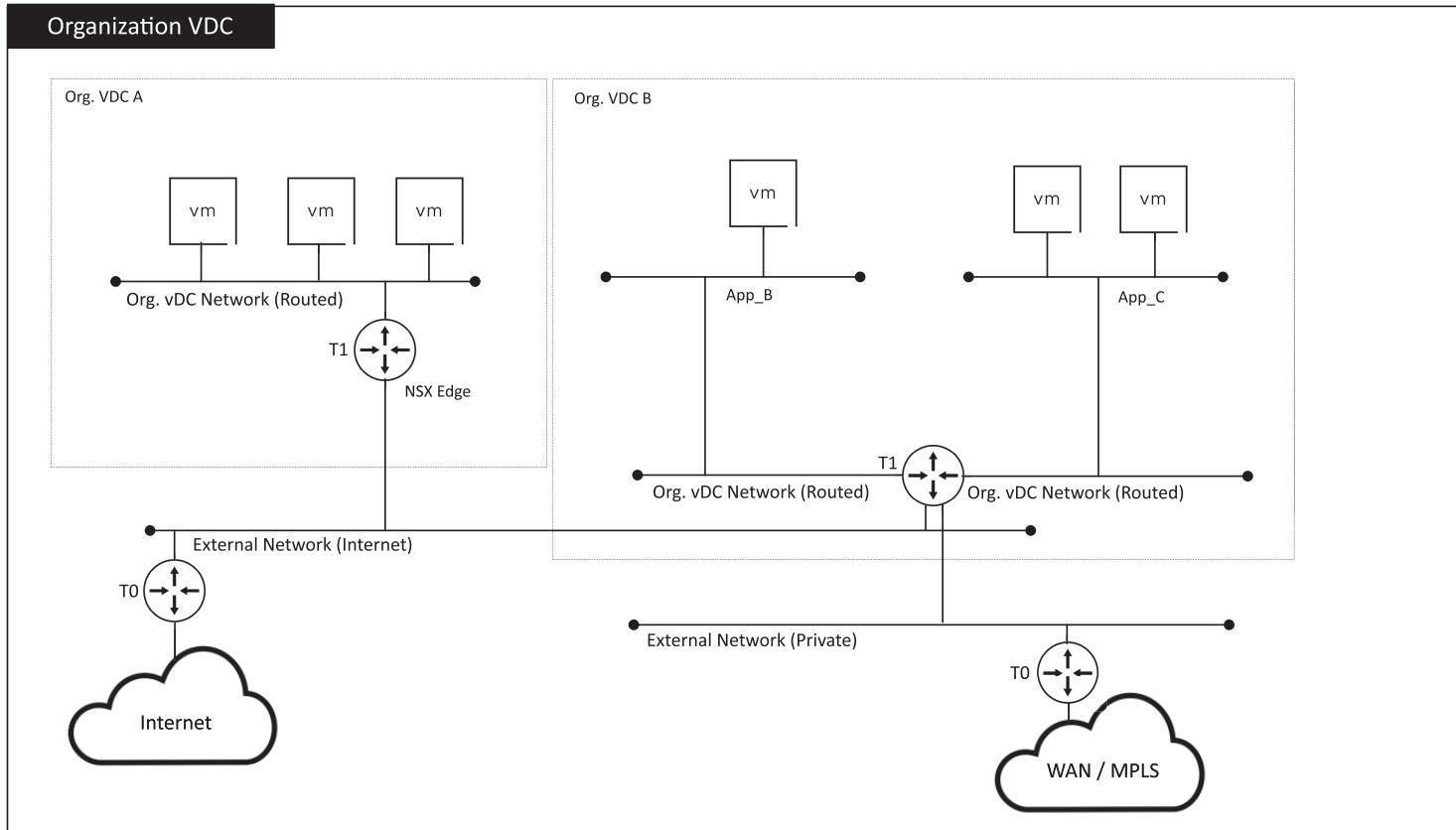


FIGURE 6.11 Cloud Director External Network Configuration

For many customers, the external networks will provide two access types, to the Internet or to a Customer WAN, a customer connection across an external firewall, which in turn is connected to a private MPLS network. Both logical networks are illustrated in Figure 6.12.

**FIGURE 6.12** Customer Internet and WAN Access

A VPN tunnel could also be optionally defined across the Internet toward a customer endpoint. This VPN connection would be managed by the Organization administrator.

Organization Virtual Data Center Networks: The next layer of tenant networking is the Organization Network. Org. VDC networks provide the organization with a private network where workloads can be connected. The Org. Network is used for intra-cloud connections, and even though an Org. Network is intended for internal traffic, as illustrated in Figure 6.12, it can be linked to an External Network to create an entry point or exit from the Org. VDC to other address space.

However, that doesn't mean it must be connected to an External Network, it could be completely isolated and used for internal traffic only. For instance, a test/dev environment where workloads will only need to communicate with each other and not externally. It should also be highlighted that the Org. Network is a required construct. Each Organization needs an Org. Network. The Org. Network is the only mandatory network object in Cloud Director, with three different options available for workload traffic to communicate with outside address space, or not, dependent on the customer use case:

- **Direct Network:** Direct networks connect an organization to an External Network by directly connecting to a Port Group that backs the External Network. A workload will use one IP address from a list of IP addresses that is configured when the network was created, illustrated in Figure 6.14. Directly connected Org. VDC networks are created by Cloud System Administrators only, as they are directly connected to an External Network
- **Routed Network:** A Routed Network is a dedicated Layer 2 network that is routed to an External Network through an NSX Edge. Cloud Director deploys an NSX Edge when this network is created and provides services that include DHCP, NAT, Router, Firewall, load balancing, and DNS relay. Routed networks use IP addresses from network pools, illustrated in Figure 6.13 and 6.15.
- **Internal Network (Isolated):** An Internal Network is internal to the Organization. It does not connect to any external networks and essentially is isolated. Workloads connected to this network can only talk with other workloads on the same network. However, internal networks can span between organizations if needed, illustrated in Figure 6.13 and 6.16.

Routed or Isolated Org. VDC networks are instantiated through network pools by Organization Administrators. Org. VDC networks can also be optionally shared among other Org. VDCs, within in the same Organization.

Routing between the Org.VDC Networks and External networks is provided by Edge Gateways. Each Edge Gateway can have up to 10 network interfaces that can be either external or internal and supports the use of trunk ports. Organization Administrators can configure multiple network services on the Edge Gateways, such as Edge Firewalling, NAT, DHCP, VPN, Static

Routing and load balancing. Edge gateways can be deployed in three sizes, depending on the scale of traffic (small, medium, and large) and also optionally in high availability mode.

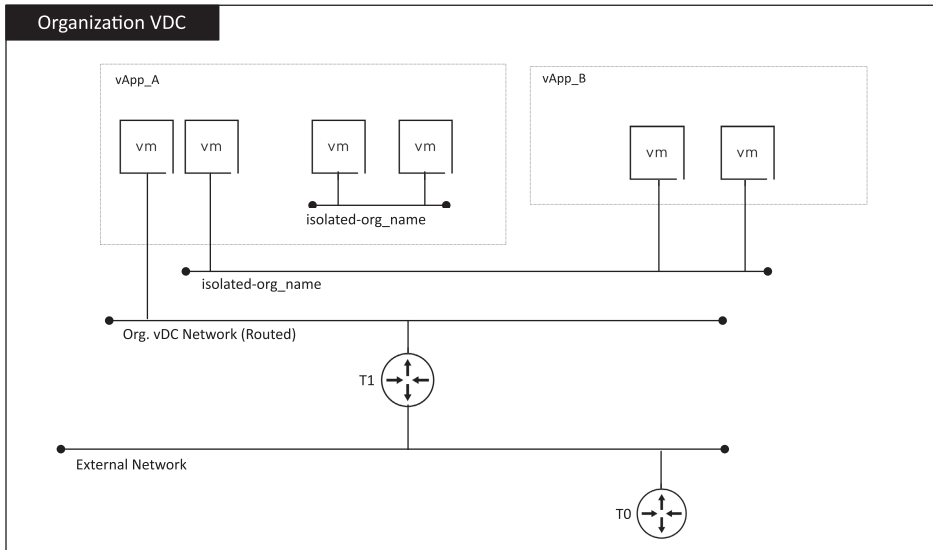


FIGURE 6.13 Organization VDC Network Configuration

Network Pools: Before we progress, we need to introduce another concept that is of paramount importance for Cloud Director operations, *Network Pools*. A Network Pool is a set of Layer 2 networks that the cloud provider has declared as *available*. Network pools are key to how we give an end user a controlled way to deploy Layer 2 networks. This concept removes all the traditional constraints of network provisioning, such as maintaining an EXCEL spreadsheet, and choosing an available VLAN. In a cloud self-service model, that approach simply can't work. If the customer needs to deploy Layer 2 networks, the network pools effectively provide an IPAM database.

A Network Pool contains network definitions used to instantiate private or routed Organizational networks. Networks that are created from network pools are isolated at Layer 2.

The following types of network pools are available in Cloud Director:

- vSphere Port Group backed are network pools that are backed by pre-provisioned Port Groups in vCenter Server. These can be created on any switch type but most likely are N-VDS-based Port Groups (typically backing external networks).
- GENEVE network pools that use a Layer 2 over Layer 3 MAC in UDP encapsulation to provide scalable, standards-based traffic isolation across Layer 3 boundaries.

- VLAN-backed network pools are backed by a range of pre-provisioned VLAN IDs. This implies all VLANs specified are trunked into the cloud environment. This option is not typically used on the Cloud Provider Platform due to scalability limitations.

vApp Networks: As highlighted previously, vApp networks are, at the time of writing, only available when implementing the Cloud Provider Platform with NSX for vSphere. Therefore, this section is provided for information only.

A vApp is a container for a distributed software solution and is the standard unit of deployment in Cloud Director. It has power on operations, consists of one or more virtual machines, and can be imported or exported as an OVF package. vApp Networks are created by cloud consumers and connect multiple virtual machines in a vApp together. vApp networks separate vApp virtual machines from the workloads in the Organization virtual data center network.

Network Use Cases

As we have addressed, Cloud Director has different layers of networking. Each of these layers has a specific purpose within the Cloud Provider Platform architecture. The external network is the connection to the outside world and the Org. Network is linked to a specific organization.

In this section we will address common use cases and architectures being deployed by cloud customers in order to illustrate several different use cases. The aim is that the examples provided will help firm up the readers understanding of the network constructs discussed.

Directly connecting virtual machines to an Org. VDC network, as illustrated in Figure 6.14, demonstrates how an application can be directly connected to an External Organization VDC network, an External Network that must be exclusive to that specific organization.

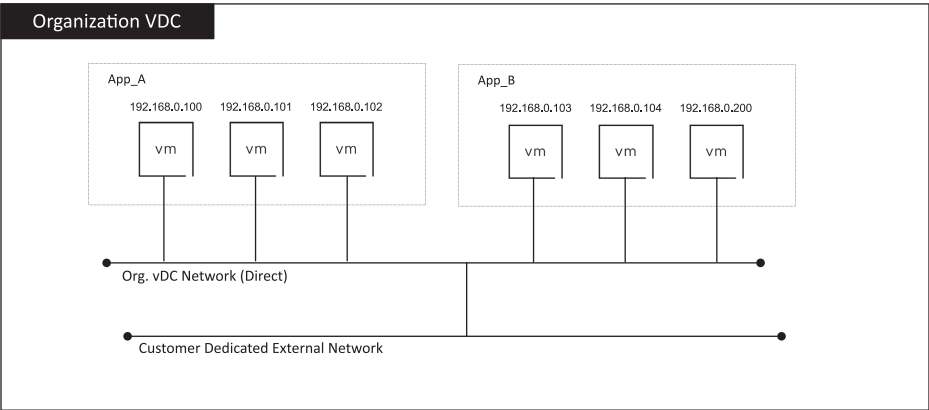


FIGURE 6.14 Directly Connected Org. VDC Network

In Figure 6,15 we add an NSX Edge to fence that network. This figure shows an application connected to a routed Org. VDC network. In this example, the NSX Edge could be providing DHCP, Firewall, NAT, and Static Routing services to the Org. VDC network.

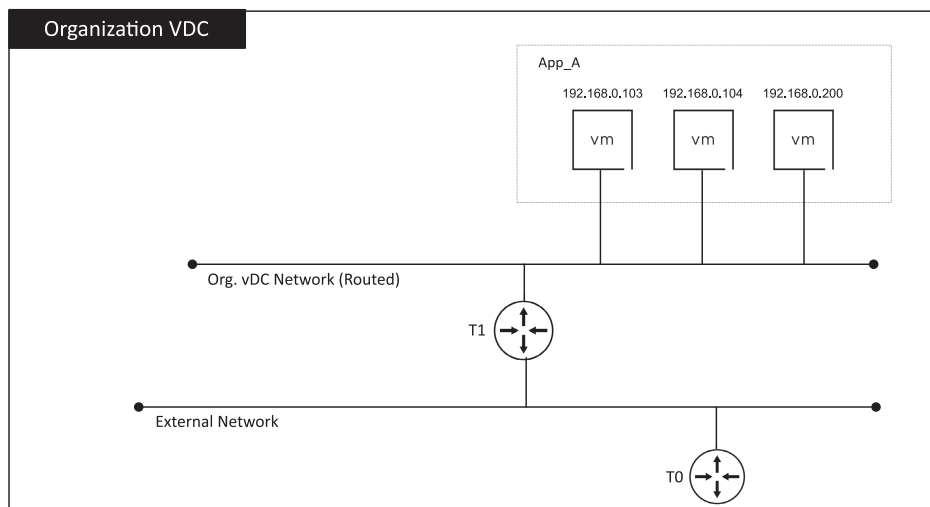


FIGURE 6.15 Routed Org. VDC Network

The next figure depicts a Org. VDC network that is isolated within its Org. VDC. An NSX Edge can be deployed if DHCP is required.

An application network configure to none is completely isolated and the virtual switch of the corresponding Port Group is the endpoint for this network. This network is isolated on Layer 2 and only intracommunication is possible.

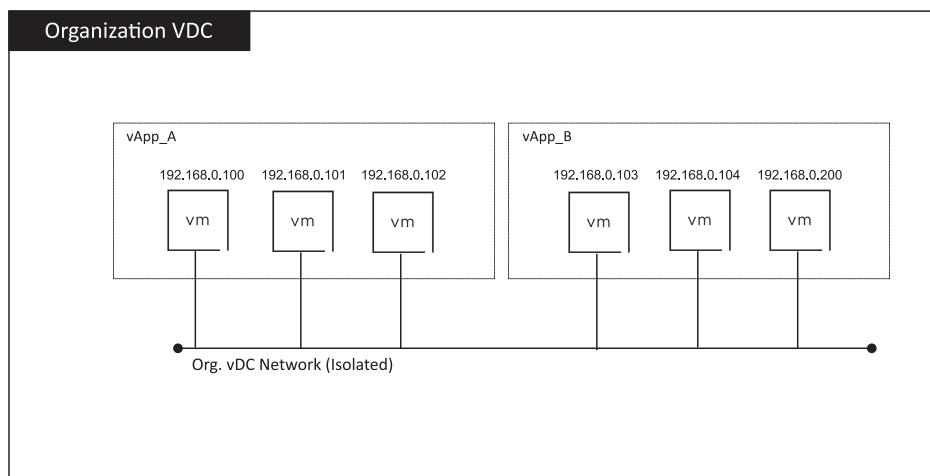


FIGURE 6.16 Internal (Isolated) Network

Static Routing: Another feature in Cloud Director is the support for static routing. This provides the ability to route between network segments without the use of NAT, and enables increased flexibility in implementing network connectivity within a complex cloud environment. This feature can help make it easier for customers to replicate their on-premises networks as closely as possible on the Cloud Provider Platform.

Though most networks will have a directly connected default gateway, it is possible for networks to have more than one route (such as when using multiple interfaces on an NSX Edge device). Static routing provides a way to manually configure routing tables so that traffic can be forwarded to these remote networks, while still using the default gateway for all remaining traffic.

This static routing can be configured at the routed Org. VDC network level. For Edge Gateway instances that are connected to multiple External Networks and Org. VDC Networks, routes can be applied on the entire NSX Edge Gateway or on any one of the external networks connected to the Edge Gateway.

Unlike NAT, static routing does not support overlapping network ranges. If there are plans to leverage static routing within the Cloud Provider Platform environment, provider operational teams need to work in conjunction with the customer to make sure that IP addresses for organizational networks are unique when allocated. The management of static routing is typically provided as a customer responsibility by most cloud providers.

For instance, illustrated in Figure 6.17, where access is required to network resources on an external network, this applies to scenarios where there is a requirement for connectivity to networks through a different next hop address, other than the default external gateway. An example might be access to a remote management network or accessing services in another Org. VDC.

Load Balancing: Load balancing is another functionality provided by Cloud Director in conjunction with NSX Data Center, which can be used to increase application availability, either through redundancy, increasing throughput, or minimizing response time by redirecting to servers based on load. The management of application load balancing can either be provided as a managed service by the cloud provider or be delivered as a customer managed solution, depending on the offering type and load balancer technology consumed.

A load balancer defines a virtual service which offers an external representation of an application. Traffic received by the load balancer can be distributed and load-balanced across the set of target hosts, with the service appearing as though it was handled by the target host itself. This model is popular as it provides benefits for application scale-out and provides high-availability particularly, but not exclusively, for web applications.

There are three options for implementing load-balanced applications on the Cloud Provider Platform. Differences in architecture will occur based on the type of load balancer chosen:

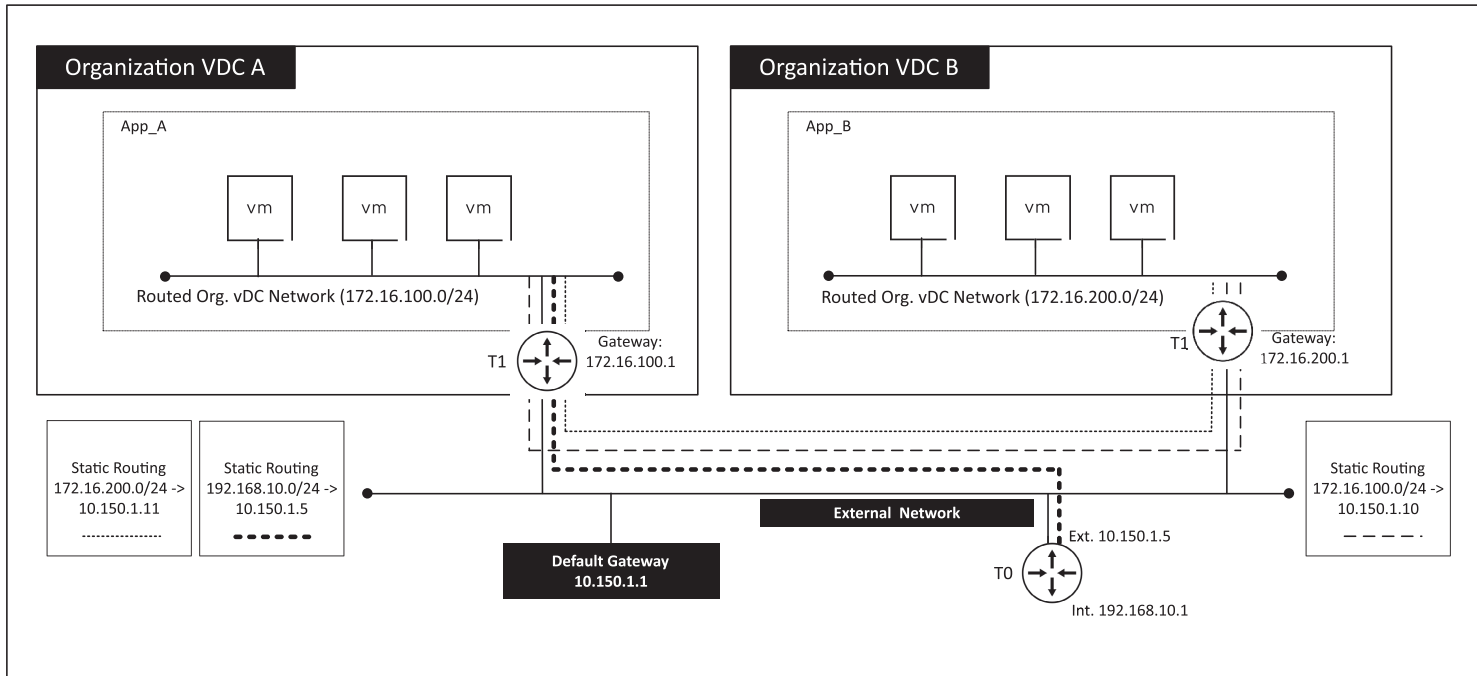


FIGURE 6.17 Organization VDC Static Route Communication

- Native Cloud Provider Platform NSX Edge Gateway as a Load Balancer
- VMware NSX Advanced Load Balancer (Avi Networks)
- Third-party Virtual Appliance Load Balancer (e.g., F5 BIG-IP Virtual Edition)

NSX Data Centre Edge Load Balancer: The first and default option is typically the native NSX Edge Load Balancer. With Cloud Director it is possible to configure this option as a self-service load-balancing solution using the NSX Edge built-in features.

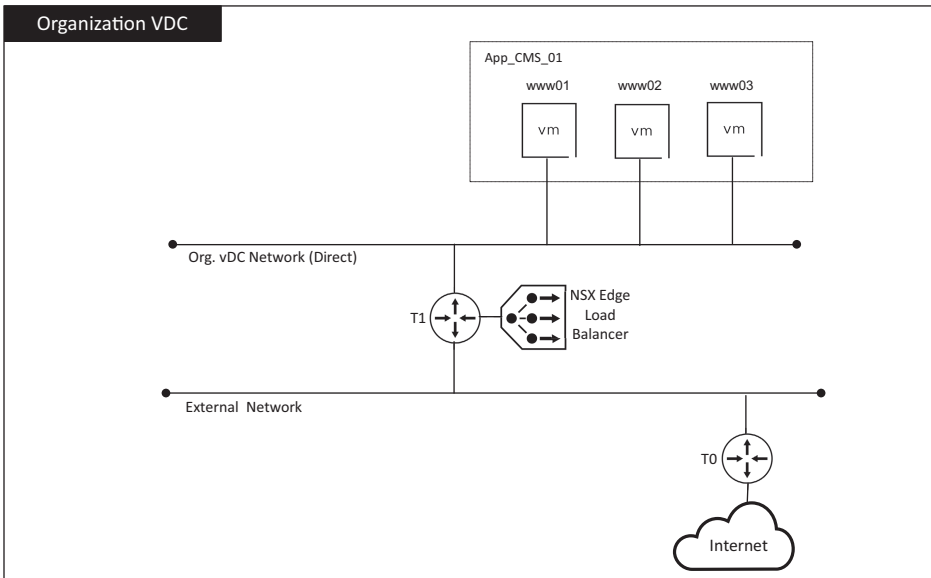


FIGURE 6.18 NSX Edge Load Balancer

The native Edge Gateway provides load balancing for TCP, HTTP, and HTTPS traffic and is capable of load balancing up to Layer 7, enabling web application auto scaling. An external or public IP address can be mapped to a set of internal servers for load balancing purposes. The load balancer accepts requests on the external IP address and decides which internal server to use.

With this native solution, it is even possible to improve the availability for the Edge Gateway load balancer by running it in native high availability mode. This facilitates the Edge Gateway with an almost instantaneous failover, with session preservation.

VMware NSX Advanced Load Balancer: The second option is the VMware NSX Advanced Load Balancer (formerly Avi Networks). More likely to be offered as an advanced managed service by cloud providers, the NSX Advanced Load Balancer provides multi-cloud Load Balancing, Web Application Firewall capabilities, and application analytics across on-premises data centers and any cloud endpoint. The software-defined platform delivers applications

consistently across bare metal servers, virtual machines, and containers to ensure a fast, scalable, and secure application experience.

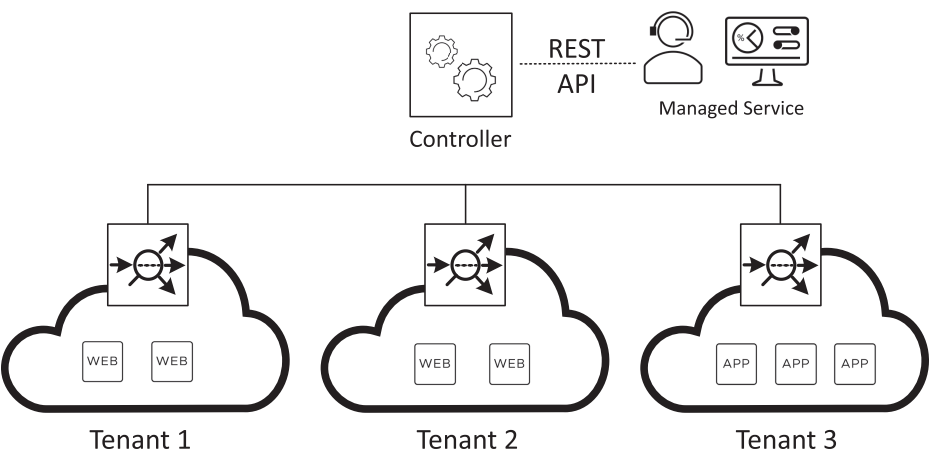


FIGURE 6.19 VMware NSX Advanced Load Balancer (Managed Service)

In addition, the underlying Load Balancing automation is surfaced through RESTful APIs that can facilitate seamless scaling of workloads, which is key to application autoscaling and bursting, one of the key use cases for hybrid cloud.

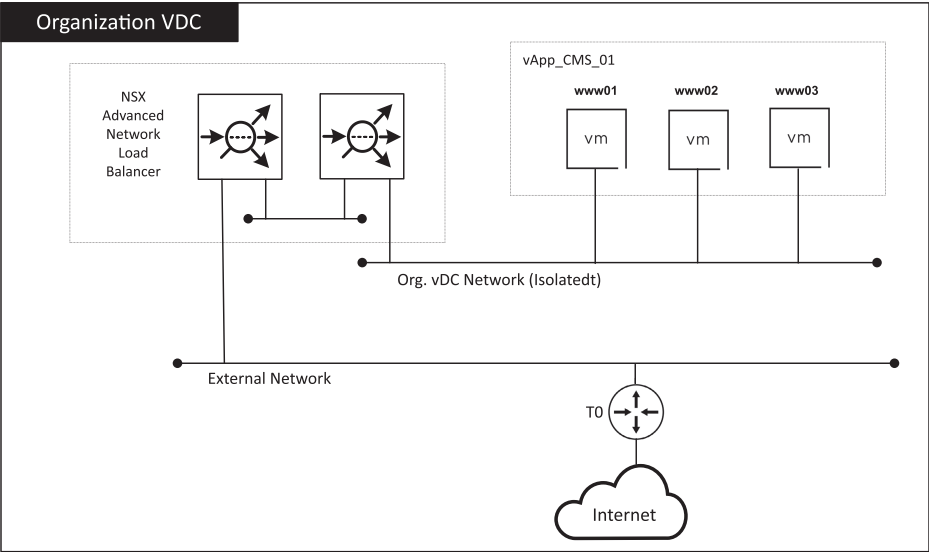


FIGURE 6.20 NSX Advanced Load Balancer and Cloud Director Example

Third-Party Virtual Appliance Load Balancer: Finally, there are many third-party companies providing virtual load balancers, with varying degrees of multiprotocol support and advanced features.

When using a virtual appliance as a load balancer it is important to protect the security of the applications by using a firewall. In this scenario, Cloud Director does not provide security or isolation for the backend workloads other than what the load balancer provides.

If not provided natively by the third-party load balancer selected, in order to provide availability, vSphere High Availability can protect against physical host failures and restart failed appliances. This provides approximately 99.9% uptime for the load balancing functionality. For a solution that provides native availability features, Cloud Director compute policies can be used to configure anti-affinity rules, ensuring that mirrored appliances do not reside on the same physical host.

While not available at the time of writing, it might be the case that in future we see Load Balancer vendors develop user interface extensions for Cloud Director, making consumable unique features and characteristics of that technology, directly from the Cloud Director user interface. When considering third-party solutions, the VMware Marketplace should typically be the first point of investigation, due to its native integration with Cloud Director, addressed in Chapter 8.

Distributed Firewall (Micro segmentation)

One of the available features on the Cloud Provider Platform, delivered through NSX Data Center, is the Distributed Firewall (DFW), a feature also often referred to as micro-segmentation.

Micro-segmentation allows customers to logically breakup the data center into distinct security zones, and provide security controls for services that run across each zone. It does this by implementing stateful packet filtering capabilities on each workload virtual network interface. This control over traffic allows the DFW to limit end-to-end the flows between workloads on the same network segment. The key purpose of micro-segmentation is to restrict the ability of a hacker to move across different segments of a data center after the perimeter network has been compromised. This feature prevents lateral movement and can stop two workloads from talking to each other, even where they sit on the same subnet, thus containing an attacker.

As illustrated in Figure 6.21, micro-segmentation allows an additional degree of control over traffic that has already passed the perimeter firewall.

The Distributed Firewall policies illustrated, which control the flow of traffic to and from each workload are configured centrally through either the Cloud Director user interface or API. These rulesets are then distributed to the ESXi hosts to implement on the running workloads.

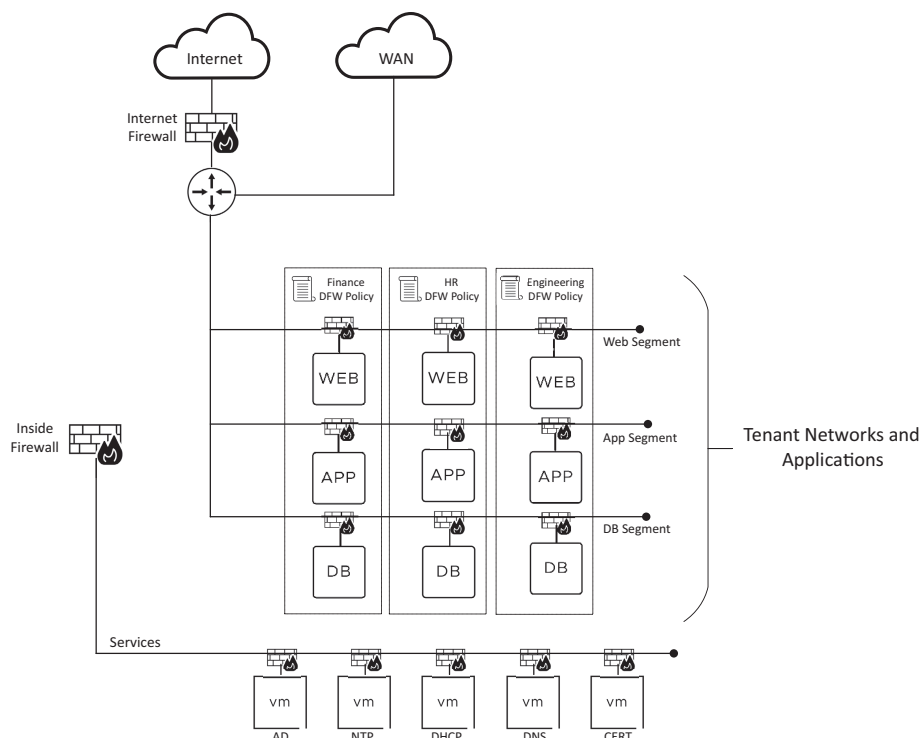


FIGURE 6.21 Distributed Firewall Architecture

Figure 6.21 illustrates a customer implementation example of a common DFW use case. The Distributed Firewall policy is being applied to traffic to and from three web servers which are also exposed to the public internet.

In the illustration, each workload is effectively separated from the network by an integrated firewall at the vNIC level. This allows the packets between workloads even on the same network to be controlled by a DFW policy, and architecture that is highly impractical in a traditional networking infrastructure.

This architecture represents an end-to-end security solution entirely managed from Cloud Director, the North/South policy on the Edge Services Gateway and the Distributed Firewall policy itself, controlling traffic flows on East/West communication. Both firewall services are managed from the Cloud Director user interface or API by either the customer, in a self-service manner or by the cloud provider, in the case of a managed offering.

On the Cloud Provider Platform, the Distributed Firewall runs as a service, in an easily consumable way. Furthermore, VMware Cloud simplifies the intelligent grouping mechanism, allowing the Tenant Organization Admin to go beyond the traditional contexts of IP Address or MAC Address and include logical constructs, such as Security Groups based on OS names, virtual

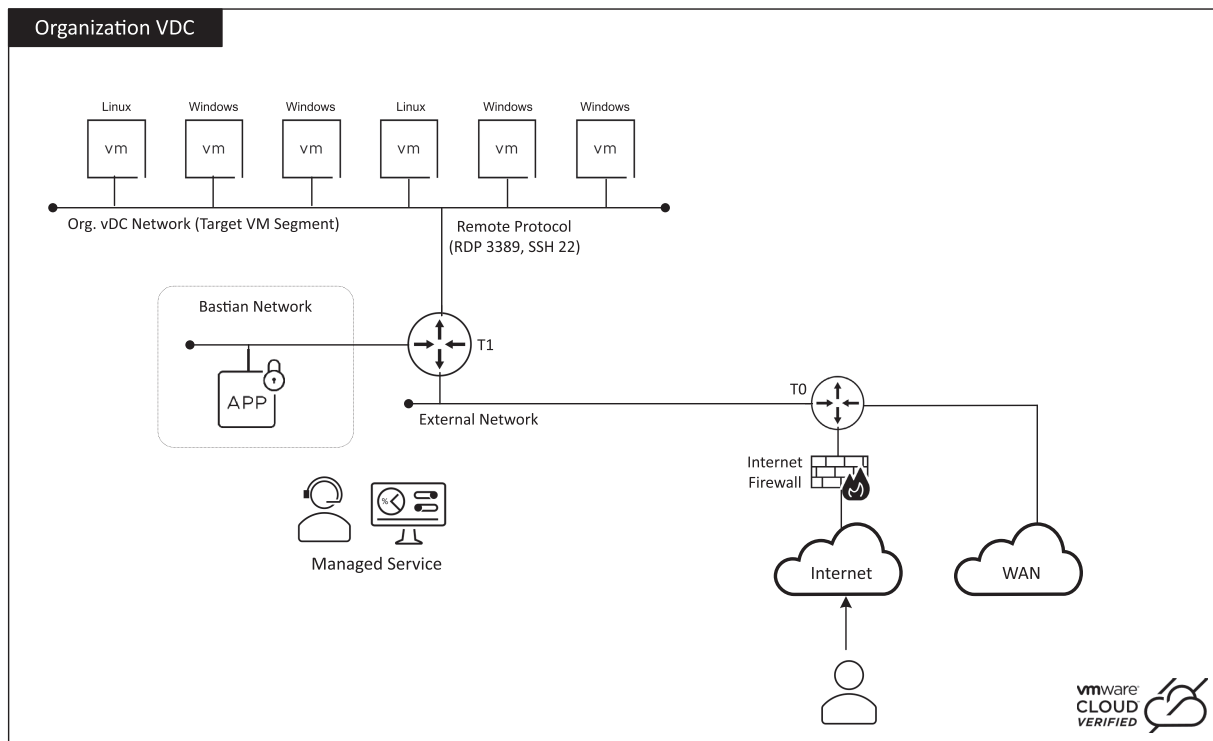


FIGURE 6.22 Cloud Provider Bastian Service

machine names, Services, Application Type etc., making it easy to implement simple and more granular security controls.

This implementation flexibility means that the Cloud Provider Platform can make micro-segmentation a reality inside an organization within a matter of hours, helping end customers improve their security and achieve their cyber-defense security requirements in an easily consumable manor.

Bastian Connectivity Service

A Bastian Service also often referred to as a *jump box* or *jump server*, is a common architectural approach to access virtual machine and application infrastructure that do not have public IP addresses. The aim of this solution is to provide secure connectivity, typically over RDP or SSH, to the customer's workloads, as a more convenient solution than the remote console service.

As illustrated in Figure 6.22, a Bastian service can help customers connect over either RDP (Port 3389) or SSH (Port 22) securely, as there is no requirement to expose these ports directly to the public internet. This architecture can therefore provide secure RDP and SSH connectivity to all workload types provisioned.

As illustrated in Figure 6.22, RDP and SSH are the most common protocols consumed by administrators to connect to workloads running in the public cloud. However, exposing these protocols publicly opens up a significant attack surface, easily identified by hackers through port scanning software. To help mitigate this threat and reduce the potential attack surface, cloud providers can offer a Bastion Service to customers on a public-facing perimeter network. This service securely allows customer connectivity to workloads, sitting on internal networks, without requiring the remote console.

Not only does this service protect against port scanning driven attacks on the public internet but can be further secured using other services, such as a Layer 7 Firewall or a third-party Web Application Firewall (WAF) to protect against attacks. When offered as a managed service, cloud providers can also ensure that the Bastian hosts are always maintained by keeping the service hardened and securely patched on behalf of the customer.

Chapter 7

Platform Automation and Standardization

VMware Cloud Service Providers deliver solutions to businesses using a range of VMware and third-party technologies. Any cloud provider will tell you, bespoke, snowflake solutions don't scale for their business and are unsustainable from an operational perspective. Building a custom solution for each customer puts an operational overhead on the provider's business that cannot be sustained. Therefore, the model most cloud providers operate under is to develop standardized solutions that meet 90 percent of prospective customer use cases, and then work with the customer to develop custom add-ons and re-architecture where necessary.

That doesn't mean that cloud providers don't evolve their service offerings to align with common customer requests, it just means that having a standard offering across all customers must be the starting point. The aim is to build a service that attracts customers, and they will come!

Cloud providers must therefore think about how they standardize hardware, service offerings, data center architecture, automation and operations, which are all key components of cloud delivery that need to be considered. At a minimum, to deliver business solutions, at scale, on the VMware Cloud Provider Platform a solution must have the following essential characteristics:

Standardized

Adaptive

Automated

Resilient

Holistic

Where possible, cloud providers must standardize infrastructure, even though services are delivered as software across pools of resources, homogeneity eliminates unnecessary complexity caused by hardware silos and the redundant tools and skills associated with them. It can also eliminate costly special purpose hardware and facilitates the flexibility to move hardware resources across different Resource Groups, Pods, or even data centers. A single scalable approach to hardware procurement also enables a simplified backup and recovery architecture, which in turn reduces operational overhead.

Infrastructure should also be adaptive and provide services on demand, unconstrained by physical topology, application, scale, or location. The aim should be to design a platform that can dynamically be reconfigured, based on workload demands, enabling maximum agility and efficiency.

Automation must be central to the solution. Built-in platform intelligence to automate provisioning, placement, configuration, and control, based on defined policies. Intelligently automated infrastructure eliminates complex management processes. Less manual intervention equates to improved operations, speed, and cost savings.

By design, the Cloud Provider Platform's software-based architecture is resilient and tolerant to hardware failure. The architecture is designed for seamless failover, redundancy, and fault tolerance to critical operations. This intelligent native automation provides resilience without the need for manual intervention.

Finally, cloud providers need to take a holistic view to develop services optimized for the entire data center fabric, providing comprehensive infrastructure services capable of supporting any and all applications. The Cloud Provider Platform aims to support any workloads, with complete flexibility to balance collective application demands, eliminating the need for diverse technology stacks.

CLOUD AUTOMATION

VMware's software-defined data center strategy has driven a trend that has led to an increased emphasis on automation among partner and customers. Once operations are automated and manual tasks eliminated, it can free up valuable time for operational teams to spend on more strategic initiatives.

This shifting focus from provisioning storage, to configuring storage rules, and cross- infrastructure needs, like performance, availability, agility, and rapid application deployment has seen the rise in use of integrated management tools. Increased attention to automation and agility across the Cloud Provider Platform has seen an emergence of new skill sets focused on automation, API integration and a broader understanding of the concepts of infrastructure as code. This has led to a new breed of operational team, commonly referred to as Site Reliability Engineering (SRE).

A SRE team provide backend support and is typically made up of system and software engineers who solve production problems, in order to keep the service available. In addition, SRE continuously aim to improve availability and reliability to deliver the efficiency of the service through the automation of common operational tasks, including lifecycle management of the stack. For SRE teams it is helpful that the cloud provider infrastructure is designed to be a tightly integrated software stack that can allow cloud providers to develop common workflows and increase automation, which helps reduce operational tasks and improve responsiveness to customer demands. These include but are not limited to:

A need to automate the delivery and management of production-ready infrastructure and application components to reduce provisioning time, improve resource utilization, and eliminate error-prone processes.

A need to reduce time to value and time to market by eliminating the operational bottlenecks caused by having to manually configure and provision networking, security policies, and infrastructure resources.

Meet customer business demands to enhance security and security policies to deliver standard and consistent applications, which protect the customer's businesses.

Meet the growing need for data storage and to rapidly deploy and managed new storage resources with quality and control, ensuring the needs of customers can be met.

vRealize Orchestrator

VMware vRealize Orchestrator (vRO) provides the drag-and-drop workflow engine that simplifies the automation and integration of a service offering. This software also allows the cloud service developers to create complex automation tasks using its workflow designer for assembling operational workflows. These developed workflows can then be executed either directly from the Cloud Director Service Library, by the provider, consumer, or reactively through a notification or blocking task.

For the Cloud Provider Platform, the key requirement of utilizing vRealize Orchestrator is to facilitate workflows integrating the cloud management platform with the underlying resources and vendor provided plug-ins, allowing daily processes to be automated, for both provider and customer, accelerating and simplifying operations. In addition, by utilizing vRealize Orchestrator to provide cloud-based services, the aim is also to reduce operational expenditure, in order to reduce the total cost of ownership of the solution.

Plug-in Support for Cloud Director: As outlined above, vRealize Orchestrator is designed to allow SRE teams and cloud operational staff to streamline tasks through automation, integrate solutions with VMware products, third-party providers, and develop custom orchestration.

To this end, vRealize Orchestrator can automate common tasks that are run against Cloud Director with a plug-in specifically developed for cloud providers. This plug-in helps vRO understand the user, admin and extension API classes for Cloud Director and provides a communication channel to talk to Cloud Director endpoints. The plug-in has out-of-the-box workflows pre-defined, which can help develop custom orchestration for cloud providers to execute common tasks such as:

1. List all enabled/disabled Organizations
2. Add/edit Provider VDC, Organizations or Org. VDCs
3. Configure networking services on Org. VDC

The primary benefit of vRealize Orchestrator is to coordinate multiple internal and external systems to achieve a composite operation, which otherwise would have taken several individual operations on different systems. It is not meant to replace the APIs that it in turn calls; rather, it is meant to bring them together when necessary. In general, if an operation uses only one underlying system, direct access to that system should be considered for efficiency and reduced complexity. For the Cloud Provider Platform, orchestration can be used to automate highly repetitive tasks to avoid manual work and prevent potential errors. There are three main categories of use cases that vRealize Orchestrator can address:

- Cloud administration operations
- Organization administration operations
- Organization consumer operations

vRealize Orchestrator can orchestrate workflows at the vSphere API, VIX API, Cloud Director API, and vRealize Operations API levels, meaning that the potential scope for automation for cloud providers is significant, it is simply about defining the use cases.

Cloud Administration Automation: These use cases are focused on infrastructure management and control over cloud resources. For instance, consider a cloud provider who wants to bring a new customer onto their Cloud Provider Platform. The key steps would be to create a new organization, add users (possibly imported from an identity source, such as Active Directory), and add networks and Org. VDCs. The provider may also want to set up a recurring chargeback report through the Tenant App so the tenant can be invoiced. In addition, as illustrated in Figure 7.1, CMDB updates may be required and the sending of an email notification to the tenant, advising them that their new cloud service is now available, might also be orchestrated.

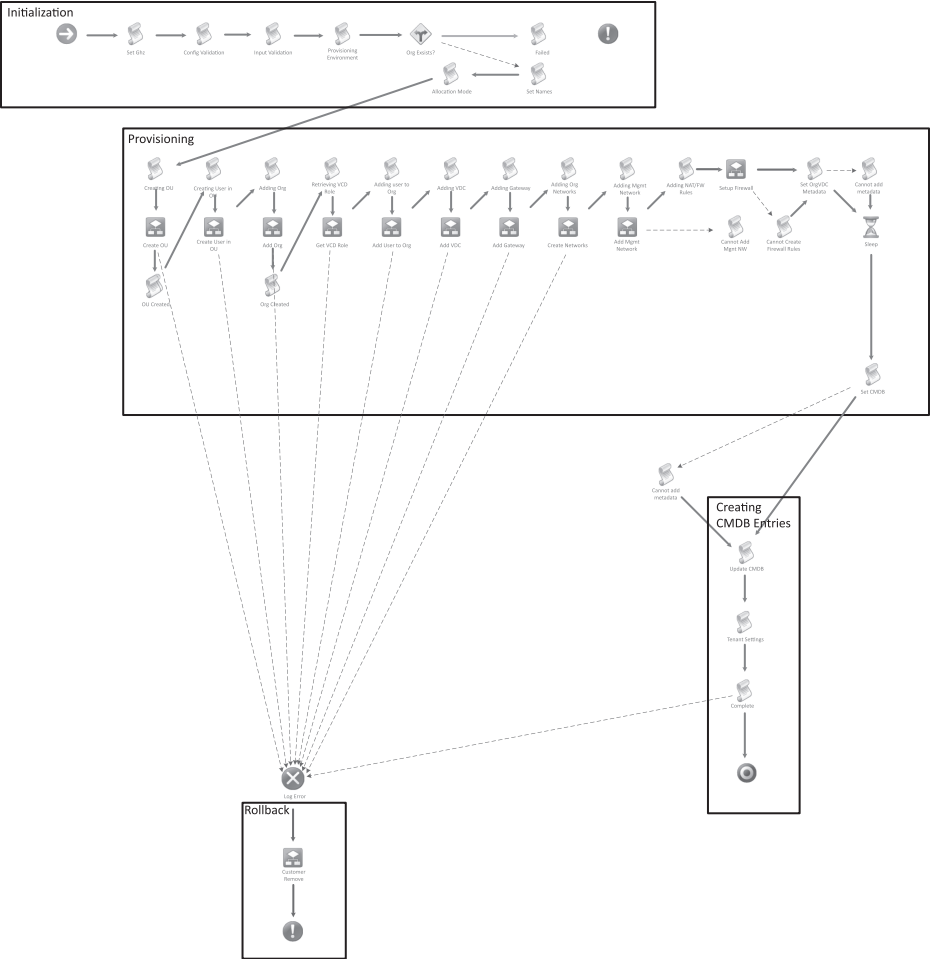


FIGURE 7.1 vRealize Orchestrator Initial Customer Provisioning Workflow

Another example would be the case of a tenant request for additional external networks. In this case, the cloud provider could automate the creation of the network, which would include name generation, allocation of an available VLAN and IP address range, configuration of the network switch and cloud perimeter firewall, creation of the external network in vCenter, and finally allocation of the external network to the customer’s organization.

Organization Administration Automation: There are also operational tasks within the customer’s organization that can benefit from automation. These are typically tasks that address the virtual machine lifecycle management, such as creation, configuration, routine maintenance, and decommissioning.

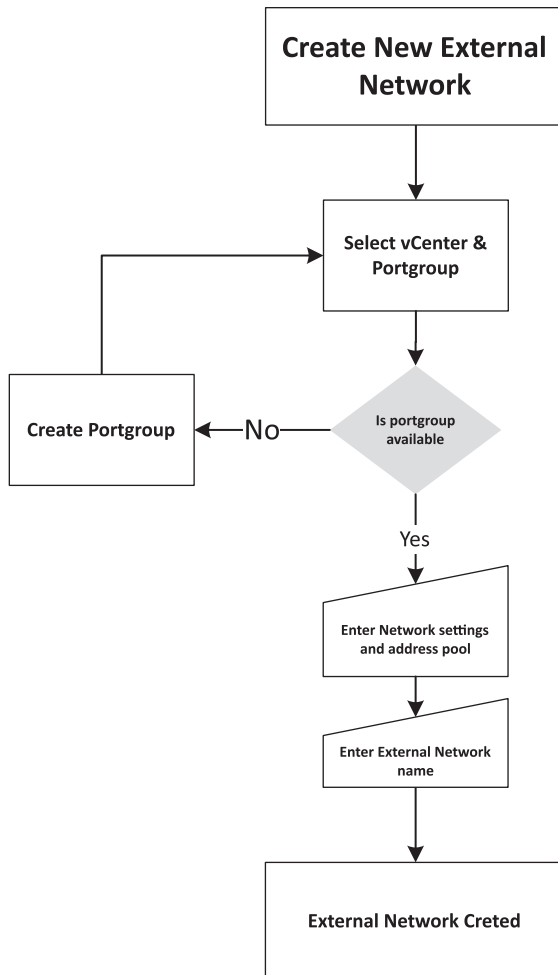


FIGURE 7.2 Request for Additional External Networks

For instance, consider the use case of virtual machine creation in an environment that uses Active Directory (AD) to identify services, such as authentication and printing. After deployment, the virtual machine must join the customer's Active Directory domain. For this, most IT AD design makes it mandatory to use an organization unit (OU) rather than the default Computers Container. vRealize Orchestrator could be used to create the virtual machine's computer account in the proper Active Directory OU, prior to virtual machine deployment, ensuring that the computer account name is unique and residing in the correct location in AD. In addition, when the virtual machine is decommissioned, the entry in the AD OU can be removed as part of the same workflow.

Cloud Consumer Operational Automation: These operations will fall into a category of tasks that the customer organization admin wants to offload as a self-service operation to developers or users. Performing the operation as a vRealize Orchestrator workflow provides an easy way to expose the task to a customer via the tenant user interface, addressed next with the Service Library, where workflows can be presented to and initiated by the customer, which leverages the Cloud Director and other API endpoints. Some examples of these types of use cases could include:

- Resetting of system or user account passwords on virtual machines using the VIX API plug-in
- Assign an IP Address from an IP Address Management tool
- Provision a tenant
- Update a CMDB
- Loading certificates into virtual machines
- Deploying instances of custom applications from the Bitnami catalog

Service Library

The Service Library feature links vRealize Orchestrator workflows to the Cloud Director tenant or provider user interface for Service Management and delivery. Effectively, this feature gives the cloud provider or customer the ability to execute vRealize Orchestrator workflows directly from within the Cloud Director portal.

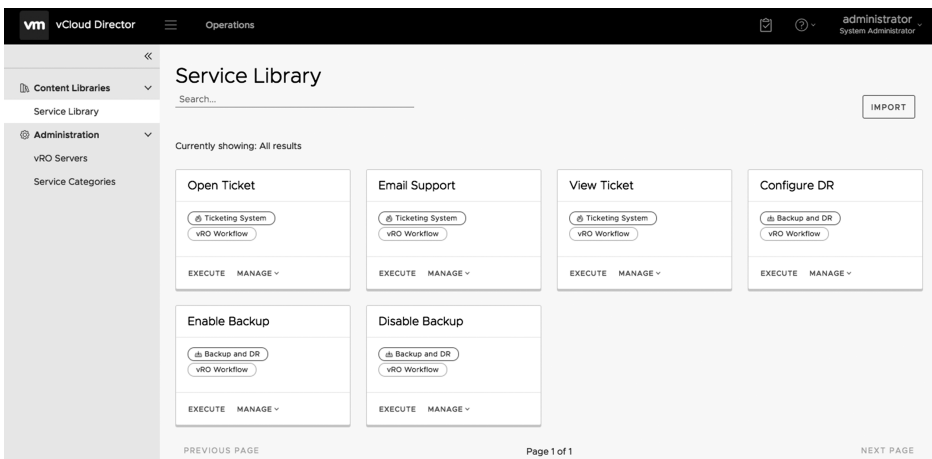


FIGURE 7.3 Service Library

This feature allows cloud providers to either leverage existing workflows or create new automation and selectively publish them to their customer Organizations. This Cloud Director and vRealize Orchestrator integration is

relatively relaxed, and effectively simplifies extensibility to publish workflows, configure parameters, which can be passed natively from vRO to VCD, that can then be executed from within the user interface.

Once the vRealize Orchestrator instances have been added to Cloud Director the cloud provider can create Service Categories, in order to organize new services. The provider is then able to import workflows from vRealize Orchestrator into its assigned Service Category. The next decision is who can access the service workflows. The Cloud Director user interface provides granular control to ensure that services are only published to the correct customers, providing the ability for the offerings to be tiered and monetized. Cloud providers can choose whether the workflows are published to only cloud provider admins, a subset of customers or everyone. In addition, Tenant Organizational Admins can control who in their organization has access to execute the services.

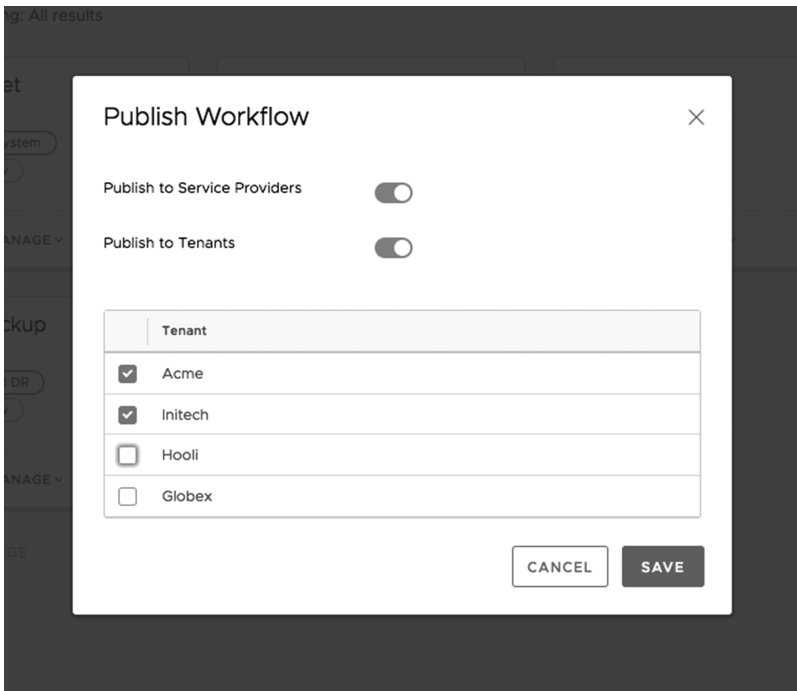


FIGURE 7.4 Access Control to the Service Workflow

A key feature of the Service Library is the ability to pass parameters from the Cloud Director user interface to vRealize Orchestrator, without which no customization of workflow objects or endpoints would be possible. For instance, consider a use case where a customer consuming an existing Org. VDC needs to assign additional storage resources for a new project.

This process could be fully automated by the cloud provider and made consumable to the customer via a self-service experience.

At a high level, the changes to the storage characteristics of the Org. VDC would require the Org. VDC to be updated. Therefore, the workflow would be required to collect several unique parameters, such as `inputOrganization`, `inputStorageProfile`, `inputVDC` and `VDC_StorageProfileLimit` and execute a number assembled operational elements:

- Extract the environment details, so that the Organization can be identified
- The customer enters the required parameters for extra storage resources
- Execute the request
- Update the cloud providers CMDB and billing DB of the change
- Send an email to the customer confirming the request has been completed

Figure 7.5 is an example of what this might look like in terms of vRealize Orchestrator workflow parameters and elements:

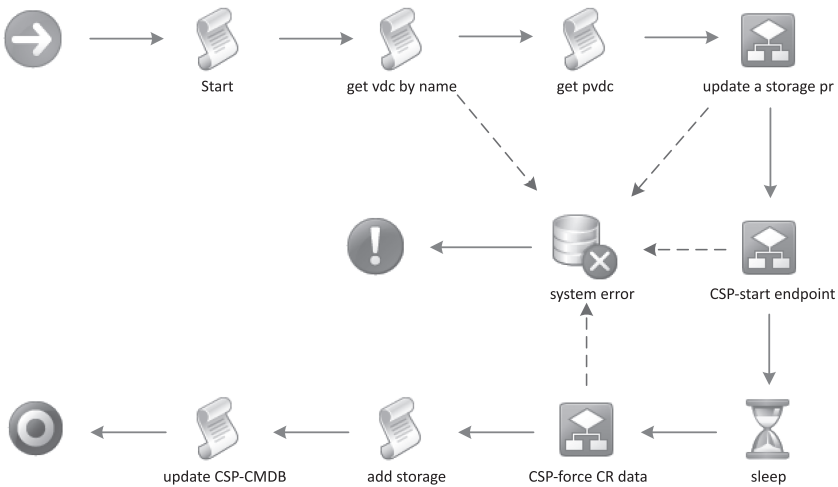


FIGURE 7.5 Change Org. VDC Storage Schema

This is a relatively basic example but demonstrates the power and flexibility that Service Libraries offer. With this and many other mechanisms to provide extensibility capabilities it has become increasingly attractive for cloud providers to expand Cloud Director from being solely based around delivering IaaS to offering a far more advanced portfolio of service offerings.

RabbitMQ

As we have seen, extensibility is an important characteristic of the Cloud Provider Platform. Giving cloud providers the ability to extend the cloud management

platform to external systems enables the end-to-end management of anticipated events that occur during the lifecycle of provisioned workloads.

Cloud Director *messages* provide the capability to connect Cloud Director with external systems by posting notifications or messages to an AMQP-based messaging broker, which can provide visibility to non-blocking and blocking notifications, allowing for end-to-end integration and is required for further extensibility of VCD, such as:

- Blocking tasks, notifications (CMDB updates, etc.)
- vCloud API Extensions (e.g., Cloud Director Availability)

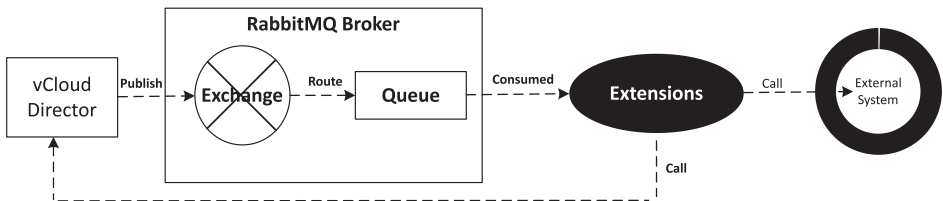


FIGURE 7.6 RabbitMQ Deployment Architecture

The RabbitMQ broker is typically deployed in a highly available architecture. In the Cloud Provider Platform model, the RabbitMQ broker is installed in an external standalone Linux virtual machine with the Pivotal RabbitMQ installation software. This is in most environments a clustered configuration with queue mirroring enabled, providing mirrored queues with SSL encryption. The AMQP service in Cloud Director points the AMQP host to a URL. A load-balanced virtual IP can be configured to distribute traffic across the nodes and provide a single address for external applications to consume AMQP messages from. If required, multiple (cross site) levels of RabbitMQ can be achieved with the ‘Shovel Plugin.’

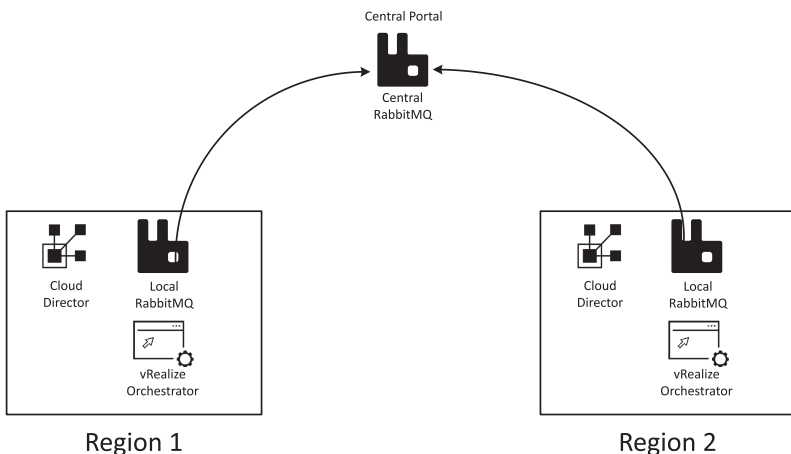


FIGURE 7.7 Distributed/Multi-Site RabbitMQ Architecture

vRealize Orchestrator acts as an extension configured to subscribe to the AMQP queue in order to receive its messages. When an appropriate message is sent, it will trigger a vRealize Orchestrator workflow. The workflow can optionally retrieve additional information about the blocked task via the Cloud API, process it, and then call the external system or again Cloud Director (via Cloud API), as illustrated in Figure 7.8.

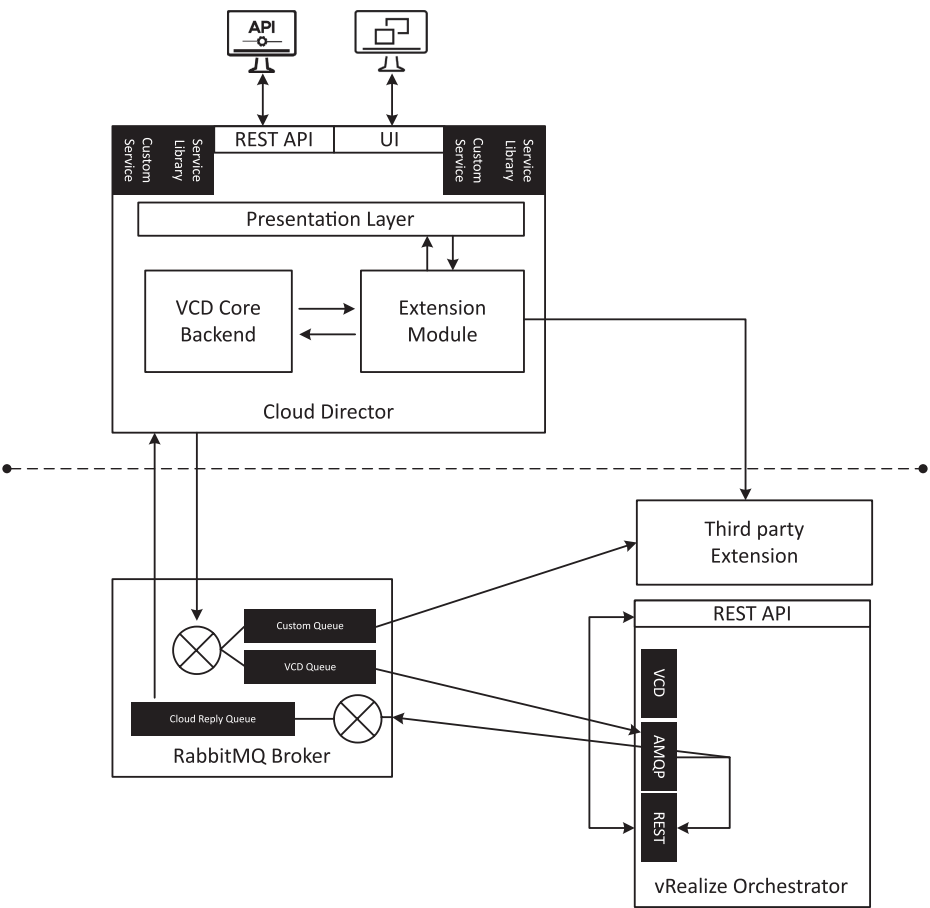


FIGURE 7.8 RMQ Message Architecture

Blocking Tasks Overview

The extensibility we are addressing in this chapter delivers many of the key benefits of the Cloud Provider Platform. This model, which gives cloud provider's the ability to deliver these additional services, allows them to go far beyond that of a traditional IaaS solution, and these mechanisms to expose these services to end consumers, are key from a service delivery perspective.

Therefore, cloud providers must design a solution that can meet the requirement of maintaining a consistent solution and ensure that events which occur at the Cloud Director level are synchronized accordingly within the cloud provider's business backend systems. For instance, if a customer deploys a new virtual machine, the provider might require a notification event that initiates the required logic to update a CMDB or billing system with the pertinent data. With the same backend logic being required for any system, such as enabling backup solutions or modifying an Org. VDCs storage resources.

The solution lays within the feature referred to as *Messages*, also sometimes referred to as *notifications*. This works as Cloud Director is capable of tracking events via the 'AMQP Message Bus,' described previously, for an external module to consume. This architecture is represented at a high-level in Figure 7.9. The logical flow of information illustrated demonstrates how Cloud Director passes the 'Message' to the AMQP Message Bus, which is then in turn picked up by the orchestration and logic layer provided by vRealize Orchestrator, which will take the appropriate action to update the back-office systems. In this example, the sequence of logic is triggered by a new virtual machine being deployed:

- Virtual Machine is deployed via the Cloud Director API
- Cloud Director passes the message on to the AMQP message bus that the VM has been created
- vRealize Orchestrator reads the message and updates the CMDB accordingly

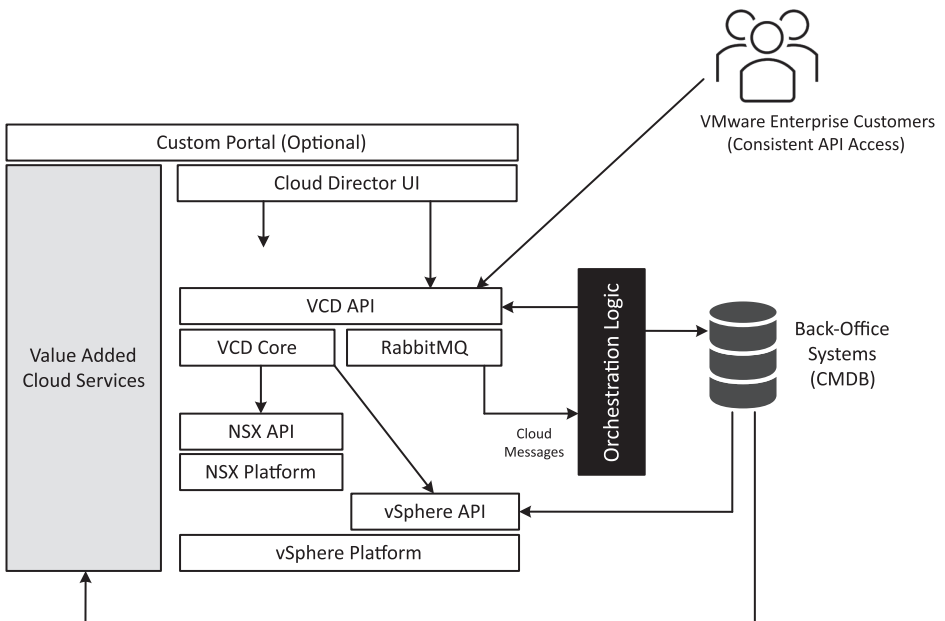


FIGURE 7.9 Blocking Task Orchestration Logic

As highlighted previously, vRealize Orchestrator can connect to multiple components and notifications can be generated in multiple places, allowing the overall solution the flexibility to go beyond simply updating database systems.

It is also worth considering, that while we are today talking about carrying out these functions via an orchestration tool, like vRealize Orchestrator, Serverless, or functions-as-a-service (FaaS) is an equally valid approach, which is something we will see far more of in these architectures going forward.

The required logic, whether provided via an orchestration layer of FaaS layer, might also be broken down in to two different parts, particularly in a complex cloud provider infrastructure, which might require interaction with several backend systems or third-party services. Figure 7.10 refers to these two systems as business logic and technical logic. Business logic is responsible for interacting with the cloud provider's back-office internal systems, while the technical logic is responsible for interacting with the customer facing and front-end component systems.

The key reason to think of these as two different areas of logic is that the internal back-office business logic is responsible for overall governance of the platform and will only typically interact with internal backend systems. However, the technical logic might interact with a wide range of VMware, third-party, and internal customer solutions and require a wide range of adaptors and connectors to perform the required actions. It is for such tasks that vRealize Orchestrator provides maximum benefit, with the wide range of plugins available. vRealize Orchestrator can provide the required interaction between multiple components and is able to extend the concept of providing additional custom cloud services to end consumers, exposed via a custom portal or directly from Cloud Director itself through the extensible user interface. The cloud provider must also consider the level of in-house development required, which can prove expensive to both initially develop and maintain.

vRealize Orchestrator has an extensive workflow library that contains pre-built tasks to solve specific challenges and perform common administrative actions. There are also hundreds of third-party plugins available at the VMware Solutions Exchange (<https://marketplace.vmware.com>). However, key vRO plugins for cloud providers include:

- vRO Plugin for Cloud Director
- vRO Plugin for NSX
- AMQP
- SSH
- PowerShell
- SQL
- Active Directory
- HTTP-Rest

However, where possible, the native functionality of Cloud Director should always be employed over vRealize Orchestrator workflow integration, as this provides for a more scalable, closely integrated solution and simplified architecture.

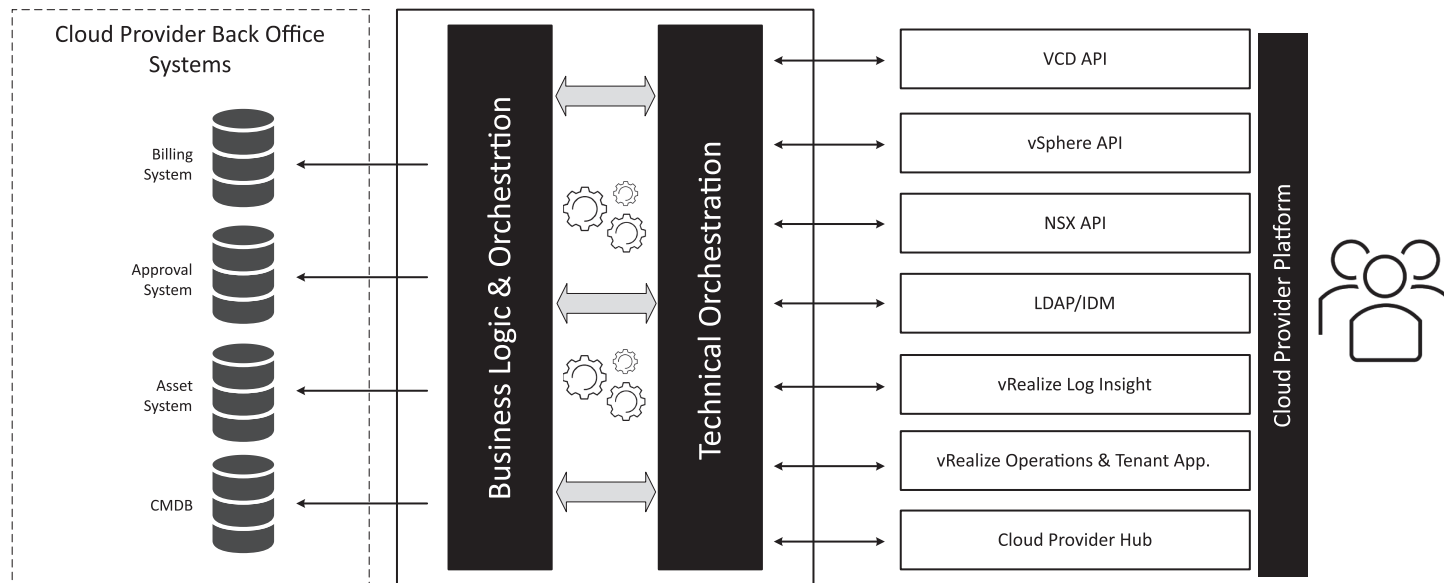


FIGURE 7.10 Separating Business Logic and Technical Logic for Improved Scalability

Terraform

VMware Cloud Director can act as a Terraform provider, and can interpret infrastructure as code, allowing customers to define and provision a data center infrastructure and application workload architecture using a high-level configuration language known as Hashicorp Configuration Language, or alternatively JSON.

Terraform can be used to create, manage, and update infrastructure resources such as Org, VDCs, virtual machines network switches, containers, and more on Cloud Director. Almost any infrastructure type can be represented as a resource in Terraform. This allows developers and DevOps engineers to automate the build, change, and versioning of their Cloud Director infrastructure as code.

The Cloud Director provider is responsible for understanding API interactions and exposing resources, such as Organizations, catalog items, networks, load balancers users etc. This approach can ensure that customers can maintain consistency across staging environments and automate deployments of multi-tier workloads, either virtual machines or containers.

Terraform also supports a number of other cloud infrastructure providers such as Amazon Web Services, IBM Cloud, Google Cloud Platform, Microsoft Azure, Oracle Cloud Infrastructure, OVH, as well as OpenStack.

Figures 8.11 and 8.12 demonstrate the capabilities of Cloud Director as a Terraform provider. The first shows how a cloud provider can automate the creation of a new VCD tenant through the use of the Terraform Provider. This code example illustrates the automated creation of VCD objects such as Organization, Org, VDC, and Edge Gateway

```
# Create Organization
resource "vcd_org" "acme_org" {
  name           = "acme_org"
  full_name      = "ACME Organization"
  is_enabled     = "true"
  stored_vm_quota = 100
  deployed_vm_quota = 100
}

#Create Org VDC
resource "vcd_org_vdc" "acme_vdc" {
  org = "${vcd_org.acme_org.name}"

  name           = "acme-vdc"
  description    = "ACME Allocation Model"

  allocation_model = "AllocationVApp"
  network_pool_name = "vc1-network-pool-VXLAN-01"
  provider_vdc_name = "vc1-PVDC-Cluster-01"

  network_quota = 5
}

# Create Edge Gateway
resource "vcd_edgeway" "acme_gw" {
  org = "${vcd_org.acme_org.name}"

  name           = "acme-gw"
  description    = "ACME Edge Gateway"
  configuration   = "compact"

  default_gateway_network = "${vcd_external_network.acme_ext_net.name}"
  external_networks      = ["${vcd_external_network.acme_ext_net.name}"]
  advanced              = true

  lb_enabled = true
}
```

FIGURE 7.11 Service Provider Onboarding a Tenant

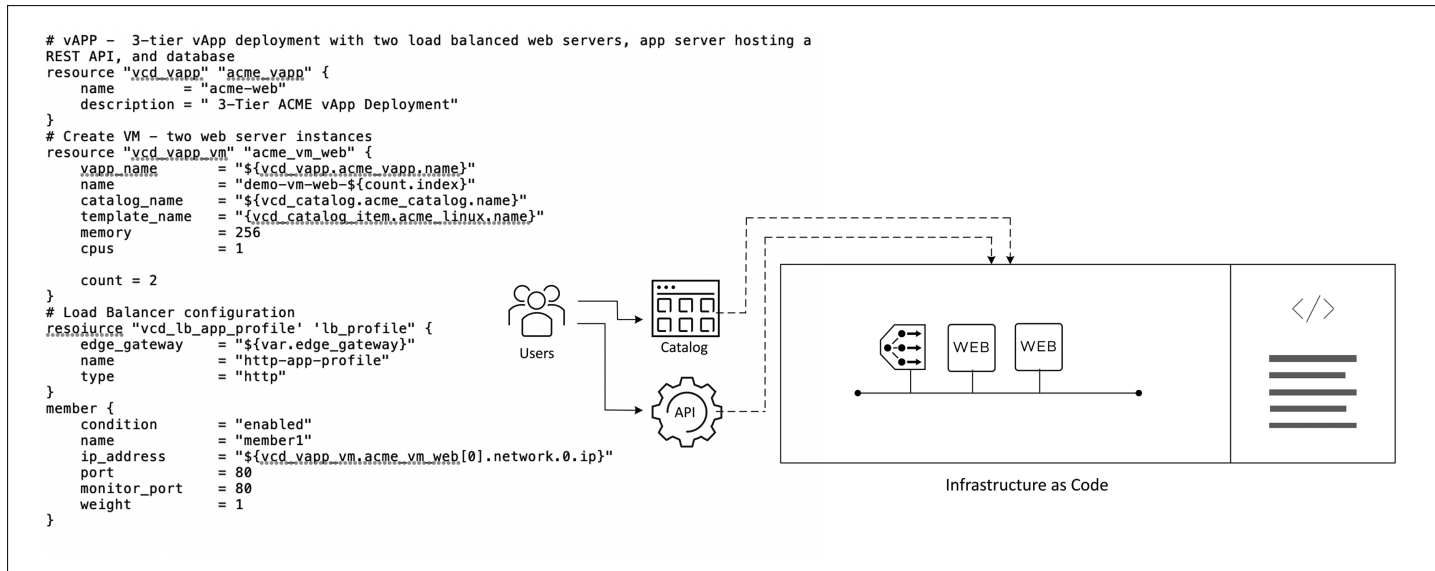


FIGURE 7.12 Terraform Application Building

The second example demonstrates how a customer could automate the deployment of a multi-tier application onto a Cloud Director Org. VDC, as illustrated in Figure 7.12. This code example shows how a customer could automate the deployment of their web app, with two load-balanced web servers.

VMWARE CLOUD FOUNDATION (VCF)

For most cloud service providers, other than those with the smallest Cloud Provider Platform deployments, VMware Cloud Foundation is recommended to simplify the deployment, operations and lifecycle management, while increasing administrator productivity and reducing overall total cost of ownership.

VMware Cloud Foundation standardizes the software-defined data center for the Cloud Provider Platform. It delivers a natively integrated software stack that can be used for shared public cloud services or dedicated private cloud solutions using the Central Point of Management (CPOM) feature to be addressed in Chapter 8, with consistent and simple operations. For the Cloud Provider Platform, VCF delivers and maintains the core components VMware vSphere, VMware vSAN, and VMware NSX Data Center.

VMware Cloud Foundation essentially makes it easier to provision complete cloud infrastructure, reducing complexity, time and cost associated with manually building the cloud resources. In a similar way that you can quickly provision workloads with Cloud Director; VCF allows cloud providers to build a full SDDC with the same level of simplicity. For this reason, Cloud Foundation provides a simple path to the Cloud Provider Platform, delivering a fully integrated SDDC, ready for consumption as resources for Cloud Director. For the cloud provider this can mean:

1. Automated Day-0 deployment and quicker time to market by helping the provider stand up SDDCs at scale, while minimizing the complexities associated with system design, testing, configuration, and provisioning
2. Upgrade, patch, and scale the SDDC infrastructure in a fully orchestrated way (lifecycle automation)
3. Reduced risk associated with new SDDC deployments, such as misconfiguration, as VCF is based on VMware Validated Designs. This means the solutions have been pre-tested and validated to help ensure success.
4. A lower TCO
5. Integrated enhanced security of a public cloud or single tenant physical infrastructure
6. VCF can be used to maintain a standardized global data center footprint of extensive cloud services by delivering VMware SDDCs in a fully automated building block manor

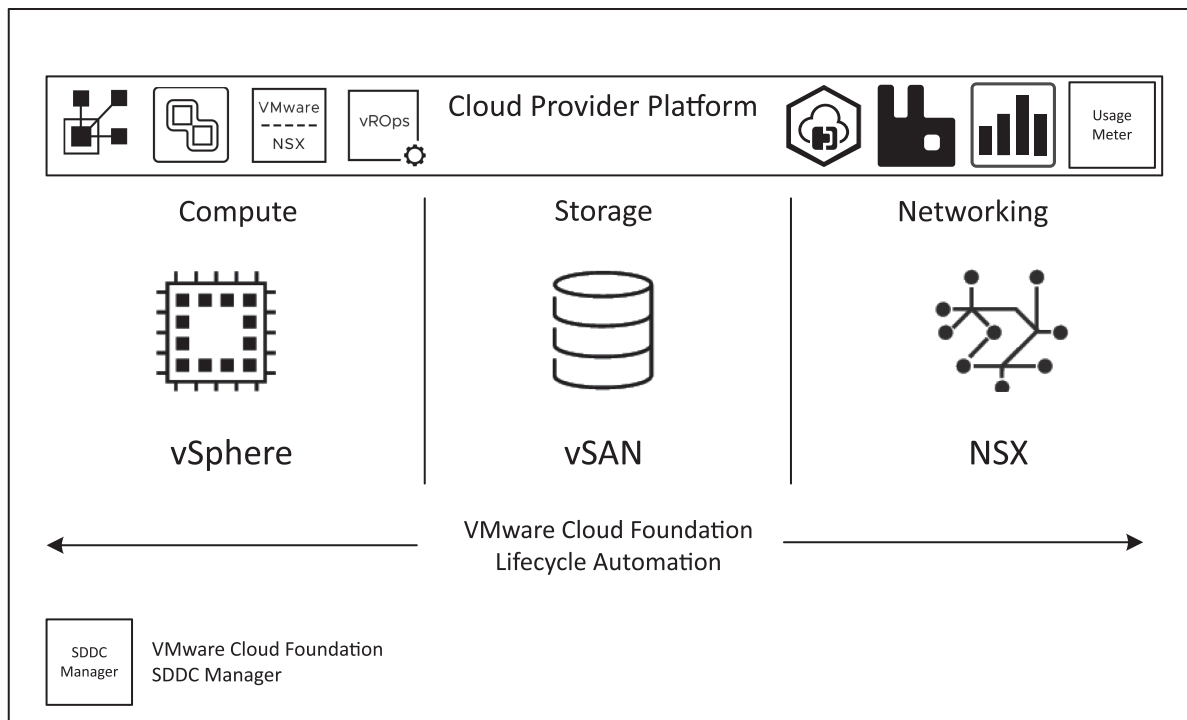


FIGURE 7.13 Software-Defined Cloud Provider Platform

The Cloud Foundation platform is controlled by SDDC Manager. SDDC manager is responsible for bringing up the Cloud Foundations system and allows the cloud provider to quickly and efficiently install vSphere, vSAN, and NSX Data Center. SDDC Manager also completes all configuration tasks and handles lifecycle management by patching and upgrading all the components from one central point of management.

VMware Cloud Foundation has the concept of Workload Domains. A Workload Domain is a logical abstraction of cloud capacity, which can automatically be provisioned and maintained by the SDDC Manager. The Workload Domain effectively provides a unit of cloud compute, storage, and network resources ready to be consumed by workloads. The SDDC Manager provides tooling that allows or the automation and expansion of an existing Workload Domain, as well as their deletion. The concept of workload domains also allows for each one to be upgraded separately, meaning each workload domain can be upgraded independently from one another.

Within the Cloud Provider Platform framework, a workload domain would be abstracted as a Resource Group. As illustrated in Figure 7.14, each Workload Domain has its own vCenter Server, which is operating multiple vSphere clusters (Resource Clusters).

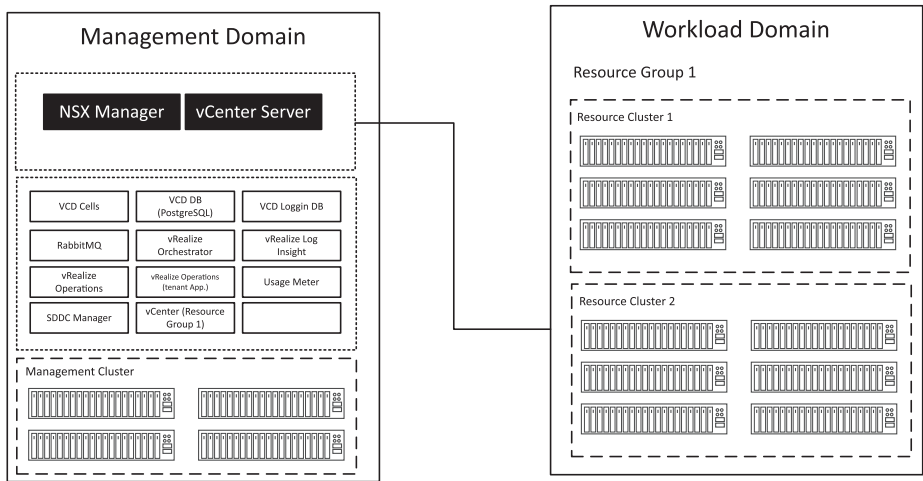


FIGURE 7.14 VMware Cloud Foundation Workload Domains

The two types of Workload Domains shown in Figure 7.14, a Management Workload Domain and a Virtual Infrastructure (VI) Workload Domain. The Virtual Infrastructure Workload Domains are dedicated to hosting customer workloads and mapped to resources consumed by Cloud Director as Provider VDCs.

The maximum number of Workload Domains per SDDC Manager is defined by the maximum number of vCenter Server instances that can be configured in enhanced linked mode.

A Management Domain is primarily dedicated to hosting Cloud Foundation management components, such as SDDC Manager, as well as those components related to VMware Virtual Infrastructure Workload Domains, such as:

- Platform Services Controller Instances
- vCenter Server Instances
- NSX Manager Instances

The Management Domain can, in smaller implementations, also house the management components of the Cloud Provider Platform, such as Cloud Director cells, vRealize Operations Manager, Tenant App and other required services. For large-scale deployments, an out-of-band Cloud Management Cluster should be configured independently from VCF.

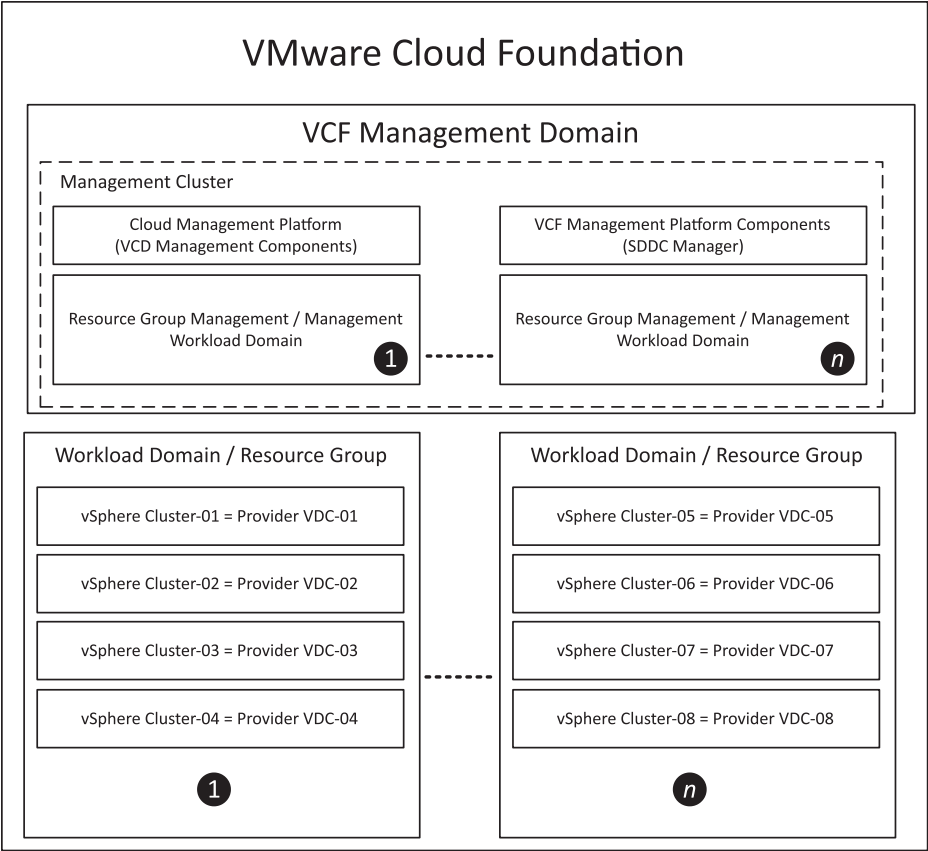
Designing a VMware Cloud Foundation with Cloud Director Solution

As outlined in the previous section, there are huge benefits in the standardized VMware Cloud Foundation approach for resources surfaced to Cloud Director.

The aim of the Cloud Provider Platform is to standardize the approach for delivering enterprise ready public cloud platforms. The use of VCF significantly simplifies day-0 deployment, lifecycle management and forces a cloud provider to take a standardized approach to their platform design by delivering the underlying SDDC platform stack, vSphere, vSAN and NSX, in an accelerated manor for Cloud Director to consume.

Before we progress, let's take a moment to address terminology between Cloud Foundation and Cloud Director, which can sometimes cause confusion. A Cloud Director Resource Group and the Cloud Foundation Virtual Infrastructure Workload Domain have multiple common characteristics. Both constructs describe a group of vSphere compute, storage and network resources, as well as their required management components. As illustrated in Figure 7.15, for all intents and purposes, for the Cloud Provider Platform delivery model, these two constructs map one-to-one.

Based on this architecture, Virtual Infrastructure Workload Domains can be used in the standard Cloud Director architectural model to represent Resource Groups, for the hosting of tenant workloads. The Workload Domain must consist, at a minimum, of one cluster consisting of three nodes, which can be scaled incrementally through SDDC Manager. Likewise, additional clusters can be added to the VI Workload Domain until maximums are achieved, before a second VI Workload Domain can be automatically provisioned onto new physical resources, to represent an additional Resource Group for Cloud Director to consume in the same way. Based on this, most cloud providers will elect to populate a VI Workload Domain to map directly to one or more Provider VDC construct. The Elastic Provider VDC construct can be used to



Single Availability Zone / Region

FIGURE 7.15 VI Workload Domain to Resource Group Mapping

provide a single Provider VDC, which consumes resources from more than one resource pool, represented as a vSphere Cluster in this architecture. Alternatively, each Resource Cluster can be configured as a single Provider VDC. The key to the correct approach will be dictated by whether the cloud provider is taking a *scale-out* or *scale-up* approach to their cluster deployment model, discussed in Chapter 3.

CLOUD PROVIDER POD

Cloud Provider Pod is predominantly designed to enable small cloud providers to deploy a standardized stack quickly and efficiently in an automated manor. The Cloud Provider Pod delivers several key things:

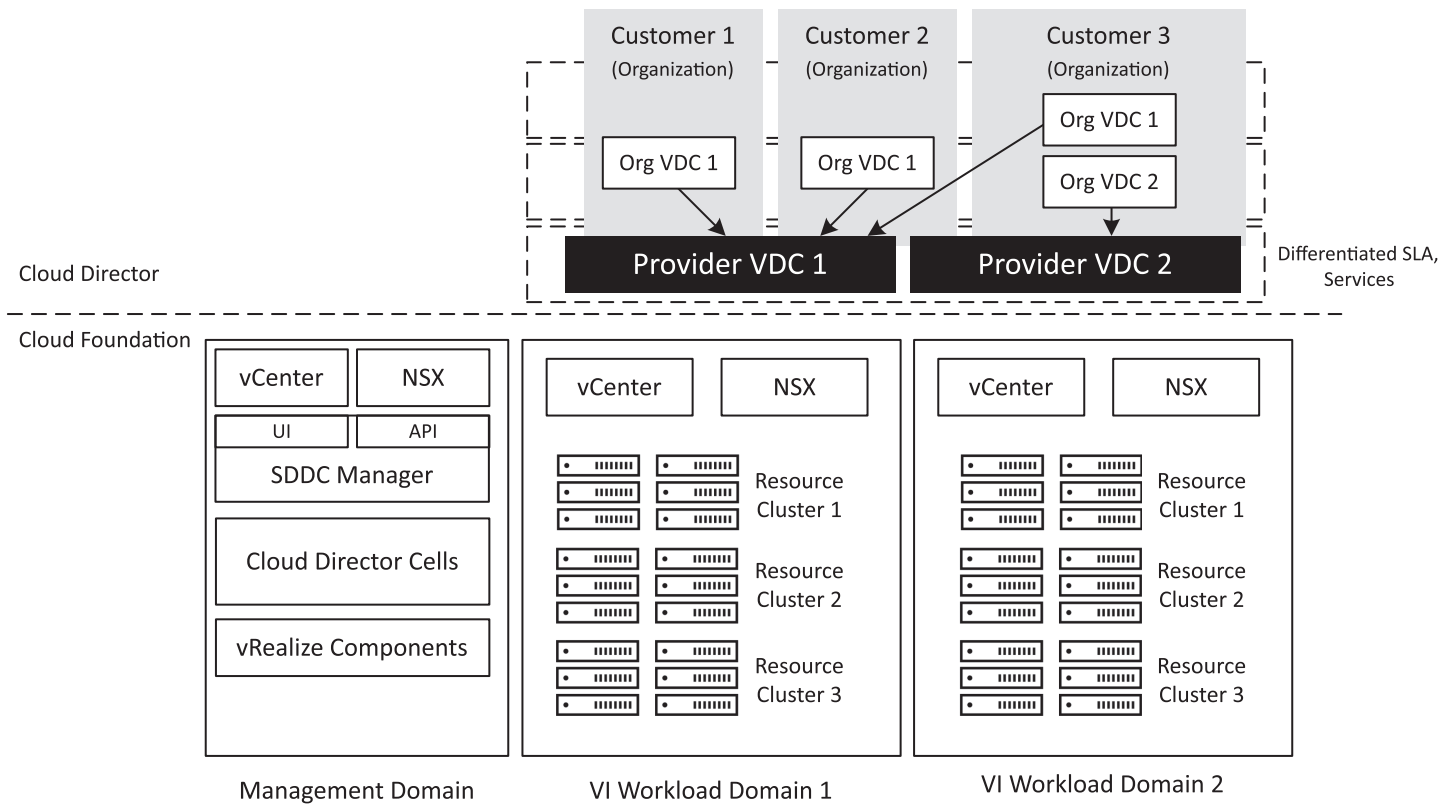


FIGURE 7.16 Cloud Director and Cloud Foundation Architecture

- It generates detailed design and operations guidance based on VMware Validated Designs for Cloud Providers
- It deploys a custom-built stack, based on the parameters provided, in a matter of hours
- It deploys a fully tested and interoperable solution, performance-tested and scale-tested from the custom inputs provided
- It brings standardization to deploying a small cloud provider environment

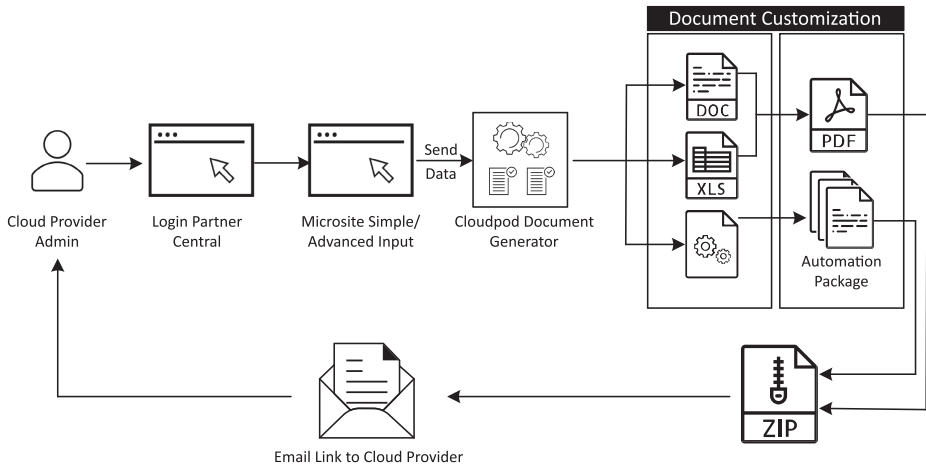


FIGURE 7.17 Cloud Provider Pod Workflow

For a cloud provider who wants to utilize Cloud Provider Pod, the cloud provider goes to the Cloud Provider Pod Designer and enters the design requirements, both simple and advanced options are available, depending on the requirements. When complete, the Cloud Provider Pod generates a package, including the document guidance on design and operations adhering, to VMware Validated Designs, and the core automation bits required to build the environment.

Once the cloud provider downloads their package, this is simply imported into the Cloud Provider Pod Deployer OVA. From there, the automated process can begin, from preparing the management cluster, deploying management cluster resources and deploying the resource cluster, the process is automated.

When comparing with the VMware Cloud Foundation architecture, previously discussed, Cloud Provider Pod offers no lifecycle management capabilities and does not address the level of scale required by many of VMware's Cloud Providers. For these reasons, Cloud Provider Pod is today targeted at small providers who wish to take advantage of the key benefits outlined above.

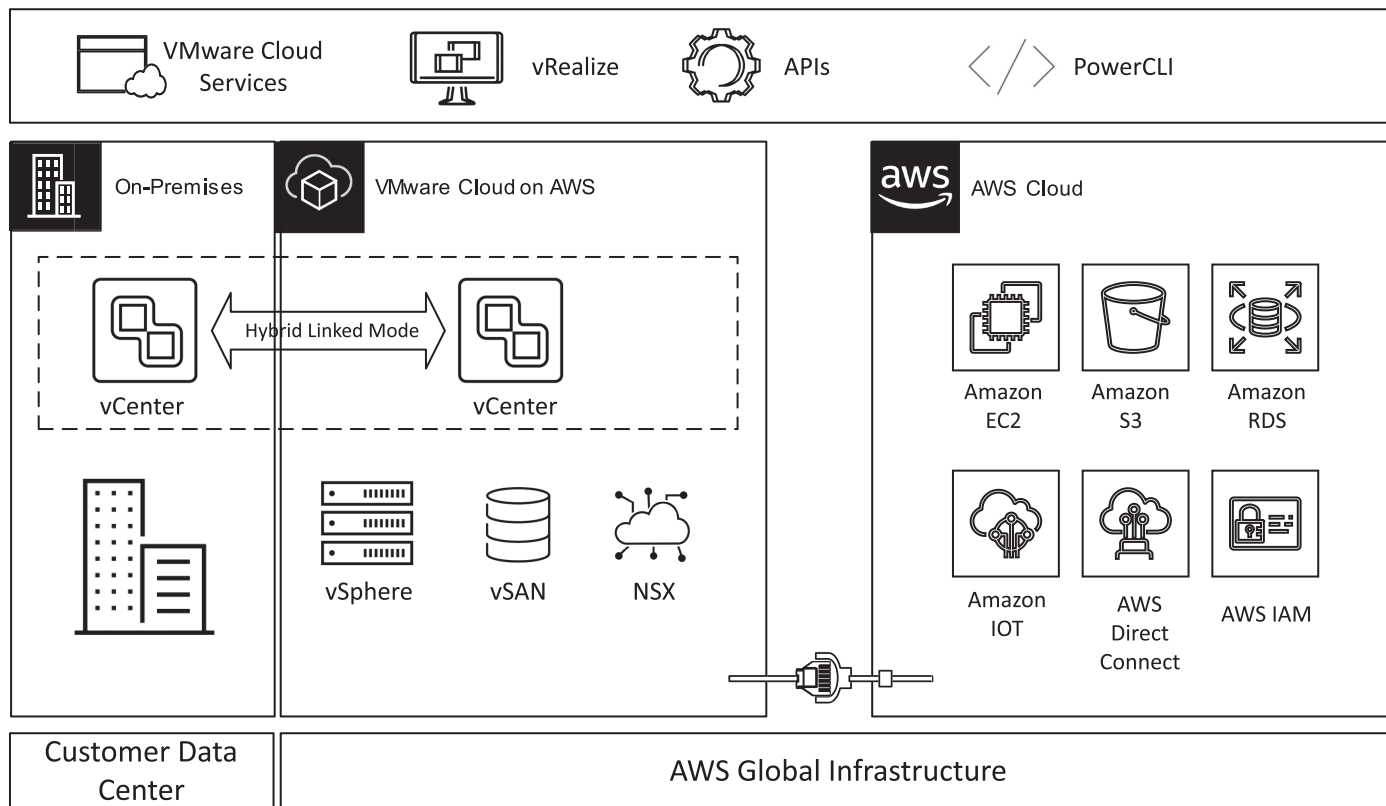


FIGURE 7.18 Anatomy of VMware Cloud on AWS

VMWARE CLOUD ON AWS (MSP)

VMware Cloud on AWS is a cloud offering that is managed, operated, and supported by VMware. The service provides dedicated Software Defined Data Center instances, which are hosted in AWS facilities with hardware which is owned and managed by Amazon. The service is dedicated, with hardware that is solely assigned to that SDDC, and delivers a private cloud operating experience with high speed low-latency access to AWS native services, offered from the customer-owned AWS account, which retains the same billing model, directly from AWS.

By its very nature, public cloud has some inherent operating difficulties for organizations large and small. VMware Cloud on AWS has been developed to deal with the challenges of hybrid cloud.

When it comes to hybrid architectures, customers typically have requirements around operational consistency, they want to be able to implement and automate operations and processes in a similar fashion in both their on-premises data centers and cloud platforms. Likewise, they want to be able to transition existing skillsets across both solutions, providing simplicity around the management of both their own assets and those that reside in the cloud with common service-level agreements. The aim is to have a level of compatibility and interoperability with applications, so they don't have to create and configure them differently, depending on where they reside.

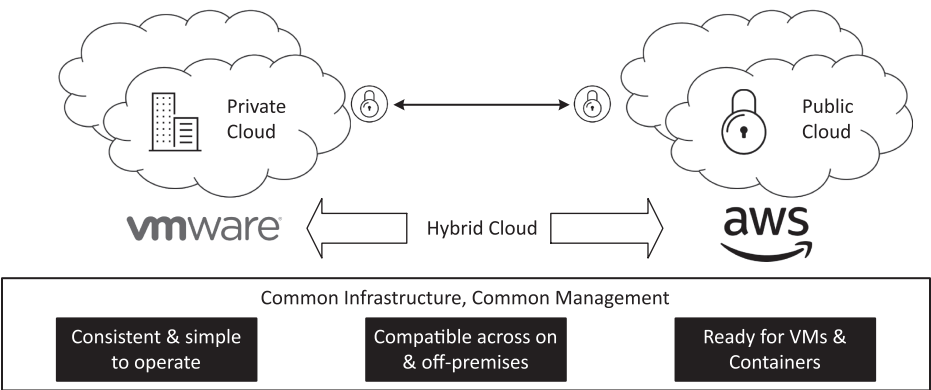


FIGURE 7.19 A Hybrid Cloud that Is Based on a Common Platform

VMware Cloud on AWS is designed specifically to address these challenges. One of the key benefits for customers is that they do not have to deal with these types of multi-cloud infrastructure provisioning and management challenges any longer, it's just part of the service. In addition, as the VMware service is built on top of AWS on-demand billing, and is therefore based on

actual consumption, it allows VMware to offer flexible consumption models, with elastic resources that can be scaled on demand, making VMware Cloud on AWS a truly unique offering in the industry.

Simple and Consistent Operations	Enterprise Grade Capabilities	Flexible and Innovative
• Operational consistency with vCenter, APIs, and third-party tools • 100% compatibility for existing and new enterprise workloads using VMware • Delivered as a service, removing management hassles	• Enterprise app SLAs for availability and performance • Enterprise security with microsegmentation and encryption • Optimized use of underlying physical resources	• Access to a broad range of AWS services • Enable flexible consumption with elastic resources, on-demand • Deliver bidirectional workload portability between VMware and AWS

FIGURE 7.20 VMware Cloud on AWS Service Differentiators

VMware Cloud on AWS brings VMware's enterprise-class SDDC software to the AWS cloud. Powered by VMware Cloud Foundation, this service enables businesses to run any application across vSphere-based private, public, and hybrid cloud environments. One of the unique points of VMware Cloud on AWS is that it allows businesses to leverage the global footprint and breadth of services of AWS, including storage, databases, analytics, and more. This service is delivered, sold, and supported by VMware, channel partners and Managed Service Providers (MSP) as an on-demand, elastically scalable service.

As highlighted above, VMware Cloud on AWS can be consumed and offered as a service through Managed Service Providers. The Managed Service Provider (MSP) model, under the VMware Partner Connect Program, allows partners to leverage various VMware Cloud Services to broaden their portfolio of offerings and provide value added features, functionality to their customer focused managed services. VMware Cloud Services consists of several services that allow for the creation, management and monitoring of public and private cloud resources, including VMware Cloud on AWS.

All VMware Cloud on AWS is exposed to cloud providers via the MSP program for commercial constructs with commitment-based contracts. When a VMware partner signs up to a commitment for services they will be on-boarded to the MSP platform, a service called Cloud Provider Hub. The Cloud Provider Hub is an integrated system which allows the cloud

provider to manage their VMware Cloud on AWS tenants. Cloud Provider Hub functionality will be described in more detail in Chapter 8.

VMWARE CLOUD DIRECTOR SERVICE (VCDS)

As many readers will be aware, VMware are offering more and more products through a SaaS delivery model, alongside the traditional product deployment models. The SaaS delivery model for Cloud Director benefits from continuous access to new features and improvements without the need for patching and upgrades. However, the main benefits are available to managed service providers that are developing an asset light approach to service delivery. This is where a Managed Service Provider's business model sees them own relatively few capital assets, such as servers, storage and data centers. These types of Managed Service Providers target customers based on its operational and service delivery model, as opposed to its ability to provide data center facilities, server and storage infrastructure at scale, the model used by the majority of traditional asset heavy cloud providers.

For this reason, the initial go to market for VCDS is aligned with delivering multi-tenancy on VMware Cloud on AWS, with out of the box integration with that platform.

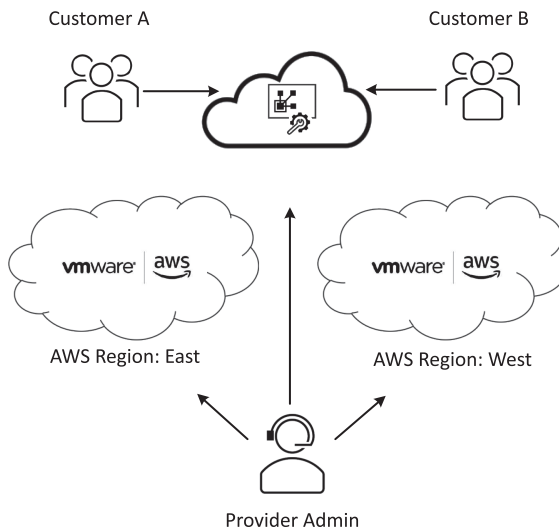


FIGURE 7.21 Multi-tenancy for VMware Cloud on AWS

However, we will quickly see VCDS evolve to support the full range of cloud endpoints and delivering the same capabilities available today from the partner deployed Cloud Provider Platform, illustrated in Figure 7.22.

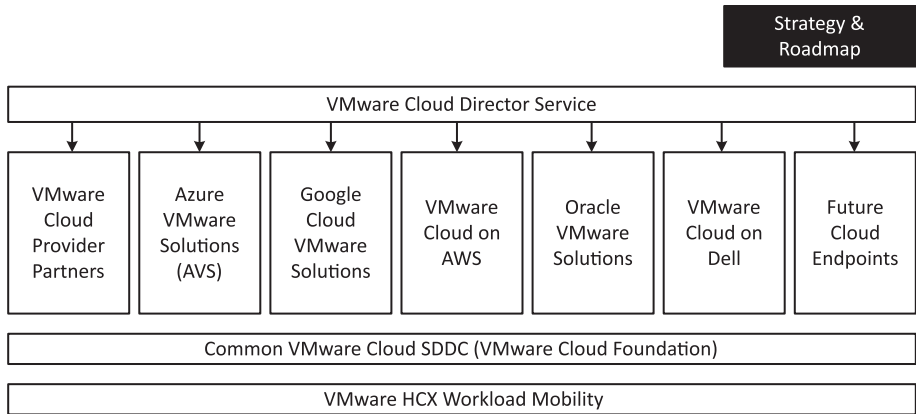


FIGURE 7.22 Multi Cloud Service Vision

VMWARE MAAS SOLUTION PARTNERS

Extending or migrating an on-premises infrastructure to the public cloud can be time consuming, costly and fraught with complex cloud migration, implementation processes and challenges. The customers must contend with differences in skill sets and tools required to manage these environments, typically leading to cloud silos. In addition, separate access, security and networking policies, leading to inconsistencies in control, and difficulty delivering enterprise application service-level agreements with consistent availability and performance across on-premises and public clouds.

VMware has committed to resolving these challenges and creating a public cloud experience that is simply about consumption, and makes cloud complexity invisible to the customer, allowing them more time to focus on transforming their business. One of the ways VMware is doing this is through the MaaS offerings.

MaaS stands for *Metal as a Service* and it enables specific partners to offer SDDC as a Service, based on vSphere, NSX Data Center and vSAN, on public cloud. MaaS is a supplemental offering which comes under the VMware Partner Connect subscription model and is the mechanism under which VMware offer solutions with global hyperscale public cloud providers, including:

1. Azure VMware Solutions (AVS)
2. Google Cloud VMware Solutions
3. Oracle Cloud VMware Solutions
4. AliCloud VMware Solutions

These solutions offer VMware SDDCs running on hyperscale public clouds, similar to what VMware has done with the VMware Cloud on

AWS solution, however, there are some key differences between these offerings:

Azure VMware Solutions	VMware Cloud on AWS
Sold by Microsoft and Azure Cloud Solution Providers	Sold by VMware, directly and via channels
Delivered and managed by VMware partners and supported by Microsoft and VMware partners	Delivered, supported, and managed as a service by VMware
Engineered by MaaS Authorized VMware partners in collaboration with Microsoft	Jointly engineered by VMware and AWS, with the roadmap owned by VMware

These MasS public cloud offerings require the operator to maintain the environment with the latest VMware codebase, with a product bundle referred to as *VMware Cloud Provider Stack* (VCPS). These hyperscale cloud providers must use a VCPS release bundle and cannot use individual generally available products. VCPS will only support 2 major releases at any time and deployed SDDC instances cannot be more than 1 release behind the most current major release of VCPS. Therefore, automation of lifecycle management of VCPS bundles is an essential component required of the VMware partners operating these solutions, alongside bring-up and autoscaling automation.

These VMware on public cloud solutions make Hybrid Application Design a reality. Hybrid Application Design describes an application topology which spans both a VMware Cloud platform and native public services, with a unique solution that delivers a 2-way high-bandwidth low-latency interconnect with native services for hybrid application architectures.

This allows customers to gain the benefits of native cloud services by incorporating them into application designs—in other words, binding VMware IaaS with public cloud services on a per workload or application basis. For instance, as illustrated in Figure 7.24, Oracle RAC running on a VMware Cloud platform with front-end web service consumption in Azure. Low latency, high-bandwidth, zero cost connectivity provides efficiency and an improved user performance.

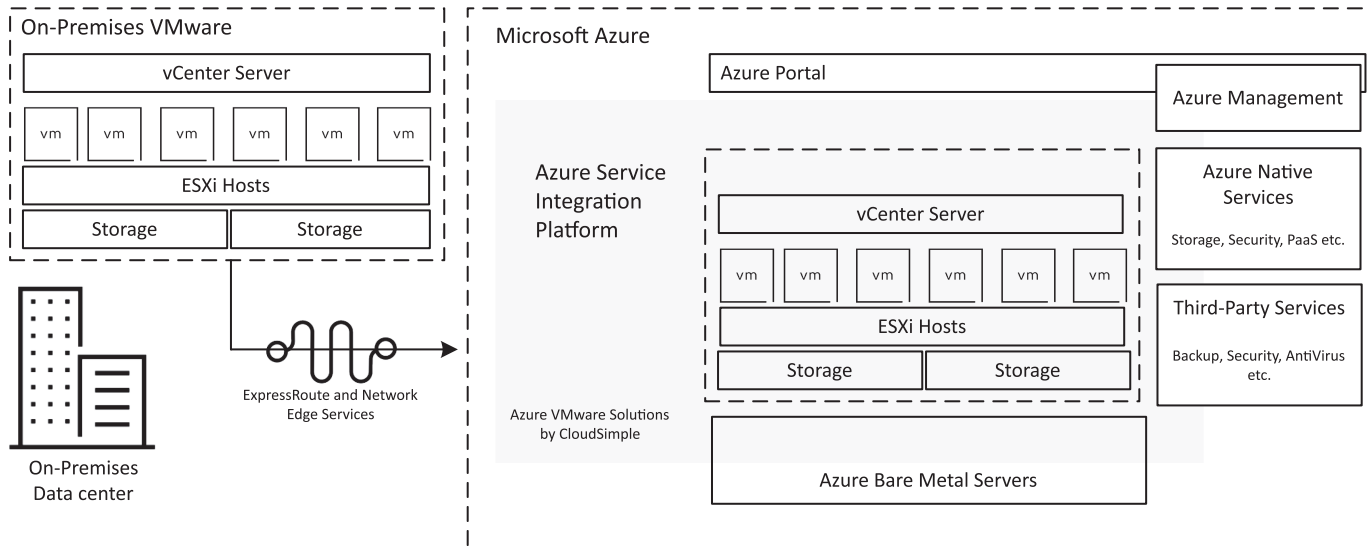


FIGURE 7.23 Azure VMware Solutions

Chapter 8

Building and Delivering Services with Extensibility

The traditional role of cloud providers simply making IaaS accessible to tenants has evolved into a competitive landscape where providers must be able to extend their portfolio to deliver PaaS, SaaS and other advanced service offerings to stay relevant. Cloud providers need to be able to do that in a way that is aware of multi-tenancy, and fully capable of tenant isolation, and they must be able to provide consistency in the experience that they are delivering to their customers.

Extensibility refers to the Cloud Provider Platform's ability to extend into additional services, that are not available natively. These 'Extensions' allow cloud providers to add new functionality into the platform, including the Cloud Director tenant and provider portals. This approach allows cloud providers to:

- **Differentiate:** Create innovative and differentiated service offerings or industry relevant service offerings to cater to the tenants' increasing cloud demands
- **Increase Revenue:** Create new revenue sources, create new value add services, expand wallet share with new business models and unlock new customer segments and increase profitability by integrating new solutions
- **Enhance Customer Experience:** Simplify service consumption, service creation, enable seamless workload migration, seamless management across clouds
- **Operational Efficiency:** Increase profit margins through process automation, tool consolidation, platform standardization, reference designs and implementation examples
- **Integrate:** With ISVs of your choice, leveraging a standard framework

As we addressed in the previous Chapter, cloud providers can automate workflows through vRealize Orchestrator (vRO) and natively integrate into Cloud Director, allowing vRO to deliver complex automated workflows. This enables cloud providers to set up a variety of different services that can be deployed in a 1-click fashion, with as much granularity and visibility for the tenant or provider admin as required. As vRO offers virtual limitless possibilities, this functionality can be considered like a swiss army knife for service automation and integration, with minimal manual development effort required. This has a myriad of benefits, including OpEx reduction, unlocking new customer segments with a more diverse portfolio, and benefits for ISVs too, who can drive business growth by integrating with VCD through vRO.

However, custom user interface and API extensibility takes this model to the next level with an SSO ready extensible framework for service delivery, allowing deeper integration with solutions for Backup, Disaster Recovery or other ISV monetizable solutions through a standardized custom cloud provider service.

This extensible framework means any ISVs or third-party cloud software provider can integrate their services onto the Cloud Provider Platform environment through an extensible framework that is SSO-ready. Naturally, this helps cloud providers offer a variety of valued-added services with standardized integration; thereby, increasing revenue opportunities. Illustrated in Figure 8.1 is one such example, of the native integration of the Rubrik Extension for VMware Cloud Director.

This Rubrik integration brings data protection services into Cloud Director and allows Organizational tenant administrators to manage backup, recovery, and replication operations, which apply policies at the Org. or workload levels, and importantly, provide these services to tenants in a self-service manner.

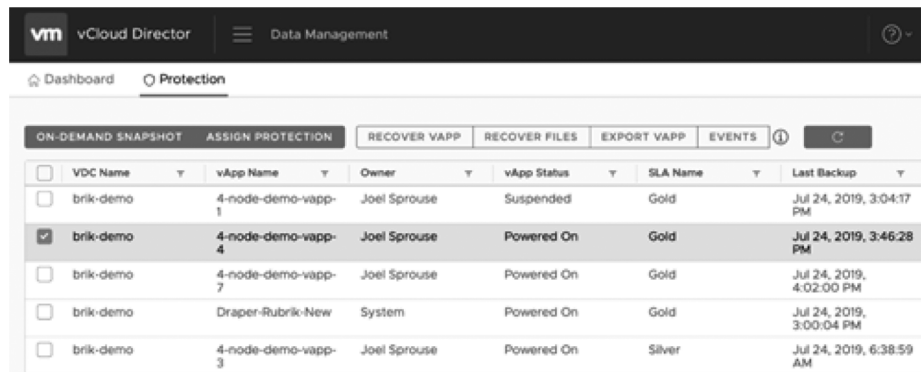


FIGURE 8.1 Rubrik User Interface Extension

The extensibility service framework allows multiple extensions from different sources to be combined with the service offering. For instance, as illustrated in Figure 8.2, extensions might be VMware written and supported, such

as the tenant and provider user interface extension for vRealize Operations or the tenant and provider user interface for Cloud Director Availability. These VMware extensions can be used in conjunction with extensions from third-party ISVs (such as Rubrik) as well as custom written extensions, built and supported by the cloud provider themselves.

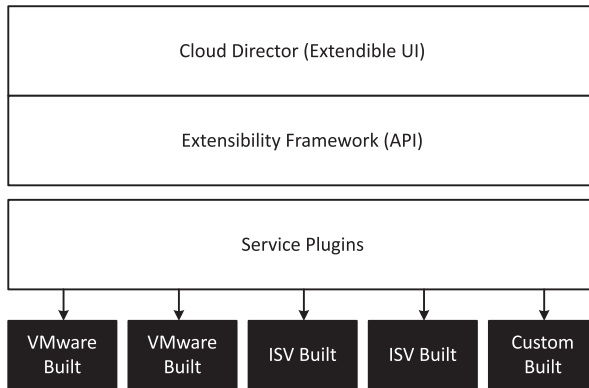


FIGURE 8.2 Extensible Service Framework

The Cloud Director extensibility framework allows cloud providers to consume Cloud Director as the single pane of glass to offer new services through, allowing them to take advantage of its multi-tenancy engine to provide tenant-aware services in a tenant-safe way.

Either way, whether it is partners who want to bring their backup, or middleware, or security or other offering to a wider user base and realize increased revenue, or it is cloud providers looking for a way to differentiate themselves in a crowded market by providing specialized functionality to their tenants, these extensions provide an entry point into the cloud provider's offering through Cloud Director with a single, consistent, and fully theme-able experience around all of the services and features offers to customers.

VMWARE CLOUD DIRECTOR AVAILABILITY

VMware Cloud Director Availability (VCD Availability), adds to the Cloud Provider Platform a set of use cases based around disaster recovery and the onetime onboarding of tenant workloads. This feature is tightly integrated into the overall solution and can offer cloud providers the ability to develop key service offerings within their portfolio.

For cloud providers, offering a simple, secure and cost-effective disaster recovery service continues to be one of the key service offerings enterprise businesses are looking for, and is a great initial use case to offer customers, many of which go on to add additional services from that provider.

VMware Cloud Director Availability offers two different types of recovery for customers. The first solution is a Disaster Recovery to Cloud Service (DR2C) which can be consumed by customers to protect their on-premises workloads in the cloud. Using a vCenter Server Plugin, end customers can replicate their workload applications into an Org. VDC. Should a disaster strike at any time, those replicated copies can be powered on through an automated workflow which can bring those applications back online quickly, making them available to the service consumers again with minimal disruption.

VMware Cloud Director Availability also offers a Cloud-to-Cloud (C2C) solution, which can be consumed by either the customer or provider to protect those workload applications that have already been migrated to the Cloud Provider Platform, making them recoverable across different data center availability zones or regional locations.

One of the key benefits of VMware Cloud Director Availability is that it provides a fully integrated experience to end customers using Cloud Director extensibility and a plugin for on-premises, with both components supporting their respective HTML5 interfaces. The VCD extension provides a fully multi-tenanted experience to both customers and providers and delivers a fully features solution ready to be consumed by customers either in a self-service manor or as part of a managed service offered by the cloud provider.

Finally, VMware Cloud Director Availability also offers a onetime onboarding feature for cost effective migration to the Cloud Provider Platform.

Architecture Overview

The deployment of both the Cloud Provider Platform components and on-premises appliances require minimal effort to install and configure. A single OVA for Provider and single OVA for tenant must be deployed using standard vSphere procedures.

For the cloud provider, the architecture consists of three appliances:

- 1. vCloud Replication Management Appliance:** The cloud replication management appliance is made up of three sub-components that provide the user interface, the VCD plugin and translates all the VCD constructs:
 - a. VCD Availability Portal:** User interface for Tenants and Providers
 - b. vApp Replication Manager and Portal:** Provides the main interface for the cloud-to-cloud replication operations. It understands the Cloud Director level concepts and works with vApps and virtual machines
 - c. Replication Manager:** A management service operating on the vCenter Server level. It understands the vCenter Server level concepts for starting the replication workflow for the virtual machines
- 2. Cloud Director Availability Replicator Appliance:** The replicator exposes the Host-Based Replication (HBR) primitives for the ESXi hosts as Restful API endpoints and proxies the replication connections between ESXi hosts.

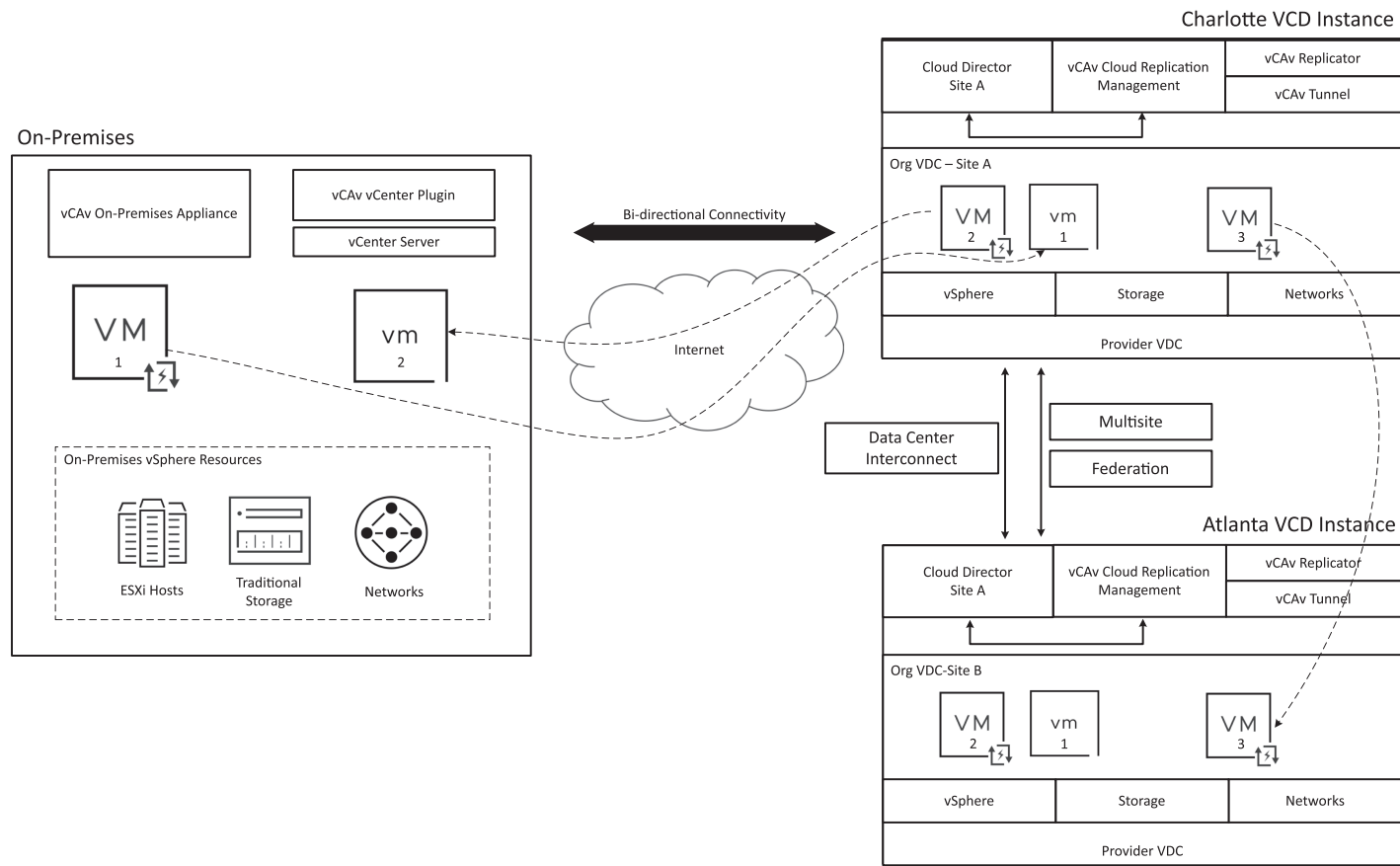


FIGURE 8.3 Cloud Director Availability Architecture

3. Cloud Director Availability Tunnel Appliance: The tunnel appliance provides secure connections between locations and simplifies the provider networking configuration by channeling all incoming and outgoing traffic for a site through a single point.

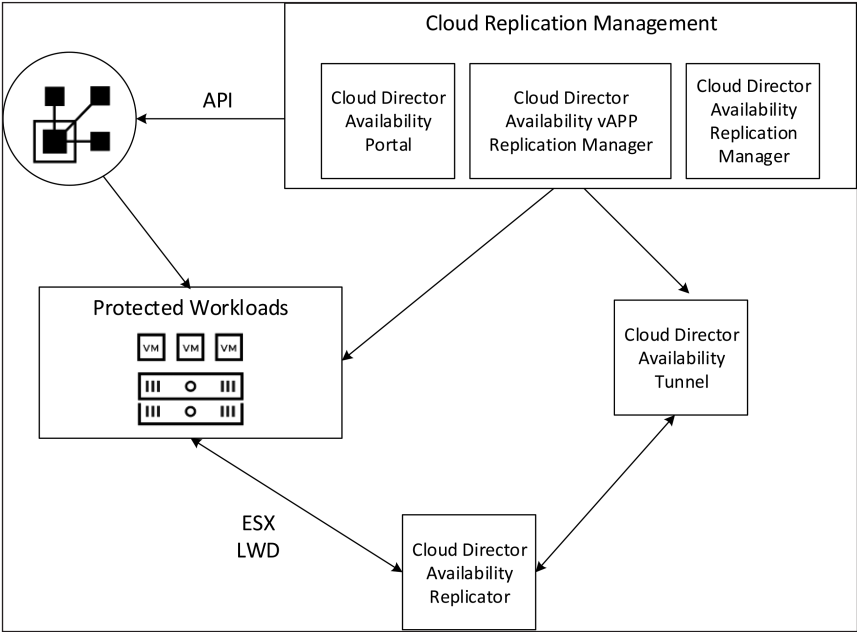


FIGURE 8.4 Cloud Provider Platform Component Overview

For testing purposes and lab environments, VMware provide a combined appliance that consolidates all three components into a single virtual appliance. Cloud providers will need to deploy and configure Cloud Director Availability for each Cloud Director server group instance, and pair sites ready for migration and disaster recovery services between cloud sites.

Disaster Recovery to Cloud (DR2C) Overview

As outlined previously, Cloud Director Availability supports multiple customer and provider use cases. Disaster recovery to cloud allows vSphere administrators on-premises to protect workloads by replicating them to a Cloud Director endpoint securely over the public internet.

Each cloud provider can support the recovery of hundreds of customer environments and can scale to handle increasing loads for each tenant in a linear manner.

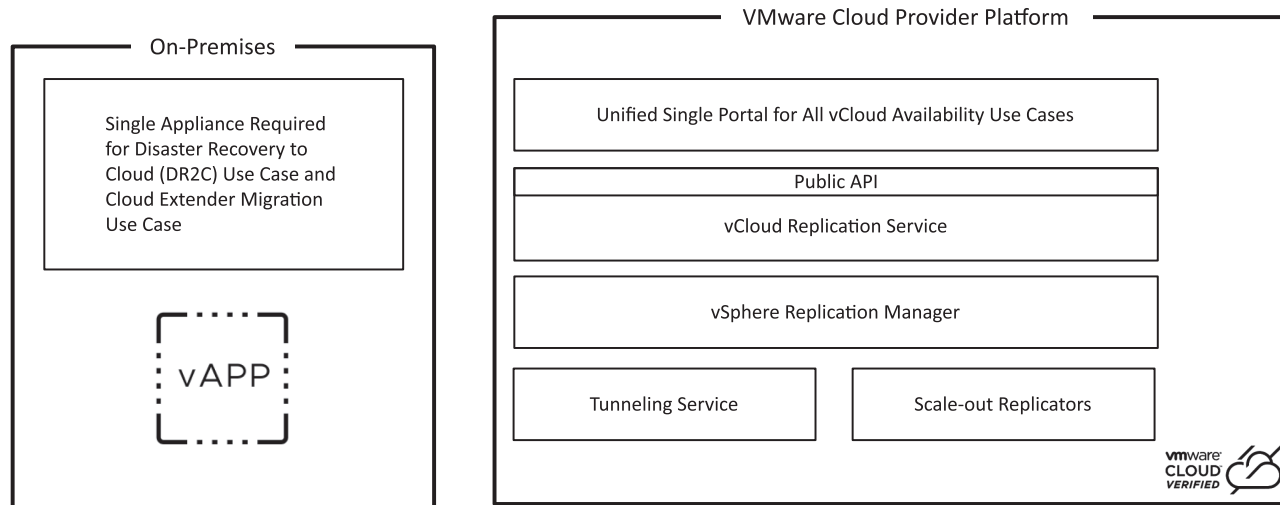


FIGURE 8.5 vSphere On-Premises Architecture

On the customer side, a single virtual appliance must be deployed in the on-premises vSphere environment. This appliance provides the replicator and a secure tunnel endpoint that is used to oversee the replication operation for each replicated workload. The replicator appliance exposes the low-level HBR primitives as Restful APIs and acts as a HBR proxy.

The on-premises appliance also provides the vCenter plugin, which can be used to manage all replication tasks after the secure tunnel is established with the cloud provider’s DRaaS endpoint.

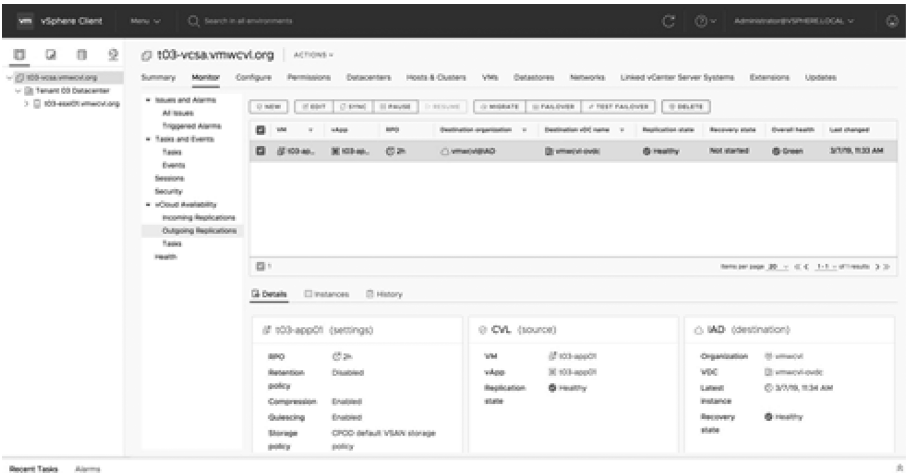


FIGURE 8.6 HTML5 vSphere Plugin Integration

As illustrated in Figure 8.6 the Cloud Director Availability vCenter plugin is directly integrated into the on-premises user interface, along with the context switching menu of Cloud Director. Once the tenant and VCD instance has been paired, workflow, replication and protection management can be managed from on-premises, via the plugin or through the Cloud Director user interface. The aim of this solution is that the non-IT user can find this experience intuitive and easy to use, making it consumable by customers without the need of a vSphere expert.

Replication traffic and all metadata communication is securely transmitted via the encrypted tunnel. This tunneling mechanism simplified provider networking set by channeling all incoming and outgoing traffic from the customer site through a single endpoint. The appliance establishes a reverse tunnel with the Tunnel Appliance at the DRaaS cloud endpoint on port 8048, allowing the Replicator to register with the cloud providers replication manager.

When deploying a Disaster Recovery to Cloud solution, there are several design factors which should be considered by the cloud provider. For instance, how do you support the customer in their deployment and

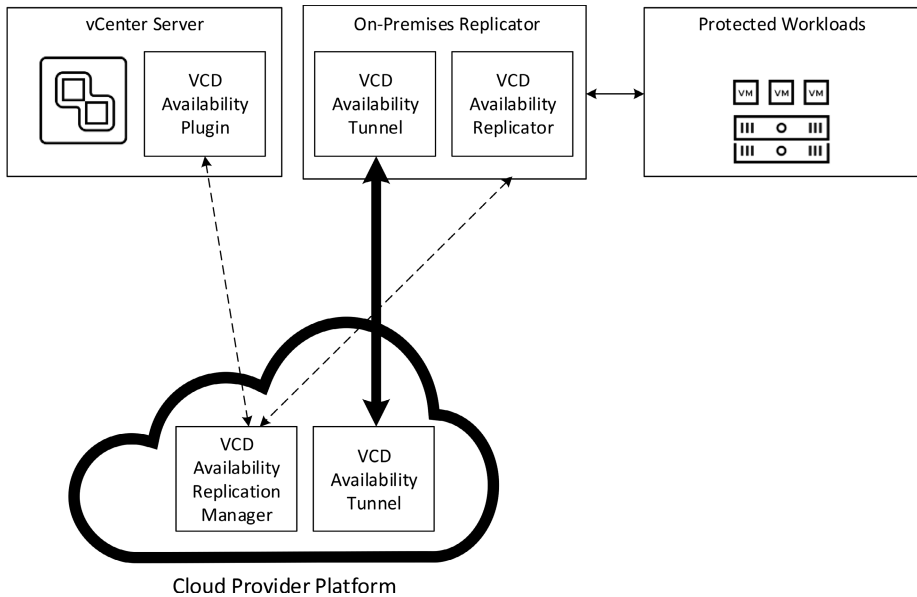


FIGURE 8.7 On-Premises vSphere Architecture

configuration of the vSphere components. In most environments the secure connection required will require changes to the firewall, although the architecture described above does not require any inbound policies to establish a connection to the DRaaS endpoint. Nevertheless, for some customers this means a security exception, which may need approval through a change control process.

To facilitate the bring-up of workloads, should a disaster scenario occur, Cloud Director Availability provides simplified workflows that are easy to utilize and provide explicit control over workload recovery. A protection workflow binds metadata for the replicated entity such as destination location, retention policy, recovery point objective (RPO), compression ratio and quiescing. In addition to these protection factors, the protection workflows also define networking configuration, which is key for the successful recovery of a protected virtual machine. These networking settings allow the customer to configure:

1. Allow for selection of Migrate/Failover and Test Network
2. Reset MAC address
3. IP Mode (ability to re-IP/re-address on failover):
 - Keep (default)
 - Static—IP Pool
 - DHCP
 - Static—manual

Protection workflows also specify storage policy and Organization owner and can be reconfigured at any time.

Another feature of the protection workflow is that they can be scheduled to start at a specific time, allowing initial replication and seeding to occur outside core business hours. In addition, they can be configured to run only during non-business hours to maximize throughput while minimizing impact to office users.

Another key component of Cloud Director Availability is Policy Management. Cloud Director Availability Policy Management allows cloud providers to define policies allowing for granular tenant controls, such as enabling and disabling incoming and outgoing replications, defining the maximum number of concurrent replications, the maximum snapshot retention and minimum RPO, with up to a 5-minute RPO and 24 point-in-time recoverable instances. Therefore, as cloud providers can define granular policies, Policy Management can be consumed to define different levels of service on a customer by customer basis.

Cloud to Cloud (C2C) Overview

VMware Cloud Director Availability *Cloud to Cloud* provides a Cloud Director to Cloud Director solution for replication and failover of Cloud Director-based workloads, for the purposes of disaster recovery or migration.

As illustrated in Figure 8.8, Cloud Director Availability's Cloud to Cloud solution replicates and recovers workloads between two different VMware Cloud Director instances for purposes of disaster recovery or planned migration, such as for maintenance periods, data center refresh or data center relocation.

Like the Disaster Recovery to Cloud feature, Cloud to Cloud is offered either through self-service customer operations, for disaster recovery, or as a provider driven service offering. This solution also offers bi-directional protection, and the ability to perform all disaster recovery runbook operations from either source or destination sites. Also, like the DR2C solution, a single, unified management portal, fully integrated into the Cloud Director user interface can be consumed by both customer and cloud provider admins, providing complete self-service capabilities to replicate and recover workloads.

Native Integration with Cloud Director

This native integration is key to delivering a fully integrated customer focused experience, with centralized management through the integration with Cloud Director's provider and customer portals, for operations such as:

- Configuring the recovery services for multiple sites (cloud and on-premises)
- Replication management from either site
- Secure tunneling through TCP proxy for both cloud and on-premises pairings
- Replication management by either provider or tenant

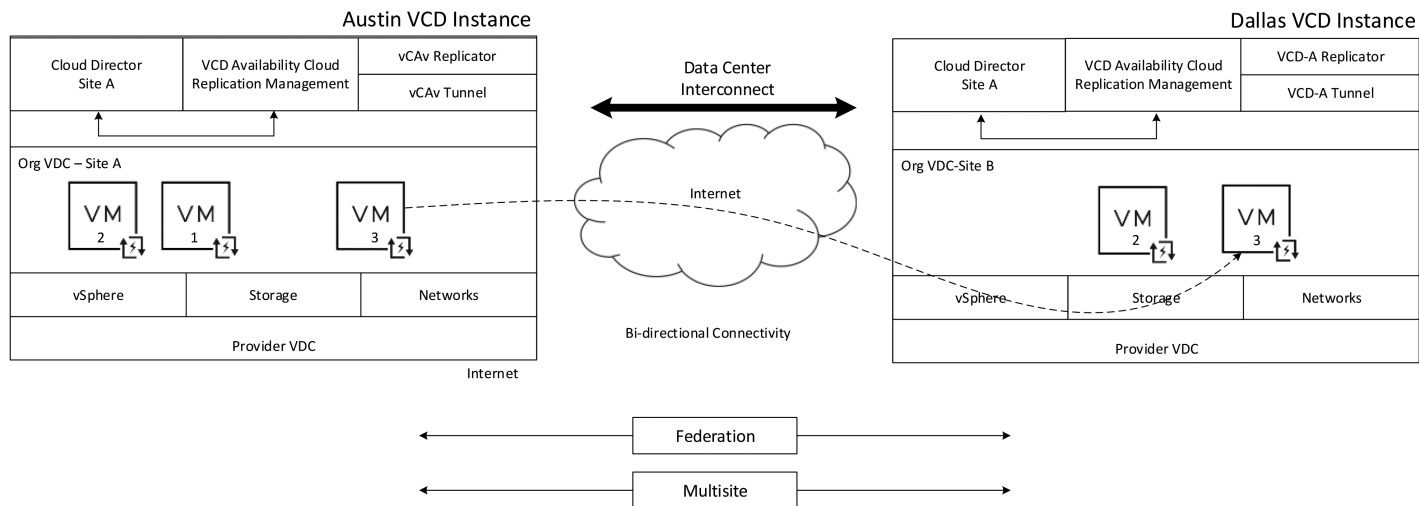


FIGURE 8.8 Cloud to Cloud Architecture

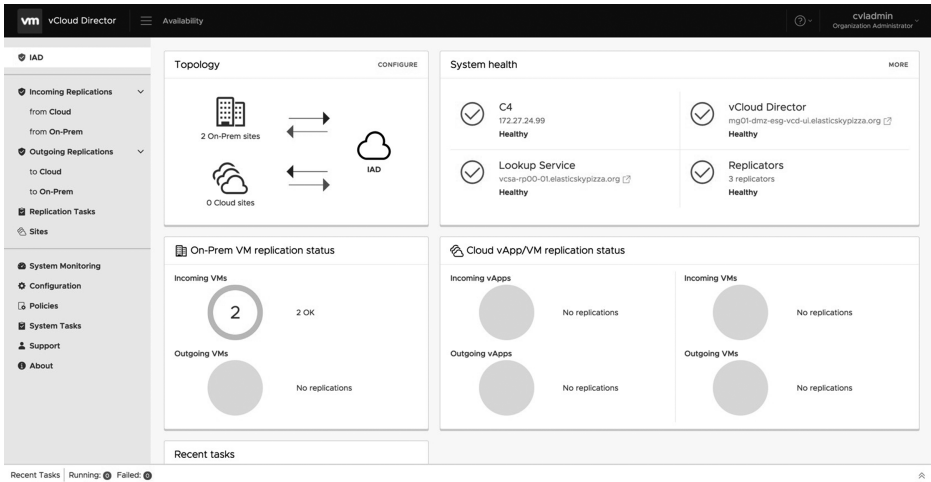


FIGURE 8.9 Cloud Director Availability and Cloud Director User Experience

Cloud Director Availability Migration verses HCX

In addition to Disaster Recovery and Cloud-to-Cloud protection-based solutions, Cloud Director Availability also provides simple, secure and cost-effective migration use cases for both:

- Cloud Director to Cloud Director migrations, for the purposes of planned migrations, such as maintenance periods, data center refresh or data center relocation.
- vSphere to Cloud Director, for the purposes of onetime small-scale onboarding migration onto the Cloud Provider Platform.

For larger scale migrations to the cloud it is typically recommended to utilize the WAN optimized SaaS offering, VMware HCX, which will be addressed next.

PLANNING FOR APPLICATION MIGRATION WITH VMWARE HCX

The migration of applications to the cloud, a place considered by many to be preferable to on-premises for a whole range of reasons, requires the end customers to be prepared for that transition. VMware cloud providers have to deal with a wide range of migration use cases, these might include, but are not limited to:

1. Onboarding new customers or additional customer workloads
2. Internal DC migration (inter and Intra data center migration)

3. Hybrid bursting and dynamic scaling of resources for temporary placement of workloads
4. Temporary relocation (for maintenance or hardware refresh)
5. Migrating customer workloads and applications to public cloud
6. Off boarding customers

As we have seen, one of the key differentiators for the VMware Cloud Provider Platform is the operational consistency and the ability to move applications to the public cloud, without the significant risk and lack of return on investment for refactoring those applications. Therefore, while many customers will build new environments and applications on the platform, for many it's the complexity and risks associated with moving existing applications to the public cloud that is a key driver for adoption.

Many of you will have a lot of experience when it comes to migration planning and be aware it is a huge topic to cover here, therefore we will just be looking at it from a mile-high viewpoint, with the key factors associated with migration to a VMware Cloud Provider Platform being addressed specifically.

Onboarding a customer to the Cloud Provider Platform is simple, and takes a matter of minutes, assuming you have properly prepared and have the information you need to hand. Once deployed and all necessary configurations have been completed and communication is verified, the next step is to begin migrating applications. However, before this can be done, there is likely to be a process of analysis and gathering of data about the existing environment, predominantly focusing on not only application requirements but also the needs of the business. Therefore, the planning of application migration and processes around it are likely to be significantly more prolonged than the actual deployment of the environment itself, a factor that must be considered in the overall migration planning.

In most organizations, technical and business discovery can be carried out using specific application dependency and discovery tooling, and a series of internal workshops and discussions with the various lines of business and application owners.

Technical data gathering might include the use of VMware tools such as vRealize Network Insight and vRealize Operations to map the characteristics of the application and its supporting component services, such as DNS or its external authentication mechanism. The aim is to understand the complexity of the application and therefore the potential complexity of its migration. In addition to its specific network communication and dependencies, other data points on some or all the following workload characteristics might also provide the insight into the source environment required to prepare for a successful migration project:

- Provisioned, Average and Peak memory
- Provisioned, Average and Peak CPU
- Provisioned, Average and Peak storage I/O

- Current latency across the environment and applications
- Raw storage provisioned and used

While this is not a comprehensive list of metrics that may be required, from the cumulative data gathered from the tools, such as those listed, it is then possible to bucketize workloads into collections of applications with similar migration characteristics. From this point it is then possible to build consensus with lines of business and application owners on the methodology required to move these applications and their supporting services to the partner's Cloud Providers Platform.

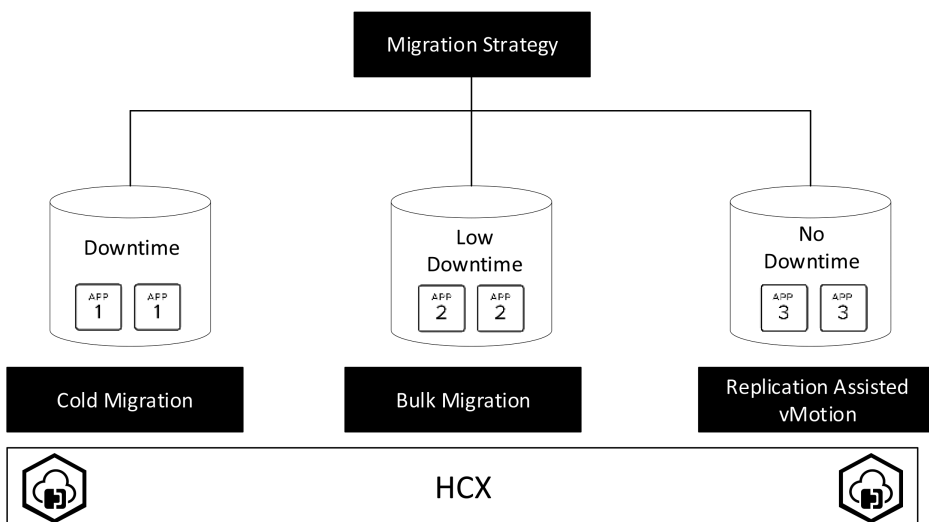


FIGURE 8.10 Bucketizing Workloads

Additionally, we must not forget to consider and similarly categorize business criteria, in how it relates to the criticality of the application, the operating model of the application, and most significantly, the potential impact to the business if a migration was to be unsuccessful. Key criteria that must be addressed include, but might not be limited to:

- The business value of migrating the applications to the cloud, or Return on Investment (ROI)
- Buy in from application owners or line of business
- Security or compliance requirements
- Licensing restrictions that might impact the cloud strategy
- Support concerns—if a third-party is managing the application, and might not agree on the strategy or future supportability of the application on the Cloud Provider Platform

As we have seen, there is a significant amount of preparation work that should be taken into consideration and completed prior to starting any migration project and moving applications. Once complete, the primary technological aim of any migration project is to move the existing applications from their existing data center to the cloud provider's environment, quickly, efficiently, cost effectively and with as little associated risk as possible.

Introduction to HCX

The primary aim of VMware Hybrid Cloud Extension (HCX) is to provide application migration capabilities without application downtime, or with minimal application downtime, and with no requirement to refactor applications, or their dependencies, in any way.

VMware HCX provides the hybrid cloud capabilities to bridge between an existing data center and the Cloud Provider Platform. The service can offer bi-directional application migration, which delivers data center mobility for live bulk migration and low downtime migration to the provider's environment. Built on Top of SilverPeak SD-WAN technology, HCX provides a high throughput network extension, with WAN optimization, traffic engineering, automated VPN with strong encryption data center interconnectivity, with built-in vSphere protocol proxies.

Through this combination of features, VMware HCX enables cloud on-boarding, without the need to upgrade the source infrastructure, supporting migration from vSphere 5.0+ to the Cloud Provider Platform stack, with a simple low risk migration SaaS platform.

It's important to highlight, while here we will be discussing VMware HCX in the context of migrating to the Cloud Provider Platform, this tool is not exclusively available for this platform, but is a part of the VMware Cloud Services portfolio, meaning that it can be leveraged by customers and partners for a wide range of migration projects.

HCX Component Overview

While VMware HCX is provided through a SaaS-based platform, migration-based connectivity is point-to-point, meaning no customer migration data traverses the SaaS-based components.

Therefore, as outlined below, HCX requires several components, in the form of virtual appliances, to be deployed at both the customer's source data center and into the customer's target Org. VDC. These various appliances work together to provide a range of management and network services, which ultimately delivers the HCX migration platform.

VMware HCX is initially deployed through the HCX Manager virtual appliance, along with a vCenter plug-in that enables the HCX feature. Once the HCX Manager virtual appliance is deployed, it is responsible for deploying the other appliances required. The VMware HCX platform is deployed

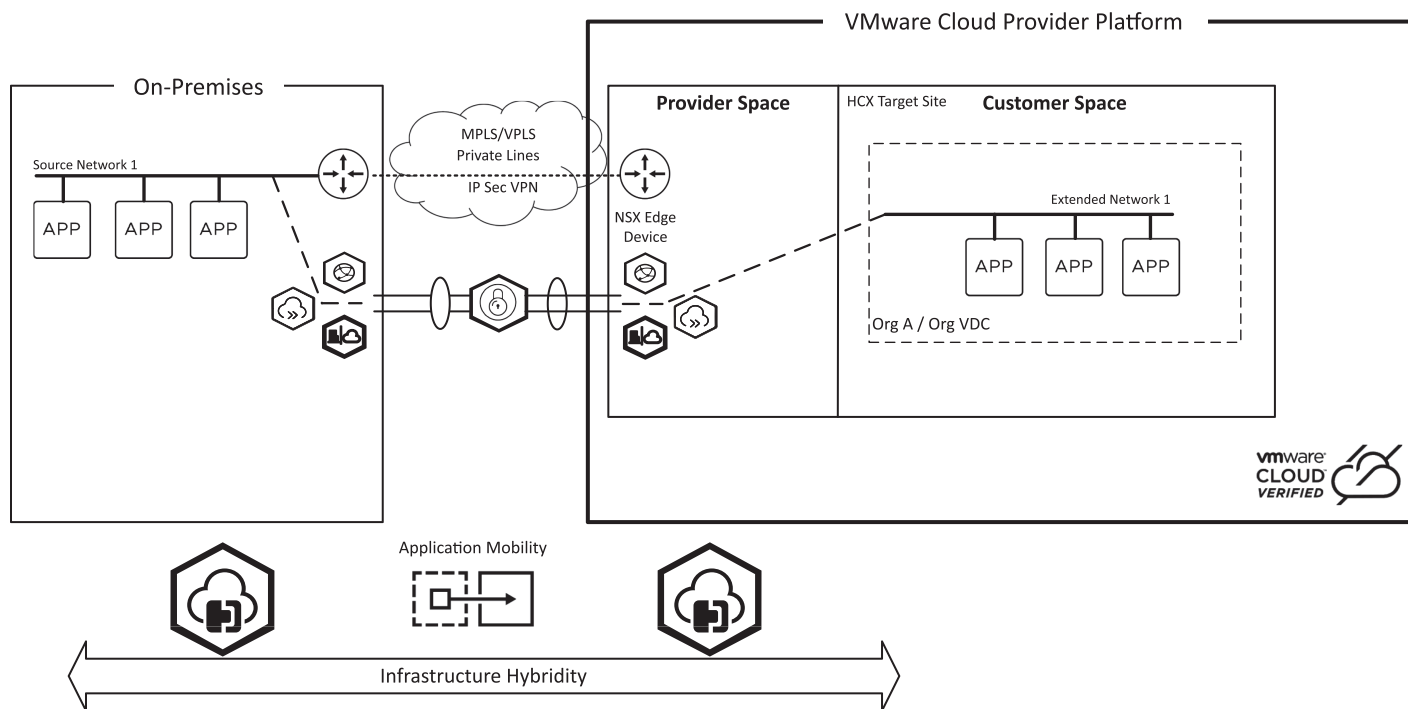


FIGURE 8.11 VMware HCX Overview

through the technologies and capabilities of the following four component appliances:



HCX Enterprise Manager Appliance: Provides the framework for deployment of the HCX service appliances. It integrates with vCenter and leverages existing SSO for authentication.



HCX WAN Interconnect Appliance: Provides migration and cross-cloud vMotion capabilities over the internet or private connectivity. The HCX WAN Interconnect Appliance also provides suite-B encryption, traffic engineering and workload mobility.



HCX WAN Optimization Appliance: Improves performance characteristics of Internet or private lines and leverages data-deduplication and line conditioning, for performance that is closer to a LAN environment.



HCX Network Extension Appliance: Provides high performance L2 extension capability, which is capable of 4-6Gbps of throughput per VLAN. This component can be scaled out to accommodate the extension of additional VLANs.

Migration Methodology

Depending on the application migration window, tolerable downtime, available bandwidth, latency and other network factors, VMware HCX offers multiple migration options to align with business and application owner requirements.

HCX Assisted vMotion: VMware HCX allows you to simply click a virtual machine and migrate it to the VMware Cloud SDDC, just like standard Long-Distance vMotion. However, on the backend, HCX is doing all the work of incorporating the integrated SilverPeak SD-WAN capabilities to provide network efficiency through deduplication, compression, WAN acceleration and traffic management, such as intelligent routing. As a result of the compression and optimization features of HCX, bandwidth requirements are significantly reduced, with HCX only requiring approximately 100Mbit of bandwidth to perform a vMotion operation (also requires Hardware Version 9), making seamless live migration to the hybrid cloud much more of a reality.

As discussed, vMotion over HCX delivers optimized bi-directional live migration over a software-based stretched Layer-2 network, running between the customer's data center, running vSphere 5.5 or newer, and the Cloud Provider Platform. This solution operates without domain boundaries or Single-Sign-On (SSO) constraints. The result is that organizations can migrate workloads into a VMware Cloud Director organization without impacting the application, based on requirements from the business. This allows customers to build truly burstable applications, which span across both public and private cloud models.

Bulk Migration: Bulk Migration is driven through pre-staged replication, for efficient low down-time migration, effectively providing a *reboot to cloud* solution. Bulk migration provides a low-downtime solution as the workload remains online during the replication process. This allows for the large-scale migration of hundreds of workloads at a time to the target VMware platform.

Many of you will already know, this is not groundbreaking technology. However, HCX adds some compelling differentiators, such as its SD-WAN optimization, scheduling capabilities and the abstraction of the physical layer, allowing for backward compatibility as far back as vSphere 5.0 (specifically ESXi 5.0 with vCenter 5.1).

For instance, if the business plans to migrate 200 workloads over the course of a single weekend, HCX could be configured to begin replicating over to the VMware Cloud SDDC well in advance, allowing the initial full replication to seed from the remote site. Once complete, only delta synchronization need occur. It is noteworthy to highlight HCX utilizes hypervisor replication, not the vSphere Replication Appliance and benefits from the same deduplication, compression and WAN optimization features applied to live migration.

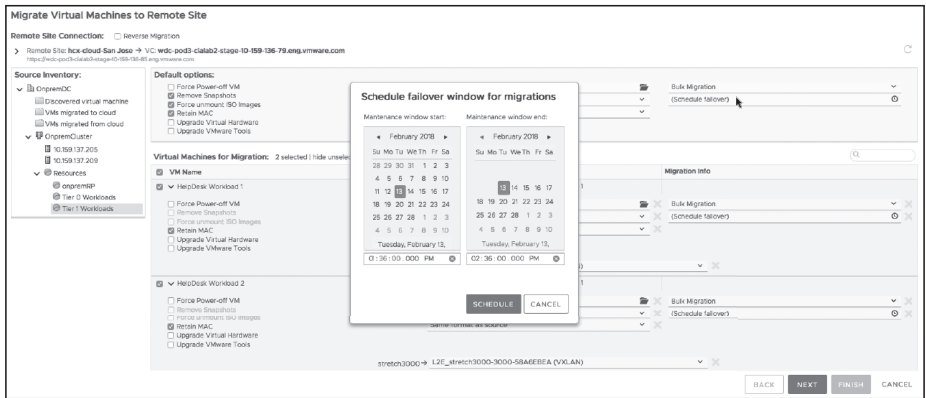


FIGURE 8.12 Scheduling a Failover for Bulk Migration

When ready, the administrator can select the day and time they wish to schedule the switchover, a maintenance window agreed by the business, such as Sunday from midnight to 6am. During the scheduled migration time, the source workload is powered down, removed from the source inventory, and a final delta synchronization performed. The final task is for the workload to be booted on the target Cloud Provider Platform, after which the migration is completed, and the administrator can begin any required testing or validation tasks.

In summary, the Bulk Migration feature of HCX utilizes hypervisor-based replication to reduce the downtime associated with migration, fundamentally by allowing the workloads to remain online during the seeding and delta replications stages. When data replication has been completed, we can either

immediately switchover to the target site or schedule the switchover during a pre-determined maintenance window. This all occurs over the HCX SD-WAN optimized fabric and is backwardly compatible as far back as vSphere 5.0, meaning there is no requirement to upgrade to a newer version of vSphere prior to the migration, which is common with many third-party migration tools. In addition, during a Bulk Migration HCX can optionally upgrade VMTools and Hardware Compatibility during the migration, helping customers who are on older versions of vSphere to take care of these upgrades at the same time.

Replication-Assisted vMotion (RAV): Replication-assisted vMotion brings together the key features from HCX assisted vMotion and Bulk Migration to provide a more efficient mechanism to carry out large-scale migrations of live workloads, by pre-staging the data prior to the final live migration. Replication-assisted vMotion is also able to leverage parallel operations, path resiliency and scheduling for the RAV switchover, where vMotion is engaged for the final migration of delta data and the virtual machine state, for the zero-downtime migration of large numbers of workloads.

The reason RAV is important comes down to the often discussed 80/20 rule of migration. When planning a large-scale workload migration, the 80/20 rule often applies as organizations often see that 20 percent of workloads are typically placed in the *No Downtime* bucket, as illustrated in Figure 8.10. These applications are so mission-critical to the business, they require zero downtime. However, the remaining 80 percent of workloads can be migrated during a maintenance window, the Low Downtime bucket, and are therefore allowed to be rebooted during the scheduled time.

Replication-assisted vMotion allows for workloads to be migrated live at scale, unlike HCX assisted vMotion operations, which can only be carried out individually. Also, like the Bulk Migration feature, Replication-assisted vMotion offers the option to schedule the migration during a pre-defined maintenance window.

In the background, RAV is utilizing host-based replication to seed data then using vMotion technology to migrate the final deltas, while workloads are still running. RAV this is not limited in numbers like the HCX vMotion and can be used to migrate large numbers of workloads in a single batch.

HCX for Hybrid Network Extension

VMware HCX can offer a permanent hybridity solution by securely connecting the customer's data center networks and seamlessly extending them to the Cloud Provider Platform with an SD-WAN optimized Layer 2 solution.

Traditionally, when connecting an existing data center to a cloud environment the network administrator built a Layer 3 VPN and then proceeded to create new networks on the cloud platform. This is how VMware Cloud providers have operated for many years. However, this approach can often be a complex process that requires interaction with security teams, application owners and can be a blocker to more agile cloud operations. In addition,

the use of Layer 3 can limit the ability of the platform to act as a true hybrid cloud, bursting out as workloads migrate to and from the platform, based on resource consumption billed hour-by-hour.

Also, while the Cloud Provider Platform is built on NSX Data Center technology, that does not necessarily mean that the customer must be operating NSX in their existing data centers. NSX on-premises is not a requirement, but might have additional benefits for workloads that either remain running in an existing data center or operate within hybrid solution.

VMware HCX can help organizations bypass these challenges by providing a Layer 2 secure bridge over the internet. The HCX virtual appliance gives you everything you need to quickly set up a secure bridge, typically over the Internet, either permanently or only during a migration period. This way workloads can move to the cloud and back again without any IP address changes or changes to the application.

HCX for Disaster Recovery

In addition to its workload migration features, HCX also offers limited functionality that can be used to provide a basic disaster recovery solution. The use of HCX for workload protection leverages the replication engine over the optimized SD-WAN network for the protection of applications, and can help address certain disaster recovery use cases, such as transient environment protection or low-cost protection for development platforms. However, HCX does not provide all the features and functionality expected from a production level disaster recovery solution, designed for the protection of critical applications. Key features provided by Cloud Director Availability, such as support for Policy Management and Protection Workflows, which provide the advanced automated recovery workflow capabilities, are not available from VMware HCX. For these and a range of other reasons, HCX is not a recommended solution for disaster recovery to the Cloud Provider Platform.

VMware Cloud Director HCX Design Considerations

- When consuming HCX with a VMware Cloud Director, target some specific design factors should be considered as part of the implementation.
- A Load Balancer is typically used to front-end end user's VCD connections, the HCX Cloud configuration (in a VCD cloud) uses the same load-balanced point of entry for HCX Cloud operations.
- HCX Cloud automatically syncs inventory changes from Cloud Director by consuming messages broadcast by VCD via RabbitMQ. HCX can leverage the same message bus as VCD.

In addition to the above, VMware Cloud Director can only be consumed as a target for HCX and not a source of workloads, as illustrated in Figure 8.13.

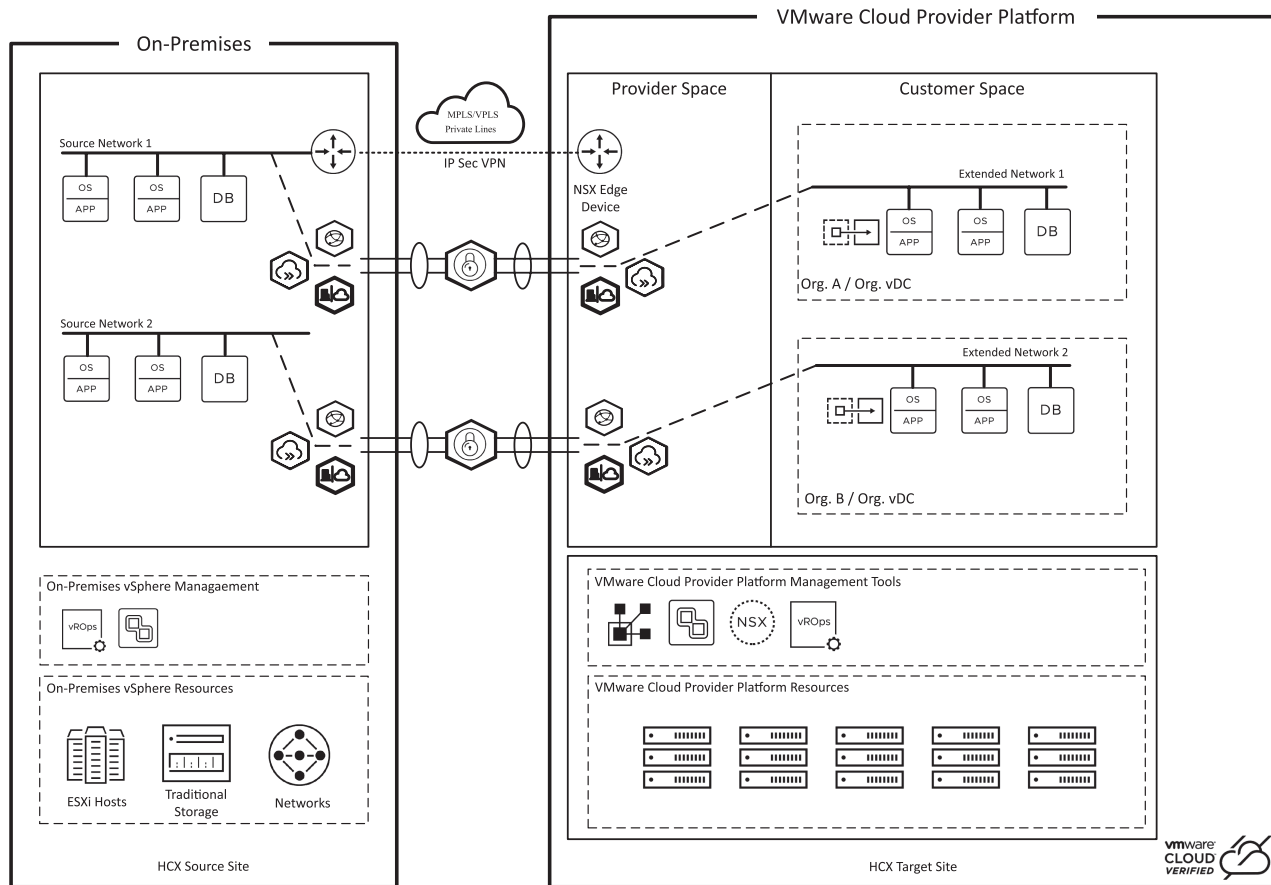


FIGURE 8.13 VMware Cloud Director HCX Design Example

Operating System Assisted Migration (OSAM)

Finally, HCX also offers the ability to migrate workloads running on a non-vSphere hypervisor to a VMware-based platform. One of the key use cases of this feature is for the repatriation of workloads from public or private clouds to a VMware Cloud Provider Platform.

For instance, workloads running on RedHat OpenStack/KVM or Microsoft Hyper-V environments can be migrated into a vSphere environment, giving the cloud provider a simple path to onboarding customers who are consuming one of the supported non-vSphere hypervisors and guest operating systems.

The guest operating system is key as OSAM requires the installation of Sentinel software on the non-vSphere workload that is targeted for migration. The Sentinel software is available for download from the HCX Management interface.

Cloud provider should always seek out the latest information, as HCX is being delivered in a SaaS-based model, code drops occur continuously and the supported guest VM operating system list will continue to grow with future releases. However, at the time of writing, the following guest operating systems are supported:

1. RHEL 6.x (64-bit)
2. RHEL 7.x (64-bit)
3. CentOS 6.x (64-bit)
4. CentOS 7.x (64-bit)
5. Windows Server 2012
6. Windows Server 2012 R2

VREALIZE OPERATIONS MANAGEMENT PACK FOR CLOUD DIRECTOR AND THE TENANT APP

An operational monitoring solution can be provided by consuming VMware vRealize Operations across the entire VMware Cloud Provider Platform. This solution collects data from vCenter instances across multiple data centers, availability zones and geographic regions to provide a single source of truth of important data points, helping operational teams make informed decisions. Likewise, vRealize Operations can also provide executives, who want a quick overview of their global VMware Cloud Provider Platform deployments, a view to keep track of important Key Performance Indicators (KPIs) or track configuration drift. vRealize Operations cloud provider use cases will be addressed in further detail in Chapter 9.

This management pack monitors the health of supported VMware Cloud Director server groups and sends smart alerts for monitored provider resources and helps plan capacity requirements for cloud providers.

The vRealize Operations Management Pack for Cloud Director Tenant App can be used in conjunction with the management pack for Cloud Director

to provides a mechanism to deliver advanced analytics tenant-facing dashboards for VMware Cloud Director tenant admins. These include providing low-level metrics on the customer Org. VDCs, workloads and alerts and metrics for troubleshooting and capacity planning.

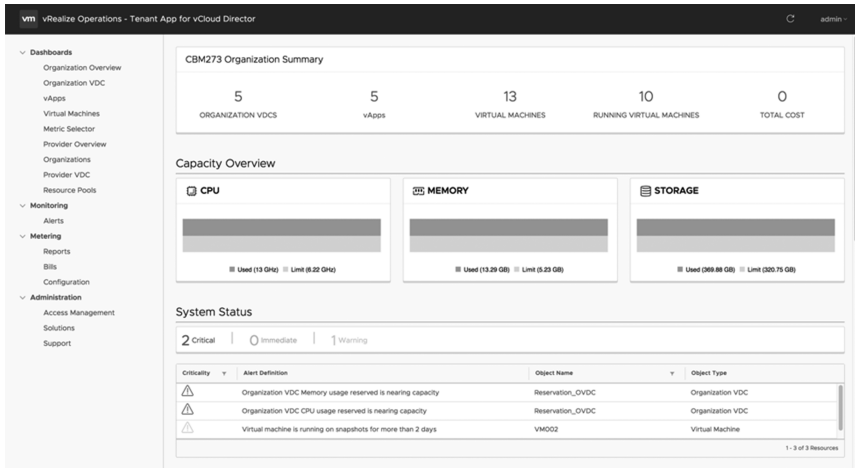


FIGURE 8.14 vRealize Operations Management Pack for Cloud Director

As with the Cloud Director Availability extension, this solution provides a simple customer experience that is fully integrated into the Cloud Director HTML5 tenant user interface, with SSO authentication with Cloud Director eliminating the need to login to Tenant App, enhancing the user experience while giving them deep insight into their running applications.

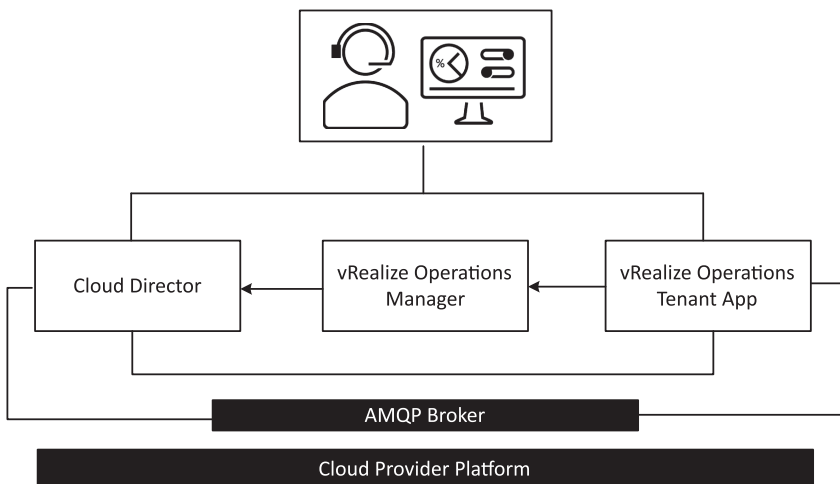


FIGURE 8.15 vRealize Operations Manager Tenant App

As we have demonstrated, the vRealize Operations Manager Tenant App allows cloud providers to deliver a view of the infrastructure to their customer, which is multi-tenanted and secured. Including these components within the Cloud Provider Platform allows customers to see utilization and performance of their resources directly within the Cloud Director user interface.

TENANT METERING

In addition to tenant metrics, the vRealize Operations Manager Tenant App also gives cloud providers the ability to publish customer-specific metering information through the integrated tenant app portal. This feature allows cloud providers to:

- **Define Pricing Policies** unique to each tenant or globally
- **Map Pricing Policies** to resources to allow customized pricing per tenant
- **Generate Bills** based on usage and pricing policy
- **View Bills**—tenants can see their own bills
- **View Reports**—only available for the cloud provider

Metering is a critical business component that enables accurate and automated usage measurement, cost calculation, analysis, and reporting to cloud providers. Through the vRealize Operations Manager Tenant App cloud providers deliver transparency with detailed billing data for each tenant, which includes the ability of every customer to review their bills using the integrated Cloud Director Tenant App interface.

In addition, this billing and chargeback support allows cloud providers to define pricing policies or rate cards globally or per tenant and map those policies to Org. VDCs. The cloud provider can then extract data via the user interface or API to generate bills or pull data for input into an existing billing system for multi-tenanted reporting, based on utilization and pricing policy. Tenant App APIs are also available for automation by using Swagger. This data can also be exposed to tenants by allowing them to view their own bills.

At time of writing there are 15 reports available for cloud providers to run or schedule based on their needs. These reports can be download in either a PDF or CSV format.

CENTRAL POINT OF MANAGEMENT (CPOM)

Central Point of Management (CPOM) allows cloud providers to use the Cloud Director user interface to securely publish dedicated vCenter Server-based services to a customer organization. This solution effectively uses the Cloud Director entity as a reverse proxy and single point of entry for service delivery of a dedicated customer vCenter Server instance.

There is an important distinction that must be made clear with this solution. The vCenter Server instances published to customers through the CPOM model are not delivering cloud resources to Provider VDCs, but are instead a dedicated private cloud, managed by a dedicated vCenter Server, which is presented to the customer by logging into the VMware Cloud Director HTML5 portal.

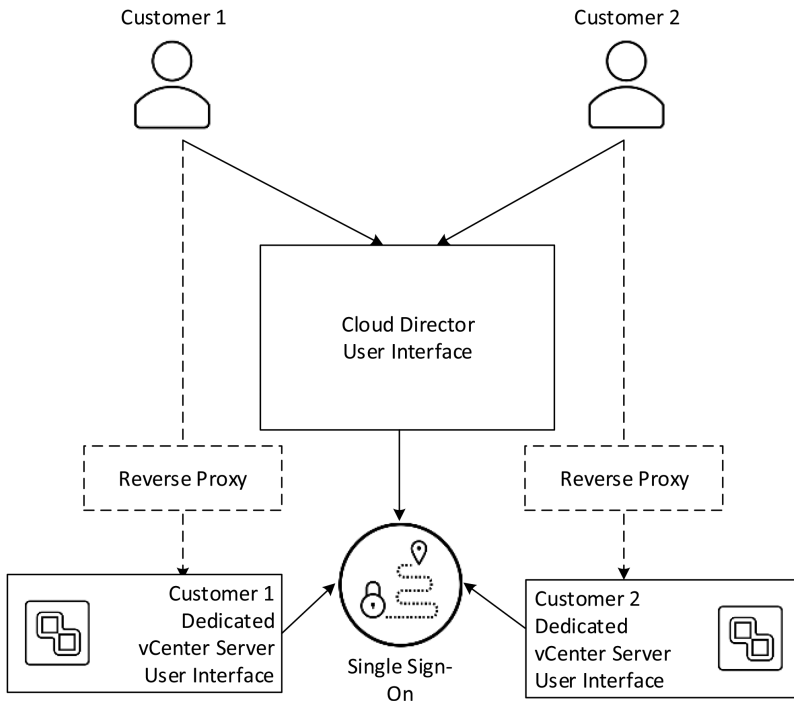


FIGURE 8.16 Central Point of Management Architecture

Dedicated vCenter Server service offerings, such as these, continue to be a very popular solution delivered by many of VMware's largest and most successful cloud service providers. These are typically dedicated SDDC services, which may encompass a whole range of VMware and third-party supported tools. The end customer might be consuming the entire SDDC stack (vSphere, vSAN and NSX), or they may just be consuming vSphere with either dedicated or shared array-based storage. None of these variables impact the CPOM feature. This capability simply allows the cloud provider to publish these dedicated vCenter Server instances to customers securely, and with integrated authentication.

The CPOM solution allows for total resource consumption flexibility. Customers can consume resources from the Cloud Director Provider VDCs by maintaining Org. VDCs alongside the dedicated vCenter instances.

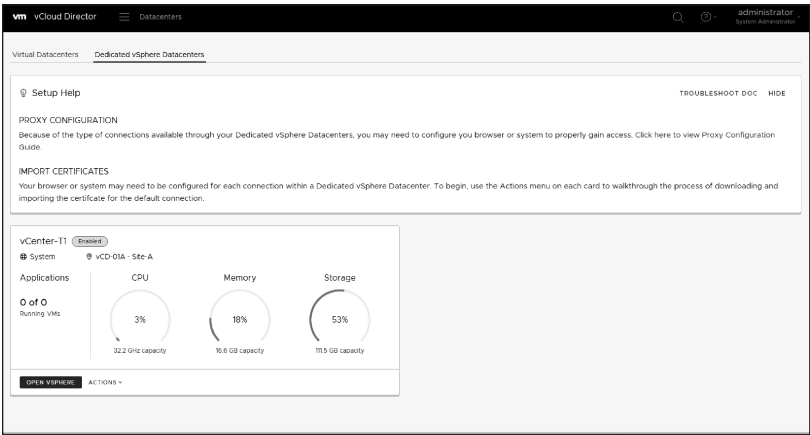


FIGURE 8.17 Central Point of Management User View

A customer could consume multiple vCenter Server endpoints from within the Cloud Director interface. Also, the dedicated vCenter Servers instances could be consumed from different cloud endpoints. For instance, a single customer could consume resources through vCenter Server, from VMware Cloud on AWS, Azure VMware Solutions, on-premises data centers, cloud provider data centers, or any combination. All that is required is a direct network connection to each vCenter Server instance. If the vCenter Server instance uses an external Platform Services Controller (PSC), Cloud Director also requires a direct network connection to the PSC instance as well.

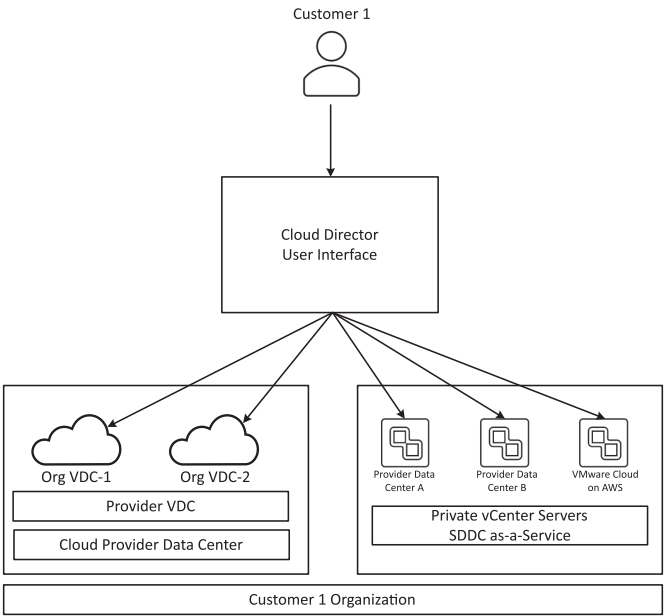


FIGURE 8.18 Multi-Endpoint Resource Consumption

SDDC AS-A-SERVICE WITH CPOM

Many of VMware's largest cloud provider partners offer dedicated private cloud platforms based on a standardized VMware SDDC stack of products. The cloud provider delivered SDDC is continuing to gain prominence due to the value it delivers through its key attributes, automation, control, and cost. Other key solution deliverables include:

- **SDDC Architecture:** The Software-Defined Datacentre service offering provides access to a 100% dedicated cloud infrastructure with *Software-Defined* technologies to not only enhance the performance, security, and management of the infrastructure, but also to facilitate an automated building-block, scale-up, and scale-out architecture.
- **Security:** A secured and dedicated cloud management platform for delivering a wide range of business applications to end consumers of varying size and profile, including firewall, load balancers, or other security services to the platform.
- **Optimized & Performant:** A dedicated and optimized vSphere platform, which provides guaranteed resource for workloads.
- **Cloud Management:** A SDDC as-a-Service could also employ the vRealize Automation Cloud management platform to deliver automated application deployment, scaling and workflow management and vRealize Operations to provide in-depth management of the entire infrastructure to increase productivity.
- **Implementation:** Low-level design and implementation can be delivered either by VMware PSO or by in-house architects and engineers—based on VMware Validated Designs (VVDs).

The software-defined data center is the ideal architecture for offering private cloud's at scale, which extends a common architecture with ease of management through a common toolset and resources while allowing the customer control and flexibility over dedicated resources.

With SDDC as a service, customers have the ability to run any application through an abstraction of a physical Infrastructure with vSphere, vSAN and NSX. With the same APIs as being consumed on-premises, automation delivers greater business value through a standardized approach. This combined with the benefits of working with an established VMware cloud provider delivers the key benefits of SDDC as a service. Having a common platform architecture makes it easy for enterprises to seamlessly build and manage their cloud environments both on-premises and in the cloud. Key advantages to service consumers include:

- **Pay-as-you-Go Model:**
 - Purchase dedicated capacity and functionality as a service, as required
 - No requirement to invest in advance for future needs (OpEx Model)
 - Leverage existing trusted provider

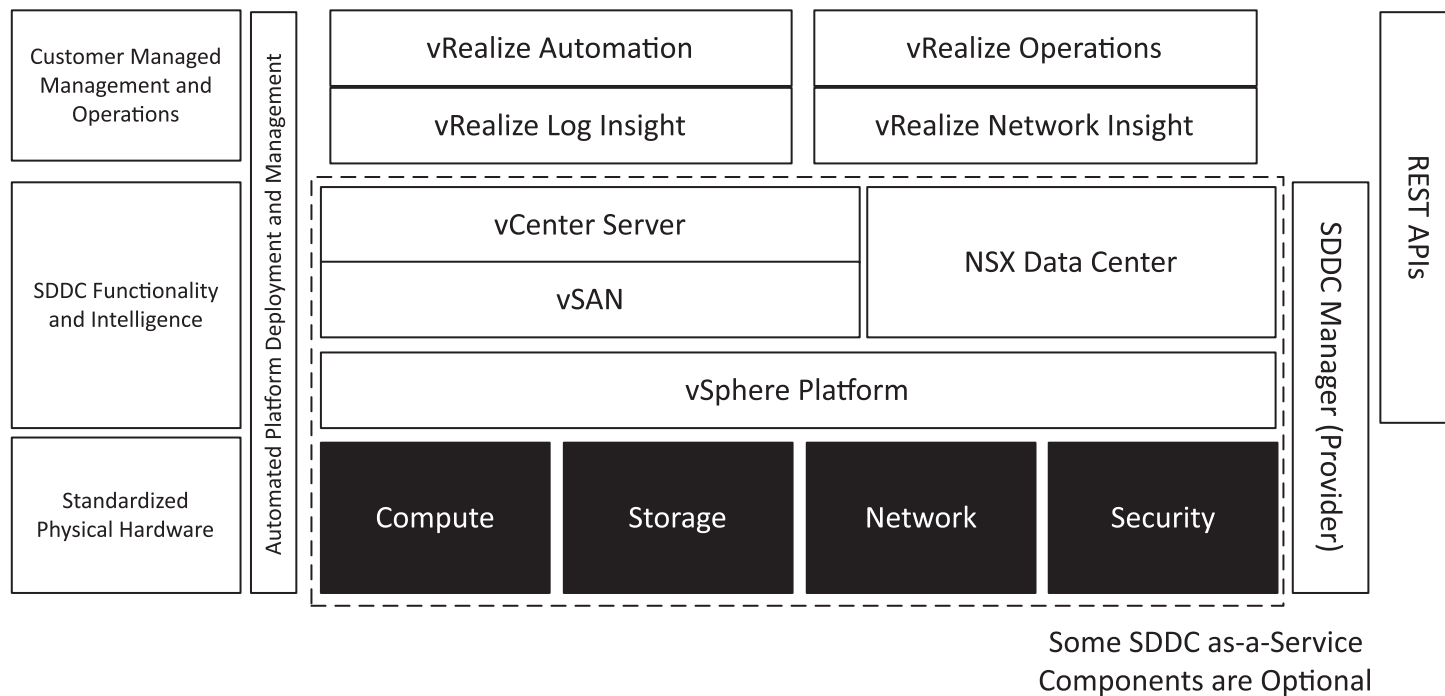


FIGURE 8.19 SDDC Private Cloud Service Offering—Application Stack

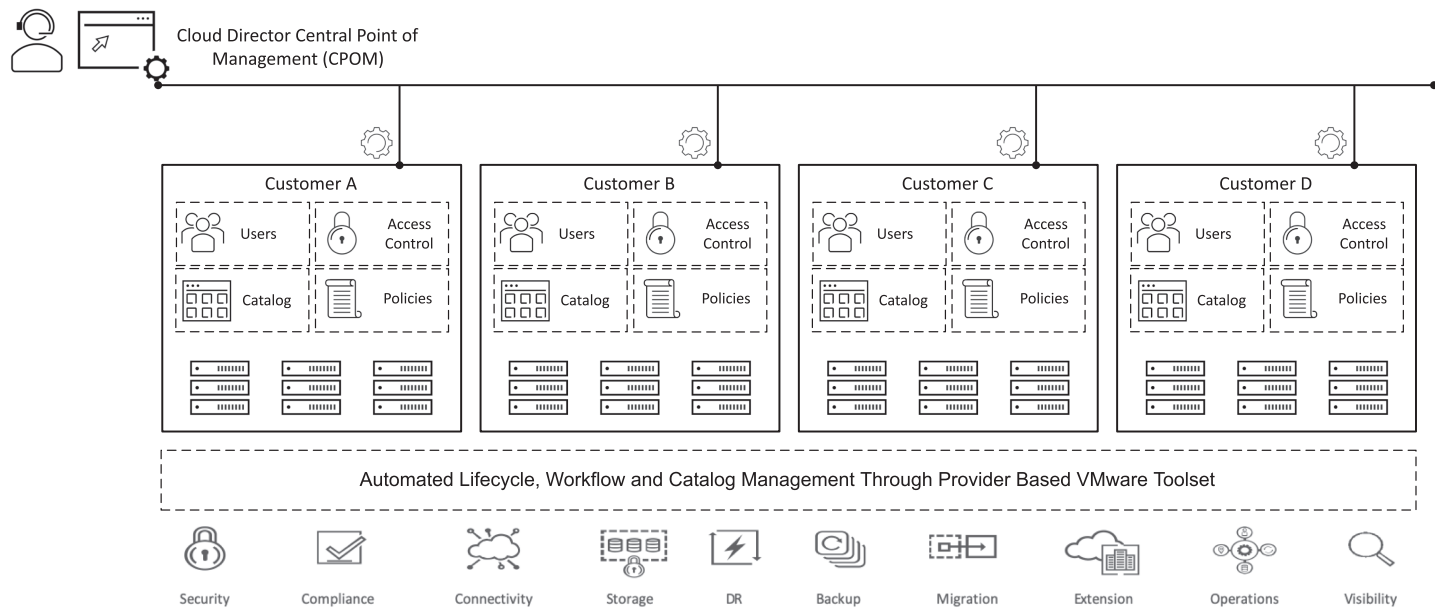


FIGURE 8.20 Provider Management & Operations

- **On-Demand (& Reduce TTM):**
 - On-demand, self-service-based provisioning of resources
 - Operational agility and cost reduction
 - Faster application development and testing lifecycle
- **Costing Transparency & Data Sovereignty:**
 - Metering and chargeback of usage
 - VMware Cloud Providers operating in over 120 countries
 - Multiple local compliance certifications and market specifications covered
- **Improved Efficiency:**
 - Saves consumer IT time, allowing them to focus on other strategic initiatives
 - Multi-tenant, dedicated, and virtual private cloud options
 - BCDR, monitoring, and compliance and operational efficiency

VMware Cloud Foundation can be consumed across private and cloud provider’s data centers with a dedicated SDDC service, typically made up of at least three hosts with vSAN for storage and NSX Data Center for deploying the private customer network segments.

Small	Medium	Large	Custom Build
• x4 Intel Hosts (4 cores /8 threads) • Clock Speed: 3.7 Ghz • Memory: 96 GB • 4 x 1Gbps NICs • 2000 Logical Networks • 3000 GB Raw Hybrid Storage • Monthly Billing • 24/7/365 Customer Support Line • Single Data Center • 1.0 Gbps Bandwidth Guaranteed • Secured Management Access Includes vSphere, NSX and Virtual SAN	• x8 Intel Hosts (10 cores /20 threads) • Clock Speed: 3.0 Ghz • Memory: 128 GB • 4 x 10Gbps NICs • 4000 Logical Networks • 6000 GB Raw Flash Storage • Monthly Billing • 24/7/365 Customer Support Line • Single or Multi-Data Center • 1.5 Gbps Bandwidth Guaranteed • Secured Management Access Includes vSphere, NSX and Virtual SAN + vRealize Operations & Log Insight	• x12 Intel Hosts (20 cores /40 threads) • Clock Speed: 3.0 Ghz • Memory: 256 GB • 4 x 10Gbps NICs • 8000 Logical Networks • 12,000 GB Raw Flash Storage • Monthly Billing • 24/7/365 Customer Support Line • Single or Multi-Data Center • 1.5 Gbps Bandwidth Guaranteed • Secured Management Access Includes vSphere, NSX and Virtual SAN + vRealize Automation, Operations, Network Insight & Log Insight	Custom Design
\$1800.00 /month	\$2100.00 /month	\$3750.00 /month	\$ TBD /month

FIGURE 8.21 Service Packages (P&P Examples)

A dedicated SDDC-as-a-service, delivered through the CPOM feature in Cloud Director, can be built with no limits, with dedicated VMware vSphere hypervisors, NSX, and optionally the vRealize management tool sets. A customer can create and manage an uncapped number of workloads, including containers, on a fully dedicated platform, while still taking advantage of the firewall, load balancing and security capabilities offered by NSX and the all-flash storage platform provided by VMware vSAN or third-party storage arrays.

CLOUD PROVIDER HUB

The VMware Cloud Provider Hub (CPH) is the Platform provided by VMware partners for delivering managed services. The aim of CPH is to provide end-to-end customer lifecycle management including the purchase, provisioning and management of VMware services.

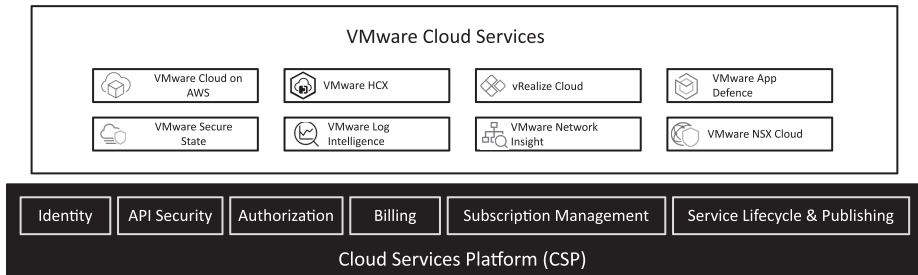


FIGURE 8.22 Cloud Provider Hub

Cloud Provider Hub delivers a self-service user interface and API for providers to transact multiple cloud services, tenant management, user management, billing and usage and support for cloud providers. In addition, it also provides a self-service user interface and API for end customers to consume the services that cloud providers have given them access to, and allows them to view usage and manage their user access to the services. However, end customers cannot view the payment method or view the commit contract(s) associated with the provider's master organization.

The VMware Cloud Provider Hub supports the delivery to end customers of a number of VMware services and is adding new cloud services on an ongoing basis. For instance, support for the VMware Cloud on AWS service, so partners can extend their managed services portfolio beyond their own data center deployments into an *asset light* VMware deployment model.

In the MSP model, on behalf of the end customer, the provider might:

1. Size and design the VMware Cloud on AWS deployment
2. Provision the Cloud SDDC
3. Configure the clusters, network, connectivity, and storage
4. Offer managed services such as:
 - a. VM Operations,
 - b. Migration
 - c. Backup
 - d. Disaster Recovery
 - e. Monitoring

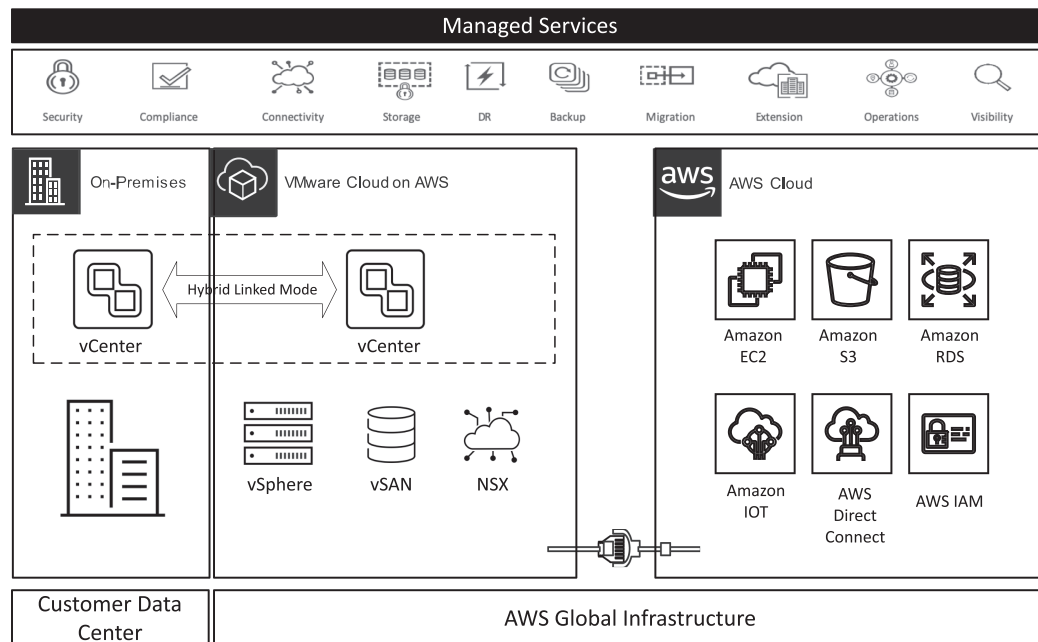


FIGURE 8.23 VMware Cloud on AWS MSP Delivery Model

Provider:

Architects the cloud-cloud deployment
 Provisions the SDDC
 Configures the clusters, network and connectivity
 Offer Managed Services: VM & Kubernetes Operations, Migration, Backup, Disaster Recovery, Monitoring, Compliance and Security
 Manages Support
 Manages billing

Customer:

Self-service consumption of SDDC

- f. Compliance
- g. Security
- 5. Manage support
- 6. Manage billing

In VMware Cloud on AWS, the demarcation of duties into a managed service provider model is facilitated by the Master Organization.

The Master Organization is a top-level construct for VMware Partners who are part of the Managed Service Provider (MSP) program and can therefore offer additional management services or other focused add-on services beyond the VMware Site Reliability Engineering (SRE) management framework. The Master Organization, illustrated in Figure 8.24, offers a Managed Service Provider an additional level of administrative control over multiple individual tenants, which links directly to service billing, operational management, and other features offered through the VMware Cloud Provider Hub.

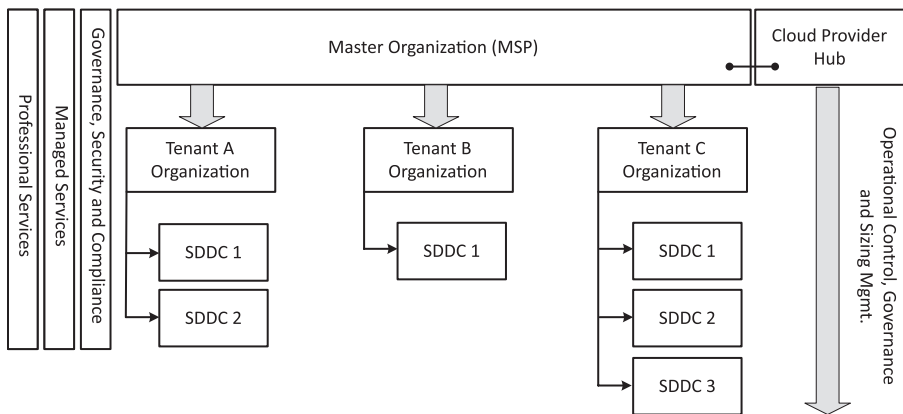


FIGURE 8.24 MSP Organization Structure

The Master Organization construct and Cloud Provider Hub are only applicable to Managed Service Providers who own the terms of service and provide tier 1 support to their customers or tenants. This additional root level Master Organization construct only exists for MSPs and is not visible or applicable to enterprise organizations consuming VMware Cloud on AWS as end customers, directly from VMware.

Cloud Provider Hub also provides a similar delivery model for multiple other VMware cloud-based services, with new services being added continuously. At the time of writing, these include:

- **VMware Hybrid Cloud Extension (HCX):** Used to connect across the hybrid cloud and deliver migration at massive scale for production workloads, even across diverse VMware and non-VMware environments.

- **VMware Log Intelligence:** a simple but powerful log collection and analytics tool that can be used to troubleshoot SDDC technologies, on-premise or in the public cloud.
- **Cloud Health Technology:** A fully integrated, multi-cloud-based business system
- **vRealize Network Insight:** Helps cloud providers build an optimize, highly available and secure network infrastructure across hybrid and multi-cloud environments.
- **vRealize Automation Cloud:** make it easy and efficient for IT and developers to get what they need to build and deploy applications.

CONTAINER SERVICE EXTENSION (CSE)

Building a *Developer-Ready Platform* is a strategic requirement for many cloud service providers as many are now being routinely asked to support DevOps activities for their end customers.

In order to facilitate this, VMware Cloud Director offers the Container Service Extension (CSE). CSE allows cloud providers to offer Kubernetes-as-a-Service to their customers, creating a simple mechanism for end customers to deploy and manage Kubernetes clusters on the Cloud Provider Platform in a secure multi-tenanted model. This is achieved by creating customized VM templates and enabling tenant or Organization administrators to deploy fully functional Kubernetes clusters in self-contained vApps.

As illustrated in Figure 8.25, CSE has several components associated with the architecture, as well as a several different personas interacting with it, Cloud Provider administrator, tenant administrator, and customer developer.

Firstly, the cloud service provider will configure the VMware Cloud Director environment, to provide the compute, storage and network resources on which Kubernetes clusters can be deployed. CSE uses the Cloud Director extensibility framework to deploy Kubernetes clusters and provide scaling operations, such as scaling-up or scaling-out.

- **CSE Server:** The cloud provider is also responsible for deploying and configuring the CSE server and VM templates. The CSE server automatically downloads and installs the required binaries such as Kubernetes, docker, weave etc. on the template. The CSE server handles client requests for the creation and deletion of Kubernetes clusters and nodes.
- **CSE Client:** Tenant Organization Administrators and users use CSE client to handle Kubernetes cluster management. This includes deploying clusters, adding worker nodes, configuring NFS storage. The CSE client runs on a virtual machine and runs as an extension of **vcd-cli**, which leverages both CSE and VCD public APIs to manage and administer the service, which includes the lifecycle management of the Kubernetes clusters by the tenant. It is from the client virtual machine that CSE commands are being issued to Cloud Director, which takes instructions by using the AMQP message bus.

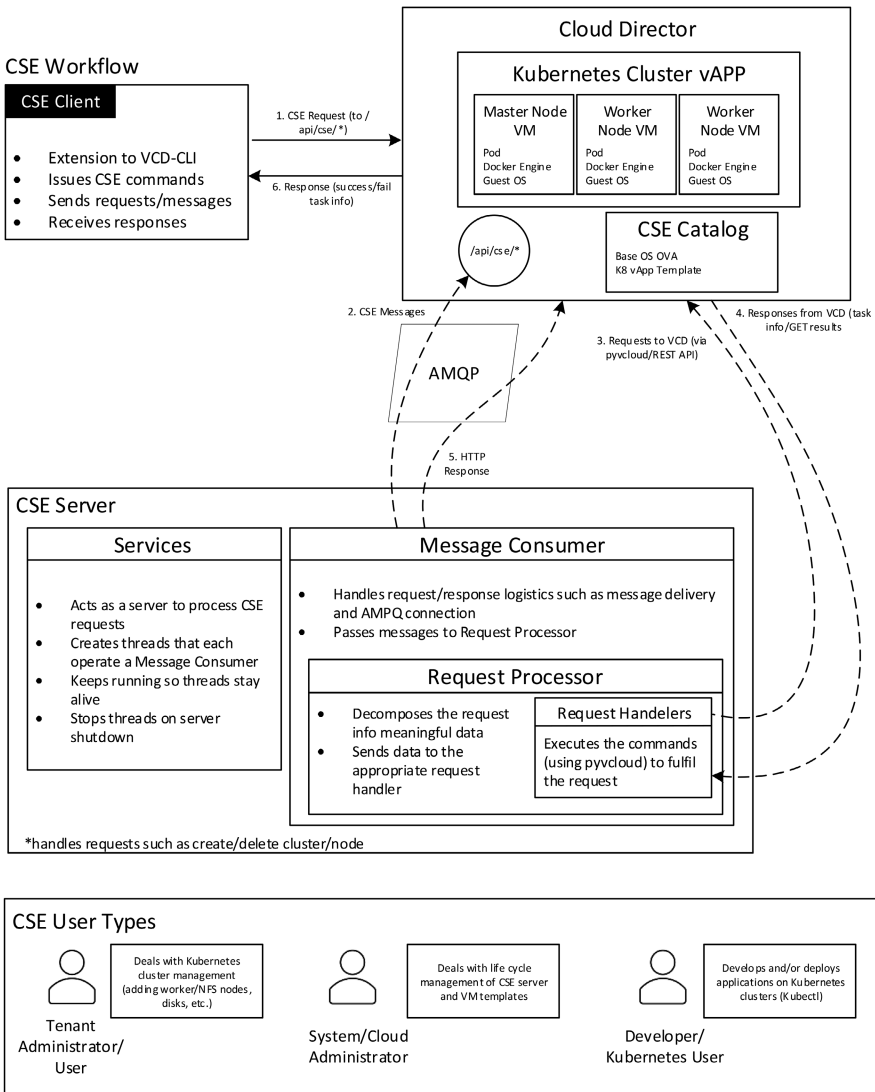


FIGURE 8.25 CSE Component Architecture and Interaction Points

- Developer Accessibility:** Kubernetes users, such as developers have accessibility to interact with the cluster using the native **Kubect1** command tool. For customer developer users, the Kubernetes clusters work like any other Kubernetes cluster implementation. There is no requirement for any additional Cloud Director specialist knowledge or CSE administration is required. The customer development teams do not even need a Cloud Director account.

From a deployment, inside the customer's Org. VDC, the basic CSE architecture on VMware cloud director is illustrated in Figure 8.26:

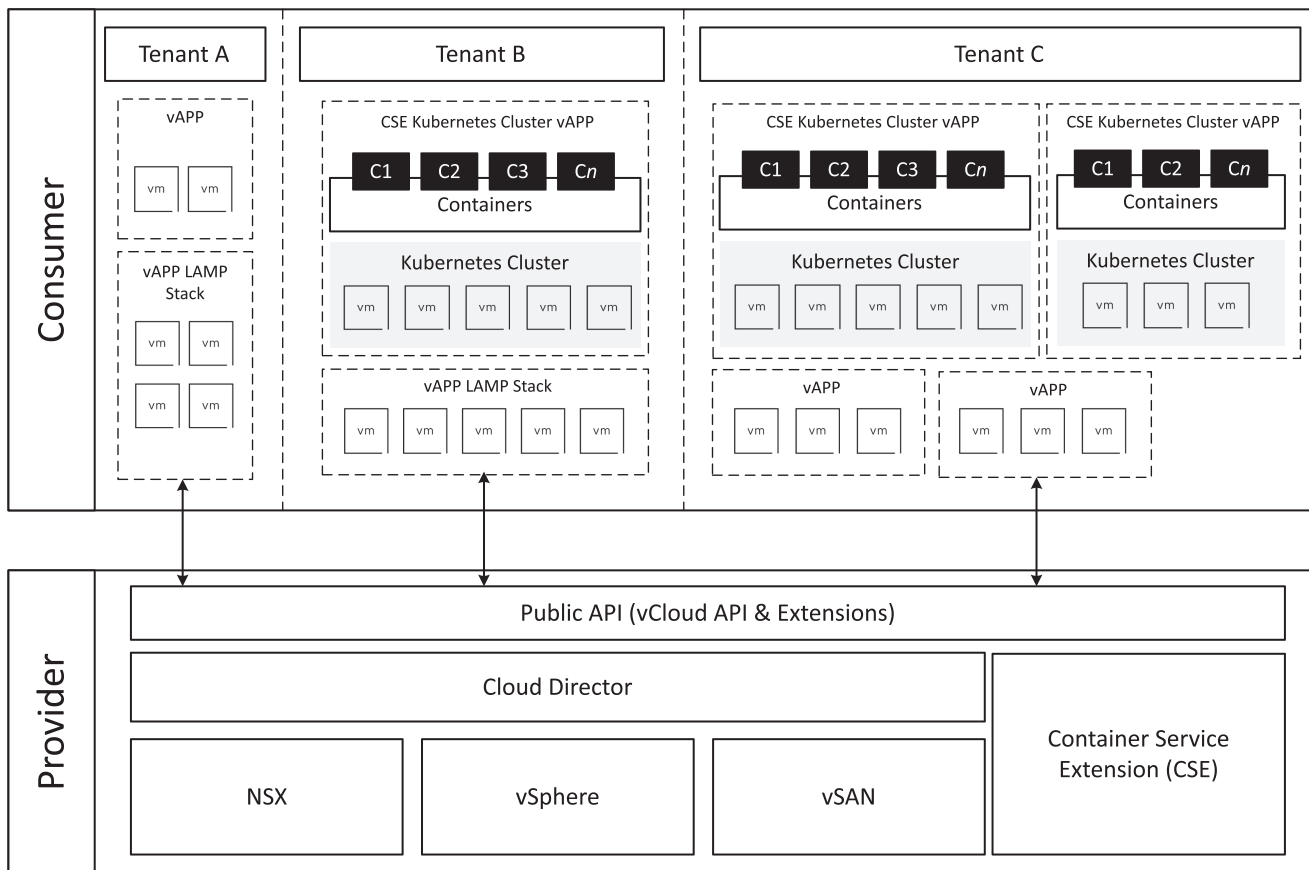


FIGURE 8.26 CSE Architecture

In addition to native Kubernetes, as discussed above, CSE also offers integration with VMware's Enterprise Kubernetes-based container services, PKS.

VMware Enterprise PKS, which is a jointly developed product offered by VMware and Pivotal, adds significantly more operational flexibility by adding:

- Improved availability
- Container level networking and automated load balancing with NSX Data Center
- Kubernetes supported by VMware
- Harbor integration

OVERVIEW OF COMMON STORAGE SERVICES

In addition to VM storage, many customers are looking to consume additional storage from cloud providers for other use cases, such as persistent storage for stateful Kubernetes clusters, storage for data archiving or extremely low-latency block storage for high-speed database applications.

VMware cloud providers can leverage new opportunities to capitalize on customer's interest in cloud storage solutions by offering a broad array of on-demand, secure, VMware-compatible cloud storage solutions, delivered by trusted providers and third parties, that can meet any business need.

	File Storage	Block Storage	Object Storage
Interface	User	Operating System	Programmatic (API)
Cost	Medium / High	High	Low
Performance	Medium / High	High	Low
Proximity	LAN	Dedicated Network	Internet
Use Case	Shared User Data	VM OS / Database	Images / Video / PDFs
Scale	Medium	Limited	High

FIGURE 8.27 Common Storage Models/Features and Costs

As illustrated in Figure 8.27, data center storage services are generally classified in the way they are consumed from the client connection. Traditionally, a shared filesystem, or *file storage*, offered to multiple clients

concurrently, provided the ability to access a single shared folder. The two most popular shared filesystem protocols in use today are NFS and SMB/CIFS.

Block storage provides network access to the equivalent of raw block devices. A client machine connects to a specific volume on the storage service and formats it as if it were a local block device. In terms of applications, multiple clients do not typically mount the same block volume, but this is different where block devices are consumed by ESXi hosts for the purpose of hosting virtual machines. Block devices are usually presented over Fibre Channel, iSCSI or AoE (ATA over Ethernet).

Object Storage

Object storage and its delivery by VMware Cloud Service Providers is a hot topic. Object storage provides applications, especially cloud style applications, with a simple and scalable solution to store and retrieve unstructured data (such as photos, videos, audio files, disk backups and more). For VMware Cloud Service Providers looking for the next stage of successful growth, providing an object-based storage offering is key to fully integrating high value-service add-ons and providing differentiated, high-margin services, such as cloud native offerings, backup services, disaster recovery solutions and big data applications, that integrate with data management software to expand data lake storage options.

Object Storage has become the *must have* service offering for cloud providers. The primary reason for this is the ever-growing mountain of data businesses must retain is matched only by the equally important desire to access it. However, the rise of cloud computing and interconnected cloud data centers makes managing this huge amount of data not only possible, but highly beneficial to businesses of all types.

In this section I will address the use cases for cloud data storage. In order to do this, we will first examine several different storage types, evaluate APIs, that enable direct and programmatic control, and address the important matter of security. We will also address other key design factors, such as latency and scalability, which will be done in part, by examining specific market leaders in cloud data storage, such as Google Cloud Platform and Amazon Web Services, as well as competitive services from other VMware partners, such as Rackspace and Virtustream.

Cloud object storage is a relatively new storage type, designed for unstructured data such as media, documents, logs, backups, application binaries and VM images. Conceptually they are like a persistent key/value store; objects are usually submitted via a REST API call, and an identifier returned. Most object stores allow attaching metadata to objects and aggregating them into containers (more commonly referred to as buckets). The most popular cloud object store is Amazon Web Services S3, and many object store implementations are compatible with it.

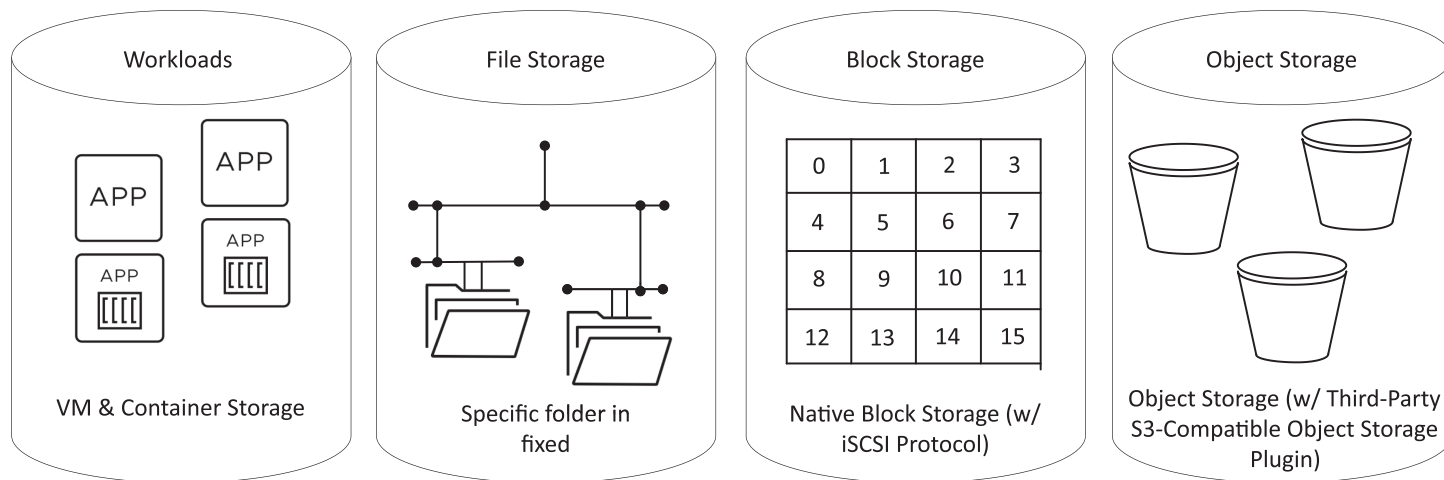


FIGURE 8.28 Cloud Storage Landscape

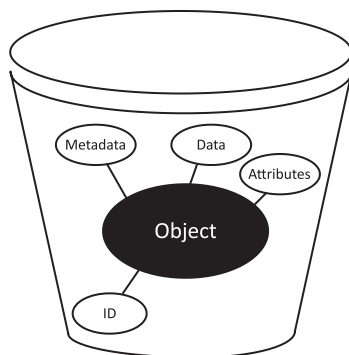


FIGURE 8.29 The Object Structure

Cloud Storage Landscape Overview: The subject matter around cloud data storage is continuously evolving and has come of age with the recent expansion of cloud-based network infrastructure as it has extended across the global market. The term Cloud Data Storage refers to some sort of repository for digital information, which exists across multiple services, in one or more geographical locations. The Amazon Web Services S3 Service was the first to bring cloud storage to the forefront of the cloud computing marketplace, but today there are a growing number of companies, such as Google Cloud Platform, Oracle, Microsoft, Alibaba and Rackspace, who offer cloud data services.

These data storage services have, in many cases, being successful, as they meet a wide variety of use cases. These include, but are not limited to:

- Application data, regardless of where the user is located
- Big data, both from the perspective of file size and quantity of records
- Routine archiving and backups
- Long-term data archiving
- Disaster recovery

However, in addition, there are numerous other benefits for end customers of storing data in the cloud. These include:

- Access from anywhere there is an internet connection
- Scalability, cloud storage is, subject to few limitations, infinite, and can grow with your business needs
- Security. Data access can be restricted to authorized users
- Availability. As storage is stored both across zones and geographies, the potential for data loss is limited
- Cost. Hosting business data in the cloud might result in significant cost savings versus maintaining your own storage systems, and will also reduce the operational costs associated with data administration

Linear Scalability	 Scales to billions of objects
Support for large files	 Object sizes are in TBs
Web friendly	 Firewall friendly, http, https, REST accessibility
No Locking	 No locks on write or create operations
Geo-scale	 Geo-scale and distributed
Metadata and extensibility	 Objects can be extended to multiple policies (immutability retention, etc.)

FIGURE 8.30 Key Object Storage Characteristics

On the other side of the argument for cloud data storage, the following topics are often highlighted as drawbacks:

- **Security.** Cloud storage providers and business users must continuously update strategies to keep the data from being compromised
- **Privacy** also goes together with security. Cloud storage providers and business users must ensure data is encrypted, both at rest and in transit, and is only accessible by authorized users
- **Data Sovereignty.** Businesses and cloud service providers must be aware of the privacy laws that govern data center locations
- **Network issues** and the accessibility to users over the internet should also be considered
- **Backup and restoration speeds**, based on bandwidth might also prove to be factors for some use cases
- The various cloud storage providers in the market offer a wide variety of services, but common among them are features such as:
 - Online management of storage, including import, export and backup/achieving operations
 - API for automated and programmatic data storage and control
 - Operational optimization methods, such as managing ACLs, for authorized users, or setting-up automated parallel data object transfers, for optimal efficiency

The cloud storage market is growing, with many players offering a range of services. Amazon Web Services provides a full end-to-end cloud platform, with services in computing, databases, analytics, applications, deployment and storage. Their primary object storage system is the Amazon Simple Storage Service, most commonly known by the acronym S3. S3 is a simple but highly robust object storage system, without any limit to quantity, and

allowing a maximum single file size of 5TB. S3 also provides a high-level of replication across multiple regional data centers. Lower cost object storage is also available through Amazon's Reduced Redundancy Option or the Amazon Glacier service. Other storage types offered by Amazon include Elastic Block Storage (EBS), used alongside EC2. EBS is like a more traditional file system, while remaining highly scalable.

Google Cloud Platform provides a similar selection of services, which leverages Google's existing global infrastructure. Cloud Storage is Google's primary object storage, and like S3, offers limitless storage, automatically replicating across many data centers located across the globe.

Microsoft Azure, as one would expect, also has a full offering of cloud data storage solutions. In Azure, objects, such as documents and media files are handled by Azure Blob storage, via a RESTful interface and client libraries for .Net, C++, Java, node.JS and Android, among others.

Security: Security is without doubt the primary concern of most potential cloud data storage consumers. The downside of being able to access your data from anywhere is, yes, that people from anywhere can potentially access the data. Previously we discussed strategies for restricting access to data to only authorized users, however, there are also a number of other security issues to be aware of. These can be broken down into two main areas of concern:

- Data in transit: Data transferring to and from the cloud storage platform hosts
- Data at rest: Data residing on the storage host at the remote facility

Firstly, there is a significant vulnerability to data while it is being transferred to and from the Cloud Provider Platform. Industrial strength encryption is typically the answer to this problem, but for this to be successful you must have a robust encryption key management system in place, in order to maintain long term accessibility. When in transit, HTTPS and SSL protocols should always be employed. More sensitive data can also be encrypted prior to transfer, via some corporate standard client-side encryption tool. A corporate solution such as this requires a key management strategy to be in place by the consumers.

Most cloud storage platforms can perform encryption, ensuring the data is secured while at rest. For instance, Amazon S3 allows you to establish bucket policies to enforce encryption, which will stop the egress transfer unless the data's header contains a request to encrypt the data server side. AWS also supports another layer of protection for server-side encryption, the key management service (KMS). KMS gives the business or organization's consumer the control over server-side encryption keys, preventing those keys from ever being exported, and providing a full audit trail of their use. Another way of securing data is by versioning. Once versioning is enabled,

the data is protected from accidental deletion or override. Versioning is typically enabled at the bucket level.

Also, from a security standpoint, it is typically a good idea to enable logging, often disabled by default. Once configured, all requests for service access are tracked, and often include requester details, bucket name, object name, request time, request action, response status and an error code, if any.

Finally, the cloud provider must be confident in their own security, when the data is at rest, to guard against data breaches and internal data theft. It is critical that this protection is kept current and is constantly evolving, as the threat to data and the mechanisms used by individuals who wish to access data unlawfully are. In addition, you should ensure that access logs are maintained, are complete and are routinely monitored.

Service Level Agreement (SLA): As with any cloud provider service, there is always a possibility of downtime. Limiting network inaccessibility to an absolute minimum should be considered a key requirement for any service provider, and one that most work diligently to address. While maintaining external networks is often beyond the control of end consumers, the key tool for dealing with such problems is the service-level agreement (SLA). The SLA should detail uptime expectations and the consequences of them not being met, typically through some sort of service credit when things go wrong.

Delivery and Latency: Not only do you want to ensure your data is accessible, but you also want to ensure it can be delivered as quickly and efficiently as possible, especially important when it comes to application data. With global networks in use, latency can be of concern, and response time and return speeds are often a critical design factor. Latency can be defined as ‘the amount of time it takes one packet of data to travel from one location to another.’ In terms of cloud data storage, it is typically measured as the length of time from when the request is received by the cloud data platform, to when the response is received by the requesting client. Latency is a key defining characteristic for various storage classes and most vendors use it as one of the aspects to determine pricing level.

When it comes to optimizing latency still further, location is extremely important, and it is best to house the data as close as possible to the users who wish to access it. Most cloud data storage vendors allow you to specify the region in which the buckets are created. Typically, however, the regions are pretty wide, such as the US, Europe or Asia. However, depending on your cloud storage vendor, you might also be able to consume services that break down regions even further, such as Eastern US, United Kingdom or Frankfurt. Therefore, if the data consumers are global, then you probably want to take advantage of a global cloud service that has a reach across multiple global data centers, that includes facilities located close to the regions required.

Understanding how latency works, and the associated options, is a key design factors for cloud provider when considering a cloud data storage solution.

Scalability: Scalability is the ability of a system to efficiently adapt to handle the current workload. The massive scale of the networks now available to cloud providers means that cloud object storage is virtually unlimited. In addition, this scalability, is for the most part effortless to consumers, as the infrastructure is already in place and being maintained by the cloud provider. As scale increases, cloud providers systems share the load among multiple servers. If the use reduces, the numbers of servers employed will shrink to meet the reduction in demand. While the providers design their own systems to scale in and out in this manor automatically, there are serval techniques which can be employed to further optimize utilization and therefore reduce consumption and cost.

Cloud data storage hosts typically will support an infinite number of buckets, with each bucket being infinitely large. When users try to store more objects in a bucket than can physically be contained in a single drive, the data will be written to other systems, with the objects that are overflowing the device being moved to another device, whilst being maintained in the same virtual bucket.

When considering scalability, we typically think about scaling-up, but the ability also for storage to scale-down as easily as it scales-up is equally an important design factor. As cloud data storage is a 'pay for what you use' service, with costs being calculated on resources consumed, this can have a positive impact on monthly use costs.

Another design factor is the use of replication. Replication, in the context of cloud data storage, is the duplication of data in real time over the network. It is common practice among cloud data storage platforms to automatically replicate object data when it's added to the buckets and store the redundant objects on different devices, usually in the same region. When the object is replicated, copies of the data keep the same bucket ID, key name, and meta-data. The primary goal of data replication is data protection, or durability, to ensure the data objects are available when required.

Durability is the probability or likelihood that data will be intact after one year. The greater the likelihood that the data will be available, the higher the durability. Therefore, 100% durability would mean that an object could not be lost, and 90% durability implies there is a 1 in 10 chance. For instance, AWS rates their standard S3 storage class as 99.999999999% durability. This means if you store 10,000 objects with them, 1 might get lost every 10 million years or so.

This automatic replication is to other devices within the same region. However, you might also want to replicate data to other regions, in order to:

- Reduce latency by placing data as close as possible to consumers
- Conform with regulatory compliance that requires distant redundancies
- Remote office that require access to identical data

Many of VMware's largest cloud providers offer object storage powered by OpenStack, which provides triple replication. These offerings typically provide block-level compression with 256-bit key encryption. However, there

is now the option to build a fully integrated Cloud Storage Service natively within the Cloud Provider Platform, addressed next.

Object Storage on VMware Cloud Provider Platform

As we have seen, business and organizations have multiple options available to them to consume objects and other storage types. However, as VMware cloud providers are aware, large numbers of organizations choose to consume services from a trusted regional service provider and for those VMware Cloud Providers, offering block storage solutions and files-based storage solutions has been a part of their mainstream service offerings for many years. For VMware cloud providers who also wish to offer object-based storage, VMware have a solution architecture that can meet this requirement. Offered through the Cloud Director-based Cloud Provider Platform.

VMware's cloud provider object storage offering is integrated with and managed via Cloud Director using an Object Storage Extension plug-in, delivering a fully S3-compatible storage endpoint consuming Cloudian Software or Appliance-based systems. Cloudian bring to the table the most widely deployed independent object storage provider with the highest S3 compatibility and interoperability giving the Cloud Provider Platform the ability to deliver the richest service provider feature set for object storage, including fully multi-tenant storage at Exabyte-scale.

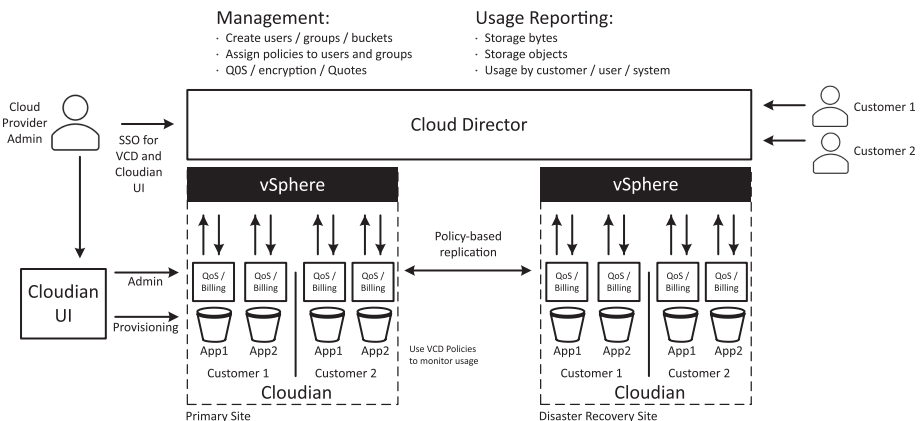


FIGURE 8.31 vCloud Object Storage Service

The Cloudian Object Storage solution for Cloud Director is an integrated storage platform that allows end customers deploy and manage S3-compatible storage within the Cloud Provider Platform environment. This allows the cloud provider to open their services to new vehicles and build revenue streams for storage services such as:

- Data Protection
- Big Data

- Media Asset Management
- Video Surveillance Storage

This solution also allows end customers to maintain compatibility with S3 APIs, which is the DeFacto standard for cloud storage. The Cloudian Object Storage solution for Cloud Director incorporates a fully compatible, native S3 API and can be deployed by cloud providers on their existing storage platform or on a separate Cloudian hardware platform, both of which are fully integrated with Cloud Director.

This solution delivers for VMware cloud providers a versatile, scalable storage pool usable for any S3-compatible application and enables a number of high-value cloud service offerings, such as:

- **Storage-as-a-Service (STaaS):** Provide S3-compatible storage capacity to help customers address their growing volumes of data.
- **Backup-as-a-Service (BaaS):** Offer a backup target that integrates easily with most popular data protection platforms.
- **Archive-as-a-Service (AaaS):** Provide a long-term data repository for customer's infrequently accessed or cold data.
- **Disaster Recovery-as-a-Service (DRaaS):** Help customers avoid the risks of business and organizational disruptions resulting from disasters by maintaining a data copy data offsite.
- **Big Data-as-a-Service (BDaaS):** Integrate with customer's big data applications to provide a limitlessly scalable data lake.
- **Containers-as-a-Service (CaaS):** Persistent object storage for new containerized applications.
- **Software Development:** The fully native S3 API provides a platform for S3-compatible application development.

Cloudian Object Storage for Cloud Director also supports storage requirements for cloud operations, such as importing/exporting and storing catalogs.

MACHINE LEARNING-AS-A-SERVICE (MLAAS)

The use of Machine Learning (ML) platforms by enterprises is growing rapidly with the regular introduction of new toolsets, making it increasingly difficult for customer's data engineers, data scientists and systems administrators to keep pace with accelerating innovation. In addition, the requirement for a large number of component toolsets, illustrated in Figure 8.32 and Figure 8.33, also creates additional complexity.

Therefore, in these fast-changing Machine Learning environments, the ability to provide flexible compute resources makes the Cloud Provider Platform

an ideal software infrastructure for running ML workloads. Organizations who are looking to GPUs to drive performance and achieve results faster with customers being able to run both GPU-based ML and non-GPU-based ML in virtual machines on various kinds of host servers.

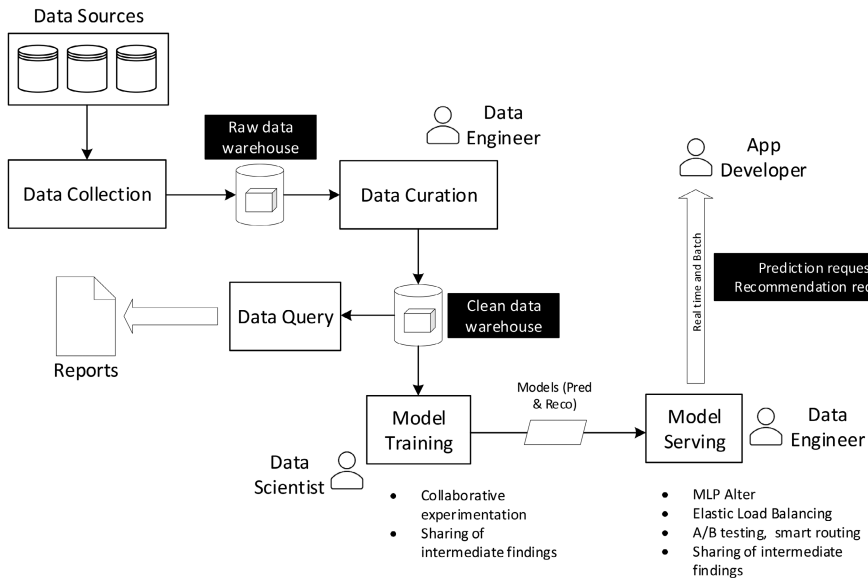


FIGURE 8.32 Distributed Machine Learning on TensorFlow for Machine Learning

As illustrated in Figure 8.32, the VMware Cloud Provider Platform can support data scientists and data engineers to provision different ML platforms, such as TensorFlow, languages, tool versions to enable modeling on a single common infrastructure with strong isolation. This gives data scientists and data engineers the flexibility to rapidly scale-up resources or scale-out resources using vApp templates and cloning automation.

The Cloud Provider Platform could even offer, as an on-demand service, General-purpose GPU capabilities, which enables customers to run machine learning environments across a vSphere-based accelerated environment. GPUs are increasingly being used for machine learning, artificial intelligence, high end analytics, 3D modeling, video editing, weather prediction, oil and gas geology, and scientific computing. Thousands of cores allow GPUs to run computations in parallel, therefore providing significantly better performance compared to CPUs.

GPU accelerated applications can run the sequential part of a workload on CPU cores while accelerating parallel processing on the GPU, as illustrated in Figure 8.34.

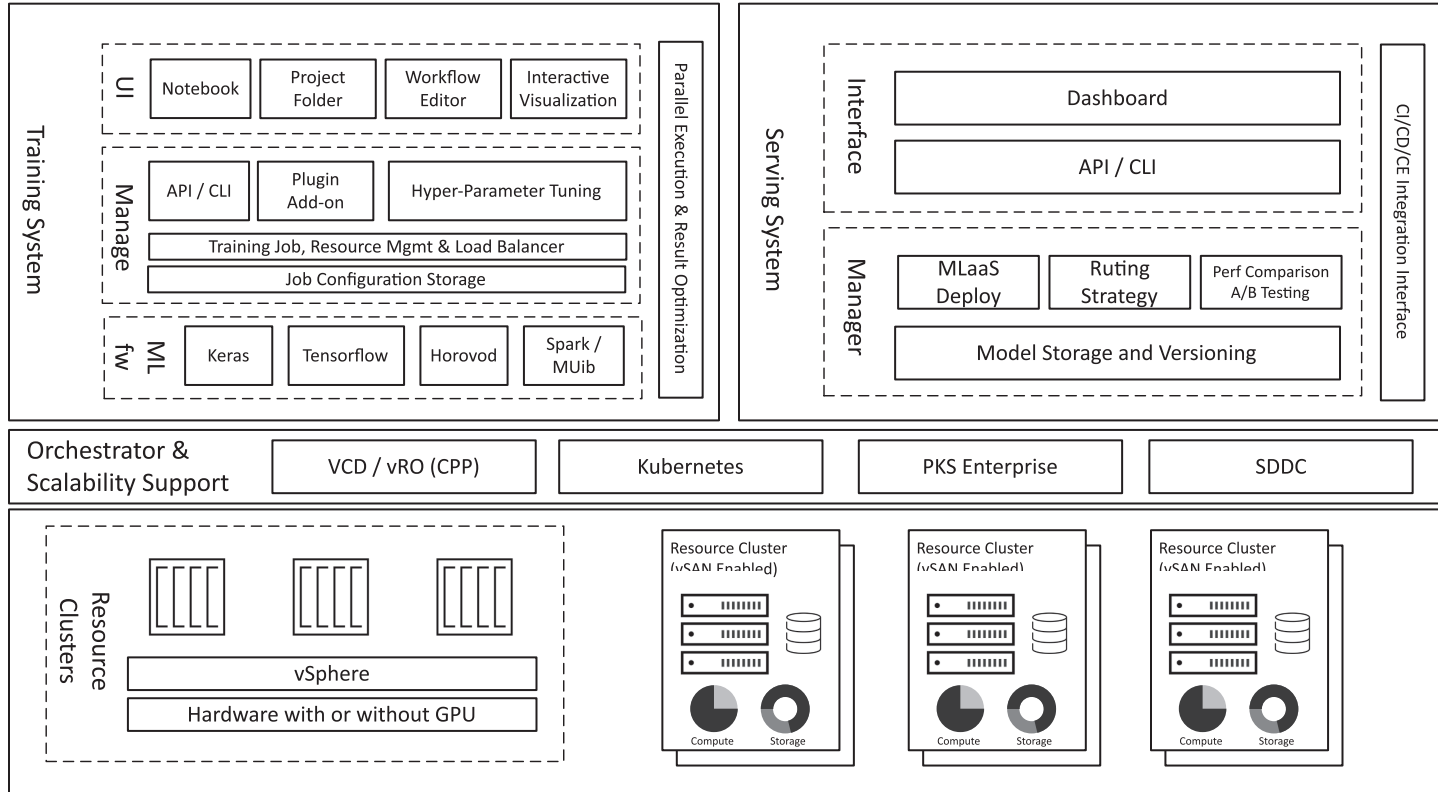


FIGURE 8.33 Distributed Machine Learning on TensorFlow Component Architecture

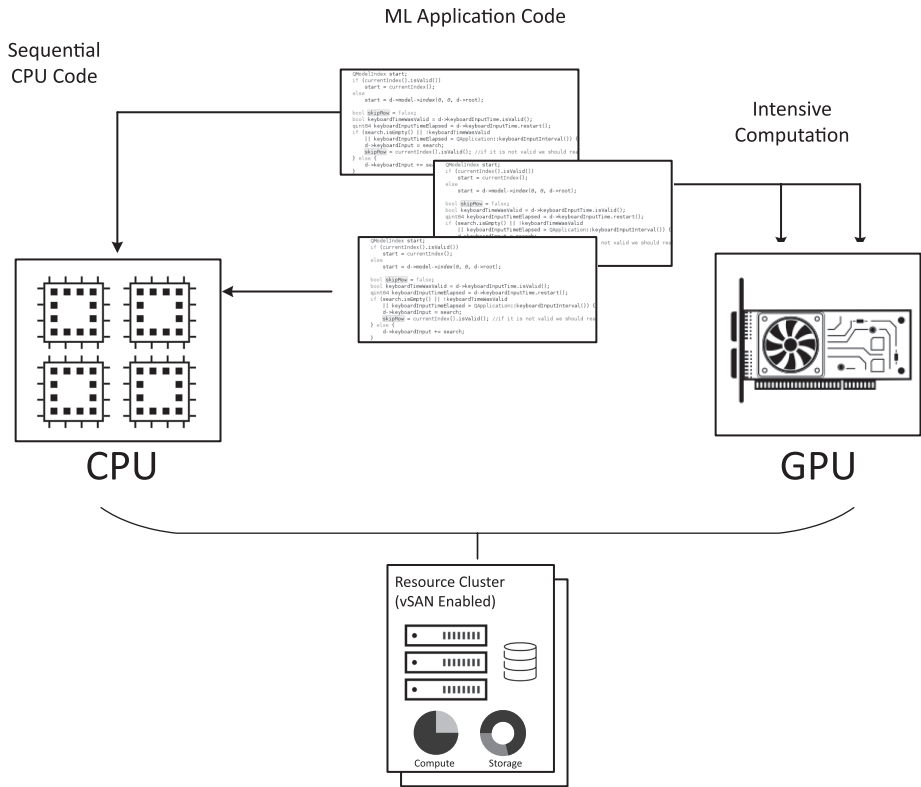


FIGURE 8.34 GPU Computing

MLaaS on the Cloud Provider Platform offers a number of benefits over customer-dedicated solutions, such as:

- Minimizes costs by running the learning phase and the production/inference phase on the same shared infrastructure platform
- Share compute and GPU resources across a variety of workloads using strict resource allocation policies
- Maximizes the ROI of GPU hardware by being able to share resources across customers, with higher utilization and greater availability
- Run Spark or other ML runtime platforms, languages, and tool versions alongside one another in total isolation on a common shared infrastructure
- Easily incorporation new ML technologies onto the platform as they become available
- Run multiple versions of ML application code at one time in complete isolation

To monetize this growing opportunity, cloud providers will need to develop the capability to support the management of these toolsets and

to provide operational support of GPU accelerated hosts and workloads, as GPU support is not configurable from within Cloud Director. These services might range from customers who are just getting started, such as using basic tabular data with an existing environment, to the more sophisticated neural network-based forms of ML.

MULTI-CLOUD MANAGEMENT-AS-A-SERVICE

Many cloud providers wish to deliver services to their customers, which are only available on the large hyperscale public cloud platforms, such as AWS, Azure, or Google Cloud.

This customer requirement can be met in an integrated manor through the Cloud Provider Platform by exposing VMware Cloud Services to customers, allowing your tenants to consume and pay for only the services required and are pertinent for their use case. VMware offer a wide range of Cloud Services to customers and partners, all of which are delivered using as SaaS-based model. For instance, public cloud host monitoring is neither possible nor is it relevant. What is relevant is application performance and ensuring optimized delivery of complex application stacks. This is where VMware Wavefront can be consumed by customers to meet a specific use case.

Another example might be the provisioning of new workloads onto the public cloud. For this use case, customers can consume vRealize Automation Cloud, surfaced using extensibility in Cloud Director, providing public cloud programmable provisioning and automation for multi-cloud through either a SaaS or hosted delivery model.

This portfolio of VMware services is exposed to customers via the VMware Marketplace and can be provisioned, billed and managed by providers through the VMware Cloud Provider Hub, addressed earlier in this chapter.

For vRealize Automation Cloud and its components, customers are able to define, aggregate, and deploy applications across clouds with a simple and integrated user experience, which is convenient for consumers, efficient to deliver, and centralized for cloud providers to manage.

Similar approaches to that shown can be taken with other VMware Cloud Services, to meet the specific use cases of the customer, providing a huge level of flexibility and control of most public and private cloud environments.

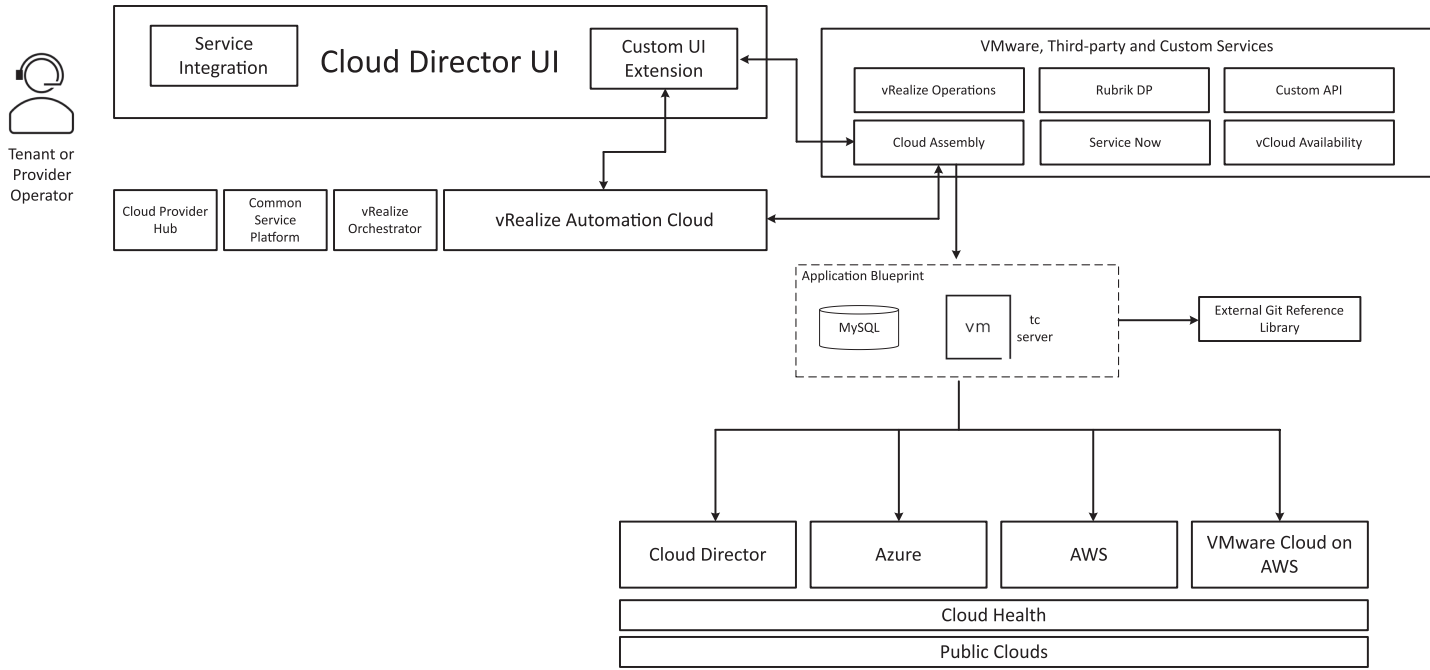


FIGURE 8.35 Multi-Cloud Application Blueprinting through Cloud Director

By allowing customers to consume the VMware Cloud Marketplace, customers can search for, filter through, and select the third-party solution that meets their use cases and can simply click to deploy these solutions directly into their Cloud Provider Platform environment, through a simple, guided interface.

This allows the wide base of VMware customers that are transitioning out of their own data centers into cloud partner environments can continue to use their trusted third-party software, even after they make that move.

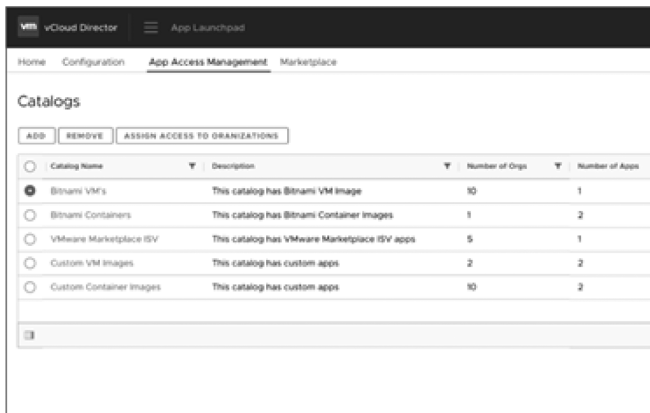
Furthermore, with VMware's acquisition of Bitnami, cloud providers can give their customers access to this rich set of workload services, including developer tools, databases, and network and security solutions. By using the building blocks from the Bitnami catalog, cloud providers can facilitate customer development teams and architects to shorten the end-to-end process of an application's lifecycle, from initial deployment to decommissioning, freeing developers to focus on application development, knowing that they are modeling their application build on trusted, compliant solutions that are production ready.

VMware Cloud Marketplace solves a core problem for cloud providers by providing a seamless way for end customers to consume third-party software. The Cloud Director tenant can access marketplace items through the Cloud Director catalog, while giving cloud providers the ability to customize catalog content and offer specific managed catalogs per customer organization.

APP LAUNCHPAD

App Launchpad is built on VMware Marketplace and provides a simple and automated application deployment mechanism for Cloud Director.

For cloud providers, App Launchpad allows them to build custom catalogs, based on the VMware Marketplace, highlight *Featured Apps*, so they are displayed prominently at the top of the tenant's catalog, and pre-populate configuration parameters for the deployment of the selected applications.



☐ Catalog Name	Description	Number of Orgs	Number of Apps
☒ Bitnami VMs	This catalog has Bitnami VM Image	10	1
☐ Bitnami Containers	This catalog has Bitnami Container Images	1	2
☐ VMware Marketplace ISV	This catalog has VMware Marketplace ISV apps	5	1
☐ Custom VM Images	This catalog has custom apps	2	2
☐ Custom Container Images	This catalog has custom apps	10	2

FIGURE 8.38 App Launchpad Catalog Management—Provider View

Giving the provider the ability to curate catalogs on a tenant by tenant bases, allows for monetization opportunities for cloud providers as well as giving tenants a clear view on which applications are supported by that specific provider. Only the provider can assign these catalogs to organizations that grant access in App Launchpad.

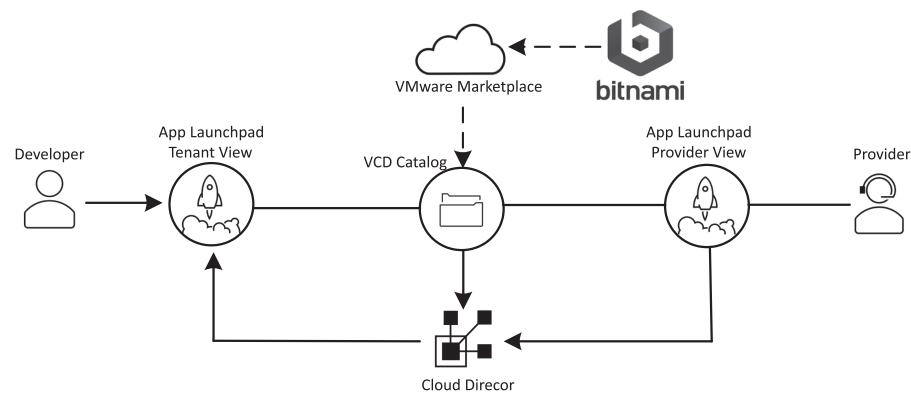


FIGURE 8.39 App Launchpad High-Level Architecture

For tenants, the App Launchpad offers a curated catalog of applications that can be deployed simply through a single click, with deployment configurations, such as networking, firewalling and public IP address parameters are all pre-defined by the cloud provider. As the customer deploying the new application, I do not require any knowledge of the underlying cloud infrastructure, I can simply provision my application and access it.

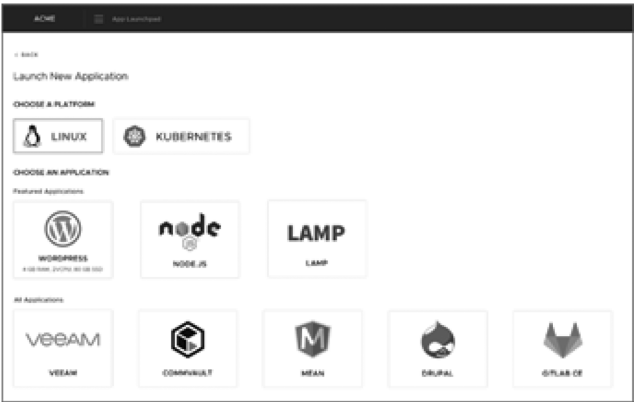


FIGURE 8.40 App Launchpad—Tenant View

App Launchpad significantly simplifies the consumption of the VMware Marketplace, including the Bitnami catalog, by automating and simplifying

application deployment into Cloud Director. This solution helps cloud providers define supported applications, opens new monetization opportunities by defining application support as part of the catalog experience and helps address developer and DevOps use cases for cloud consumers.

DATA PROTECTION & ISV SERVICE ECOSYSEM

Protecting workloads is a critical component of any production environment, therefore, Backup-as-a-Service (BaaS) is a key offering for cloud providers. On the Cloud Provider Platform, this function is driven by partner integrated solutions, developed by Independent Software Vendors (ISVs).

Data protection is the process of safeguarding important information from corruption, compromise or loss, which today must be addressed alongside newer challenges, such as massive data growth, the growth of container-based workloads, and the pivot toward storing data in the cloud.

However, these are just some of the challenge's organizations are facing today. The ever-increasing costs for greater storage capacity, demands for shorter Recovery Point Objectives (RPO), the increasing expectations for the rapid provisioning of backup services, and the complexity of managing backup workflows alongside the sheer volume of data, makes the data protection challenge greater than ever.

Data protection for the VMware Cloud Provider Platform is provided through partner solutions, with many of the most common solutions being supported. In many cases, these solutions also include application-aware backups to provide application-consistent restores, that are a key part of the business continuity plan, data protection, and disaster recovery solution. This level of integration is often not achievable with snapshot-based solutions offered in the public cloud. In addition, we are now seeing multiple vendors of data protection solutions integrate natively into the user interface and API of Cloud Director, in order to provide a simplified self-service experience to customers. This approach allows cloud providers to offer either managed or self-service solutions for data protection, allowing the customer to maintain and manage their own data backups, if that is what they want to do.

For instance, through user interface extensibility, Dell EMC IDPA has provided customers with natively integrated data protection capabilities. This allows the tenant admin to perform tasks such as:

- Policy template definition
 - Schedule and frequency
 - Retention period
- Org. VDC-level per-policy quotas on protected capacity
- Org. VDC-level quota on ad-hoc backups with globally configurable retention period
- Backup repository assignment
- Backup replication configuration



FIGURE 8.41 Dell EMC IDPA Extension for Data Protection

In addition to Dell EMC, other ISVs, such as Rubrik and Cohesity, have also followed the same model of delivering a user interface extension for Cloud Director. The true value of such extensions is that it tightly integrates the data protection solution into the Cloud Director user interface, simplifies the deployment, and provides multiple enhancements to cloud providers and their tenants, including:

- **Customer Experience:** Extends the VCD tenant portal and REST API to provide tenants a single management endpoint for their virtual data centers
- **Lower TCO:** Reduced deployment time and customer management

As illustrated in Figure 8.38, these Cloud Director data protection extensions provide a simple, self-service data protection framework through an integrated tenant portal user interface extension to deliver to customers a multi-tenant data protection service, which can meet a range of customer business requirements, including:

- **Backup:** A periodic data copy meant as a source of last resort for recovery of an environment. Typically held on different media, frequently in a different geography. Often replicated to a secondary site, just in case.
- **Image-Level Backup:** Taken at the VM level, typically based on a snapshot of the virtual machine.
- **Guest-Level Backup:** Taken at the OS or Application level, can include app-specific intelligence such as DB log truncation, quiesce, and memory flush.
- **Disaster Recovery:** Ability to recover from a disaster affecting a fault domain (server, room, floor, site, region). measured in terms of RPO, RTO, and cost.

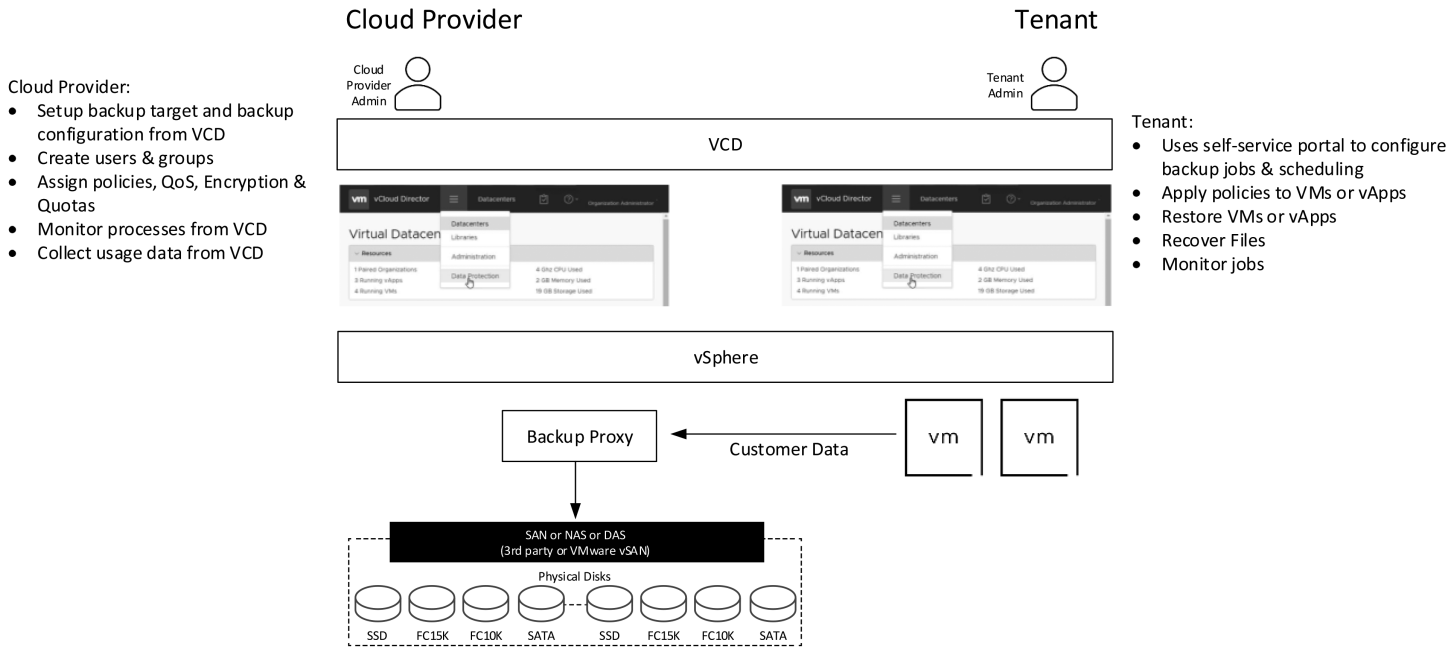


FIGURE 8.42 Multi-Tenant/Multi-Tier Data Protection Service

Chapter 9

Operations and Security

The ability to manage and monitor resources within the Cloud Provider Platform environment is key to a successful operational strategy for cloud providers. Cloud providers are required to implement a range of management tools to address requirements for both the physical elements of the platform in addition to its resources provided to customers, including storage, compute and network layers of abstraction. For cloud providers, delivering management platforms to achieve these objectives is key to allowing their business to maintain a sufficient level of operational efficiency, flexibility and the agility required for their customers to realize the benefits of their enterprise public platform.

It is beyond question that for cloud providers there is now an ever increasing need to provide faster response times and more flexibility in order to support customer's business operations, to ensure their needs are met for a more dynamic and varied platform. The Cloud Provider Platform architecture offers not only the benefit of resource and infrastructure consolidation, but also addresses the challenges with deploying and coordinating multiple different public cloud offerings and allows customers to deliver resources as cohesive environments, while maintaining customer access and service levels. This requires cloud providers to effectively manage shared storage, network and compute resources on their enterprise public cloud platform. In addition, cloud providers must deliver the ability to not only effectively manage resources, gain high levels of control, but also maintain extensive visibility into their infrastructure to provide security and service levels for their customers.

In order to deliver a secure cloud architecture that provides end-to-end management, monitoring, and accounting solutions, which can enable the

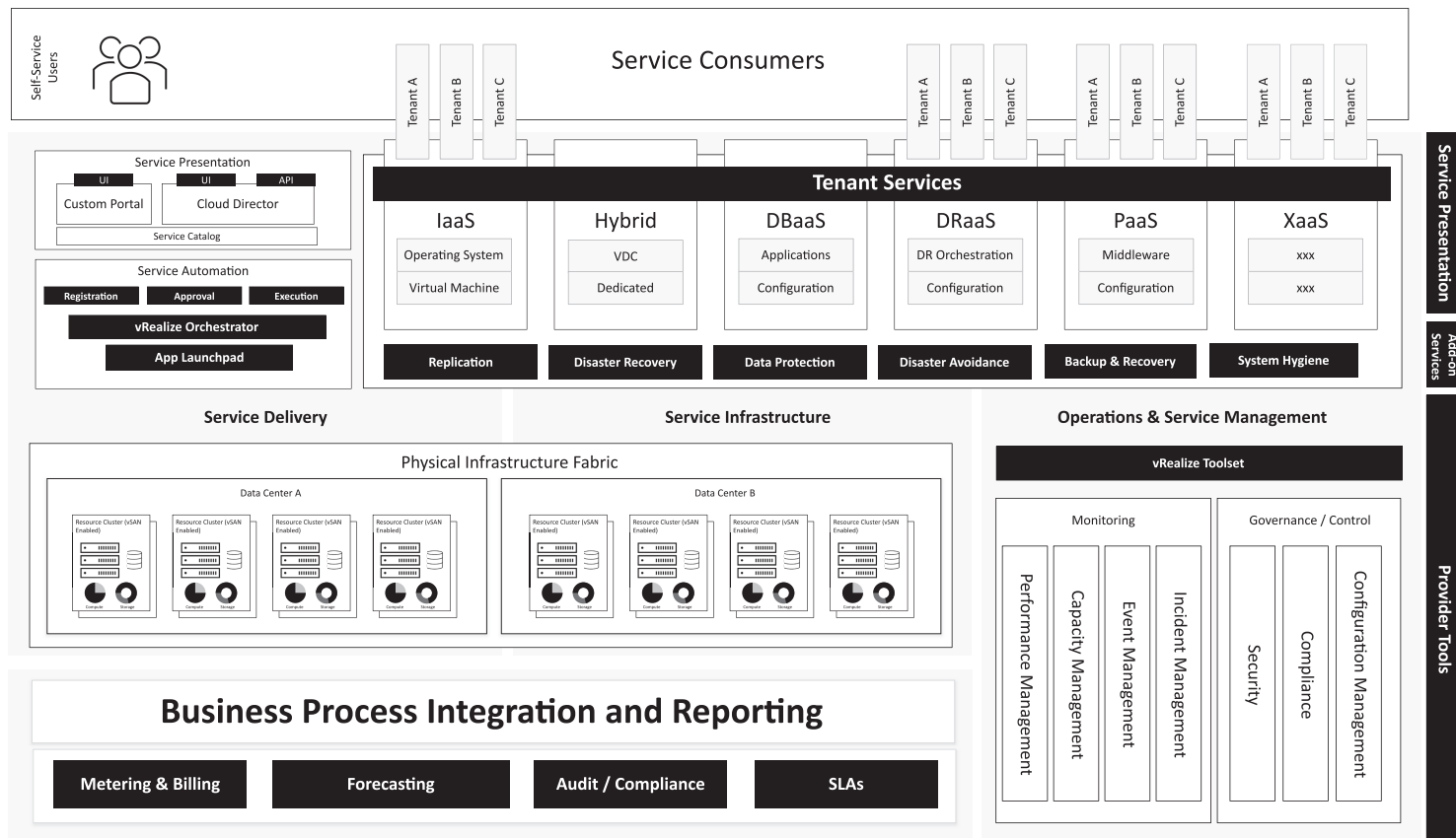


FIGURE 9.1 Cloud Operational Architecture Example

desired level of efficiency, utilization, and availability of cloud service delivery, cloud providers must plan a comprehensive operational model with toolsets that are wide ranging in their scope while maintaining a consolidated view of all aspects of the platform. This must typically be done while also:

- Reducing CapEx and OpEx—achieved through improved data center performance, health, operations and capacity management
- Delivering a simplified and optimized infrastructure
- Providing proactive analysis, through customized monitoring and altering

DEVELOPING A CONSISTENT OPERATIONAL AND GOVERNANCE MODEL

For monitoring, alerting, capacity management and operational automation, vRealize Operations Manager (vROps) and vRealize Log Insight (vRLI), can provide the core solution for the Cloud Provider Platform architecture.

VMware vRealize Operations Manager and VMware vRealize Log Insight offer a *single pane of glass* view to monitor the state of the Cloud Provider Platform, and analyze the cloud infrastructure in real time. vRealize Operations connects to the Cloud Provider Platform environments via vCenter Server and Cloud Director to provide hierarchical information on all the constructs in the multi-tenanted infrastructure, from vCenter Servers and ESXi hosts to customer virtual machines and their storage and network tiers.

It can monitor utilization, provide performance reports and run analytics. For instance, Cloud providers can:

- Monitor the collective health of resource pools mapped to tenants
- Determine when to add additional resources to scale Provider VDCs
- Set alerts for when resources are depleting and reach a certain threshold
- Get a bird's eye view of all the inventory in their infrastructure

vRealize Log Insight collects application and system logs via syslog and provides analytic capabilities to cloud providers within a dashboard. In addition to syslog, Log Insight can import logs from systems, applications, and operating systems via SMTP or web, before parsing and indexing them.

Cloud provider can then query indexed logs to determine patterns that can provide insight into the behavior and state of the Cloud Provider Platform. From these logs, cloud providers can capture alerts that can be difficult to be identified via other operational monitoring mechanisms. This allows cloud providers to capture issues that are missed by vRealize Operations alerts, because they are occurring between application and infrastructure layers. The indexed logs can also help run post incident and root-cause analysis and are critical for security teams for the purposes of auditing and verifying

compliance requirements. Cloud providers can generate dashboards for various use cases, including:

- Unsuccessful login attempts by tenant/system users
- Database connectivity alerts
- The number and source of rejected connections on an Edge firewall

The following two sections will drill into the details of each of these tools in order to demonstrate their value as components within the Cloud Provider Platform.

vRealize Operations Manager: vRealize Operations Manager can be used within the Cloud Provider Platform to collect and analyze all metric information from a range of VMware and integrated third-party components. Based on the millions of metrics collected, it will proactively identify immediate problems, provide long-term analytics, and identify trending and patterns across the entire infrastructure. vRealize Operations Manager can include the following key functionality for cloud providers:

- **Operational Dashboards:** Providing at-a-glance views into health, risk, and efficiency of the virtual infrastructure
- **Health and Workload Views:** Helping to identify anomalies, faults, and stressed workloads that can impact the infrastructure performance and health
- **Workload Details View:** Provides in-depth analysis of the causes of performance and health issues on virtual machines, hosts, data stores, and vSphere clusters
- **Adapters:** Provide visibility into other applications and products

vRealize Operations Manager can also be used to provide custom dashboards, as well as custom alerting based on specified conditions or incidents. As outlined in Figure 9.2, in a pod architecture, such as that described in Chapter 4, vRealize Operations can employ Remote Collectors in each pod to process data in that pod, in order to reduce the load on the analytics cluster. If necessary, remote collectors can be grouped for high availability. This information is presented to the cloud provider's operations users through views, reports, and dashboards. With the Tenant App, described in detail in Chapter 8, these metrics can be presented to tenants via the Cloud Director HTML5 user interface. If required, alerts can be forwarded via SNMP to an upstream operational solution.

vRealize Operations Manager Design: vRealize Operations Manager is deployed in the form of a virtual appliance. Each instance of the appliance can be configured to perform one of the following roles within the

complete vRealize Operations Manager instance/cluster. These building block roles are:

- **Master Node:** Mandatory first node in the cluster, or single standalone node in smaller deployments.
- **Master Replica Node:** Used to provide high availability to the master node role.
- **Data Node:** Used for scalability.
- **Remote Collector Node:** Used to overcome data collection issues, such as limited/poor network performance across the enterprise network.

A single instance of vRealize Operations Manager puts some load on vCenter Server when it collects virtual infrastructure information. While these numbers vary based on the size of the inventory, the following provides some indicative metrics around impact on vCenter Servers:

- Increase in CPU load on the vCenter Servers by 5 to 10 percent once every 5 minutes for approximately 30 to 60 seconds (by default, vRealize Operations Manager collects performance data every 5 minutes).
- Increase in transient memory (40 to 50 MB) used on the vCenter Servers, once every 5 minutes.

vRealize Operations Federated Deployment: As illustrated in Figure 9.3, vRealize Operations allows cloud service providers to scale operations into multiple geographies and run federated operations across regional or global environments.

With the vRealize Operations Federation Management Pack you can unify your multi-site vRealize Operations Manager deployment into a single pane of glass. You can instantiate a deployment of vRealize Operations Manager with the capability of receiving key metrics for specified objects from other vRealize Operations Manager instances in the environment. The management pack allows you to unlock the following use cases:

- Provide a summary of performance, capacity, and configuration to Senior Executives and Administrators across all Cloud Provider Platform environments in a multisite deployment.
- Provide a unified view of events triggered across the Cloud Provider Platform into a *single pane* for making it easier for NOC or the Helpdesk to initiate action.
- Ability to create a data warehouse where KPI metrics can be stored for data archiving and reports.
- Ability to provide summarized views of health and configuration of the Cloud Provider Platform stack. This includes core applications such as VMware vCenter Server, VMware NSX, VMware Cloud Director and VMware vSAN.

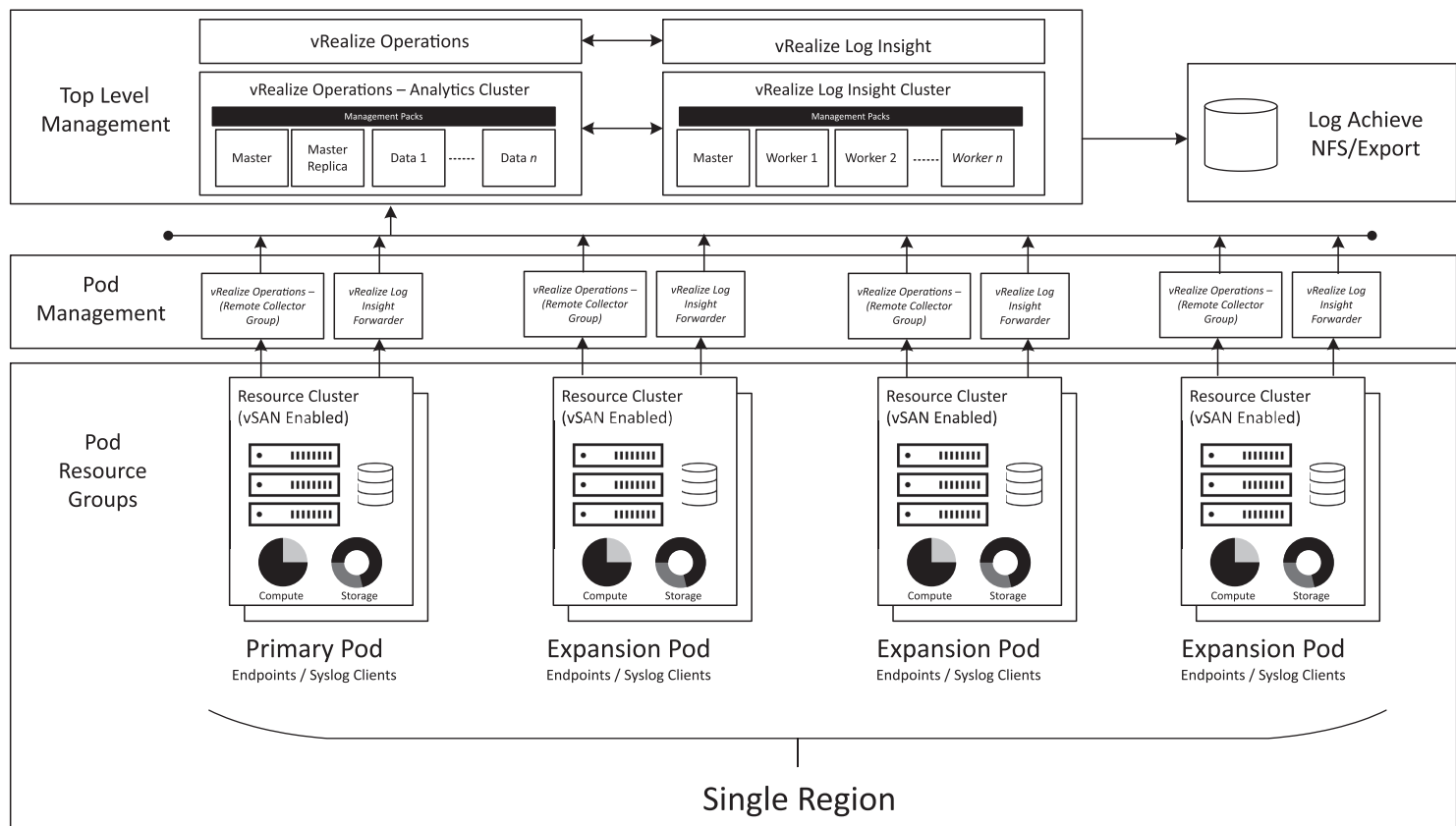


FIGURE 9.2 High-Level Monitoring and Altering Architecture

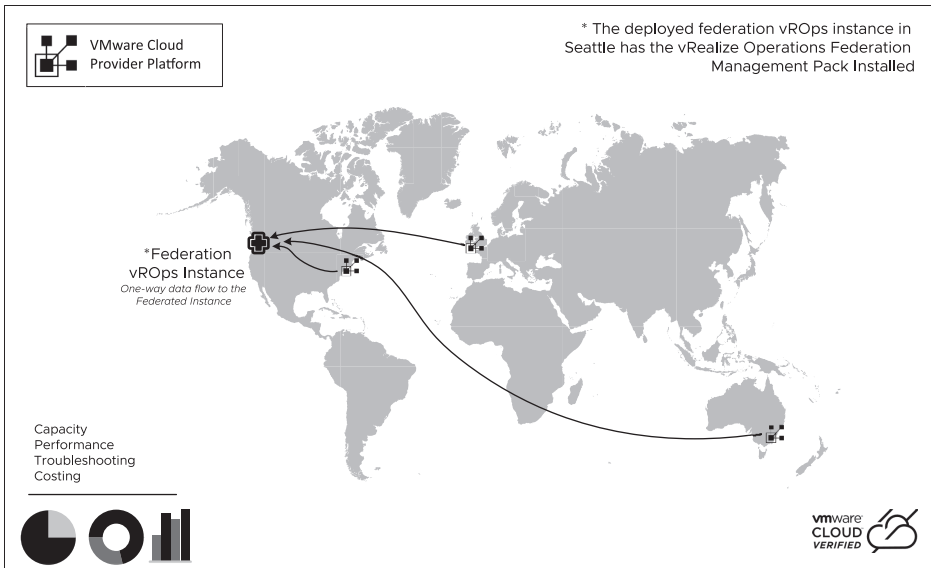


FIGURE 9.3 Cloud Provider Platform Federated Deployment Architecture

As illustrated in Figure 9.3, the use of federated vRealize Operations allows for single screen reporting across multiple platform deployments in multiple geographic regions, reducing the complexity of monitoring these deployments individually. This approach simplifies management for global organizations who have geographically distributed deployments, facilitating a follow the sun support model with a central single point of truth.

BUILDING A LOGGING SERVICE

In addition to vRealize Operations Manager, the Cloud Provider Platform will also need to provide a Logging Service for internal provider operational purposes, security auditing and compliance and also, for customers who require access to logs as part of their service requirements.

vRealize Log Insight gives cloud providers the ability to consolidate logs, monitor and troubleshoot Cloud Provider Platform and third-party infrastructure, and perform security auditing, compliance testing, log querying, aggregation, correlation, and retention. The vRealize Log Insight virtual appliance includes a syslog server, log consolidation tool, and log analysis tool that will work for any type of device that can send syslog data. vRealize Log Insight administrators can also create custom dashboards based on saved queries that can be exported, shared, and integrated with Cloud Director, through extensibility, and VMware vRealize Operations Manager, to provide a uniform approach to dashboard monitoring and operational management.

vRealize Log Insight provides a syslog collector for the entire Cloud Provider Platform, including, but not limited to ESXi, vCenters and Cloud Director via the appropriate content packs.

Security Monitoring

The monitoring of user activities is an important part of the overall system security. The cloud provider will have rules governing who is allowed to access and make changes to systems. The utilization of an audit log will enable security teams to verify compliance with rules, detect any violations and initiate remediation activities as well as detect attempts to gain illegitimate access to the system.

A more comprehensive view of the audit log, as well as, its long-term retention is to be achieved through the use of the vRealize Log Insight Cluster and archiving of log data to NFS mounts, as illustrated in Figure 9.2.

Logging Service

As an example, by default an ESXi host, logs on an average day with default logging settings, ~250 MB of data per host. Therefore, even with a relatively small number of hosts, querying this log data when troubleshooting a problem quickly becomes very difficult, and correlating the information even more so. Then, as a cloud provider, and you are maintaining hundreds or even thousands of hosts across multiple geographically dispersed data centers in different regions, effective log management becomes paramount.

On the Cloud Provider Platform, all actions, including parent actions taken on vCenter and child actions run on ESXi hosts for user CORP\smithj (for instance), can be tracked in vRealize Log Insight. Matching a username to actions provides accountability, auditing, and forensics capabilities and is a key requirement of compliance objectives.

Data Archiving

vRealize Log Insight supports the use of NFS mounts for data archiving, which might be necessary to include in a design if the cloud provider has requirements for long-term auditing, compliance, or data retention. The process of data archiving preserves old logs that would otherwise be removed from the vRealize Log Insight appliance due to storage limitations. Where the customer's business requirements are likely to demand it, plan to include vRealize Log Insight data archiving in the design.

Under normal operating conditions, vRealize Log Insight should never run out of disk space, because it checks the current storage status every minute, and if there is less than three percent of disk space remaining, it retires old *data buckets*. If archiving is enabled, vRealize Log Insight archives the data buckets before retiring them.

Part of any vRealize Log Insight design should include archiving calculations that outline the required NFS disk space needed each year to meet the provider's regulatory, compliance, or data retention requirement.

vRealize Log Insight itself does not manage the NFS mount used, because this is maintained through the storage management tools typically provided by the cloud provider's storage vendor. To receive system notifications, you must configure vRealize Log Insight so it can send out an email alert when the NFS mount is getting low on available space. More importantly, if the NFS mount runs out of space or is not available for a period of time, which is longer than the retention period of the vRealize Log Insight appliance, data ingestion will stop until the NFS mount has been restored, and enough free space becomes available for archiving. Alternatively, you can disable archiving temporarily.

It is noteworthy that log events that have been archived are no longer searchable. To search archived logs, you must first import the logs into a vRealize Log Insight instance. Typically, you should not import archived data into the existing production vRealize Log Insight instance. This is since data from the archive import will force the oldest ingested data to be deleted to make room, which might affect the ability to maintain the desired retention period for events. For this reason, VMware recommends that you import archives into a dedicated vRealize Log Insight instance that is not ingesting message events, and is dedicated to this function.

After data archiving is configured, it is the operational team's responsibility to verify that the archive destination is cleaned up and maintained and does not run out of space. vRealize Log Insight does not have any mechanism to manage or monitor the NFS destination, and will continue to attempt to archive data, even after the destination becomes full. Typically, a simple cleanup script can be leveraged to manage this.

Extending vRealize Log Insight Services

vRealize Log Insight can accept, process, and query logs from non-VMware sources in addition to those from VMware infrastructure sources. vRealize Log Insight Content Packs are plug-ins with information about syslog messages from external sources such as storage systems, Server Systems, Windows Servers, or Windows applications such as Active Directory or SQL.

Out of the box, vRealize Log Insight includes only the vSphere Content Pack. VMware or third parties can create Content Packs with information about specific events from other syslog sources that can be used by system administrators, operational teams, and engineers. The content pack itself is made up of information that includes dashboards, fields, aggregations, alerts, and queries.

To forward logs from Microsoft Windows Servers or Microsoft Windows-based applications such as Active Directory or Exchange, the

vRealize Log Insight Windows agent must be installed on each source operating system, allowing messages from Windows event channels and log files to be forwarded to the vRealize Log Insight server. Linux operating systems typically include a syslog agent. If not, you can usually install one. The most common syslog agents found on Linux operating systems are `rsyslog` and `syslog-ng`.

Management Environment Logging

As we addressed in Chapter 3, the management pod will host the virtual machines that provide management services to the customer workloads within the Cloud Provider Platform. For example, services such as Active Directory, vCenter Server, VMware vRealize Operations Manager or VMware Cloud Director components can all reside in the management pod.

Figure 9.4 illustrates a VMware Cloud Provider Platform-based network operations center (NOC).

Cloud providers offer a wide variety of different service offerings to their customers. This extensive list of potential offerings changes regularly in the fast-moving cloud space and the market will continually develop and evolve to meet the needs of a whole range of industries. For example, one cloud provider might specialize in providing services exclusively designed for the gaming industry or financial services, while another might target secure government customers for their service offerings.

That being the case, the design requirements for a specific service offering to provide access to host and other device syslog information directly to the service consumer for troubleshooting, security, or simply informational purposes will vary. The specific design required to meet these use cases would depend very much on other design factors for the explicit use case. For example, an architect would need to consider:

- Does the customer have a dedicated hardware infrastructure?
- Does the customer have a dedicated vCenter Server?
- Does the architecture also include a requirement for vRealize Operations Manager?

Based on questions like these, it is possible to conceive of a number of different design scenarios. The remainder of this section details two possible scenarios for which a cloud provider delivers services.

Scenario 1 employs the built-in functionality provided by most devices, that is, ability to send syslog messages to two or more independent targets. In this example, ESXi hosts are configured to send syslog messages to two separate instances, consumer and provider, with the consumer instance possibly residing within the customer's Org. VDC. Note that some devices support only a single syslog target.

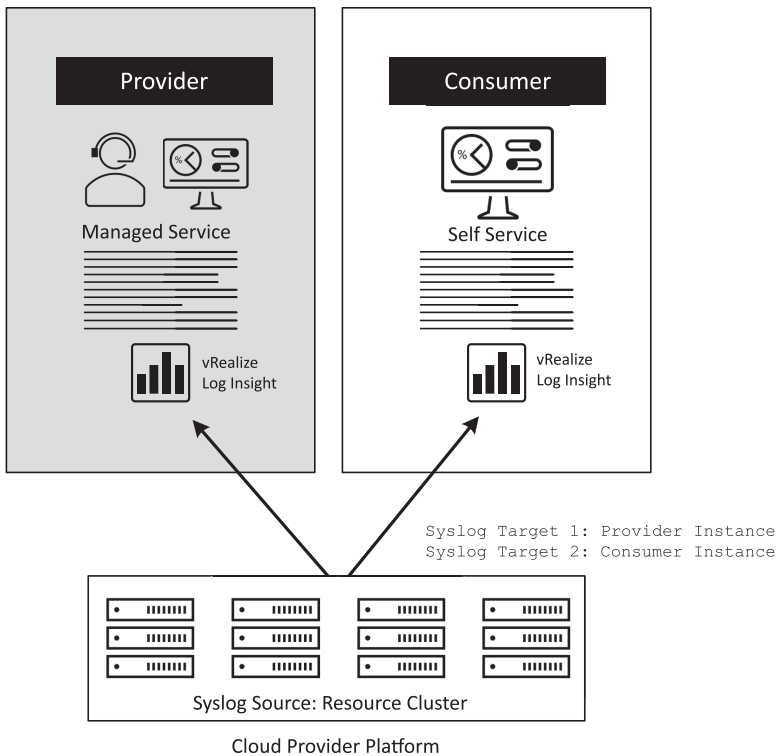


FIGURE 9.5 Multitenant Scenario 1

In Scenario 2, illustrated in Figure 9.6 vRealize Log Insight Forwarders are being employed to forward syslog event messages to the secondary consumer instance. This has the advantage of allowing the provider to filter messages before they are forwarded.

In addition to these two scenarios, vRealize Log Insight supports multi-tenant logging, in which you can control who sees what data. Through the built-in role-based access control (RBAC) mechanism, it is possible to control access to data sets, dashboards, interactive analytics, and administrative tasks. By configuring data sets, a static filter can be assigned to roles. Defining one or more static filters to restrict data helps provide the user interface restrictions that control what people can see and do in vRealize Log Insight.

Cloud Provider Platform Management Component Logging

As with any cloud infrastructure, the Cloud Provider Platform is made up of a number of integrated components. As previously highlighted, vRealize Log Insight is able to analyze data from a wide variety of sources out-of-the box through its content packs. It provides a centralized means of monitoring an

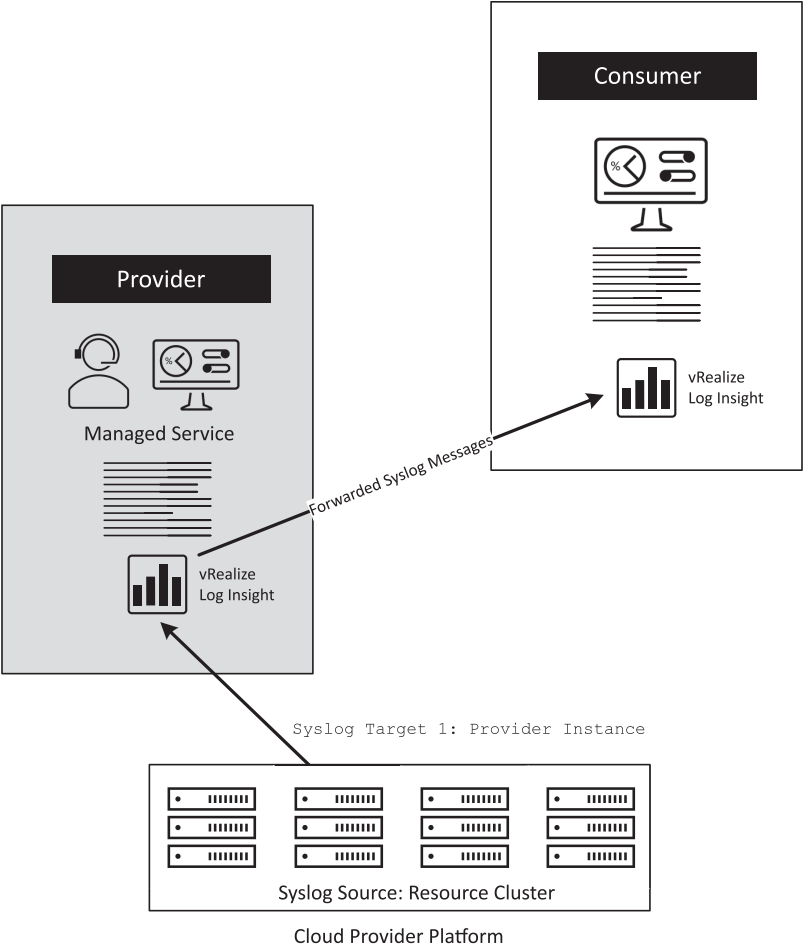


FIGURE 9.6 Multitenant Scenario 2

entire cloud stack made up from components not only from VMware, but also from a wide variety of other vendors.

NSX Edge Gateway Logs: In the Cloud Provider Platform architecture, the IP addresses of syslog servers for NSX Edge gateways are configured centrally through Cloud Director. This is the same for all edges that might be NSX Edge gateways or edges on virtual application networks. In Figure 9.7, one of the external NSX Edge gateway interfaces is connected to the syslog network that routes to the syslog servers. In this architecture a tenant can also deploy their own vRealize Log Insight appliance to a dedicated Org on a virtual data center network. The subnet and IP address of the tenant's syslog server must match the secondary syslog address of the NSX Edge gateways. This allows

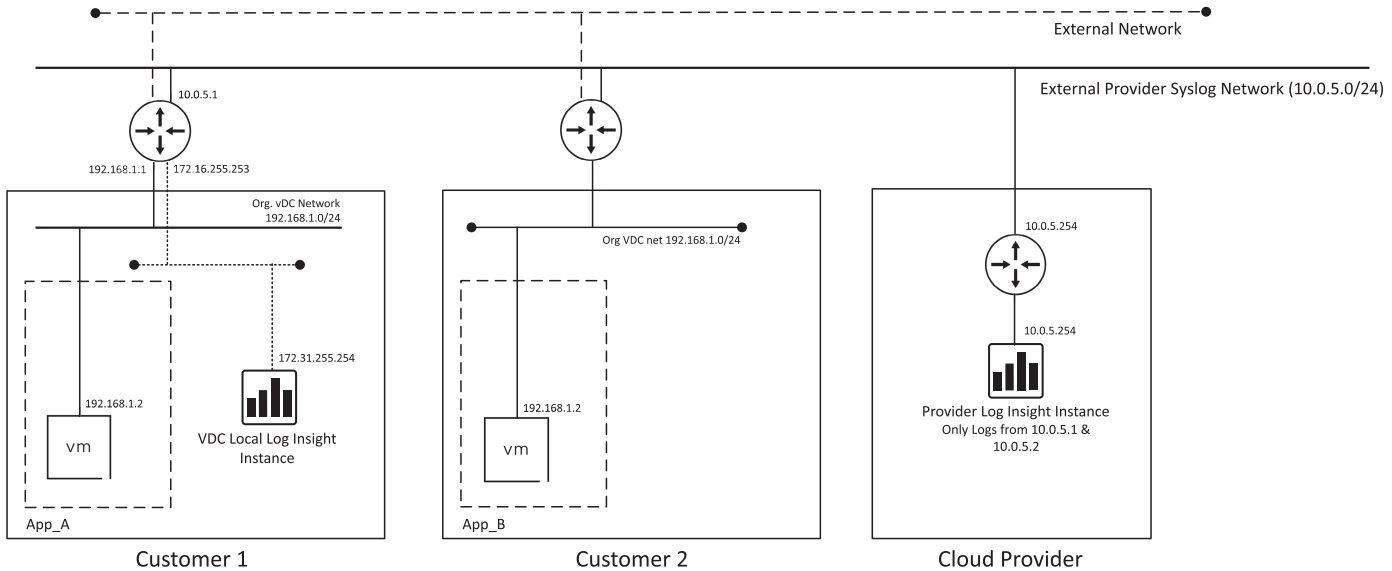


FIGURE 9.7 Edge Gateway Provider and Tenant Syslogs

the tenant to see the NSX edge logs in real time, which is very useful for troubleshooting.

The vRealize Log Insight server adds the IP address of the syslog sender, which is the NSX Edge gateway network logging external interface, making filtering of log data straight forward.

Sample Syslog Design Scenario

This design scenario illustrates just a single possible deployment configuration. In the example design the NOC is located in London alongside one of the data centers. The two other primary and secondary data centers are located at remote geographically dispersed sites. In this design, syslog data must be secured end-to-end between source and destination. Sensitive syslog messages are secured on the local network and MPLS network using SSL, and secured across the public internet using a VPN tunnel connection.

The cloud provider has more than 1500 devices logging and requires a two-node vRealize Log Insight cluster to meet the ingestion requirements of the infrastructure. To allow for host failure, without causing interruption to log message ingestion, a three-node vRealize Log Insight cluster is configured.

The three data centers send log messages to a three-node vRealize Log Insight cluster that is employed as part of the Cloud Provider Platform, including vRealize Operations Manager components.

Each data center houses a single syslog aggregator to forward logs to the central NOC in London. The external load balancing is provided by an NSX Edge HA pair device for syslog messages to be distributed evenly across the vRealize Log Insight nodes.

Log Intelligence

For those cloud providers taking an *asset light* approach to service delivery, consuming cloud-based management components such as Cloud Director Service and VMware Cloud on AWS, a suitable SaaS-based logging offering can be consumed alongside or independently of vRealize Log Insight or third-party syslog solution.

VMware Log Intelligence offers unified visibility across the Cloud Provider Platform, including infrastructure provided via VMware Cloud on AWS. This architecture can provide deep operational insights and faster root cause analysis across the various integrated platform components. Log Intelligence adds structure to unstructured log data, provides dashboards and delivers machine learning-based intelligent grouping for faster troubleshooting in the same model as vRealize Log Insight.

Audit logs can be collected in Log Intelligence for all components within the platform, it is just matter of enabling them. For instance, for the NSX Data Center firewall, it's matter of enabling it in the Log Intelligence user interface

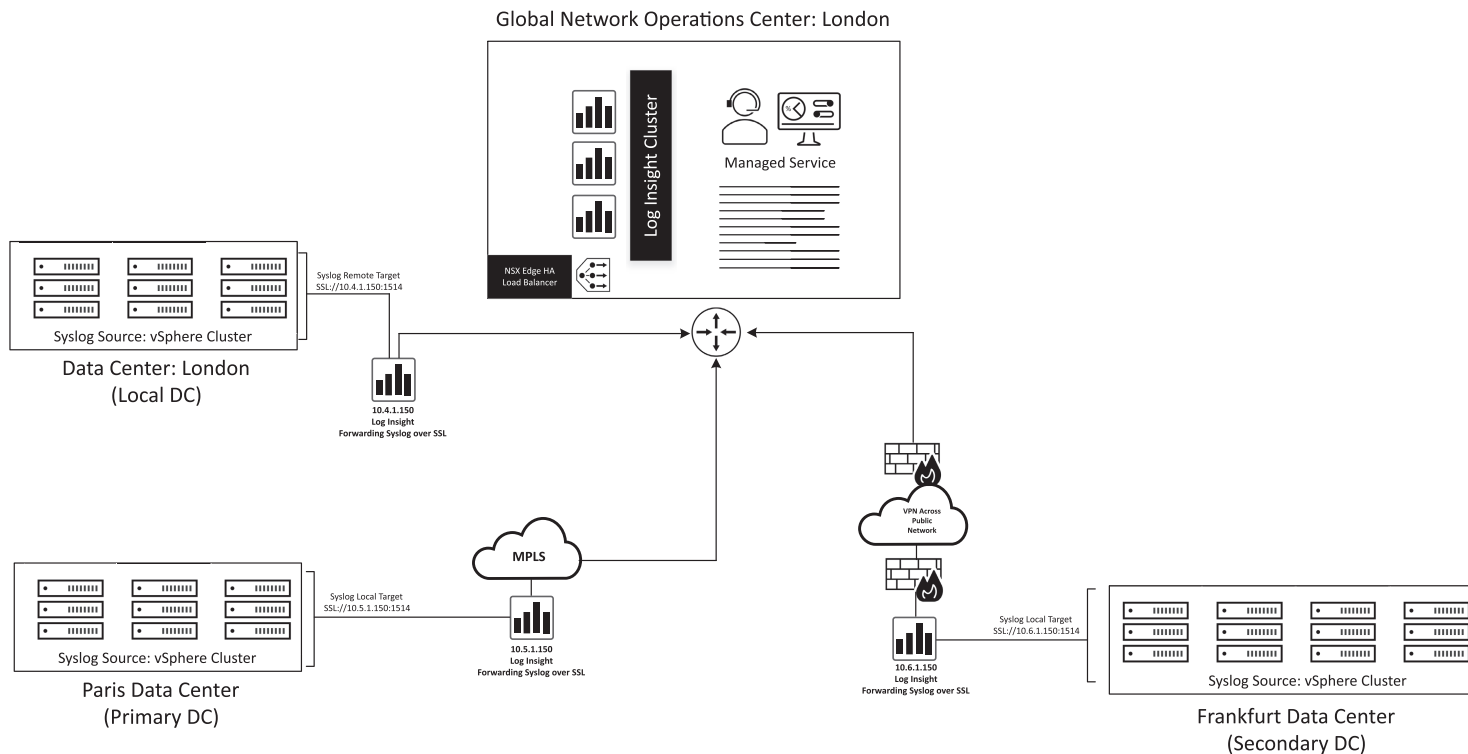


FIGURE 9.8 Design Scenario

and admins will automatically start seeing firewall logs, allowing them to audit, monitor and troubleshoot the environment.

You can save the queries on the Shared or Private Dashboards or enable Alert Definitions to send email or webhook notifications. Note that any On-Premises endpoint will need a Cloud Proxy deployed in the Cloud Provider Platform environment, which log intelligence communicates with to forward logs.

SECURE MULTITENANCY

As with any new IT implementation, particularly where it is a public-facing, security is absolutely critical. Any security vulnerability or risk exposed by the Cloud Provider Platform infrastructure could have a serious impact on the reputation of the cloud provider and the IT teams charged with implementing such platforms.

Security is one of the key components of the Cloud Provider Platform architecture, with tenant isolation being a key deliverable of the platform. Therefore, security is applied to the architecture at the storage, compute and network layers, creating 360 degrees of protection around the platform. Every aspect of the platform puts security considerations at the heart of each design decision, with the key security design factors providing:

- A secured cloud provider cloud platform, for delivering a wide range of business applications to end customers of varying size and profile
- Native Cloud Director Tenant Isolation
- Cloud Director Advanced Networking through native NSX integration

Another aspect of cloud provider operation's is ensuring and maintaining secure multitenancy within a shared platform. Enabling cloud providers to develop a platform to provide secure isolation while still delivering the management and flexibility benefits of shared resources is the key value proposition for the Cloud Provider Platform. Cloud providers must ensure all customer data, communication, and application environments are securely isolated from other tenants. This separation must ensure complete and secure tenancy with customers having no visibility of each other.

In the following sections, we address some of the security measures that can provide protection in complex Cloud Provider Platform, multi-tenanted environments.

The Cloud Provider Platform requires comprehensive ESXi Host Hardening. To provide an ESXi security baseline, cloud providers should consider their requirements for hardening the hypervisor. VMware guidance on security hardening and the recommendation level depends on the rating that corresponds to the operational environment in which it is to be applied. Each cloud

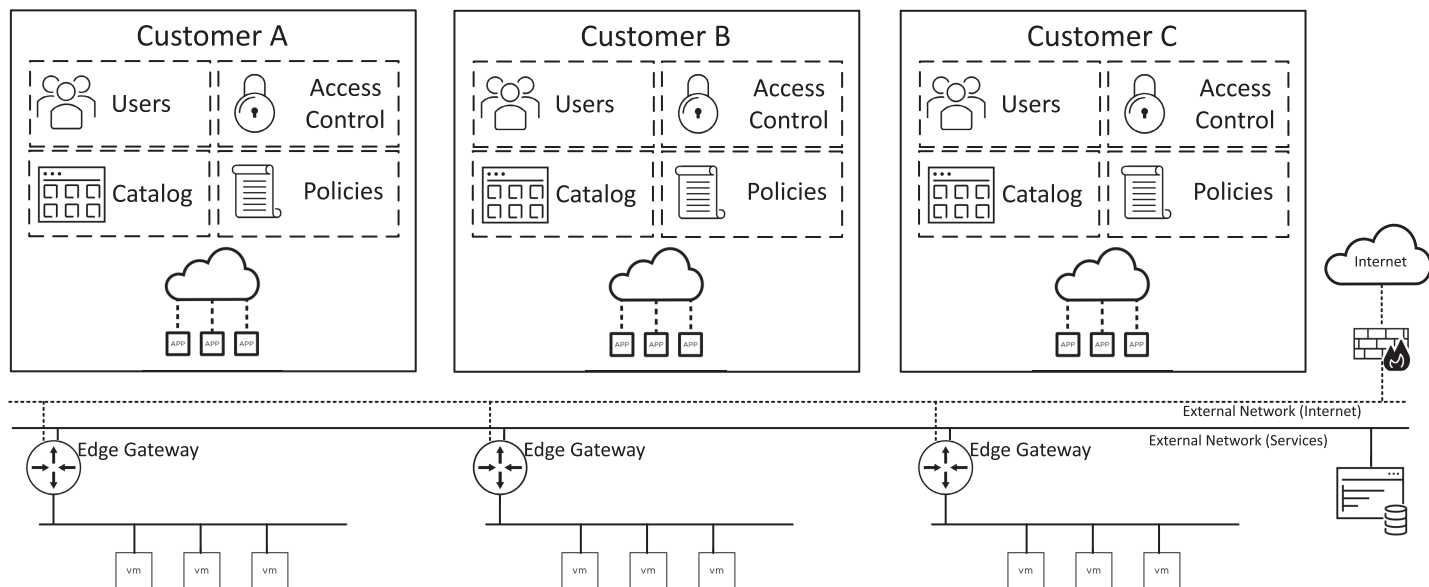


FIGURE 9.9 High-Level Secure Multitenancy Design

provider will need make their own determination as to the applicability of each level. VMware provides the following three baseline levels for hardening:

- **Enterprise Security Level 1 (Enterprise L1):** This includes most enterprise production environments. The recommendations are meant to protect against most security attacks and provide protection of confidential information to the level required by all major security and compliance standards.
- **Specialized Security Level 2 (SSL2):** This includes environments that are particularly susceptible to targeted attacks. Examples include: Internet-facing hosts and internal systems with highly confidential or regulated data.
- **Specialized Security Level 3 (SSL3):** This represents unique and specialized environments that have some aspects that makes them especially vulnerable to attacks. Recommendations at this level might result in loss of (ease-of-use) functionality or purposefully cause the inability to use certain features. Careful consideration must be given to determining the applicability of these recommendations, including the possibility of using alternate compensating controls.

The VMware security hardening guides provide detailed guidance for customers on how to deploy and operate VMware products in a secure manner. For more information on ESXi host hardening, refer to the VMware Security Hardening Guides.

The Cloud Provider Platform addresses concerns around secure cloud architecture by developing a secure, isolated multi-tenant solution at every layer: storage, network, and compute. At the compute layer, vSphere provides built-in security for protecting data, infrastructure and access. This works in conjunction with policy-driven security to protect unauthorized data access both at-rest and during vMotion operations.

At the network layer, NSX Data Center delivers micro-segmentation and granular security to the individual workload, enabling a fundamentally more secure data center. In addition, workload security policies travel with the workloads, independent of where it resides in the environment.

At the storage layer, VMware vSAN offers data-at-rest encryption at the vSphere cluster level. this storage encryption is built for compliance requirements and offers simple key management with support for all KMIP compliant key managers, such as CloudLink, Hytrust, SafeNet, Thales, and Vormetric. Storage is fully abstracted to customers into pools of resources, without the ability to browse datastores.

This architecture enables cloud providers to securely offer their users unprecedented control over their entire application environment. The combination of isolation technologies, combined with extensive management flexibility, delivers all of the benefits of cloud computing for cloud providers to confidently provide high levels of security and compliance focused services for multi-tenant customer and consolidated application environments.

This secure cloud architecture, the result of years of VMware product development and engineering, offers the best products and features to create a public cloud service that can realize the efficiencies of shared infrastructure while still offering the security, flexibility, resilience, and ease of management of a traditional IT deployment. As previously outlined, at the core of the Cloud Provider Platform is VMware Cloud Director, which delivers a number of the key constructs required by cloud providers to enable the secure multitenancy outlined above.

VMware Cloud Director Security

Cloud Director was developed and tested with multi-tenancy, scalability, and security concerns in mind, but the way it is architected, designed and deployed can have a significant impact on the security. Cloud Director does not provide direct visibility or access to most system-level resources that include physical host information, such as CPU type, memory locations, physical storage locations, vSphere access, IP addresses, and MAC addresses, to any of the non-administrator users of the Cloud Provider Platform.

Cloud Director requires HTTPS encryption for all server to client communication. Port 80 is open but will only redirect the connection to the secured connection port. The Cloud Director user interface and API is secured by SSL over HTTPS (port 443/TCP).

Designing Host Security for Multi-Tenanted Clouds

The hypervisor is the construct underlying the integrity of the tenant's virtual machines. Protecting it effectively is of the utmost importance. Any compromise of the hypervisor could seriously affect the virtual machines it hosts, leading to performance issues, data corruption, data loss, or even data exposure. In addition, because these attacks can occur below the guest operating system, they can be challenging to detect and have a much larger impact.

For instance, the effect of a Denial of Service (DoS) attack against an ESXi host is magnified because it affects all the virtual machines running on it. On a classic physical server, a DoS attack that monopolizes the system's CPUs affects only that host. However, in a virtual infrastructure, that same attack against an ESXi host can starve the physical CPUs of resources, affecting all virtual machines hosted on that hypervisor. In addition, if that host is part of a vSphere cluster and virtual machines are relocated as a result of the attack, this might impact all hosts, and in turn, all virtual machines in the cluster, leading to serious performance degradation of the entire infrastructure.

The next consideration is protection of the virtual machines. Protecting just the hypervisor is not enough. The virtual machines themselves must be secured in the same way as a classic physical server. While on the Cloud Provider Platform, virtual machines are isolated in terms of having a separate guest instance on top of a dedicated VMM, they still communicate with each other over the network, the same way other hosts do. Unfortunately,

traditional mechanisms, such as physical firewalls, would have limited effectiveness, because much of the intercommunication takes place on the hypervisor itself. Fortunately, VMware NSX mitigates many of these risks.

Secure Communication

Several of vSphere capabilities require components to communicate over management, or other dedicated networks. Because networks that provide management communication provide direct access to core functionality, VMware recommends that ESXi host management communication only occur over a dedicated and isolated network segment, illustrated in Figure 9.10.

Communication between vCenter Server and each host is encrypted by default using standard X.509 version 3 certificates. However, other traffic types described in this section are not encrypted by default, making it necessary to evaluate any requirement on data traversing the network for these components to be made secure and unreadable.

Other traffic flows that should be given due consideration from a security perspective include vSphere vMotion traffic (which is not necessarily tied to a Layer 2 data center LAN), and storage traffic, such as iSCSI, NFS, replication, or vSAN.

These traffic types should be isolated and strongly secured from all other traffic flows going to and from virtual machines. The most typical way of achieving this is by isolating networks through the creation of separate VLANs for each traffic type. This way, virtual and physical switches can be shared as long as traffic remains logically isolated for tenant virtual machines.

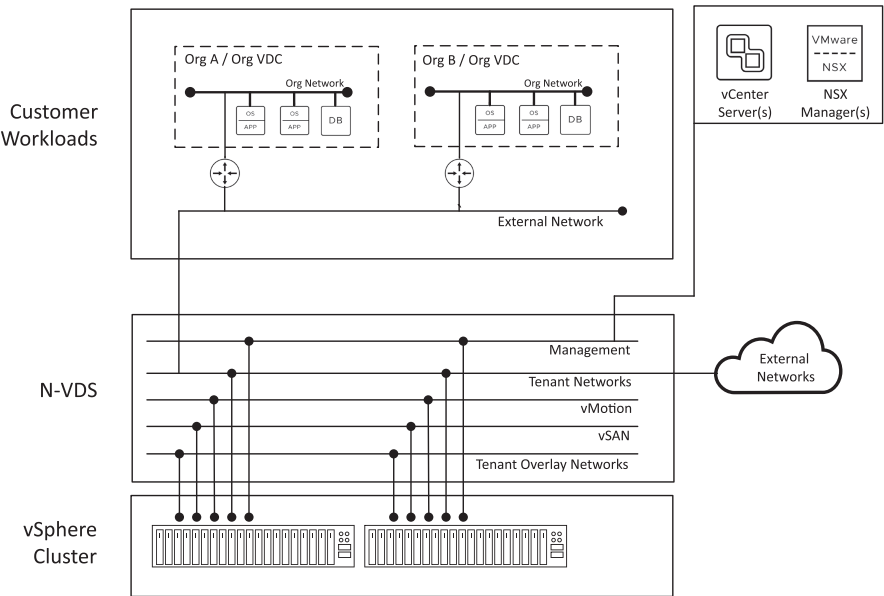


FIGURE 9.10 Network Segmentation

The principle goals of the cloud service provider in securing hypervisor communication include:

- Verifying that tenants or external attackers cannot gain privileged access to the hypervisor through services running on the management, backup, or other secure networks.
- Ensuring tenants or external threats cannot sniff vSphere vMotion, vSphere FT, vSAN, or other privileged traffic to obtain memory or file system contents of a virtual machine or other data that could assist in the staging of a man-in-the-middle attack.
- Making sure that all replicated storage or storage access traffic, which is typically not encrypted, cannot be viewed by anyone.

Despite the need to protect the VMware Cloud Provider Platform from attacks, appropriate operational teams still need access to the vCenter Servers and ESXi hosts. However, rather than configuring direct access to these protected networks, consider limiting access only through a virtual private network (VPN) or through the configuration of “jump boxes” that reside on the management network.

As highlighted previously, a management cluster also provides resource isolation and can satisfy the requirement to have physical isolation between management and tenant workloads. This further protects access to the management virtual machines running monitoring, management, and cloud platform services.

Certificate Configuration and Usage

VMware Cloud Provider Platform environments (including Cloud Director, vCenter and vRealize Orchestrator) use certificates for secure communication between these and other IaaS components. During the implementation phase, the team must obtain or generate certificates for the various management components. A certificate that is trusted by a web client should be used.

On any production platform, the self-signed certificates generated when you install the components should always be replaced with those from a verifiable trusted Certificate Authority (CA) to enable and ensure certificate verification on all internal vSphere and external client connections.

Self-signed certificates are vulnerable to man-in-the-middle attacks and are unlikely to comply with your organization’s security policy.

Local Account Management

It is likely that occasionally, some advanced configuration and troubleshooting of an ESXi host might still require local privileged access through the console, the DCUI or SSH. While better managed through Active Directory integration, local host user accounts are often required, if for no other reason than *in case of emergency* accounts.

In vSphere you can manage local accounts on the hypervisor using new ESXCLI commands, providing the ability to script and inject such configuration during deployment. The ESXCLI commands allows administrators to add, list, remove, and modify accounts across all hosts in a cluster from the vCenter Server. Setting, removing, and listing local permissions on ESXi servers can also be centrally managed.

However, despite this simple administration of local host accounts, which strengthens the security of the Cloud Provider Platform, the number of local user accounts created on the hypervisors should be limited to those that are essential.

Host Active Directory Configuration Status

ESXi hosts can be joined to Active Directory, or more precisely, can use Active Directory for authenticating users, which allows for assigning permissions to domain users at host level. The advantage of this is that you can manage user accounts using Active Directory for authentication, authorization, and compliance, which is significantly easier and more secure than trying to manage local accounts.

VMware recommends taking advantage of this functionality by employing the use of a tenant administrative or federated domain as a security repository for permitting easy authentication and authorization with unique administrative credentials. This model is simplified by using the pre-created Active Directory group “ESXi Admins” providing root access to authorized administrators as well as a way to audit direct access to the ESXi hosts.

Transparent Page Sharing Security

Transparent Page Sharing (TPS) is a mechanism used by the VMkernel for more effective use of physical memory resources by only storing once memory pages that are identical on two or more virtual machines. Each virtual machine has only read access to shared memory pages, and as soon as a virtual machine tries to modify a shared page, a new private copy is created by the VMkernel.

Shared pages are most commonly seen when a host is running multiple virtual machines with the same guest operating system. However, the advent of more and more x64-based guest operating systems, where the guest leverages large page tables, has seen the effective benefits of TPS being significantly reduced.

Although VMware believes the risk of TPS being used to gather sensitive information is low, to ensure that ESXi ships with default settings that are as secure as possible, TPS management options have been introduced and inter-virtual machine TPS not enabled by default. Administrators are able to easily revert this new default behavior to enable TPS if they want to do so.

SNMP Hardware Monitoring

In addition to all the monitoring options available to hosts, ESXi includes an SNMP v3 daemon that can both send SNMP and receive polling requests to send SNMP traps for monitoring the server's hardware and virtual machines. In addition, vCenter Server can be configured to send SNMP traps, which are typically sent to another management system.

Nothing in either ESXi or vCenter SNMP provides you with anything that you cannot already discover through the VMware vSphere API, PowerShell, vCenter CIM monitoring, or syslog. However, the primary situation in which SNMP might still be implemented is when a pre-existing SNMP management application within the service provider's infrastructure is already in place, and the organization wants to continue with its use.

Host Lockdown Mode

Enable lockdown mode to increase security of ESXi hosts and to further mitigate the risk of unauthorized access to the ESXi console by limiting it to only the appropriate operational team through vCenter Server. Two different lockdown mode options exist.

With "normal lockdown mode," no users other than the `vpxuser` account have authentication permission. So other accounts cannot perform any actions against a host directly. Normal lockdown mode forces all operations to be carried out through the vCenter Server, although DCUI access is not stopped, and users on the `DCUI.Access` list can access the DCUI. However, with "strict lockdown mode," the DCUI access is stopped.

vSphere also provides functionality called "Exception Users." These can be local accounts or Active Directory accounts with permissions defined locally on the host, where these users have host access. Exception Users are not recommended for general user accounts but are recommended for use only by third-party applications. For example, Service Accounts, where the host needs access when either normal or strict lockdown mode is enabled. Permissions on these accounts should be configured as the absolute minimum required for the application in question to carry out its task and with an account that needs only *read-only* permissions to the host.

ESXi Firewall

The ESXi hypervisor has included a firewall that is enabled by default to only allow the incoming and outgoing connections that are necessary for managing virtual machines. The ESXi firewall allows for low-level control over network access, provides the ability to restrict access to specific network segments, and is centrally configurable through host profiles to reduce the operational overhead of managing changes globally within the virtual infrastructure.

Role-based Access Control

Role-based access control (RBAC) can be employed in a number of different scenarios for the Cloud Provider Platform. Internally, from an operational perspective, to ensure only the appropriate operational or administrative teams have access to internal and tenant environments. Also, from the perspective of giving tenants access to either shared or dedicated vCenter Servers or hosts directly, allowing for self-managed hosting use cases. No matter what use cases are defined by the cloud provider, if proper role-based access controls are not in place within the environment, virtual machines will be vulnerable.

The following is a list of common and best practices for designing RBAC for the Cloud Provider Platform:

- Any vSphere administrative permission granted to an administrator account should use a privileged account (PA).
- All permissions should be assigned to AD groups and not individual user accounts. No local accounts or groups should be used for day-to-day administration of the Cloud Provider Platform.
- Grant permissions only where needed for job role, using the minimum number of permissions makes it easier to understand and manage the permissions structure.
- Create new groups for vCenter Server administrators. Avoid using Windows built-in groups or other pre-existing groups.
- If you assign a restrictive role to a group, check that the group does not contain the Administrator user or other users with administrative privileges. Otherwise, you could unintentionally restrict administrators' privileges in parts of the inventory hierarchy where you have assigned that group the restrictive role.
- Use caution when granting permission at the root vCenter Server level. Administrators with permissions at the root level have access to global data on vCenter Server, such as roles, custom attributes, vCenter Server settings, and licenses. Changes to licenses and roles propagate to all vCenter Server systems in a linked mode group, even if the administrator does not have permissions on all of the vCenter Server systems in the group.
- In most cases, enable propagation on permissions. This ensures that when new objects are inserted into the inventory hierarchy, they inherit permissions and are accessible to administrators.
- Use the No Access role to mask specific areas from the hierarchy that you do not want particular users to have access to.
- Certain privileges can be harmful to hosts and should be assigned to administrators only when required. This includes any privilege that allows a user to delete, rename, remove, or create items that can cause data loss or datastores to be filled up. This can cause a denial of service attack on your VMs (e.g., prevent snapshot creation).

- Create roles that are customized to a administrator/tenant's requirements. For example, to create a role for an operations team that is responsible for monitoring VMs, create one that allows only VM interactions (e.g., power on, power off, reset, and console interaction). This allows team members to look at the console of a VM to see what is happening and power cycle a VM.
- A privilege that should be assigned sparingly is the datastore low-level file operations privilege, which allows users to upload and download files to a host datastore. This privilege can create a security risk.
- Other potentially dangerous privileges are in the network and virtual switch categories, which can allow a user to move a VM to any available virtual LAN that is configured on your virtual switches. This can be particularly risky if you have public and private network virtual switches on a host where you definitely do not want a VM moved between them or connected to both at the same time (e.g., edge cluster hosts). Assigning the network privileges to your network admins and denying them to everyone else is a good practice.

CHANGE AND CONFIGURATION MANAGEMENT

The change and configuration management processes are important when maintaining any production cloud platform. Anything that can affect the tenants service-level agreements becomes in-scope for change and configuration management. In addition to software version maintenance and configuring hardware and software changes, there should be documented processes and procedures necessary to maintain, support, and administer the infrastructure on a daily basis, these should include:

- Ensuring that the appropriate people who can approve changes are in place for each component
- Ensuring that all business decision makers for shared systems are represented
- Ensuring that appropriate documentation and approvals for all changes to configurations are recoded
- Maintaining of documentation for default component settings as well as those explicitly set
- Mechanisms for the monitoring of configuration drift
- Procedures for the regular backup of all configuration settings are in place

PATCH AND UPDATE PRACTICES

Maintaining an up-to-date IT infrastructure is critical for the health, performance and security of the entire environment. Maintaining an up-to-date infrastructure can be a daunting task for IT administrators, but if not performed

dependably and routinely, the infrastructure is at risk. Through a simplified solution architecture, the Cloud Provider Platform, particularly when implemented with VMware Cloud Foundation, aims to deliver a simpler platform to address cloud providers previous challenges with respect to this aspect of operations.

Consideration is required for both scheduled and emergency patches and updates to protect systems from security vulnerabilities, as well as help ensuring stability of performance. Just as with other systems, the Cloud Provider Platform components should be maintained, and devices patched and updated in line with the cloud provider's internal policies. The patch and update process should include the following practices:

- Documentation on the version of each hardware and software component within the environment
- Documentation on risk acceptances for patches delayed or not installed in a timely manner
- Research done to mitigate or to reduce the risk when patches cannot be installed
- The following of change management procedures to ensure appropriate documentation and internal approvals
- Establishing regular patch cycles for high and for low priority patches (for example, weekly and monthly)
- Establishing and testing of processes for emergency out-of-cycle patching
- Ensuring that virtual machines are re-patched if restored from a snapshot prior to a scheduled patch date

Final Word

In recent years, as enterprise IT environments have dramatically grown in scale, multi-cloud complexity and services consumption, customers have often found themselves deploying applications in silos of dedicated infrastructure, finding the best cloud for the job. These silos are typically built around specific applications, business organizations, operational requirements and regulatory compliance such as Sarbanes-Oxley, HIPAA, PCI, or to address specific proprietary data confidentiality, like Healthcare organizations ensuring patient record confidentiality, or universities partitioning student user services from business operations.

As this cloud landscape rapidly changed, the focus for cloud providers was to develop innovative strategies and solutions that could reduce costs, quicken time to market and empower customers. The aim of cloud providers was to ensure a customer's approach to delivering IT services could be focused on meeting business objectives and growth and not be encumbered by IT service delivery. In order to deliver this, enterprise public cloud services typically require the provider to pursue many of the following objectives:

- Increase operational efficiency through cost-effective use of infrastructure
- Drive economies of scale through shared resourcing
- Rapid and agile deployment of customer environments, applications, and middleware services
- Improve service quality and accelerate delivery through high levels of standardization
- Promote green computing by maximizing efficient use of shared resources to lower energy consumption

Achieving these goals can have a profoundly positive impact on profitability, productivity, and product quality. However, leveraging shared infrastructure and resources in a cloud-services architecture introduces additional challenges for cloud providers who are faced with demands from customers who require secure isolation for their application environments but also require highly flexible management and control.

The cloud provider must protect all customer data, communication, application environments and ensure secure isolation from other tenants at all costs. The VMware Cloud Director model enables complete public cloud

separation to secure tenants and guarantee customers have no visibility of each other. This allows VMware cloud providers to build enterprise public cloud environments that can meet this demand to provide secure isolation, while still delivering the management and flexibility benefits of shared resources.

The VMware Cloud Provider Platform builds on Cloud Director's capabilities to create a unique infrastructure solution that incorporates comprehensive storage, network and computing technologies that facilitate dynamic, shared-resource management while maintaining a secured and isolated environment, which delivers a powerful solution to fulfill the demanding requirements for secure isolation and flexibility in cloud deployments of all scales. This secure cloud architecture further extends user control end to end throughout the environment with broader control over resources, with infrastructure as-a-service just being the start.

The Cloud Provider Platform and its components also offer extensibility, enabling both third-party vendors and cloud providers themselves to evolve the platform, while keeping the services readily accessible and available to the immediate needs of the customer applications stacks. As a result, as demand for cloud resources hopefully grows, the Cloud Provider Platform environment has the ability to scale to meet the customer's needs. While individual physical and software components may have capacity limitations, with a properly designed solution, the result is an infinitely scalable infrastructure that is able to meet the scalability demands of the most successful cloud providers.

The Cloud Provider Platform also helps remove the traditional application silos within the providers data center and operational teams and introduces a new level of flexibility and scalability to manage resources. This flexibility helps address the challenges facing enterprises and helps cloud providers rapidly accelerate their customer's digital transformation program through cost reduction and focus on time to market.

Designing an enterprise class public cloud infrastructure is a daunting task. In the past, no single vendor could provide the full set of capabilities required across the entire environment in a way that the Cloud Provider Platform is able to. This secure cloud architecture, the result of years of VMware engineering effort, offers the best products, services, and features to create a cloud platform that can realize the efficiencies of shared infrastructure, while still offering the security, flexibility, resilience, and ease of management of a traditional IT solution.

Cloud Provider Platform Bill of Materials

Dependent on use case and architecture, the following components make up the Cloud Provider Platform

LOGO	PRODUCT	DESCRIPTION
	vCenter Server	VMware vCenter Server is advanced server management software that provides a centralized platform for controlling your VMware vSphere environments, allowing you to automate and deliver a virtual infrastructure across the hybrid cloud.
	ESXi Hosts	ESX hosts are the servers/data storage devices on which the ESXi hypervisor has been installed. ESXi is highly efficient, as one host device can support multiple workloads.
	vSAN	VMware vSAN is a software-defined, enterprise storage solution that supports hyper-converged infrastructure (HCI) systems. vSAN is fully integrated with VMware vSphere, as a distributed layer of software within the ESXi hypervisor.
	Cloud Director	Appliance or Installable: Appliance: PhotonOS OVA appliance. Includes Highly Available Database (3 Nodes minimum required) Installable: Supported OS for installable CentOS 7 and Red Hat Enterprise Linux 7
 PostgreSQL	PostgreSQL	Cloud Director Database. Already embedded in the if using the VCD appliance. High Availability requires RepMgr/PgPool
	RabbitMQ (AMQP Messaging Bus)	RMQ Erlang Supported OS CentOS 7 and Red Hat Enterprise Linux 7

LOGO	PRODUCT	DESCRIPTION
	Cloud Director ad-hoc VM Metrics DB	Cassandra Supported OS CentOS 7 and Red Hat Enterprise Linux 7
	vIDM	Identity Manager—Authentication layer for Cloud Director at tenant and provider level
	vRealize Orchestrator	vRO Plugins for: • Cloud Director • VCD Availability • NSX
	Usage Meter	Required for VMware Partner Connect subscription licensing contract
	NSX	NSX for vSphere or NSX Data Center (formally NSX-T)
	HCX	VMware HCX is an application mobility platform designed for simplifying application migration, workload rebalancing and optimized disaster protection across data centers and clouds
	vRealize Network Insight	VMware vRealize Network Insight delivers intelligent operations for software-defined networking and security, with converged visibility across virtual and physical networks, planning and recommendations for microsegmentation and operations management for NSX.
	Cloud Director Availability	Disaster Recovery & Migration as a Service (featuring both types of availability: On-prem to Cloud (DR2C) & Cloud Director to Cloud Director (C2C)
	Site Recovery Manager	Suitable to recover the Management Cluster to a secondary site.
	vRealize Operations (vROps)	Chargeback and Cloud Provider Platform Monitoring (tenant/provider)
	vROPS Tenant App for Cloud Director	The vRealize Operations Tenant App provides metering capabilities that allow the service providers to chargeback their customer base and individual customers to review their billing data. It requires the Cloud Director MP vROPS Management Pack.

LOGO	PRODUCT	DESCRIPTION
	vRealize Log Insight	vRLI Plugins for: • VCD • Core SDDC (vSphere, vSAN, NSX-V, NSX-T, etc)
	Container Service Extension for VCD	Container Service Extension (CSE) is a VMware Cloud Director extension that helps tenants create and work with Kubernetes clusters. CSE brings Kubernetes-as-a-service to VCD by creating customized VM templates and enabling tenant/organization administrators to deploy fully functional Kubernetes clusters in self-contained vApps.
	Python SDK for Cloud Director	<i>pyvcloud</i> is the Python SDK for VMware Cloud Director VCD.
	Cloud Director CLI	VCD-CLI is the Command-line Interface for VMware Cloud Director.
	Hashicorp's Terraform for Cloud Director	Terraform is a tool for building, changing, and versioning infrastructure efficiently on cloud infrastructure and works with configuration files for connecting to Cloud Director.
	vRealize Automation Cloud	OPTIONAL: vRAC might be part of the public VCD Catalog to deliver private cloud experience—suitable for NOC operations. (1 vRA instance per VCD tenant)
	Cloud Provider Hub	Cloud Provider Hub is the portal for Managed Server Provider, featuring: • VMC on AWS • Cloud Health • NSX SD-WAN by Velocloud • HCX • Workspace ONE • Horizon Cloud

This page intentionally left blank

Index

A

- Active Directory (AD), 245–246, 336, 350–351
- Allocation Models, 67–69
 - Flex, 68
 - ‘Pay-As-You-Go’ model, 68, 91
 - Reservation Pool, 90–91
- ‘Allocation Pool’ model, 68
- AlwaysOn Availability Groups feature, 31, 33
- Amazon S3, 312
- Amazon Simple Storage Service, 311
- Amazon Web Services, 23, 254, 308, 310, 311
- AMQP Message Bus, 251
- AMQP service in Cloud Director, 249
- API, 9, 11, 17, 44, 45, 47, 49, 61, 66, 77, 105–107, 191, 216, 294, 297, 301, 308
- App Launchpad, 323–325
- Apple Releases, 73
- Application cluster, 69
- Archive-as-a-Service (AaaS), 316
- Array-based replication (ABR), 201, 210
- Automatic failover option, 33
- AWS, 8, 77, 107
- VMware Cloud on, 264–266
- Azure, 11, 77, 267, 269, 312
- Azure VMware Solutions, 267–269
 - Intelligent Application Delivery, 270

B

- Backup-as-a-Service (BaaS), 316, 325
- Bastion connectivity service, 239
- Big Data-as-a-Service (BDaaS), 316
- BIOS settings, 124
- Bitnami Catalog, 55, 112, 246, 322–323, 324

- Blade servers compared with rackmount servers, 118–120
- Blocking tasks, 250–253
- Bulk Migration, 40, 288–289
- Business strategy, 3–5, 18

C

- Capacity management for cloud service providers, 126–127, 170, 330
- Cassandra cluster, 59
- Cassandra database for storing historic metric data, 59–60
- Cassandra metric data, 60
- Catalogs, 72, 90, 186, 194–195
 - sharing of, 54
- Central Control Plane (CCP), 215
- Central management stack, multi-site architecture with, 96–97
- Central Point of Management (CPOM), 87, 90, 108, 215, 256, 294–296, 300
 - SDDC as-a-Service with, 297–300
- Certificate configuration and usage, 349
- Change and configuration management, 353
- Class of Service (CoS), 169
- Cloud, next evolution of, 2–3
- Cloud administration automation, 243–244
- Cloud API, 65–66, 93
- Cloud automation, 241–242
 - blocking tasks, 250–253
 - RabbitMQ, 248–250
 - Service Library feature, 246–248
 - Terraform provider, 254–256
 - vRealize Orchestrator (vRO), 242
 - Cloud administration automation, 243–244

- Cloud automation (*continued*)
 - Cloud Consumer operational automation, 246
 - Organization administration automation, 244–245
 - plug-in support for Cloud Director, 242–243
- Cloud Consumer operational automation, 246
- Cloud Director, 2, 3, 6, 8, 9–10, 11, 12, 13, 35, 36, 37, 38, 41, 43, 44, 45, 66, 334
 - custom branding, 104–105
 - database, 58–59
 - network layers, 218
 - and NSX integration, 216–217
 - storage design with vSAN, 186
 - impact of vSAN on Cloud Director compute and network resources, 193–194
 - multiple tiered resource clusters as dedicated provider VDCs (model), 187–190
 - single provider VDC hosting multiple tiered resource clusters (model), 190–193
 - storage profiles with external storage devices, 198
 - transfer storage, 60–61
- Cloud Director Abstraction Layers, 9, 185
- Cloud Director APIs, 13, 65–66, 105, 110, 251
- Cloud Director Availability Replicator Appliance, 274
- Cloud Director Availability Tunnel Appliance, 276
- Cloud Director cells, 56–58, 60–61, 64, 80, 85, 195, 259
- Cloud Director Central Point of Management (CPOM), 90, 108
- Cloud Director External Networks, 225–228
- Cloud Director HTML5 user interface, 293, 331
- Cloud Director Linux, 63
- Cloud Director Management Cluster, 161
- Cloud Director NFS storage requirement, 194–196
- Cloud Director Provider VDC, 193, 295
- Cloud Director Remote Console, 64–65
- Cloud Director Resource Group, 114, 259
- Cloud Director security on the internet, 62–64
- Cloud Director Server Group, 58, 59, 66, 96, 99, 195, 276, 292
- Cloud Director Service, 76, 77, 101, 342
- Cloud Foundation. *See* VMware Cloud Foundation (VCF)
- Cloud Foundation Virtual Infrastructure Workload Domain, 259
- Cloud Health Technology, 10, 304
- Cloudian Object Storage solution, 315–316
- Cloud Management Platforms (CMP), 45, 56, 62, 74, 93, 214, 242, 248, 297
- Cloud Platform (vSphere) patch management, 142
- Cloud Platform edge cluster, 131–133
- Cloud Provider Hub (CPH), 8, 10–11, 40, 265, 301–304
- Cloud Provider Platform (CPP), 1, 2, 3–4, 6, 7, 45, 49, 77, 80, 90, 93, 94, 101, 111–112, 114, 139, 143, 163, 196, 218, 241
 - architecture, 328–329
- Cloud Provider Hub, 10–11
- Cloud Provider Platform core use cases, 35
 - data protection and recovery, 41–42
 - Kubernetes service, 42–44
 - VMware Cloud Provider Platform networking and security overview, 35–36
 - workload migration and initial onboarding, 36–41
- Cloud Verified program, 14–15
- component architecture, 8
- enterprise class public/hybrid, designing, 17–20
- hybrid cloud use case example, 31–35
- key pillars of, 47
- on-premises-to-provider cloud connectivity, 20
 - private circuits, connectivity over, 23
 - public internet, connectivity over, 20–23

- Wide-Area Networking (WAN)
 - platforms, evaluation of, 24–31
- provider data center, 8–10
- service definition, building, 16–17
- Service-Level Agreement (SLA),
 - building, 17
- Software-Defined Data Center (SDDC), 13–14
- vCloud Usage Meter and VMware
 - vCloud Usage Insight, 13
- VMware Validated Designs (VVD) for Cloud Providers, 15–16
- vRealize Operations and Tenant App, 11–13
- Cloud Provider Platform core use cases, 35
- data protection and recovery, 41–42
- Kubernetes service, 42
 - Container Service Extension (CSE) Enterprise PKS enablement, 44
- networking and security
 - overview, 35–36
- workload migration and initial onboarding, 36
 - HCX, 40–41
 - vCloud Availability migration, 37–39
- Cloud Provider Platform management
 - component logging, 339–340
 - NSX Edge gateway logs, 340–342
- Cloud Provider Pod, 260–263
- Cloud Provider use cases, 48, 145, 292
- Cloud resource design, 77
 - dedicated hosted cloud, 90–93
 - management cluster, 79–80
 - Resource Clusters, 80–81
 - scalability, 81–82
 - Expansion Pods, 85–87
 - hyperscale growth, 88–89
 - pod architecture, 82–84
 - pod evolution and addressing technology changes, 87–88
 - pod sizing, 84
 - Primary Pod, 84–85
- Cloud service providers, 3, 4, 6, 8–9, 10, 17, 19, 23, 24, 29, 43, 48, 49, 61, 73, 90, 93, 111, 112, 304, 308, 349
- Cloud Storage, 307, 310–312
- Cloud to Cloud (C2C) solution, 274, 280
- disaster recovery solutions, 37
- Cloud usage meter data collection
 - workflow, 14
- 'Cloud Verified' cloud services, 3, 4
- Cloud Verified* compute platform, 9, 14, 111–113
 - automating host deployment, 127–128
 - capacity management for cloud service providers, 126–127
- Cloud Platform (vSphere) patch management, 142
- Cloud Platform edge cluster, 131–133
- comparing blade servers with rackmount servers, 118–120
- determining appropriate virtual CPU-to-physical CPU ratio, 122–124
- determining host CPU and memory requirements, 121–122
- Enhanced vMotion Compatibility (EVC), 139–140
- host sizing, 120–121
- infrastructure protection & availability, 118
- non-uniform memory access (NUMA),
 - performance tuning with, 124
- vNUMA, 126
- planning a scalable vSphere host
 - clusters architecture, 129–131
- planning for server failure with resilience, 133–135
 - vSphere High Availability, 135–136
- resource balancing and transparent maintenance, 136
 - Distributed Resource Scheduler (DRS) automation, 137–139
- scalability and designing physical resources, 113–117
- vCenter High Availability architecture, 140–142
- Cloud Verified* network platform,
 - designing, 211
- NSX data center, Cloud Director
 - networking with, 217–218
- Bastian connectivity service, 239
- Distributed Firewall (DFW), 236–239
- load balancing, 232–236
- physical network architecture, 218–220

- Cloud Verified* network platform
 - (continued)
 - static routing, 232
 - VMware Cloud Director tenant networking, 223–230
 - vSphere network constructs, 220–223
- NSX data center architecture
 - overview, 214
 - Cloud Director and NSX integration, 216–217
 - Control Plane, 215
 - Data Plane, 216
 - Management Plane, 215
- NSX for vSphere (NSX-v) versus NSX-data center, 212–214
- ‘Cloud Verified’ partners, 3
- Cloud Verified* storage platform, 143
 - external array storage, consuming, 196
 - Cloud Director storage profiles with external storage devices, 198
 - considerations for use of vSphere storage and arrays features, 201
 - datastore clusters with external storage devices, 198–199
 - IOPS management with external storage devices, 199–201
 - traditional VMFS storage considerations for Cloud Director, 198
- key vSAN design factors, 177
- Virtual Volumes (vVols), 201–203, 207–210
 - Protocol Endpoint (PE), 205
 - VASA Provider (VP) / storage provider, 205–207
 - Virtual Datastore and a Storage Container, 203–205
 - vVol Storage Policies and Storage Policy Based Management, 207
- VMware vSAN architecture, 147
 - application support, 160
 - data integrity, 152–153
 - data locality, 155–156
 - data replication, 157–158
 - host and vSphere cluster failure tolerance, 150–152
 - quality of service, 154–155
 - rack awareness, 153–154
 - space efficiency, 147–150
 - swap efficiency, 154
 - vSAN Encryption, 159–160
 - vSAN Stretched Clusters, 156–157
- vSAN design, 144, 160
 - application requirements, 174
 - disk controllers, 164–165
 - disk group design, 162–164
 - failures-to-tolerate policy, impact on sizing of the number of, 164
 - flash-based devices, 166
 - host CPU overhead, 162
 - host memory requirements, 161–162
 - network requirements, 166–167
 - recoverability, vSAN and, 174–175
 - vSAN cluster design, 160–161
 - vSAN default policy, 170–174
 - vSAN policy capabilities, 170
- vSAN storage and cloud director, 180
 - Cloud Director NFS storage requirement, 194–196
 - Cloud director storage design with vSAN, 186–194
 - designing a highly scalable HCI platform, 196
 - vSAN storage resource isolation, 183–186
- vSAN stretched cluster deployments, 177–180
- Coarse Wavelength Division Multiplexing (CWDM), 24–27
- Cohesity, 326
- Component-level scrubbing, 152
- Compression, 147–150, 160
- Compute Policies, 69
 - placement, 69–71
 - sizing, 71–72
- Consistent hybrid cloud experience, 19
- Consistent VMware cloud platform, 19
- Containers-as-a-Service (CaaS), 316
- Container Service Extension (CSE), 43, 304–307
- CSE Enterprise PKS enablement, 44
- Control Plane, 215
- CPU Ready Time, 123
- CRC32 algorithm, 152
- Credit card data, security of, 4

- Cross-vCenter NSX configuration, 101
- Cross-VDC Networking, 101
- Customer associated sites, multi-site
 - architecture with, 97
 - federation, 99–101
 - load balancer requirements, 101
 - multisite networking, 101–104
 - Organization user and group identities, 99
- Custom portal (optional) design factors, 105–110
- Custom portal architecture, 109
- Custom UI/API Extensibility, 272

D

- Dark fibre, 24, 27
- Data archiving, 335–336
- Databases, availability of, 31
- Data encryption key (DEK), 159
- Data integrity, 152–153
- Data locality, 155–156
- Data Plane, 216
- Data protection, 41–42
 - and ISV Service Ecosystem, 325–327
- Data replication, 157–158
- Data sovereignty, 3
- Datastore clusters with external storage
 - devices, 198–199
- DCUI .Access list, 351
- DDOS attack, 64
- Deduplication, 147–149
- Dell EMC IDPA cloud director integrated solution, 42
- Dell EMC IDPA extension for data protection, 326
- Denial of Service (DoS) attack, 347
- Dense Wavelength Division Multiplexing (DWDM), 24–27
- DevOps, xxi, 14
- Direct Connect (DX) service, 23
- Direct networks, 228
- Disaster recovery, HCX for, 290
- Disaster Recovery-as-a-Service (DRaaS), 316
- Disaster recovery protection, 33
- Disaster recovery to cloud (DR2C), 37, 274, 276–280
- Disasters, recovery from, 31

- Discrete Administration, 94
- Discrete management, multi-site
 - architecture with, 94
- Disk controllers, 164–165
- Disk group design, 162–164
- Disk group sizing, 163
- Distributed Firewall (DFW), 73, 236–239
- Distributed Machine Learning
 - on TensorFlow component architecture, 318
 - on TensorFlow for Machine Learning, 317
- Distributed Resource Scheduler (DRS)
 - automation, 136, 137–139
- Distributed Resource Scheduling, 81
- Durability, 314
- Dynamic routing (DLR), 73

E

- Edge Cluster services, 133
- Edge gateway provider and tenant
 - syslogs, 341
- Edge Gateways, 228–229, 232, 234
- Edge Services Gateway (ESG), 21, 73, 131, 259
- Elastic Block Storage (EBS), 312
- Elastic Provider VDC construct, 259
- Elastic VDCs, 190, 191
- End-to-end cloud services, 62, 311
- End-to-end security, 211, 237
- Enhanced vMotion Compatibility (EVC), 139–140
- Enterprise class public/hybrid,
 - designing, 17–20
- Enterprise Public Cloud services, 46, 377
- ESXCLI commands, 350
- ESXi 6.7, 127
- ESXi Admins, 350
- ESXi firewall, 351
- ESXi Host Hardening, 344–346
- ESXi hosts, 49, 57, 64, 69, 72, 80, 122, 127, 129, 135, 136, 142, 175, 178, 205, 212, 236, 330, 335, 337, 347, 349, 350
- ESXi hypervisor, 113, 125, 127, 161, 217, 351
- ESXi security baseline, 344
- ESXi servers, 64, 70, 130, 350
- Exception Users, 351

Expand Managed Services on Public Cloud, 8

Expand Service Opportunities at Edge, 8

Expansion Pod, 82, 84, 85–87

ExpressRoute circuits, 23

Extensibility, 272

VMware HCX, planning for application migration with, 282

HCX component overview, 285–287

HCX for disaster recovery, 290

HCX for hybrid network extension, 289–290

migration methodology, 287–289

Operating System Assisted Migration (OSAM), 292

VMware Cloud Director HCX design considerations, 290–291

VMware vCloud Availability (vCAv), 273

architecture overview, 274–276

Cloud to Cloud (C2C) overview, 280

disaster recovery to cloud (DR2C) overview, 276–280

native integration with Cloud Director, 280–282

vCAv migration versus HCX, 282

vRealize Operations Management Pack for Cloud Director and Tenant App, 292

App Launchpad, 323–325

Central Point of Management (CPOM), 294–296

Cloud Provider Hub (CPH) is, 301–304

common storage services, 307–315

Container Service Extension (CSE), 304–307

data protection & ISV Service Ecosystem, 325–327

Machine Learning-as-a-Service (MLaaS), 316–320

Multi-Cloud

Management-as-a-Service, 320–322

SDDC as-a-Service with CPOM, 297–300

tenant metering, 294

VMware Cloud Marketplace and Bitnami for Cloud Providers, 322–323

Extensibility, 41, 48, 272

Extensible service framework, 273

External array storage, consuming, 196

Cloud Director storage profiles with external storage devices, 198

datastore clusters with external storage devices, 198–199

IOPS management with external storage devices, 199–201

traditional VMFS storage considerations for Cloud Director, 198

vSphere storage and arrays features, considerations for use of, 201

F

Failures-to-tolerate (FTT) policy, 176, 191
impact on sizing of number of, 164

Faster Provisioning, 53

Fault tolerance method, 150

Fibre Channel over Ethernet (FCoE), 205

Financial services, 4

Firewall services, 36, 131, 237

Flash-based devices, 166

Flex allocation model, 68

Flexible mixed storage layer, 144

Flexible pod architecture, 87

Flex user interface, 48

Functions-as-a-service (FaaS), 252

G

GENEVE (Generic Network Virtualization Encapsulation) overlay, 36, 222–223

Google Cloud Platform, 11, 308, 312

GPU Computing, 319

Guest Operating Systems, 72, 182

H

HCI Cloud Director Resource Group, 114

HCI Clusters, 114–115

HCX, 40–41, 285

vCloud Availability Migration
verses, 282

HIPAA, 4, 355

- Host Active Directory configuration
 - status, 350
 - Host and vSphere cluster failure
 - tolerance, 150–152
 - Host CPU overhead, 162
 - Host deployment, automating, 127–128
 - Host lockdown mode, 351
 - Host memory requirements, 161–162
 - HTML5 user interface, 48, 49, 61, 62
 - HTML5 vSphere Plugin Integration, 278
 - HTTPS encryption, 347
 - HTTPS protocol, 64
 - Hub. *See* Cloud Provider Hub (CPH)
 - Hybrid cloud architecture, 37
 - Hybrid Cloud Extension (HCX), 282–285
 - component overview, 285–287
 - design considerations, 290–291
 - for disaster recovery, 290
 - for hybrid network extension, 289–290
 - migration methodology, 287
 - Bulk Migration, 288–289
 - HCX assisted vMotion, 287
 - replication-assisted vMotion (RAV), 289
 - vCloud Availability migration
 - verses, 282
 - Hybrid Cloud model, 18, 19
 - Hybrid cloud solutions, 1–2, 4
 - Hybrid cloud use case example, 31–35
 - Hybrid network extension, HCX for, 289
 - Hybrid vSAN cluster, 155
 - Hybrid WAN architecture, 30
 - Hyper-Converged Infrastructure (HCI)
 - solution, 143
 - Hyperscale growth, 88–89
 - Hyperscale regional architecture, 89
 - Hypervisor, 347
- I**
- IaaS engine, 90
 - IaaS functionality, 35
 - Identity provider (IDP), 77, 99
 - Independent Software Vendors (ISVs),
 - 41, 46, 194–195, 325, 326
 - Infrastructure as a service (IaaS), xix
 - Internal Network, 228
 - Internet Protocol Security, 30
 - iopsCapacity, 201
 - IOPS management with external storage
 - devices, 199–201
 - IPsec, 20, 30
 - iSCSI, 205, 308
 - Island clusters, 69, 71
 - concept of, 69
 - use cases for, 69
 - ISO 27001, 4
- J**
- JMX service, 13
 - JSON network profile, 133, 134
 - Jumbo frames, 168, 220
- K**
- Key encryption key (KEK), 159
 - Key Management Service (KMS) server,
 - 159, 312
 - Key Performance Indicators (KPIs), 292
 - kubectl, 44
 - Kubernetes, 5, 42–44, 304, 305
 - Kubernetes-as-a-Service, 43, 304
 - Kubernetes Cluster, 44, 47, 73
 - Kubernetes service, 42
 - Container Service Extension (CSE)
 - Enterprise PKS enablement, 44
- L**
- L2VPN, 21
 - LAG, 221
 - Latency, 313
 - Lifecycle management (LCM), 58,
 - 128, 142
 - Linked clones, 56
 - Linux distributions, 73
 - Linux operating systems, 337
 - Linux servers, 56
 - Load balancer, 36
 - Load balancing, 232
 - NSX Data Centre Edge Load
 - Balancer, 234
 - third-party virtual appliance load
 - balancer, 236
 - VMware NSX Advanced Load
 - Balancer, 234–235
 - Local account management, 349–350
 - Local Control Plane (LCP), 215
 - Locality of reference. *See* Data locality

- Local VMware cloud service provider, working with, 3
- Logging Service, building, 334
 - Cloud Provider Platform management component logging, 339–340
 - data archiving, 335–336
 - extending vRealize Log Insight Services, 336–337
 - logging service, 335
 - Log Intelligence, 342–344
 - Management Environment Logging, 337–339
 - sample syslog design scenario, 342
 - security monitoring, 335
- Log Insight, 330, 334, 336
- Log Intelligence, 342–344

M

- MaaS, 267–268
- Machine Learning-as-a-Service (MLaaS), 316–320
- Managed Kubernetes, 46
- Managed Service Provider (MSP) model, 4, 6, 265
- Management cluster, 79–80
- Management Domain, 259
- Management environment logging, 337–339
- Management Plane, 215
- Master Organization, 303
- Messages, 251
- Metering, 294
- Metropolitan area network (MAN), 24
- Microsoft Azure, 23
- Microsoft SQL Server 2019 AlwaysOn Availability Groups, 31, 33
- Multi-cloud application blueprinting through Cloud Director, 321
- Multi-Cloud Management-as-a-Service, 320–322
- Multi-cloud service delivery, 6
- Multi-endpoint resource consumption, 296
- Multi-Path Plug-ins (MPP), 201
- Multiprotocol Label Switching (MPLS) Network, 27–29
- Multi-site architecture, 93
 - with central management stack, 96–97

- with customer associated sites, 97
 - federation, 99–101
 - load balancer requirements, 101
 - multisite networking, 101–104
 - Organization user and group identities, 99
- with discrete management, 94
- with stretched resource clusters (vSAN Stretched Cluster(s)), 94–96

- Multisite system, 76
- Multi-Tenancy, 52
- Multi-tenanted clouds, designing host security for, 347–348
- Multi-tenant/multi-tier data protection service, 327

N

- Native AWS, 11
- Native VMware Hybrid Cloud model, 18
- Network interface cards (NIC), 167
- Network I/O Control, 167, 168
- Network operations center (NOC), 338
- Network Pools, 229–230
- Network QoS and network I/O control, 168–169
- Network requirements, 166
 - jumbo frames, 168
 - network interface cards (NIC), 167
 - network QoS and network I/O control, 168–169
 - virtual switches, 167–168
 - VLANs, 168
 - VMkernel network port group, 168
- NFS mount, 335
- NFS protocols, 205
- NFS share, 61
- Non-uniform memory access (NUMA), 124, 125
 - vNUMA, 126
- Notifications, 251
- NSX, 21, 36, 41, 48, 211
 - and cloud director logical architecture, 213
- NSX APIs, 13
- NSX Data Center, 212, 346
 - architecture overview, 214
 - Cloud Director and NSX integration, 216–217

- Control Plane, 215
- Data Plane, 216
- Management Plane, 215
- and Cloud Director API integration, 217
- and Cloud Director integrated architecture, 223–225
- high level architecture, 215
- NSX data center, Cloud Director networking with, 217–218
- Bastion connectivity service, 239
- Distributed Firewall (DFW), 236–239
- network use cases, 266
 - load balancing, 232–236
 - static routing, 232
- physical network architecture, 218
 - jumbo frames, 220
 - network transport, 219–220
 - top of rack connectivity, 219
- VMware Cloud Director tenant networking, 223
 - Cloud Director External Networks, 225–228
 - Network Pools, 229–230
 - NSX Data Center and Cloud Director integrated architecture, 223–225
 - Organization Virtual Datacenter Networks, 228–229
 - vApp Networks, 230
- vSphere network constructs, 220–221
 - GENEVE overlay, 222–223
 - N-VDS teaming policies, 221–222
- NSX Data Centre Edge Load Balancer, 234
- NSX Distributed Firewall, 64
- NSX Edge devices, 131
- NSX Edge firewall, 55
- NSX Edge Gateway, 21, 340–342
- NSX Edge load balancer, 74
- NSX for vSphere (NSX-v) verses NSX-data center, 212–214
- NSX-Manager, 49, 215
- N-VDS teaming policies, 221–222
- cloud storage landscape overview, 310–312
- delivery and latency, 313
- scalability, 314–315
- security, 312–313
- Service Level Agreement (SLA), 313
- on VMware Cloud Provider Platform, 315–316
- On-premises-to-provider cloud connectivity, 20
- private circuits, connectivity over, 23
- public internet, connectivity over, 20–23
- Wide-Area Networking (WAN)
 - platforms, evaluation of, 24
 - dark fibre, 24
 - Multiprotocol Label Switching (MPLS), 27–29
 - Virtual Leased Line (VLL), 27
 - Virtual Private LAN Services (VPLS), 29
 - wavelength division multiplexing (WDM), 24–27
- On-premises vSphere architecture, 279
- OpenAPI, 66
- Open Virtualization Format (OVF)
 - standard, 73
- Operating System Assisted Migration (OSAM), 292
- Operating system templates, 54
- Oracle Cloud Managed Service Provider, 4
- Organization, 52–53
- Organization Administrators, 98, 99, 228
 - automation, 244–245
- Organizational Virtual Data Centers (Org. VDCs), 47, 51, 52, 59, 62, 67, 69, 72, 90, 91, 101, 133, 183, 189, 191, 228, 230–232, 247
- Organization unit (OU), 245
- Organization Virtual Datacenter Networks, 228–229
- OrgIdapSettings element, 101

O

- Object-based storage, 144
- Object-level scrubbing, 152
- Object storage, 5, 308

P

- Partners, VMware offerings for, 10
- Pass-through mode, 165
- Patch and update practices, 353–354

- 'Pay-As-You-Go' model, 68, 90–91
- Payment card systems, 4
- PCI, 355
- PCI-DSS compliant services, 4
- Personas & role based access control, 74–77
- Photon OS, 56
- Physical CPU (pCPU), 122
- Platform as a service (PaaS), 6, 35, 53
- Platform automation and standardization, 240
 - Cloud automation, 241–242
 - blocking tasks, 250–253
 - RabbitMQ, 248–250
 - Service Library feature, 246–248
 - Terraform provider, 254–256
 - vRealize Orchestrator (vRO), 242–246
 - Cloud Provider Pod, 260–263
 - VMware Cloud Director Service (VCDS), 266–267
 - VMware Cloud Foundation (VCF), 256–259
 - designing a VMware Cloud Foundation with Cloud Director solution, 259–260
 - VMware Cloud on AWS (MSP), 264–266
 - VMware MaaS Solution Partners, 267–268
- Plug-in support for Cloud Director, 242–243
- Pod architecture, 82–84
- Pod evolution and addressing technology changes, 87–88
- Pods, 82–84
- Pod sizing, 84
- Policy Management, 280
- PostgreSQL database, 57, 58, 59, 63
- Primary Pod, 82, 84–85
- Private circuits, connectivity over, 23
- Private cloud solution, 1, 4
- Private cloud use cases with Cloud Director, 49
- Proof of concept (PoC), 174
- Protocol Endpoint (PE), 205
- Provider constructs, 51
 - Cassandra database for storing historic metric data, 59–60

- Cloud API, 65–66
- Cloud Director cells, 56–58
- Cloud Director database, 58–59
- Cloud Director Remote Console, 64–65
- Cloud Director transfer storage, 60–61
- fast provisioning (linked clone), 56
- Organization, 52–53
- provider catalogs, vAPPs and virtual machines, 53–55
- Provider Virtual Data Center, 51
- securing Cloud Director on the internet, 62–64
- Support Management Organization, 53
- user interface, 61
- Provider data center, 8–10
- Provider Roles, 76
- Provider VDC Compute Policy, 69–70
- Provider Virtual Data Center (Provider VDC), 9, 49, 51, 52, 61, 67, 77, 81, 90, 91, 96, 181, 182, 260
- Public cloud solution, 1, 4
- Public cloud strategy, 19
- Public internet, connectivity over, 20–23
- Python SDK (pyvcloud), 66

Q

- Quality of Service (QoS), 145, 154–155, 168

R

- RabbitMQ, 248–250
- Rack awareness, 153–154
- Rackmount servers, blade servers compared with, 118–120
- RAID 0 mode, 165
- RAID-1 (mirroring), 150
- RAID-5/6 (erasure coding), 150
- RDP, 239
- Real Application Cluster (RAC), 175
- Recovery Point Objectives (RPOs), 42, 94, 157, 325
- Recovery Time Objective (RTO), 96
- Red Hat Linux kickstart script, 128
- RedHat OpenStack/KVM, 292
- Remote collectors, 331
- Remote Console Proxy, 64
- Remote Console Traffic, 64
- Replication, 348
- Replication-assisted vMotion (RAV), 289

- Reservation Pool Allocation model, 90
- Reservation Pool Model, 91
- Resource balancing and transparent maintenance, 136
 - Distributed Resource Scheduler (DRS) automation, 137–139
- Resource Clusters, 80–81
- RESTful APIs, 65, 66, 215–216
- Retail organizations, 4
- Return on Investment (ROI), 2
- Rights bundle, 76
- Role-based access control (RBAC), 36, 46, 74, 352–353
- Round Trip Time (RTT), 94
- Routed Network, 228
- `rsyslog`, 337
- Rubrik, 272, 326
- Run VMware Cloud, 7–8

S

- SaaS (Software as a Service), 6, 10, 35, 77
- SAP Cloud Platform Partner, 4
- Sarbanes-Oxley, 355
- Scalability, 314–315
- Scalable vSphere host clusters
 - architecture, planning, 129–131
- Scale-out clusters, 129
- Scale-up clusters, 129
- Secure communication, 348–349
- Secure multitenancy, 344
 - certificate configuration and usage, 349
 - designing host security for multi-tenanted clouds, 347–348
- ESXi firewall, 351
- Host Active Directory configuration status, 350
- host lockdown mode, 351
- local account management, 349–350
- role based access control (RBAC), 352–353
- secure communication, 348–349
- SNMP hardware monitoring, 351
- Transparent Page Sharing (TPS) security, 350
- VMware Cloud Director security, 347
- Securing communication, 30–31
- Security Groups, 73–74

- Self-signed certificates, 349
- Server failure, planning for
 - with resilience, 133–136
- Service Accounts, 351
- Service definition, building, 16–17
- Service Level Agreements (SLAs), 17, 80, 82, 136, 264, 313
- Service Library feature, 246–248
- Service Provider Admin Portal, 61
- SilverPeak SD-WAN technology, 40
- Single operating environment, 4–5
- Single sign-on, 101
- Site Recovery Manager (SRM), 80, 85, 96
- Site Reliability Engineering (SRE), 241
- SNMP hardware monitoring, 351
- Software-defined architecture, 37
- Software-defined compute resource, 113
- Software Defined Data Center (SDDC), 8, 13–14, 264
 - SDDC as-a-Service with CPOM, 297–300
 - SDDC Manager, 258
- Software-Defined Networking (SDN), 23
- Software-Defined Wide-Area Networking (SD WAN), 21–23
- Space efficiency technologies, 147–150
- Spanning Tree Protocol (STP), 220
- Spatial locality, 156
- SSH, 239
- Standalone Remote Console, 64
- Static routing, 232
- Storage-as-a-Service (STaaS), 316
- Storage Policy Based Management (SPBM) mechanism, 145, 150, 191, 207
- Storage services, 307–316
- Stretched cluster architecture, 179
- Stretched resource clusters, multi-site architecture with, 94–96
- Support Management Organization, 53
- Support Organization, 53
- Swagger UI, 66
- Swap efficiency, 154
- Synchronous-commit mode, 33
- Syslog agents, 337
- Syslog design scenario, 342
- `syslog-ng`, 336–337
- System Administrators, 76, 77
- System Organization environment, 53

T

Technical data gathering, 283
 Temporal locality, 156
 Tenant
 Predefined Roles, 77
 roles, 77
 Tenant App, 11–13, 60, 193, 331
 Tenant constructs, 66–67
 Allocation Models, 67–69
 Compute Policies, 69
 placement, 69–71
 sizing, 71–72
 Organization Virtual Data Center (Org. VDC), 67
 tenant application catalogs and libraries, 72
 tenant networking, 73–74
 workload types virtual machine verses vApp Blueprint, 72–73
 Tenant layer 2 (L2) VPN access, 74
 Tenant metering, 294
 Tenant Organization Admin, 237
 Tenant Portal, 61, 97
 Terraform provider, 254–256
 Third-party virtual appliance load balancer, 236
 Tier 0 CPU, 194
 Top of Rack (ToR) VLAN, 219
 Transparent Page File Sharing (TPS), 70, 121–122, 350
 Transport mode, 30–31
 Tunnel mode, 30–31

U

Uninterruptable power supplies (UPS), 118
 Universal Logical router, 101
 Universal Logical Switch, 101
 Universal Transport Zone, 101
 Up-to-date IT infrastructure, maintaining, 353–354
 Usage Insight Service, 13
 Usage Meter appliance, 13

V

vApp Blueprint, 72–73
 vApp Networks, 230
 vAPPs, 53–55

VASA Provider (VP) / storage provider, 205–207
 vcd-cli, 44, 66
 vCenter, 9, 335
 vCenter High Availability architecture, 140–142
 vCenter Server, 13, 49, 57, 80, 205, 212, 337
 vCenter Server Plugin, 274
 vCenter Servers, 330, 349
 vCloud API, 66
 vCloud Availability, 157
 vCloud Availability migration, 37–39
 vCloud Availability migration verses HCX, 282
 vCloud Director. *See* VMware Cloud Director (VCD)
 vCloud elastic provider VDC vSAN logical architecture, 190
 vCloud Replication Management Appliance, 274
 vCloud Usage Insight, 13
 vCloud Usage Meter, 13
 VDC Compute Policy, 69
 VDDK (Virtual Disk Development Kit) support, 42
 VeloCloud, 21
 Virtual CPUs (vCPUs), 122, 125, 126
 vCPU-to-pCPU ratio, 122–123
 Virtual Data Centers (VDC), 45, 180, 181, 188
 Virtual Datastore and a Storage Container, 203–205
 Virtual Hardware Versions, 72
 Virtual Leased Line (VLL), 27
 Virtual Machine Console, 13
 Virtual machines (VMs), 53–55
 protection of, 41, 347
 Virtual non-uniform memory access (vNUMA), 126
 Virtual Private LAN Services (VPLS), 29
 hybrid WAN architecture, 30
 securing communication, 30–31
 WAN interconnect high availability, 30
 Virtual private network (VPN), 36, 349
 Virtual switches, 167–168
 Virtual Volumes (vVols), 201–203, 207–210

- Protocol Endpoint (PE), 205
- VASA Provider (VP) / storage provider, 205–207
- Virtual Datastore and a Storage Container, 203–205
- vVol Storage Policies and Storage Policy Based Management, 207
- VLAN-backed network, 21
- VLANs, 168
- VMDK files, 56, 170, 193
- VMkernel, 350
- VMkernel network port group, 168
- VMware Cloud Architecture Toolkit 5.0, 143
- VMware Cloud Director (VCD), 45–48, 337
 - Cloud Director custom branding, 104–105
 - Cloud Provider use cases, 48
 - Cloud resource design, 77–79
 - dedicated hosted cloud, 90–93
 - management cluster, 79–80
 - Resource Clusters, 80–81
 - scalability, 81–89
 - custom portal (optional) design factors, 105–110
 - extension plug-in integration, 48
 - hosted private cloud with, 49
 - multi-site architecture
 - with central management stack, 96–97
 - with customer associated sites, 97
 - with discrete management, 94
 - with stretched resource clusters (vSAN Stretched Cluster(s)), 94–96
 - personas & role based access control, 74–77
 - provider constructs, 51
 - Cassandra database for storing historic metric data, 59–60
 - Cloud API, 65–66
 - Cloud Director cells, 56–58
 - Cloud Director database, 58–59
 - Cloud Director Remote Console, 64–65
 - Cloud Director transfer storage, 60–61
 - fast provisioning (linked clone), 56
 - Organization, 52–53
 - provider catalogs, vAPPs and virtual machines, 53–55
 - Provider Virtual Data Center, 51
 - securing Cloud Director on the internet, 62–64
 - Support Management Organization, 53
 - user interface, 61
 - security, 347
 - tenant constructs, 66–67
 - Allocation Models, 67–69
 - Compute Policies, 69–72
 - Organization Virtual Data Center (Org. VDC), 67
 - tenant application catalogs and libraries, 72
 - tenant networking, 73–74
 - workload types virtual machine versus vApp Blueprint, 72–73
- VMware Cloud Director Container Service Extension, 73
- VMware Cloud Director Service (VCDS), 266–267
- VMware Cloud Director tenant networking, 223
 - Cloud Director External Networks, 225–228
 - Network Pools, 229–230
 - NSX Data Center and Cloud Director integrated architecture, 223–225
 - Organization Virtual Datacenter Networks, 228–229
 - vApp Networks, 230
- VMware Cloud Foundation (VCF), 81, 128, 256–259, 300
 - designing with Cloud Director solution, 259–260
- VMware Cloud Marketplace and Bitnami for Cloud Providers, 322–323
- VMware Cloud on AWS, 10, 11, 264–266
- VMware Cloud on AWS MSP Delivery Model, 302
- VMware Cloud Provider Hub, 108
- VMware Cloud Provider Platform object storage on, 315–316

- VMware Cloud Provider Program (VCP), 13, 265
- VMware Cloud Provider Stack* (VCPS), 268
- VMware Integrated OpenStack (VIO), 108
- VMware Marketplace, 10, 55
- VMware NSX, 348
- VMware NSX Advanced Load Balancer, 234–235
- VMware NSX Edge, 114
- VMware Remote Console (VMRC), 64
- VMware Remote HTML5 Console or Standalone VMRC, 64
- VMware Validated Designs (VVD) for Cloud Providers, 15–16
- VMware vCloud Availability (vCAv), 273
 - architecture overview, 274–276
 - Cloud to Cloud (C2C) overview, 280
 - disaster recovery to cloud (DR2C) overview, 276–280
 - vCloud Availability migration verses HCX, 282
- VMware vRealize Operations Manager, 337
- VMware Wavefront, 320
- VMware 'Cloud Verified' services, 18
- VMware's Validated Designs (VVD), 111–112
- vRealize Automation Cloud, 320
- vRealize Log Insight (vRLI), 330, 334, 336, 342
 - extending, 336–337
- vRealize Operations, 11
 - and Tenant App, 11–13
- vRealize Operations Federated Deployment, 332–334
- vRealize Operations Federation Management Pack, 332
- vRealize Operations Management Pack, 59
 - for Cloud Director, 12
 - for Cloud Director and the Tenant App, 292
 - App Launchpad, 323–325
 - Central Point of Management (CPOM), 294–296
 - Cloud Provider Hub (CPH) is, 301–304
 - common storage services, 307–315
 - Container Service Extension (CSE), 304–307
 - data protection & ISV Service Ecosystem, 325–327
 - Machine Learning-as-a-Service (MLaaS), 316–320
 - Multi-Cloud
 - Management-as-a-Service, 320–322
 - SDDC as-a-Service with CPOM, 297–300
 - tenant metering, 294
 - VMware Cloud Marketplace and Bitnami for Cloud Providers, 322–323
- vRealize Operations Manager (vROps), 330, 331
 - design, 331–332
 - high-level monitoring and altering architecture, 333
- vRealize Operations Tenant App, 66
 - for Cloud Director, 60–61
- vRealize Orchestrator (vRO), 242, 245, 247, 250, 252, 272
 - Cloud administration automation, 243–244
 - Cloud Consumer operational automation, 246
 - integration, 48
 - Organization administration automation, 244–245
 - plug-in support for Cloud Director, 242–243
 - service delivery, 62
- vSAN
 - all-flash architecture, 148
 - architecture, 147
 - application support, 160
 - data integrity, 152–153
 - data locality, 155–156
 - data replication, 157–158
 - host and vSphere cluster failure tolerance, 150–152
 - quality of service, 154–155
 - rack awareness, 153–154
 - space efficiency, 147–150
 - swap efficiency, 154
 - vSAN Encryption, 159–160
 - vSAN Stretched Clusters, 156–157

- cluster design, 160–161
 - datastore abstraction verses storage
 - array volumes, 184
 - design factors, 177
 - encryption framework, 159
 - hyper-converged based storage, 118
 - policy design, 169–170
 - application requirements, 174
 - vSAN default policy, 170–174
 - vSAN policy capabilities, 170
 - and recoverability, 174
 - absent failures, 175–176
 - degraded failures, 176
 - understanding failure impact, 175
 - storage resource isolation, 186
 - stretched cluster architecture, 157
 - stretched cluster deployments, 177–180
 - vSAN design, 144–146, 160
 - disk controllers, 164–165
 - disk group design, 162–164
 - failures-to-tolerate policy, impact on
 - sizing of number of, 164
 - flash-based devices, 166
 - host CPU overhead, 162
 - host memory requirements, 161–162
 - network requirements, 166
 - jumbo frames, 168
 - network interface cards (NIC), 167
 - network QoS and network I/O
 - control, 168–169
 - virtual switches, 167–168
 - VLANs, 168
 - VMkernel network port group, 168
 - vSAN and recoverability, 174
 - absent failures, 175–176
 - degraded failures, 176
 - understanding failure impact, 175
 - vSAN cluster design, 160–161
 - vSAN policy design, 169–170
 - application requirements, 174
 - vSAN default policy, 170–174
 - vSAN policy capabilities, 170
 - vSAN Encryption, 159–160
 - vSAN Object-Based Storage System, 145
 - vSAN storage and cloud director, 180
 - Cloud Director NFS storage
 - requirement, 194–196
 - Cloud director storage design with
 - vSAN, 186
 - impact of vSAN on Cloud Director
 - compute and network resources, 193–194
 - multiple tiered resource clusters as
 - dedicated provider VDCs (model), 187–190
 - single provider VDC hosting multiple
 - tiered resource clusters (model), 190–193
 - designing a highly scalable HCI
 - platform, 196
 - vSAN storage resource isolation, 183–186
 - vSAN Stretched Clusters, 156–157
 - vSphere, 2, 3, 4, 9, 11, 346
 - vSphere clusters, 135
 - vSphere Content Pack, 336
 - vSphere Distributed Resource Scheduler (DRS), 136, 145
 - vSphere Distributed Switch (VDS), 216, 220
 - vSphere High Availability (HA), 135–136, 145, 161
 - vSphere network constructs, 220–221
 - GENEVE overlay, 222–223
 - N-VDS teaming policies, 221–222
 - vSphere on-premises architecture, 277
 - vSphere Standard Switch (VSS), 220
 - vSphere storage and arrays features,
 - considerations for use of, 201
 - vSphere Update Manager, 142
 - vSphere vMotion, 145
 - vSphere vMotion traffic, 348
 - vStorage APIs for Data Protection (VADP), 42
 - vVol Storage Policies and Storage Policy Based Management, 207
 - VXLAN, 131
 - VXLAN-based workloads, 21
- W**
- WAN interconnect high availability, 30
 - Wavelength division multiplexing (WDM), 24–27
 - Web application, 13
 - Web Application Firewall (WAF), 239

- Wide-Area Networking (WAN) platforms,
 - evaluation of, 24
 - dark fibre, 24
 - Multiprotocol Label Switching (MPLS), 27–29
 - Virtual Leased Line (VLL), 27
 - Virtual Private LAN Services (VPLS), 29
 - hybrid WAN architecture, 30
 - securing communication, 30–31
 - WAN interconnect high availability, 30
 - wavelength division multiplexing (WDM), 24–27
 - Wide Area Networking, 21
 - Windows Server 2019 Enterprise Edition
 - operating system platform, 31
 - Windows Server Failover Clustering (WSFC), 175
 - Workload Domain, 258
 - Workloads, bucketizing, 284
 - Workload types virtual machine verses vApp Blueprint, 72–73
- X**
- XaaS, 41, 48, 55, 62

Your Guide to Designing and Building the VMware Cloud Provider Platform

The VMware Cloud Provider Platform is a ready-to-run enterprise class public cloud solution that enables cloud providers, and enterprises that choose to operate in a similar manner, to quickly and securely build a true hybrid cloud platform to modernize data center infrastructures.

Built on a VMware software-defined data center architecture, the cloud provider platform uses the same tools, technologies, and skills that VMware customers already use on-premises, while delivering new cloud capabilities that allow IT teams to protect, extend, or replace workloads in the cloud. Unlike many public cloud offerings, the cloud provider platform supports any application, new or existing, virtual machine or container, without the need for application changes. In addition, built-in native disaster recovery mechanisms to ensure business continuity offer a security model that encompasses the dynamic nature of the infrastructure and boundaries to protect consumer assets.

The VMware's Cloud Provider Platform drives the idea of customer control and provides a robust solution architected for high availability. This together with the ability to optimize resources and ensure service levels ensures an overwhelmingly positive experience for both cloud provider and consumer.

Including:

- Modernizing the Enterprise Public Cloud with the VMware Cloud Provider Platform
- Helping customers adopt new technologies and use cases
- Leveraging the latest compute, storage, and networking technologies, so you can run any app, with an architecture designed with the future in mind, you can support new use cases like microservices and cloud-native apps
- Increasing efficiency and lowering operating risk
- Reducing complexity and troubleshooting time with built-in monitoring, proactive alerts, and a modular architecture that enables you to grow-as-you-go

[Learn how to implement a VMware Verified Cloud Provider Platform](#)

[Discover how to build a cloud management platform for operational efficiency](#)

[Find out how to leverage vRealize Orchestrator for cloud automation across the platform](#)

Martin Hosken is a Chief Technologist and part of Cloud Technical Services at VMware. Martin consults and designs cloud solutions with international customers and partners and serves as a trusted adviser in the design and transition of legacy infrastructure onto VMware software-defined cloud platforms.

vmware PRESS

