Hyperconverged Infrastructure Data Centers

Demystifying HCI

ciscopress.com

Sam Halabi

FREE SAMPLE CHAPTER

SHARE WITH OTHERS

Hyperconverged Infrastructure Data Centers

Demystifying HCI

Sam Halabi

Cisco Press

Hyperconverged Infrastructure Data Centers

Sam Halabi

Copyright© 2019 Bassam Halabi

Published by:
Cisco Press

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without written permission from the publisher, except for the inclusion of brief quotations in a review.

01 18

Library of Congress Control Number: 2018959195

ISBN-13: 978-1-58714-510-0
ISBN-10: 1-58714-510-3

Warning and Disclaimer

This book is designed to provide information about hyperconverged infrastructure data centers. 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 is provided on an "as is" basis. The author, Cisco Press, and Cisco Systems, Inc. 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 or from the use of the discs or programs that may accompany it.

The opinions expressed in this book belong to the author and are not necessarily those of Cisco Systems, Inc.

Trademark Acknowledgments

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Cisco Press or Cisco Systems, Inc., cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

Feedback Information

At Cisco Press, our goal is to create in-depth technical books of the highest quality and value. Each book is crafted with care and precision, undergoing rigorous development that involves the unique expertise of members from the professional technical community.

Readers' feedback is a natural continuation of this process. If you have any comments regarding how we could improve the quality of this book, or otherwise alter it to better suit your needs, you can contact us through email at feedback@ciscopress.com. Please make sure to include the book title and ISBN in your message.

We greatly appreciate your assistance.

Editor-in-Chief: Mark Taub

Alliances Manager, Cisco Press: Arezou Gol

Product Line Manager: Brett Bartow

Managing Editor: Sandra Schroeder

Development Editor: Marianne Bartow

Senior Project Editor: Lori Lyons

Copy Editor: Gill Editorial Services

Technical Editors: Mallik Mahalingam,
Aaron Kapacinskas

Editorial Assistant: Cindy J. Teeters

Cover Designer: Chuti Prasertsith

Composition: codemantra

Indexer: Erika Millen

Proofreader: Lori Eby

About the Author

Sam Halabi is a well-known industry figure with many years of experience in the field of information technology, multicloud, hyperconvergence, enterprise software, and data networking. Sam is a trusted advisor, capable of establishing close relationships with customers at the executive level, linking complex technologies with business benefits. Sam has worked at major companies in the United States and international markets, where he led sales, presales, consulting, marketing, and business development efforts targeting enterprises building scalable data centers. Sam is the founder of VirtuService (www.virtuservice.com), a provider of customer service, and IT consulting in private, hybrid, public cloud, and multicloud.

Sam has authored many Cisco Press books, including the bestsellers *Internet Routing Architectures* and *Metro Ethernet*.

Follow Sam Halabi on Twitter @VirtuService.

About the Technical Reviewers

Aaron Kapacinskas attended the University of California at San Diego for his undergraduate education in Mechanical Engineering. He subsequently attended Stanford University where he obtained his master's degree.

He is currently a Senior Technical Marketing Engineer at Cisco Systems, Inc., focusing on hyperconverged infrastructure as it relates to stretched deployments in synchronous replication, as well as security-related items for complex HCI deployments. He has written the *Cisco HyperFlex Hardening Guide* along with numerous white papers covering topics from encryption, native replication, and metro-level DR deployments, and is currently under patent review related to Cisco's HX replication technology.

Mallik Mahalingam is a Distinguished Engineer and CTO of HyperFlex product at Cisco, where he leads the technological and production direction for HyperFlex product-line.

Prior to Cisco, Mallik co-founded Springpath Inc., in 2012. Springpath built the software that was branded as Cisco HyperFlex, and he ran the company as its CEO and CTO until its acquisition by Cisco in 2017.

Prior to founding Springpath, Mallik led the core vSphere networking team at VMware, and built several features including VXLAN. As an early member of the VMware R&D team, which he joined in 2002, he built several networking and storage products during his decade of stay at VMware.

Mallik has developed several server technologies and products for enterprise markets over his 28 years of industry experience. He holds close to 80 US Patents and six published papers in leading academic conferences and journals.

Dedication

I would like to dedicate this book to my wonderful family, who continues to endure the agonizing process of book writing.

To my beautiful wife Roula: Thank you for your support and constant encouragement and for taking on double chores so I could finish this book.

To my wonderful sons, Joe and Jason: I am very proud of your academic accomplishments in the field of computer engineering, and I thank you for keeping me plugged into the latest technologies.

Last but not least, to my loving mother Josephine: Thank you for your persistence in asking me every day whether the book was done—that by itself was a huge incentive to finish writing it.

Acknowledgments

Special thanks to Mallik Mahalingam, CTO of the HyperFlex product and formerly cofounder and CEO of Springpath. Mallik's deep knowledge of hyperconvergence was insightful, especially in the architecture of software-defined storage and distributed file systems. I would like to thank Aaron Kapacinskas from the Cisco HyperFlex Product Management team for taking the time to thoroughly review the book and keep me up to speed on the latest in the HyperFlex development. Many thanks to Michel Khouderchah from the Cisco security business group for his input on hyperconvergence and for encouraging me to write the book. Also, special thanks to my brother Mitri Halabi for his support in the areas of data networking, hyperconvergence, and multicloud; your deep knowledge and guidance were very instrumental. Much research for this book was done by sifting through heaps of design guides, specifications, and videos, so many thanks go to all of the technology professionals out there. I do reference their work where appropriate.

This book couldn't have been possible without the support of many people on the Cisco Press team. Brett Bartow, product-line manager Pearson IT Professional Group, was instrumental in sponsoring the book and driving it to execution. Sandra Schroeder, managing editor and design manager, was masterful with book graphics. Marianne Bartow, development editor, did a wonderful job in the technical review cycle; it was a pleasure working with you. Lori Lyons, senior project editor, thank you for leading the book to success through the production cycle. Finally, many thanks to the numerous Cisco Press unknown soldiers working behind the scenes to make this book happen.

Contents at a Glance

Introduction xxiv

PART I Basics of Data Center Networking and Storage 1

Chapter 1 Data Networks: Existing Designs 3

Chapter 2 Storage Networks: Existing Designs 19

PART II Evolution in Host Hardware and Software 63

Chapter 3 Host Hardware Evolution 65

Chapter 4 Server Virtualization 77

Chapter 5 Software-Defined Storage 95

PART III Hyperconverged Infrastructure 105

Chapter 6 Converged Infrastructure 107

Chapter 7 HCI Functionality 117

Chapter 8 HCI Business Benefits and Use Cases 135

PART IV Cisco HyperFlex 151

Chapter 9 Cisco HyperFlex 153

Chapter 10 Deploying, Provisioning, and Managing HyperFlex 197

Chapter 11 HyperFlex Workload Optimization and Efficiency 211

PART V Alternative HCI Implementations 221

Chapter 12 VMware vSAN 223

Chapter 13 Nutanix Enterprise Cloud Platform 255

Chapter 14 Open Source—Compute and Storage 289

PART VI Hyperconverged Networking 305

Chapter 15 Software-Defined Networking and Open Source 307

Chapter 16 VMware NSX 335

Chapter 17 Application-Centric Infrastructure 351

PART VII Public, Private, Hybrid, and Multicloud 393

Chapter 18 The Public Cloud 395

Chapter 19 The Private Cloud 423

Chapter 20 Hybrid Cloud and Multicloud 439

Glossary 451

Index 475

Contents

Introduction xxiv

PART I Basics of Data Center Networking and Storage 1

Chapter 1 Data Networks: Existing Designs 3

Information Technology Equipment of a Data Center 4

Network Equipment 4

Networking Services 4

Traffic Redirection and Optimization 5

Security 5

Multitier Data Networking Architecture 6

Logical Server Grouping 8

Challenges of Existing Designs 9

Oversubscription Between the Tiers 9

Large Flat L2 Networks with Stretched VLANs 10

Traffic Hopping Between Tiers, Inducing Latency 11

Inter-VLAN Routing via SVI 12

Complexity of Mechanisms Used for IPv4 Address Scarcity 12

Private VLANs 13

Flooding of Broadcast, Unknown Unicast, and Multicast (BUM) Traffic 15

Loop Prevention Via Spanning Tree 16

Firewall Overload 17

Looking Ahead 18

Chapter 2 Storage Networks: Existing Designs 19

The Storage View of Multitier Designs 20

Types of Disk Drives 21

Hard Disk Drives 22

SATA Drives 22

SAS Drives 23

SAS Self-Encrypting Drives 23

FC Drives 23

Solid-State Drives 23

Disk Performance 23

Throughput or Transfer Speed 24

Access Time 24

Latency and IOPS 24

RAID 26

RAID 0 26

RAID 1 26

RAID 1+0 26

RAID 0+1 27

RAID 5 28

RAID 6 29

Storage Controllers 30

Logical Unit Numbers 31

Logical Volume Manager 33

Block-, File-, and Object-Level Storage 35

Block-Level Storage 35

File-Level Storage 35

Object-Based Storage 36

Storage-to-Server Connectivity 37

Direct-Attached Storage (DAS) 38

Network-Attached Storage 39

Storage Area Networks 40

Fibre Channel SANs 41

Fibre Channel Addressing 42

Fibre Channel Zoning 43

LUN Masking 43

Fibre Channel Switching 44

Multipathing 44

iSCSI SANs 46

iSCSI Addressing 47

iSCSI Network Segmentation 48

iSCSI LUN Masking 48

iSCSI Multipathing 48

Fibre Channel over Ethernet SANs 49

Storage Efficiency Technologies 50

Thin Provisioning 50

Snapshots 51

Traditional or Physical Snapshots 52

Copy on Write Versus Redirect on Write Logical Snapshots 52

Cloning 55

Replication 55

Deduplication 55

File-Level or Object-Level Deduplication 56

Deduplication at the Block Level 56

Target- Versus Source-Based Deduplication 57

Inline- Versus Post-Process Deduplication 58

Local Versus Global Deduplication 58

Data Compression 58

Disk Encryption 59

Storage Tiering 59

Caching Storage Arrays 60

Looking Ahead 61

PART II Evolution in Host Hardware and Software 63

Chapter 3 Host Hardware Evolution 65

Advancements in Compute 65

x86 Standard Architecture 66

Single-, Multi-, and Many-Cores CPUs 66

Physical Cores Versus Virtual Cores Versus Logical Cores 67

Virtual CPU 68

Evolution in Host Bus Interconnect 70

Non-Volatile Memory Express 71

Emergence of Flash-Based Products 72

Enhancement in Flash Technology 73

New Breed of Storage Arrays Falls Short 73

All-Flash Arrays 73

Hybrid Arrays 74

Host-Based Caching 75

Looking Ahead 75

Chapter 4 Server Virtualization 77

The Virtualization Layer 78

Type 1 Hypervisor 79

Type 2 Hypervisor 80

Docker Containers 80

Datastores 82

Virtual Machine Creation 84

LUN Configuration 84

Datastore Configuration 84

New VM Creation 84

Guest OS Assignment 85

vCPU Assignment 85

Memory Allocation 85

Adding the Network Interface Card and SCSI Controller 85

Assigning a Virtual Disk 85

Power On and OS Installation 86

Virtualization Services 86

Clusters of Servers or Nodes 86

VM Migration 87

High Availability 88

Fault Tolerance 89

Compute Load Balancing 89

Storage Migration 90

Storage Load Balancing 90

Provisioning and Management 90

Virtual Switching 90

Looking Ahead 93

References 93

Chapter 5 Software-Defined Storage 95

SDS Objectives 96

Preserving the Legacy and Offering New Features 97

vSphere APIs for Storage Awareness (VASA) and VVols 99

Creating More Granular Volumes with VVols 100

Learning Storage Array Capabilities Through VASA 102

Integration with Storage Policy–Based Management 103

Looking Ahead 104

PART III Hyperconverged Infrastructure 105

Chapter 6 Converged Infrastructure 107

Cisco UCS—The First Step in Convergence 108

The Converged Systems 112

Pros of Converged Systems 114

Converged Systems Cons 114

Looking Ahead 116

Chapter 7 HCI Functionality 117

Distributed DAS Architecture 118

Distributed Controllers 119

Scale-Out Architecture 120

HCI Performance 120

Resiliency Against Hardware Failures via Replication 121

File Systems 122

Change in the Provisioning Model 124

Hardware Acceleration 125

Networking Integration 125

Networking Policies 126

Networking Automation 126

Advanced Data Storage Functionality 127

Deduplication and Compression 128

Erasure Coding 128

Replication and Backup for Disaster Recovery 129

HCI Security 130

HCI Provisioning, Management, and Monitoring 131

Looking Ahead 133

Chapter 8 HCI Business Benefits and Use Cases 135

HCI Business Benefits 136

Fast Deployment 136

Easier-to-Scale Infrastructure 136

Enhanced IT Operational Model 137

Easier System Management 138

Public Cloud Agility in a Private Cloud 138

Higher Availability at Lower Costs 139

Low-Entry Cost Structure 139

Reduced Total Cost of Ownership 140

HCI Use Cases 140

Server Virtualization 140

DevOps 141

Virtual Desktop Infrastructure 141

Problems of VDI with Legacy Storage 143

How HCI Helps VDI 144
Remote Office Business Office (ROBO) 144
Edge Computing 146
Tier-1 Enterprise Class Applications 146
Data Protection and Disaster Recovery 148
Looking Ahead 149

PART IV Cisco HyperFlex 151

Chapter 9 Cisco HyperFlex 153
HyperFlex Physical Components 154
Cisco HyperFlex Hybrid Nodes 156
Cisco HyperFlex All-Flash Nodes 156
Cisco HyperFlex Edge Nodes 157
Cisco HyperFlex Compute-Only Nodes 157
Cisco USC Fabric Interconnect 157
Cisco UCS 6200 and 6300 Fabric Interconnect 158
Cisco UCS 6248UP 48-Port FI 158
Cisco UCS 6296UP 96-Port FI 158
Cisco UCS 6332 32-Port FI 158
Cisco UCS 6332-16UP 40-Port FI 158
Cisco C220/C240 M4/M5 Rack Servers 158
Cisco VIC MLOM Interface Card 159
Cisco UCS 5108 Blade Chassis 159
Cisco UCS B200 Blade Server 159
Cisco UCS XP Fabric Extender 159
HyperFlex Performance Benchmarks 160
Integration with UCS 162
Logical Network Design 162
Service Templates and Profiles 164
vNIC Templates 166
HyperFlex Integration with External Storage 167
Cisco's HX Data Platform 168
HX Data Platform Controller 169
HyperFlex in VMware ESXi Environment 170
HyperFlex in Hyper-V Environment 171
Docker Containers Support and Volume Driver 172

HyperFlex Data Distribution 174

HyperFlex Data Striping 176

Data Protection with Replication Factor 178

Cluster Resiliency 179

Logical Availability Zones 181

Details of Read and Write Operations 181

Log-Structured File System 181

Data Virtualization 182

The Lifecycle of a Write Operation 183

Destaging and Data Optimization 185

The Lifecycle of a Read Operation 186

Advanced Data Services 187

Deduplication and Compression 187

Snapshots 188

Cloning 189

Asynchronous Native Replication for DR with Remote Clusters 189

Synchronous Native Replication for DR with Stretched Clusters 190

Integration with Third-Party Backup Tools 191

HyperFlex Security 192

Securing Data-at-Rest 192

Operational Security 193

Complying with Security Regulations and Certifications 194

Management Security 194

Looking Ahead 195

References 195

Chapter 10 Deploying, Provisioning, and Managing HyperFlex 197

Installation Phase 197

HyperFlex Workload Profiler 199

HyperFlex Sizer 199

Management Provisioning and Monitoring 199

Cisco HyperFlex Connect HTML5 Management 200

Dashboard 200

Monitor 200

Analyze 200

Protect 201

Manage 202

VMware vSphere Management Plug-In 203

Summary Tab 203

Monitor Tab 203

Manage Tab 204

Cisco Intersight 204

Claiming the Nodes 205

Looking Ahead 209

References 210

Chapter 11 HyperFlex Workload Optimization and Efficiency 211

Enterprise Workload Issues 211

HyperFlex with Cisco Tetration 212

Data Collection 214

Tetration Analytics Cluster 214

Open Access 215

Using the Data 215

Cisco Workload Optimizer 216

Cisco AppDynamics 217

Looking Ahead 219

References 219

PART V Alternative HCI Implementations 221

Chapter 12 VMware vSAN 223

vSAN Physical Components 224

vSAN Hyperconvergence Software 225

The Object File System 226

vSAN Datastore 228

vSAN Storage Policies 228

Caching 232

I/O Operation Details 232

vSAN Advanced Functionality 233

Data Integrity 234

Data Encryption 234

Deduplication and Compression 235

Erasure Coding 236

Snapshots 236

Cloning 236

vSAN Replication for Disaster Recovery via Stretched Clusters 238

vSAN Backup for Disaster Recovery 241

Integration with Legacy SAN and NAS 243

vSAN iSCSI Target 243

vSAN and VVols 243

SMB and NFS Support 244

Persistent Storage for Containers 244

vSAN Management 244

Graphical Interfaces 244

Ease of Installation 245

Cloud-Connected Health Checks 245

Performance Diagnostics 245

VMware Update Manager 245

vSAN vRealize Operations and Log Insight 245

Thoughts on vSAN Versus HyperFlex 246

Hardware Comparison 247

Scaling Up 247

vSAN In-Kernel Versus Controller-Based Solutions 248

Distributed Versus Not-So-Distributed File System 249

One-to-One Versus Many-to-Many Rebuild 250

Implementations of Compute-Only Nodes 250

Advanced Data Services 251

Management Software 252

Networking 252

Looking Ahead 253

References 253

Chapter 13 Nutanix Enterprise Cloud Platform 255

Nutanix Enterprise Cloud Platform 256

ECP Hyperconvergence Software 257

Distributed Storage Fabric 257

Nutanix Cluster Components 258

Physical Drive Breakdown 260

I/O Path 261

Write I/O 261

Read I/O 262

Data Protection 262
- Metadata 263
- Availability Domains 263
- Data Path Resiliency 264

Nutanix Advanced Functionality 264
- Deduplication 264
- Data Compression 265
- Erasure Coding 266
- Disk Balancing 266
- Storage Tiering 267

Snapshots and Clones 268
- Shadow Clones 269
- Era Database Services 269
- Backup and Restore, Replication, and Disaster Recovery 270

Metro Availability: Stretch Clustering 271

Data At Rest Encryption 272

Nutanix Acropolis Block Services 272

Nutanix Acropolis File Services 273

Support for Hyper-V 274

Docker Containers 275

Provisioning, Managing, and Monitoring 275
- Infrastructure Management 276
- Operational Insight 277
- Nutanix Tools 278
- Calm Orchestration Tool 278

The Nutanix Competitive Landscape 279
- Hardware Comparison 280
- Distributed Architecture 281
- Log-Structured Versus Write-in-Place File System 282
- Data Tiering 282
- Deduplication 283
- *Arguments Against Post-Process Deduplication 283*
- *Arguments Against Inline Deduplication 284*
- Data Locality 285
- *Pros of Data Locality 285*
- *Cons of Data Locality 286*

Looking Ahead 287

Reference 287

Chapter 14 Open Source—Compute and Storage 289

OpenStack 290

Nova 294

Cinder Block Storage 296

Swift 297

Proxy Server 298

Ceph 300

Looking Ahead 303

References 303

PART VI Hyperconverged Networking 305

Chapter 15 Software-Defined Networking and Open Source 307

The SDN Background 308

The Overlay and Microsegmentation Edge 309

Host-Based Networking 310

Switch-Based Networking 312

The Switching Fabric 313

The Underlay Network 315

The Overlay Network 315

Microsegmentation in the Data Center 319

Networking Open Source Initiatives 320

Neutron 320

Neutron Server 320

Neutron-APIs 321

Neutron Plug-Ins 321

Agents 321

OVS Architecture 322

OVN—The Open Source SDN 324

Open vSwitch 325

OVN 326

OVN Functionality 326

OVN High-Level Architecture 327

Improvements over OVS and Neutron Integration 328

State of Vendors with Open Source 331

Looking Ahead 332

References 333

Chapter 16 VMware NSX 335

Setting and Enforcing Policies in NSX 336

Security Groups 336

Security Policy 337

Policy Enforcement 338

The NSX Manager and Controller Cluster 339

NSX Manager 339

The NSX Controller Cluster 340

Enhancements for vDS 341

Flooding Avoidance 342

NSX L2 Switching and L3 Routing 343

NSX L2 Switching 343

NSX IP Routing 343

Distributed Logical Router 343

Handling of Multidestination Traffic 346

Basic Multicast Model with No Optimization 347

Unicast Method of Replication 348

Looking Ahead 350

Reference 350

Chapter 17 Application-Centric Infrastructure 351

Cisco Application-Centric Infrastructure 352

ACI Microsegmentation Constructs 353

The Endpoint Groups 353

Application Network Profile 355

Service Graphs 358

ACI Tetration Model 360

Cisco Application Policy Infrastructure Controller 360

ACI Domains 362

Virtual Machine Manager Domain 362

Physical and External Domains 364

Physical Domains 365

External Domains 365

The ACI Fabric Switching and Routing Constructs 365

Tenant 366

VRF 366

Bridge Domain 366

EPG 366

Virtual and Physical Connectivity to the ACI Fabric 367

Virtual Connectivity to the Fabric 367

Physical Connectivity to the Fabric 368

The ACI Switching and Routing Terminology 369

The ACI Underlay Network 371

Handling External Routes 372

ACI Fabric Load Balancing 373

The ACI Overlay and VXLAN 373

The VXLAN Instance ID 376

L2 Switching in the Overlay 378

L3 Switching/Routing in the Overlay 380

Multicast in the Overlay Versus Multicast in the Underlay 383

ACI Multi-PoD 383

ACI Multi-Site 384

ACI Anywhere 386

High-Level Comparison Between ACI and NSX 386

Policy Setting 387

Policy Enforcement 388

Performance Requirement for VXLAN 388

Control Plane 389

Performance of Data Forwarding 390

Automation and Visibility in the Fabric 390

Networking Learning Curve 391

Looking Ahead 392

References 392

PART VII Public, Private, Hybrid, and Multicloud 393

Chapter 18 The Public Cloud 395

The Cloud Services 395

Infrastructure as a Service 396

Platform as a Service 397

Software as a Service 397

Amazon Web Services 398

AWS Global Infrastructure with Regions and Availability Zones 399

Networking 401

Storage 404

AWS S3 Storage 405

Amazon Elastic Block Storage 406

Launching Multitier Applications in AWS 409

Compute Instances 409

Amazon Machine Images 410

Security Groups 410

Identity and Access Management 411

Launching an EC2 Instance 412

Adding Volumes 414

Auto Scaling 414

Cloud Monitoring 416

Cloud Automation 416

Infrastructure as a Code 417

Software Development Kits 420

Looking Ahead 421

References 421

Chapter 19 The Private Cloud 423

What Is a Private Cloud? 423

Convergence and Hyperconvergence 425

Automation and Orchestration 425

Cisco UCS Director 427

UCS Director Policies 427

Virtual Data Center 429

Orchestration Concepts 430

Catalogs 431

Integration Between UCSD and HyperFlex 431

UCSD Interfaces 433

APIs and Infrastructure as a Code 434

Looking Ahead 437

References 437

Chapter 20 Hybrid Cloud and Multicloud 439

Why Hybrid Cloud? 440

Why Multicloud? 442

Cisco CloudCenter 446

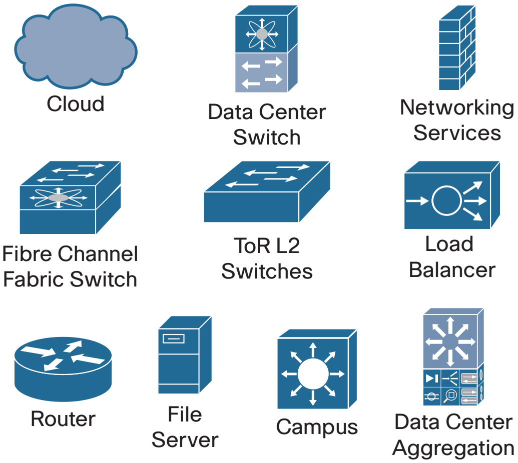
Looking Ahead 450

References 450

Glossary 451

Index 475

Icons Used in This Book



Command Syntax Conventions

The conventions used to present command syntax in this book are the same conventions used in Cisco's Command Reference. The Command Reference describes these conventions as follows:

- **Boldface** indicates commands and keywords that are entered literally as shown. In actual configuration examples and output (not general command syntax), boldface indicates commands that are manually input by the user (such as a **show** command).
- *Italics* indicate arguments for which you supply actual values.
- Vertical bars (|) separate alternative, mutually exclusive elements.
- Square brackets [ ] indicate optional elements.
- Braces { } indicate a required choice.
- Braces within brackets [{ }] indicate a required choice within an optional element.

Note This book covers multiple operating systems, and a differentiation of icons and router names indicates the appropriate OS that is being referenced. IOS and IOS XE use router names like **R1** and **R2** and are referenced by the IOS router icon. IOS XR routers will use router names like **XR1** and **XR2** and are referenced by the IOS XR router icon.

Introduction

Hyperconverged infrastructure (HCI) is the integration of different technologies such as compute, storage and storage networking, virtualization, data networking, and automation, all under the umbrella of software-defined storage (SDS) and software-defined networking (SDN). Legacy data centers are normally built by system administrators, networking engineers, storage administrators, software virtualization engineers, and network management engineers, all working within their own area of expertise. This creates different silo groups within the same IT organization trying to optimize their own set of tasks without much visibility into the other groups. In essence, the IT professionals are segmented and find pride in being the experts in their own tasks. So here comes HCI with a paradigm shift that integrates all the different technology fields into one product—hence, the need for cross-technology knowledge.

This book approaches the HCI topic from the point of view that any individual working this field needs to have enough knowledge in all the different areas, including storage, storage networking, compute, virtualization, switching and routing, and automation. The book explains each area in the context of a legacy data center design, detailing the problem statement for the particular technology and how HCI solves the problem and to what extent. The *Hyperconverged Infrastructure Data Centers* book will be the bible for IT professionals, technical folk, and management in all technology areas. It will guide them through the decision process to move in the HCI direction.

Despite this being a Cisco Press book, this book tries to be vendor neutral by objectively comparing and contrasting the Cisco HyperFlex approach with other vendor implementations. The SDS and the distributed storage file systems that the Cisco HyperFlex HX Data platform has adopted, VMware vSAN and Nutanix Enterprise Cloud software, are described. A compare and contrast between the different SDN solutions—such as Cisco application-centric infrastructure (ACI), VMware network virtualization and security platform (NSX), and the open source Open vSwitch (OVS)/Open Virtual Network (OVN) and OpenStack are also described. This gives you enough ammunition to ask the right questions when choosing an HCI solution.

Aside from describing HCI in detail, a key aspect of this book is a comparison between HCI and public cloud services from Amazon Web Services (AWS). This book helps IT professionals, chief information officers (CIOs), and information technology (IT) managers in their decision to move into an on-premise HCI deployment versus a public cloud deployment. It describes in detail and from a technical and business aspect the pros and cons for building on-premise versus outsourcing. This book also covers products such as the Cisco CloudCenter, which facilitate the migration into a multicloud environment.

Last but not least, this book goes into detail about automation and software management by gathering information from different HCI vendors and approaching the topic from a vendor-neutral angle, allowing the reader to make a decision on existing and needed automation functionality.

Goals and Methods

CIOs and IT professionals who want to simplify their IT and networking environment are now challenged with the decision of whether to move fully into the cloud, build their own HCI data centers, or both. Making such decisions depends on factors that include the scale and complexity of the CIO's and IT professionals' existing setup, the level of control over their own resources, security, availability of IT and networking resources, level of expertise, and overall fixed and recurring costs.

Because many new vendors are introducing products that offer HCI and are challenging the existing network design, the new technologies are becoming confusing to IT professionals who are trying to move into next-generation architectures while maintaining a current setup that is generating revenue. This book walks the reader step by step through the existing data center setups and the issues faced in scaling current designs. The book explains all HCI elements in detail and how to achieve high availability (HA) and scale in dealing with compute, storage, networking, and virtualization requirements. As automation becomes a key benefit of HCI, this book compares the automation functions and tools of HCI with cloud offerings from the likes of AWS and others. The objective is to provide IT administrators with a solid knowledge base across all technology areas and compare and contrast HCI offerings from different vendors. This gives you the tools to ask the right questions when you embark on the transformation of your data center into private and hybrid clouds.

Who Should Read This Book?

HCI is the integration of compute, storage, networking, virtualization, and automation that is driven by software. In general, IT professionals are divided in their areas of expertise. Individuals are spread into focus areas that overlap:

- Servers and virtualization
- Storage
- Backup and disaster recovery (DR)
- Storage networking
- Switching and routing
- Security
- Software applications
- Automation
- DevOps

Because HCI offers a unified platform that combines all these areas in one system, the coverage between the areas will blur. Although HCI will simplify the implementation of each area, cross-knowledge between these areas is essential. This means that the audience of this book is the sum of all system administrators, storage administrators, networking engineers, software virtualization engineers, and network management engineers. Also, because the book touches on the business aspects of HCI and the pros and cons of moving from private clouds to public clouds, IT managers and CIOs will benefit from understanding the impact that HCI has on the transformation of their data centers and the speed of deploying highly available applications.

How This Book Is Organized

For those readers who are familiar with the author's writing style from his previous bestseller books, including *Internet Routing Architectures*, Sam Halabi emphasizes easy reading and making the difficult look easy. This book goes through a smooth progression of the topics in a storytelling style. Many of the basic concepts are laid out in advance, so you do not miss a beat and feel comfortable progressing through the chapters. It is recommended that you read the chapters in order so that you get the full benefit of this book.

Networking, storage, compute, virtualization, and automation are not easy topics and are getting more complex every day. Not only are system administrators, networking engineers, storage engineers, and virtualization engineers asked to become multifunctional, they also need to become programmers. The learning curve is huge, and many people aren't sure where to start.

Sam Halabi has put a lot of effort into putting you on the right track and giving you the launch pad into tackling HCI. His many years of experience in both vendor and system integration tracking across different technology areas make a difficult topic such as HCI sound simple. The advantages you see from this book follow:

- An easy reading style with no marketing fluff or heavy technical jargon
- Progression through the chapters from easy to advanced topics
- Comprehensive coverage of the topic at both technical and business levels
- First book to address storage, compute, virtualization, networking, and automation in detail under one umbrella to bridge the technology gap between the different IT departments
- Benefit to IT professionals trying to evaluate whether to move in the HCI direction
- Benefit to IT management, CIO, and chief technology officer (CTO) in evaluating HCI versus the public cloud
- Coverage of the latest HCI functionality
- Discussion of automation as it compares to cloud offerings such as AWS
- Comparing and contrasting of different implementations objectively and with vendor neutrality

Book Structure

The book is organized into seven parts.

PART I: Basics of Data Center Networking and Storage

Chapter 1, "Data Networks: Existing Designs": This chapter describes the different networking equipment that constitutes the data center and discusses the challenges of the existing designs in meeting the needs of application-aware data centers. It covers challenges with the three-tier architecture, including oversubscription between the tiers, stretching VLANs over large L2 networks, latency of traffic crossing tiers, flooding of broadcast traffic, complexity in dealing with IPv4 address scarcity, loop prevention, and firewall overload.

Chapter 2, "Storage Networks: Existing Designs": This chapter presents an essential background of the different storage technologies and terminology. It describes the three-tier storage network and discusses fundamental topics such as disk drives, disk speed throughput and latency, and input/output operations per second (IOPS). The chapter familiarizes you with redundant array of independent disks (RAID) systems, storage controllers, and logical unit numbers (LUNs). Block-, file-, and object-level storage are discussed, as are the different storage architectures of storage area network (SAN) with iSCSI and fibre channel, network-attached storage (NAS), and direct-attached storage (DAS). You will see an overview of storage efficiency technologies, such as thin and thick provisioning, deduplication and compression, snapshots replication and cloning, caching, disk encryption, and storage tiering.

PART II: Evolution in Host Hardware and Software

Chapter 3, "Host Hardware Evolution": Advances in processing power make an individual server more powerful than the traditional storage controllers. This chapter covers the terminology of central processing unit (CPU), virtual cores, logical cores, and virtual CPUs. It discusses the latest advancement in host bus interconnect with technologies such as Peripheral Component Interconnect express (PCIe) and Non-Volatile Memory express (NVMe). The evolution in flash memory and cache and the emergence of a new generation of flash-based storage products are covered as well.

Chapter 4, "Server Virtualization": This chapter gives an overview of software virtualization, hypervisors, and the difference between virtual machines (VMs) and containers. It discusses the concepts of datastores and logical volumes and shows the steps in the creation of a VM. Different virtualization services such as VM migration are explored to provide HA and fault tolerance and to provide load distribution. The chapter also explores the concepts of standard and distributed virtual switches in providing VMs with all the networking attributes of a physical switch. The definition of networking and storage policies at the VM level puts the application at the center of attention.

Chapter 5, "Software-Defined Storage": This chapter introduces SDS and its objectives in decoupling storage software from hardware. Some early implementations of SDS are discussed, but they do not meet the goals of hyperconvergence. You will also learn about some important topics such as vSphere APIs for storage awareness (VASA) and Virtual

Volumes (VVols). Such architectures, although fit in a legacy converged model with storage arrays, pave the way for giving the application better visibility into the storage capabilities and putting the application in the driver seat.

PART III: Hyperconverged Infrastructure

Chapter 6, "Converged Infrastructure": This chapter introduces the Cisco Unified Computing System (UCS) and how it changed the server, compute, and networking landscape. It led to the emergence of what is called converged infrastructure (CI). The objective of CI is to simplify data center rollouts and management with better integration between the different products that make up the converged solution. The pros and cons of deploying CI are presented to give you a feel for why the data center transformation into hyperconvergence is a must.

Chapter 7, "HCI Functionality": This chapter defines HCI and delves into its functionality. A detailed description of the HCI physical and logical distributed architecture is covered. Distributed data controllers create a scale-out architecture that moves away from the legacy centralized storage architecture. Data replication provides hardware resiliency and data protection. The chapter introduces the log-structured file system (LFS) and its benefits for HCI. Advanced HCI functionalities are introduced, and you learn how services such as backup and disaster recovery are native to the architecture. You also get a feel for the new provisioning, deployment, and management model that simplifies the deployment of applications and setting of policies.

Chapter 8, "HCI Business Benefits and Use Cases": This chapter discusses HCI business benefits as seen by CIOs and IT managers who want justification for moving from a legacy SAN and converged environment to hyperconverged. The chapter discusses the multitude of HCI use cases ranging from simple server virtualization to more complex environments. It includes details about sample applications such as DevOps, virtual desktops, remote office business office (ROBO), edge computing, tier-1 enterprise-class applications, backup, and disaster recovery.

PART IV: Cisco HyperFlex

Chapter 9, "Cisco HyperFlex": The chapter gives an overview of the Cisco HyperFlex platform and its physical components. It discusses the integration between HyperFlex and UCS through the use of service profiles and templates. A detailed description of the HX Data Platform is presented, which covers the data platform controller, support for VMware ESXi, Microsoft Hyper-V, and Docker containers. The chapter also covers data distribution in the cache and capacity layers, and the life of read and write input/output (I/O). HyperFlex Advanced Data Services including deduplication and compression, snapshots, clones, synchronous replication for backup and disaster recovery, and integration with third-party backup software vendors are described. Also discussed is HyperFlex security with the use of self-encrypting drives and the adoption of industry security standards.

Chapter 10, "Deploying, Provisioning, and Managing HyperFlex": This chapter covers the deployment provisioning and management of the HyperFlex platform. It includes reference tools to help with sizing the platform depending on the workload. Different software products are used to help in managing all aspects of deploying and monitoring the data services from inside the private data center as well as from the public cloud.

Chapter 11, "HyperFlex Workload Optimization and Efficiency": This chapter describes the different issues that face enterprise workloads, including reactive mode to growth, lack of visibility into their applications, overprovisioning, and cloud creep. Better visibility into the traffic flows and the service chain between the applications is done using the Cisco Tetration platform. The Cisco Workload Optimizer (CWOM) automation tool monitors the performance consumption of applications and matches the resources with the needs of the application. Cisco AppDynamics is an application and business monitoring platform that allows enterprises to monitor their business applications and transactions and make sure they are delivering the best performance. The platform gives root cause analysis of performance issues at the code level.

PART V: Alternative HCI Implementations

Chapter 12, "VMware vSAN": vSAN is VMware's hyperconverged software product. This chapter describes vSAN's hardware implementation, including ready nodes and integrated systems. It also introduces the vSAN hyperconvergence software, including the object file system and the input/output (I/O) operation within the cache and capacity layers. The vSAN advanced data services, such as deduplication, compression, erasure coding (EC), snapshots and clones, disaster recovery, and backup are discussed as well. In addition, the chapter covers the integration of vSAN with legacy SAN and NAS. You will also see a high-level comparison between HyperFlex and vSAN, pointing out architectural differences and giving you enough ammunition to ask the right questions from your vendors.

Chapter 13, "Nutanix Enterprise Cloud Platform": This chapter describes the Nutanix Enterprise Cloud Platform software components. The distributed storage fabric, read/write I/O path, and data protection techniques are discussed. Similar to HyperFlex and vSAN, you will see a detailed description of the advanced data services, including deduplication, compression, EC, and support for backup and disaster recovery. The chapter also covers a competitive landscape, comparing the Nutanix architecture with HyperFlex, highlighting pros and cons in different areas.

Chapter 14, "Open Source—Compute and Storage": This chapter gives an overview of the open source approaches to hyperconvergence. It presents a description of OpenStack and the different components that are relevant to HCI, including Nova for Compute, Cinder for block storage, and Swift for object storage. Also, a description of important open source initiatives such as Ceph for combined block, file, and object services is presented.

PART VI: Hyperconverged Networking

Chapter 15, "Software-Defined Networking and Open Source": This chapter discusses SDN and its background and adoption in today's networking implementations. Host-based and switch-based networking options are presented, with a discussion of how they compete for the networking landscape. The chapter covers the benefits of the new leaf/spine 2-tier networking architecture. The Overlay/Virtual Extensible LAN (VXLAN), and Underlay networking models are covered, as is the important topic of microsegmentation. Open source networking via implementations from OpenStack Neutron, open source OVS, and Open Virtual Network (OVN) are also discussed.

Chapter 16, "VMware NSX": This chapter introduces VMware Network Virtualization and Security platform (NSX). This is VMware's solution for bringing automation, policy, and security into the networking environment. You will see a description of the NSX components, including the NSX manager and controllers cluster. In addition, you will read about the vSphere distributed switch (vDS) enhancements in support of VXLAN. Alternatively, VMware introduced a host-based networking solution by implementing IP routing using the concept of a distributed logical router (DLR) and an edge services gateway (ESG).

Chapter 17, "Application-Centric Infrastructure": The Cisco ACI is a measure to introduce a level of automation into setting and enforcing policies at the application level as well as configuring the switch fabric to support the connectivity requirements of the applications in the data center. This chapter describes how ACI works and presents the microsegmentation constructs, including endpoint groups (EPGs), application network profiles (ANPs), and bridge domains (BDs). The chapter covers the Application Virtual Switch (AVS) and the ACI switching and routing constructs, including overlay/VXLAN and underlay. ACI Multi-Point of Delivery (Multi-PoD), ACI Multi-Site, and ACI Anywhere concepts are also discussed. The chapter ends with a comprehensive comparison between Cisco ACI and VMware NSX, highlighting similarities and differences.

PART VII: Public, Private, Hybrid, and Multicloud

Chapter 18, "The Public Cloud": This chapter defines the different cloud models, such as infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS). It introduces services from AWS and highlights the AWS networking, storage, and compute capabilities. A description of how to launch a multitier application in AWS is presented, including initiating compute instances, identity and access management (IAM), security groups, storage, and monitoring. The chapter also covers the topic of cloud automation and the notion of infrastructure as a code.

Chapter 19, "The Private Cloud": This chapter describes, through the use of automation and orchestration, how to transform hyperconverged data centers into private clouds. It covers the characteristics and different models of a private cloud, distinguishing between automation and orchestration. Examples are drawn from the Cisco UCS Director (UCSD) and how it defines storage, compute, and networking policies to create service catalogs. Also discussed are the interaction between UCSD and HyperFlex and the creation of infrastructure as a code.

Chapter 20, "Hybrid Cloud and Multicloud": This chapter concludes the book with a definition of the hybrid cloud and multicloud and a description of the use cases and benefits. It draws on products from Cisco Systems to allow the ease of migration of applications and services between the different clouds. The chapter briefly presents the Cisco CloudCenter and its ability to decouple the application from the underlying cloud and orchestrating applications over private as well as public clouds. Finally, this chapter summarizes the HyperFlex hyperconvergence ecosystem and how it integrates with CloudCenter.

Figure Credits

Figures 9-3 and 9-4 “Hyperconverged Infrastructure with Consistent High Performance for Virtual Machines.” Tony Palmer and Kerry Dolan, March 2017, The Enterprise Strategy Group, Inc.

Figures 9-5, 9-9 © Cisco Systems, Inc.

Figures 10-1 to 10-3 © Cisco Systems, Inc.

Figure 10-4 VSphere Web Client © 2018 VMware, Inc.

Figures 10-5 to 10-8 © Cisco Systems, Inc.

Figure 11-1 © Cisco Systems, Inc.

Figures 17-4, 17-5, 17-10, 17-14 to 17-16, © Cisco Systems, Inc.

Figures 18-1 to 18-8 Amazon Web Services, Inc.

Figure 19-3 © Cisco Systems, Inc.

PART I

Basics of Data Center Networking and Storage

Chapter 1 Data Networks: Existing Designs

Chapter 2 Storage Networks: Existing Designs

Data Center Networking and Storage went through many evolutions. However, when discussing the latest and greatest technologies, many assume that enterprises already went through major overhauls and that the network is on par with the latest on the market. The fact is that marketing always superseded engineering by leaps. While it is easy to draw the most efficient architectures in PowerPoint, it takes months of planning before network, compute, and storage professionals can attempt a network change. And when the decision is made, it takes many more months, or maybe years, before a network goes through a major overhaul. Add to this the many justifications to management on why new equipment is needed and why tweaking the current network does not do the job.

The point is that many networks are still built with older technologies and are not to the point of adopting newer technologies. So before getting into hyperconvergence, it helps to go back to the basics of data center designs—from both a networking and storage perspectives—to understand some of the existing challenges and why you need to move to newer architectures.

Chapter 1 discusses the data networks, existing designs, starting by listing the networking components, discussing multitier architectures, and exploring the existing networking challenges. While this might be basic for a networking engineer, virtualization and storage engineers will appreciate some of the networking intricacies.

Chapter 2 covers the storage multitier architecture and storage basics from types of disks, storage networking alternatives, the different storage file systems, and the data services deployed by enterprises. While this chapter might be familiar to storage engineers, some of the basics are not so obvious to virtualization and networking engineers.

This page intentionally left blank

Chapter 1

Data Networks: Existing Designs

This chapter covers the following key topics:

- **Information Technology Equipment of a Data Center:** A highlight of networking equipment that constitutes the data center. This includes network equipment such as switches and routers, network services equipment including traffic optimization with load balancers, wide area network (WAN) optimizers, and security equipment such as firewalls.
- **Multitier Data Networking Architecture:** Describes the three-tier data networking architecture from access to aggregation and core. It discusses the placement of networking services and the logical server grouping for multitier applications.
- **Challenges of Existing Designs:** Discusses challenges with the three-tier architecture, such as oversubscription between the tiers, stretching virtual local area networks (VLANs) over large L2 networks, latency of traffic crossing tiers, flooding of broadcast traffic, complexity in dealing with IPv4 address scarcity, loop prevention, and firewall overload.

The existing networking designs in today's data centers, with all their flaws, served their purpose providing robust infrastructures. Some of the largest data centers were built based on networking architectures that are now called *legacy*. However, the shift in traffic patterns, the introduction of server virtualization, and the introduction of multitier applications challenge the existing designs. This chapter discusses the existing network equipment and services and how they are deployed. It also goes into detail about the challenges of the current designs in scaling to meet the needs of application-aware data centers.

Information Technology Equipment of a Data Center

There are different types of data centers depending on their particular usage. Enterprise data centers normally house customer-facing applications such as web servers as well as the applications and services needed for the day-to-day operation of the enterprise. Such applications include email servers, enterprise resource planning (ERP), customer relationship management (CRM), and relational databases. Other larger data centers operated by cloud service providers (CSPs) house many software applications that are sold as a service to enterprises and consumers that connect to the data center over the Internet or private lines. Although data centers differ in size and functionality, the basic blocks for compute, storage, networking, and software applications remain the same. Although data centers contain many elements, such as facility management, physical security, power, cooling, and so on, this book covers only the information technology (IT) equipment of the data center in the context of data center convergence and hyperconvergence, which is defined later. The basic IT equipment of a data center falls under the categories of network equipment and networking services.

Network Equipment

Network equipment encompasses the basic switching and routing equipment of two layers:

- **Layer 2 (L2):** Switches that work at the Media Access Control (MAC) and VLAN levels
- **Layer 3 (L3):** Switches and routers that work at the Internet Protocol (IP) level

This chapter does not go into the details of switching and routing, but Chapter 15, “Software-Defined Networking and Open Source,” discusses underlay and overlay networks that essentially combine L2 and L3 networking to allow the mobility of applications between different servers and virtual machines.

It is worth noting that the word *networking* normally describes both data networking and storage networking. The difference between data networking and storage networking becomes clearer in the next sections.

In the context of data networking, switches and routers have different types of interfaces—ranging from Fast Ethernet to 1 Gigabit Ethernet (GE), 10 GE, 25 GE, 40 GE, 50 GE, and 100 GE—that connect the local area network (LAN). In the context of storage networking, switches also have 2 Gbps, 4 Gbps, and 8 Gbps fibre channel (FC) interfaces that connect the storage area network (SAN).

Networking Services

Networking services in the data center are standalone appliances, software running on hardware modules inside network equipment, or software running inside servers. Because data centers are moving toward a hyperconverged space, most of these services will eventually move from standalone appliances and become more integrated into the hyperconverged equipment. Networking services are grouped as either traffic redirection and optimization or security.

Traffic Redirection and Optimization

Traffic redirection involves redirecting the traffic to a certain target based on certain criteria, such as port numbers inside the Transmission Control Protocol (TCP) / Internet Protocol (IP) packets, or based on actual traffic content, or other. There are a lot of products in this space that offer different functionality such as load balancing, WAN optimization, content switching and caching, TCP optimization, Secure Sockets Layer (SSL) offload, data compression, and so on. A sample of such products includes load balancers and WAN optimizers.

Load Balancers

Server load balancing (SLB) distributes traffic across multiple servers to get better server utilization and higher availability for applications. Applications are normally spread over multiple servers in the same data center or multiple data centers for higher fault tolerance. The client trying to reach an application points only to one target—such as a URL or an IP address that is directed to the load balancer. Once the traffic reaches the LB, it distributes the traffic to different servers according to the rules or criteria. LBs work at layer 4 (L4) to balance traffic based on information in transport and session layer protocols, such as IP, TCP, File Transfer Protocol (FTP), and User Datagram Protocol (UDP), or they work at layer 7 (L7) based on information in application layer protocols, such as on Hypertext Transfer Protocol (HTTP) headers and cookies. LBs balance traffic based on various algorithms, such as round robin, response time, and sending traffic to healthy servers only.

WAN Optimizers

WAN optimizers efficiently carry traffic over WAN links. Because most data centers are connected through the WAN to service providers or to other data centers in the case of large enterprises or CSPs, it becomes crucial to make the most of the WAN link bandwidth. The WAN optimizers make sure that traffic is prioritized and bandwidth is adequately allocated. They also perform data compression, traffic shaping, and data deduplication.

Security

Securing the data and transactions between clients and a data center is extremely important, whether the access to the data center is done from the Internet toward an enterprise or from within the enterprise itself. The security terminology and functionality in the industry are overwhelming. Security is done via standalone appliances or software that covers these different areas:

- Packet firewalls
- Proxy firewalls
- Stateful inspection firewalls
- Next-generation firewalls that work at the application level

- VPN with encryption and decryption
- Network address translation (NAT)
- Intrusion detection systems (IDSs)
- Intrusion prevention systems (IPSs)
- Access key and token management
- Protection against denial of service (DoS) attacks

The term *unified threat management (UTM)* combines some of the functionality just mentioned into one product. For the purposes of this book, let's refer to the security functions as firewall (FW), which can encompass one or many of the functionalities. Now let's examine a sample description of the main FW functionality.

Firewall

FWs constitute the first entry point from the Internet into the data center, and they allow only authorized users to enter the data center while blocking unauthorized users. They are applied within the enterprise campus to secure traffic between the different departments of the same enterprise. They are also applied at the server level to allow appropriate access between clients, applications, and databases.

There are different types of firewalls. The most basic ones are packet-based firewalls that allow or deny traffic based on source and destination IP address, TCP port numbers, and direction of traffic. These are comparable with the access control lists (ACLs) on routers that basically deny or allow inbound or outbound traffic from an IP address and port. With the sophistication of attacks based on IP addresses and port numbers, more advanced firewalls do stateful inspection and track the progress of the full TCP session, not just the port number. Also, next-generation firewalls (NGFWs) work at the application level and track traffic based on the application itself. For example, an FW can track a Structured Query Language (SQL) session toward a database and allow or deny access based on an IP address and application type being SQL. Other applications of such firewalls are filtering traffic based on HTTP headers and the content itself.

As you move toward a hyperconverged data center where servers, storage, and network equipment are collapsed into one product, the layer where you apply networking services becomes important because it makes or breaks the network. Things can get complicated with virtualized environments where applications move around between different servers and security policies have to follow. As you'll see later, networking services are moving toward being applied at the application level.

Multitier Data Networking Architecture

When networking (switching and routing) vendors talk about three-tier and two-tier designs, they are referring to how switches and routers are deployed in a data center. The traditional switching and routing layers in the legacy data center are as shown in Figure 1-1.

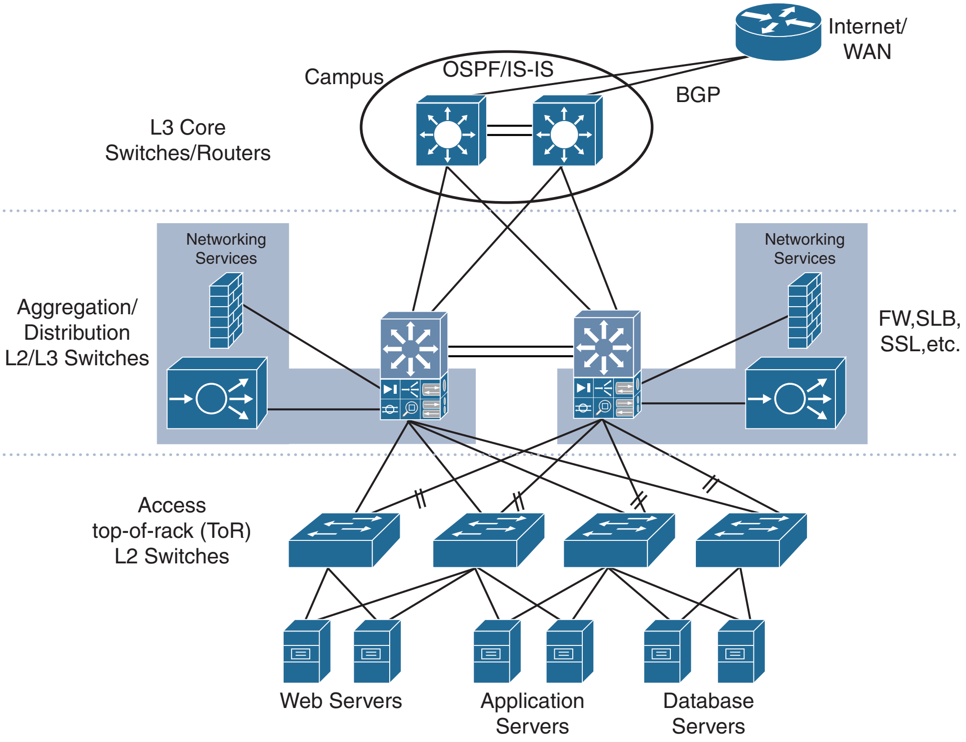


Figure 1-1 *Networking View of Three-Tier Data Center*

Access switches normally connect to the server network interface cards (NICs) and form the first level of data multiplexing. These are also called top-of-rack (ToR) switches because they sit on top of the rack connecting all Ethernet NICs coming from the servers within the same rack. Access switches are usually L2 switches, but they can also be L2/L3. The choice depends on cost and particular designs. L2 switches are a lot cheaper than L2/L3 switches because they deliver less functionality.

Aggregation switches, also called distribution switches, aggregate toward the core all Ethernet interfaces coming from the access switches. The aggregation switches normally work at L2. Traditionally, this layer existed because it reduced the number of "expensive" interfaces at the core router, it created a certain level of redundancy by having dual connections between the access switch and the aggregation layer (keeping in mind that aggregation switches are cheaper than core routers), and it shielded the core layer from L2 functionality. As shown later in this book, the newest designs can skip this layer and go directly from the access (now called leaf), to the core (now called spine). In this case, L3 functionality must start at access switches. The aggregation layer also offers connectivity to the networking services such as firewalls, load balancers, and others, as seen in Figure 1-1. The networking services are either standalone appliances or embedded inside the aggregation switches.

Core switches/routers collect all interfaces coming from the aggregation layer and multiplex them toward the WAN or the campus LAN. Most layer 3 functionality—such as routing between subnets, running routing protocols such as Open Shortest Path First (OSPF), Intermediate System to Intermediate System (IS-IS), and Border Gateway Protocol (BGP)—is done at this layer. As discussed earlier, in the newest leaf and spine designs where there is no aggregation, the core switches connect directly to the access layer and support L2 and L3 functionality. The access switches have to do some L3 functionality as well.

Logical Server Grouping

Servers are connected to access switches and grouped logically into multiple tiers, such as a web tier, an application tier, and a database tier. This distributes functionality over multiple servers, which gives higher availability to the application. Failure in one of the tiers does not affect the others. Also, a more distributed approach spreads each tier over multiple servers; for example, the web server actually runs on multiple servers located in the same data center or multiple data centers. If a web server is down, other servers absorb the load. The same distributed approach applies to application servers and database servers. As discussed earlier in this chapter, load balancers distribute the traffic between different servers based on L4 or L7 and different criteria. Figure 1-1 shows a web server, an application server, and a database server design, with load balancers distributing traffic between different servers and firewalls protecting traffic at the entry point to the data center and between the different application components.

Multitier designs require the existence of multiple servers within the same tier talking to each other and to the other tiers. This communication is normally done at L2 using a logical separation via VLANs. In the simplest form, an IP subnet is associated with a VLAN, and servers within the same subnet share the same VLAN ID and talk to each other via their MAC addresses. The VLAN is considered a broadcast domain and is isolated from other broadcast domains. This means that broadcast packets generated within a certain VLAN are contained within that VLAN. This is important in scaling the network as the number of servers and applications increase. In Figure 1-2, web servers, application servers, and database servers are grouped in different subnets, and each subnet is associated with a VLAN. There could be many web/app/database servers in the same subnet, but for simplicity, only one of each is illustrated here.

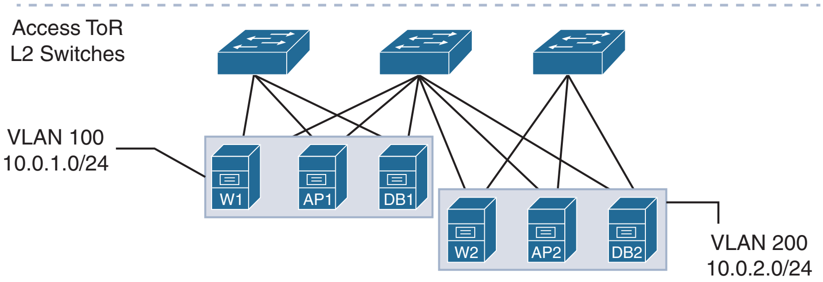


Figure 1-2 *Logical Server Grouping*

Servers house one or many NICs, and the NICs contain one or more Ethernet, 1 GE, 10 GE, or higher speed interfaces. The NICs are dual homed to the access switches for better redundancy and fault tolerance, and NIC ports work in active-active or active-passive modes. Servers within the same VLAN do not need to be connected to the same access switch or in the same rack; they can be connected to different access switches as long as the VLANs are extended across the network. Note that VLANs 100 and 200 stretch across the access and distribution switches.

As a simple example, subnet 10.0.1.0/24 is associated with VLAN 100 and contains web server 1 (W1), application server 1 (AP1), and database server 1 (DB1). Subnet 10.0.2.0/24 is associated with VLAN 200 and contains web servers, application servers, and database servers W2, AP2, and DB2, respectively.

Challenges of Existing Designs

There are many challenges for the three-tier designs in scaling to meet the demands of today’s web applications. Today’s traffic patterns in the data center have changed. Traditionally, the bulk of the traffic was north-south, meaning from the Internet to the data center and from the data center to the Internet. This affects where the firewalls are placed and how VLANs are designed. With the continued growth in data and storage and with the introduction of virtualization and three-tier web/app/database architectures, traffic is shifting toward an east-west pattern that challenges the three-tier design. This presents the following challenges:

- Oversubscription between the tiers
- Large flat L2 networks with stretched VLANs
- Traffic hopping between tiers, inducing latency
- Complexity of mechanisms used for IP subnet scarcity
- Flooding of broadcast, unknown unicast, and multicast (BUM) traffic
- Loop prevention via spanning tree
- Firewall overload

These issues are described next.

Oversubscription Between the Tiers

One of the challenges of the three-tier architecture is due to oversubscription between the tiers. For example, 20 servers can be connected to an access switch via 1 GE interface, while the access switch is connected to the aggregation switch via a 10 GE interface. This constitutes a 2:1 oversubscription between the access and the aggregation layer. Traditionally, this created no problems because most of the traffic is north-south toward the Internet. Therefore, it was assumed that traffic is limited by the Internet WAN link, and oversubscription does not matter because there is ample bandwidth in the LAN.

With the shift to east-west traffic, the bulk of traffic is now between the servers within the data center, so oversubscription between the access layer and the distribution layer becomes a problem. Access layer switches are normally dual homed into two aggregation layer switches. With the addition of servers, more access switches need to be deployed to accommodate the servers. In turn, more aggregation switches must be added to accommodate the access switches. Usually this is done in pods, where you duplicate the setup repeatedly. This is seen in Figure 1-3.

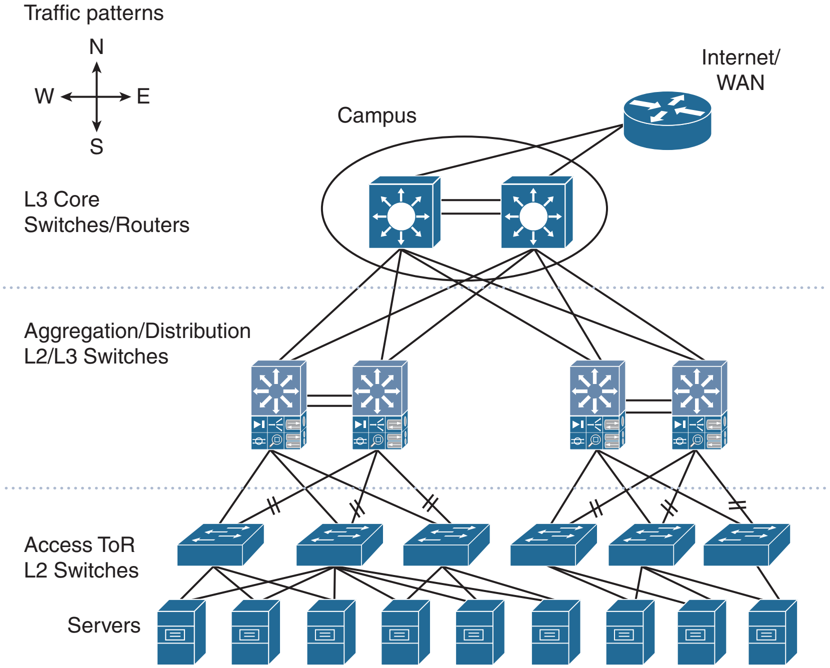


Figure 1-3 *The Challenge of Scaling Three-Tier Designs*

As you see, the addition of aggregation switches shifts the problem of oversubscription from access and aggregation to aggregation and core.

Large Flat L2 Networks with Stretched VLANs

With multitier applications, traffic moves around between web servers, application servers, and databases. Also, with the introduction of server virtualization, virtual machines move around between different servers. When virtual machines move around, they have to maintain their IP addresses to maintain connectivity with their clients, so this movement happens within the same IP subnet. Because the virtual machines could land anywhere in the data center, and because the IP subnet is tied to the VLAN, VLANs must be stretched across the whole data center. Every access and distribution switch must be configured with all VLANs, and every server NIC must see traffic of all VLANs. This increases the L2 flat domain and causes inefficiencies as broadcast packets end up touching every server and virtual machine in the data center. Mechanisms to limit L2 domains and flooding must be implemented to scale today's data centers.

Traffic Hopping Between Tiers, Inducing Latency

Latency is introduced every time traffic crosses a switch or a router before it reaches the destination. The traffic path between nodes in the data center depends on whether the traffic is exchanged within the same VLAN (L2 switched) or exchanged between VLANs (L3 switched/routed).

Traffic exchanged within the same VLAN or subnet is normally switched at L2, whereas traffic exchanged between VLANs must cross an L3 boundary. Notice the following in Figure 1-4:

- Intra-VLAN east-west traffic between W1, AP1, and DB1 in VLAN 100 is L2 switched at the access layer. All are connected to switch 1 (SW1), and switch 2 (SW2) traffic is switched within those switches depending on what ports are blocked or unblocked by spanning tree.
- Intra-VLAN east-west traffic between W2, AP2, and DB2 in VLAN 200 is L2 switched at the access layer. All are connected to switch 2 (SW2), and switch 3 (SW3) traffic is switched within those switches depending on what ports are blocked or unblocked by spanning tree.
- Inter-VLAN east-west traffic between W1, AP1, and DB1 and W2, AP2, and DB2 goes to the aggregation layer to be L3 switched because traffic is crossing VLAN boundaries.

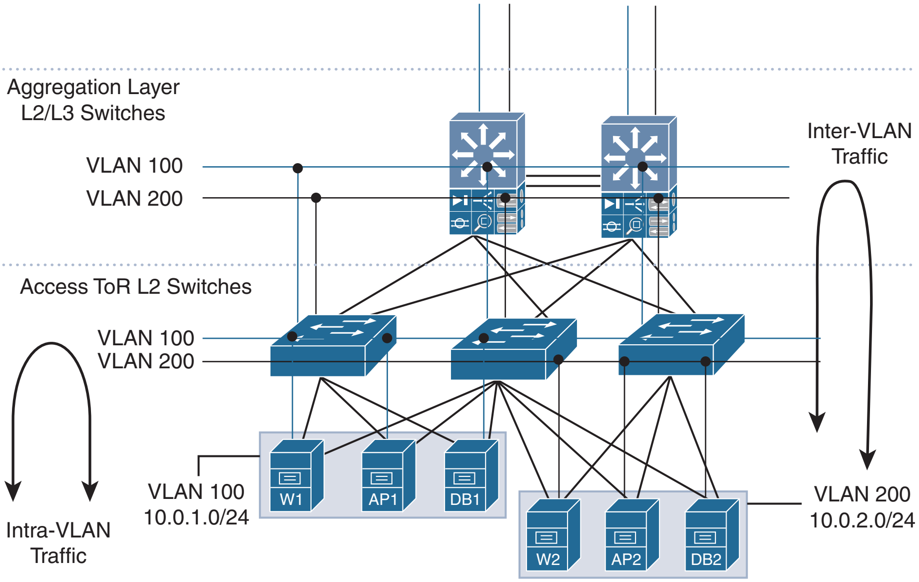


Figure 1-4 *Traffic Hopping Between Tiers*

Every time the traffic crosses a tier, latency is introduced, especially if the network is heavily oversubscribed. That's why it is beneficial to minimize the number of tiers and have traffic switched/routed without crossing many tiers. Legacy aggregation switches used to work at layer 2 and offload layer 3 to the L3 core; however, the latest aggregation switches support L2/L3 functions and route traffic between VLANs via mechanisms such as a switch virtual interface (SVI), which is described next.

Inter-VLAN Routing via SVI

An SVI is a logical interface within an L3 switch that doesn't belong to any physical port. An SVI interface is associated with a specific VLAN. L3 switches have IP L3 switching/routing between VLANs. Think of them as logical routers that live within the switch and that have their connected SVI interfaces associated with the VLAN. This allows inter-VLAN routing on an L3 switch (see Figure 1-5).

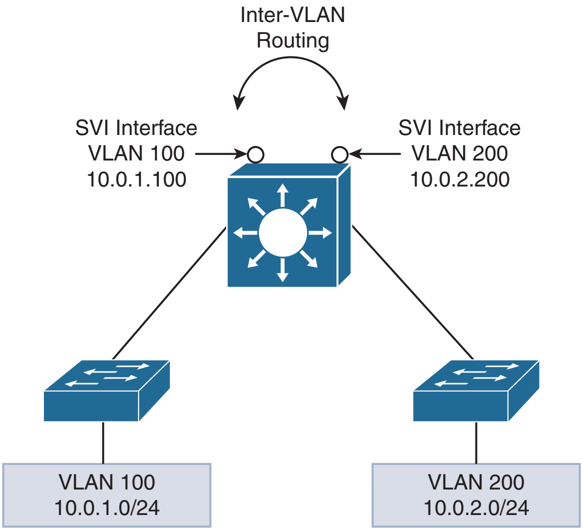

Inter-VLAN Routing
SVI Interface VLAN 100 10.0.1.100
SVI Interface VLAN 200 10.0.2.200
VLAN 100 10.0.1.0/24
VLAN 200 10.0.2.0/24

Figure 1-5 *Inter-VLAN Routing via SVI*

As shown earlier, traffic between VLAN 100, subnet 10.0.1.0/24, and VLAN 200 10.0.2.0/24 was L3 switched at the aggregation layer. To do so, two SVI interfaces must be defined: SVI interface for VLAN 100 with an IP address 10.0.1.100, and SVI interface for VLAN 200 with an IP address 10.0.2.200. When routing is turned on the L3 switch, traffic between the two subnets is routed. Servers within VLAN 100 use SVI 10.0.1.100 as their default gateway, and servers within VLAN 200 use SVI 10.0.2.200 as their default gateway.

Complexity of Mechanisms Used for IPv4 Address Scarcity

Defining many subnets in a data center easily consumes the IP space that is allocated to an enterprise. IPv4 subnetting is a complex topic, and this chapter does not go into detail, but if you are starting to get lost with what 10.0.1.0/24 represents, here is a quick review.

A 10.0.1.0/24 indicates classless interdomain routing (CIDR). An IP address is 32 bits, such as a.b.c.d., where a, b, c, and d are 8 bits each. The /24 indicates that you are splitting the IP address, left to right, into a 24-bit network address and an 8-bit host address. A /24 means a subnet mask of 255.255.255.0. Therefore, 10.0.1.0/24 means that the network is 10.0.1 (24 bits) and the hosts take the last 8 bits. With 8 bits, you can have 2 to the power of 8 (2^8 = 255) hosts, but you lose hosts 0 and 255, which have special meaning for local loopback and broadcast, respectively, and you end up with 253 hosts. 10.0.1.1 inside subnet 10.0.1.0/24 indicates host 1 inside subnet 10.0.1. You can try on your own what 10.0.1.0/27 results in.

IPv4 addresses are becoming scarce. There is a shift toward adopting IPv6 addressing, which provides a much bigger IP address space. However, so far not everyone is courageous enough to dabble into IPv6, and many enterprises still use IPv4 and mechanisms such as NAT. NAT allows the enterprise to use private IP addresses inside the enterprise and map to whatever public IP addresses they are allocated from a provider.

If a provider allocates the enterprise 128.0.1.0/27 (subnet mask 255.255.255.224), for example, the enterprise has practically one subnet 128.0.1.224 with 32 (2^5) hosts, but you lose 0 and 255, so the address range is from 128.0.1.225 to 128.0.1.254. Therefore, the maximum public IP addresses the enterprise has are 30 addresses inside one subnet.

If the enterprise chooses to divide the network into more subnets, it can use /29 to split the /27 range into 4 subnets (29–27 = 2, and 2^2=4) with 8 (32–29 = 3, and 2^3=8) hosts each. The subnets are 128.0.0.224, 128.0.0.232, 128.0.0.240, and 128.0.0.248. Each subnet has 8 hosts, but for every subnet, you lose 0 and 255, so you practically lose 4×2 = 8 IP addresses in the process.

As mentioned, most implementations today use private IP addresses internally and map these IP addresses to public IPs using NAT. The Internet Assigned Numbers Authority (IANA) reserved the following three CIDR blocks for private IP addresses: 10.0.0.0/8, 172.16.0.0/12, and 192.168.0.0/16. With the 128.0.1.0/27 allocation, an enterprise saves the 30 public addresses for accessing their servers publicly and uses as many private IP addresses and subnets internally.

However, early designs used what is called private VLANs to save on the IP space. Although this method is not straightforward and needs a lot of maintenance, it is still out there. This topic is covered briefly because it is revisited later in the book when discussing endpoint groups (EPGs) in Chapter 17, “Application-Centric Infrastructure.”

Private VLANs

This is a mechanism that allows one VLAN, called a primary VLAN (PVLAN), to be split into multiple sub-VLANs, called secondary community VLANs. The IP address space within the PVLAN can now be spread over multiple secondary VLANs, which can be isolated from each other. In a way, this removes the restriction of having one subnet per VLAN, which gives more flexibility in IP address assignment. In the traditional north-south type of traffic, this allows the same subnet to be segmented into sub-VLANs, where each sub-VLAN is its own broadcast domain, and flooding is limited. The secondary community VLANs allow hosts to talk to one another within the community. Any host from one secondary community VLAN that needs to talk to another secondary community VLAN must go through an L3 router for inter-VLAN routing. This is illustrated in Figure 1-6.

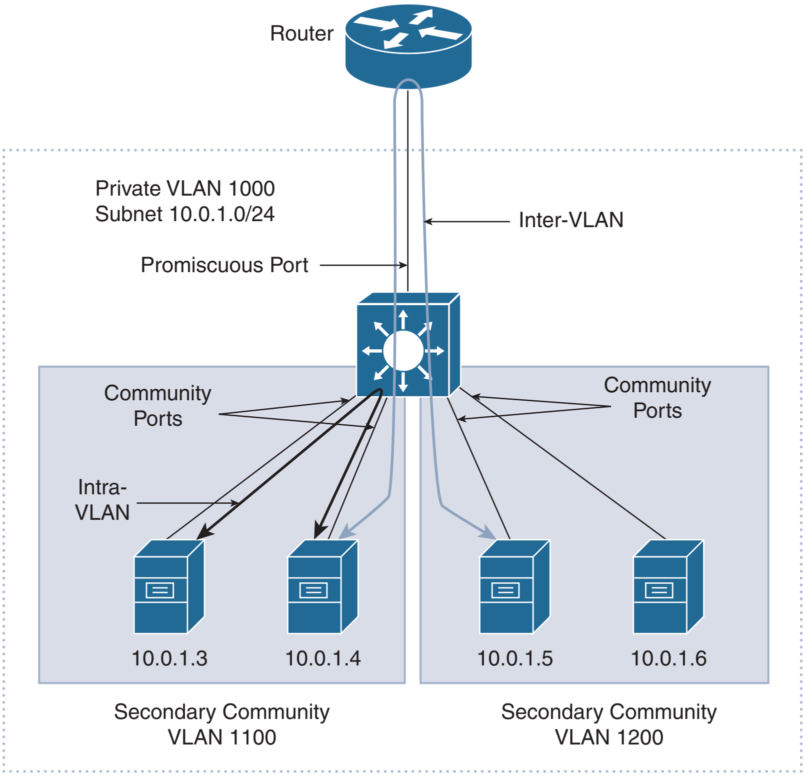


Figure 1-6 *Primary and Secondary VLANs*

The switch ports that connect to the hosts are community *switch ports*, whereas switch ports that connect to the router are called *promiscuous ports*. Ports within the same community talk to each other and to the promiscuous port. Ports between communities must go through the router. Note that hosts 10.0.1.3 and 10.0.1.4 belong to secondary community VLAN 1100, and hosts 10.0.1.5 and 10.0.1.6 belong to secondary community VLAN 1200. All of these hosts belong to the same IP subnet, but they are segmented by the secondary VLANs. So, although 10.0.1.4 talks directly to 10.0.1.3 through intra-VLAN switching, it needs to go through an L3 router to reach 10.0.1.5 or 10.0.1.6 through inter-VLAN routing.

The discussed example is simplistic. As you add multiple access switches and aggregation switches, you need to decide whether L3 routing is done in the aggregation layer or in the core layer. Also, as more and more sub-VLANs are created, such VLANs need to be carried across the network and maintained on every switch. Although these methods give flexibility in decoupling subnets from VLANs, they also add extra overhead. When Cisco's implementation of EPGs is discussed, the decoupling of IP subnets from VLANs will become much clearer, and the whole deployment model will become automated as complexities are hidden from the end user.

Flooding of Broadcast, Unknown Unicast, and Multicast (BUM) Traffic

A main issue of L2 networks is flooding of BUM traffic. Each station needs to have an IP address–to–MAC address mapping to send traffic to a certain IP address on a LAN. If a source station is trying to reach another destination station and it does not know its IP-to-MAC address mapping, an address resolution protocol (ARP) packet is sent to a broadcast address. If a VLAN is configured, then the broadcast address is flooded by a switch to all switch ports that belong to that VLAN. Whenever a device sees the ARP request for its own IP address, it responds to the source station with its MAC address. The source station that sent the original ARP request stores the IP-to-MAC address mapping in its ARP table, and from then on it uses the MAC address it learned to send traffic to the destination station. This is seen in Figure 1-7.

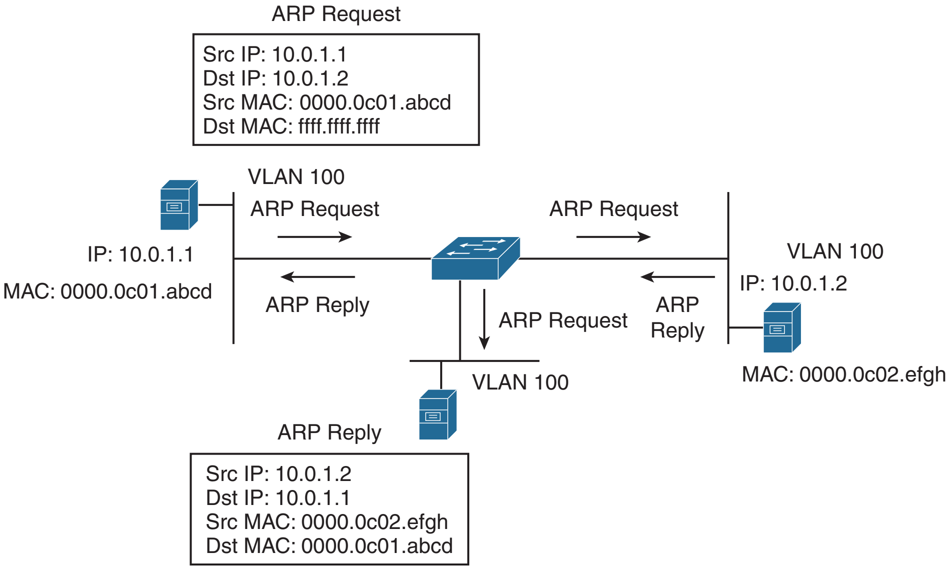


Figure 1-7 *ARP Flooding*

As seen in Figure 1-7, server 1 (S1) with IP 10.0.1.1 must send traffic to server 2 (S2) with IP 10.0.1.2, which is in the same subnet. The switch is configured to have subnet 10.0.1.0/24 mapped to VLAN 100. If S1 does not know the MAC address of S2, it sends an ARP request with a broadcast MAC address of ffff.ffff.ffff. If the switch does not have the IP/MAC address mapping in its ARP table, it floods the ARP to all the ports that belong to VLAN 100. Once S2 sees the ARP, it responds with its MAC address 0000.0c02.efgh directly to S1 (0000.0c01.abcd). After that, S1, S2, and the switch update their own ARP tables.

ARP broadcast is one type of packet that is flooded. Other types could be unknown unicast or multicast. Say, for example, that S1 knows the MAC address of S2 and sends a packet to 0000.0c02.efgh; however, the switch flushed his ARP table and does not know the mapping. In that case, the switch floods that packet to all of its ports on VLAN 100 until a station replies; after that, the switch updates its ARP.

The problem of flooding consumes bandwidth, and many measures are taken to limit it. A side effect of flooding is broadcast storms that are created by loops, as discussed next.

Loop Prevention Via Spanning Tree

The problem with L2 networks is in the potential of broadcast storms occurring in case of loops. Say that an ARP packet is flooded over all switch ports, and that packet finds its way back to the switch that flooded it because of a loop in the network. That broadcast packet circulates in the network forever. Spanning trees ensure that loops do not occur by blocking ports that contribute to the loop. This means that although some ports are active and passing traffic, others are blocked and not used. This is seen in Figure 1-8. Note that an ARP packet that is flooded by SW1 and then flooded by SW4 could return to SW1 and create a loop. To avoid this situation, spanning tree blocks the redundant paths to prevent loops from occurring and hence prevent potential broadcast storms.

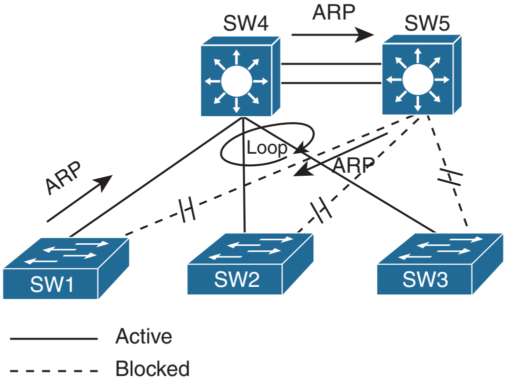


Figure 1-8 *Loop Prevention Via Spanning Tree*

The drawback is that expensive resources such as high-speed interfaces remain idle and unused. More efficient designs use every link and resource in the network.

Firewall Overload

Another issue with the existing three-tier design is that the firewalls that are traditionally connected to the aggregation layer become a catch for all traffic. These firewalls were originally meant to enforce policies on north-south traffic between the Internet and the data center. Because the traffic in the data center dramatically increased as east-west between the multiple application tiers, securing the data center from the inside is now essential. As such, policy enforcement for east-west traffic now must go through the same firewalls. This is seen in Figure 1-9. With practically all traffic going to the firewall, the firewall rules are enormous. Normally, administrators define such policies for a specific service, and when the service disappears, the policies remain in the firewall. For anyone who has worked with ACLs and setting firewall rules, it is well known that nobody dares to touch or delete a rule from the firewall for fear of breaking traffic or affecting an application. Moving into hyperconverged infrastructures and a two-tier design, you see how policy enforcement for applications shifts from being VLAN centric to application centric. Cisco's ACI and VMware's networking and security software product (NSX), covered later in this book, describe how policies are enforced in a hyperconverged environment.

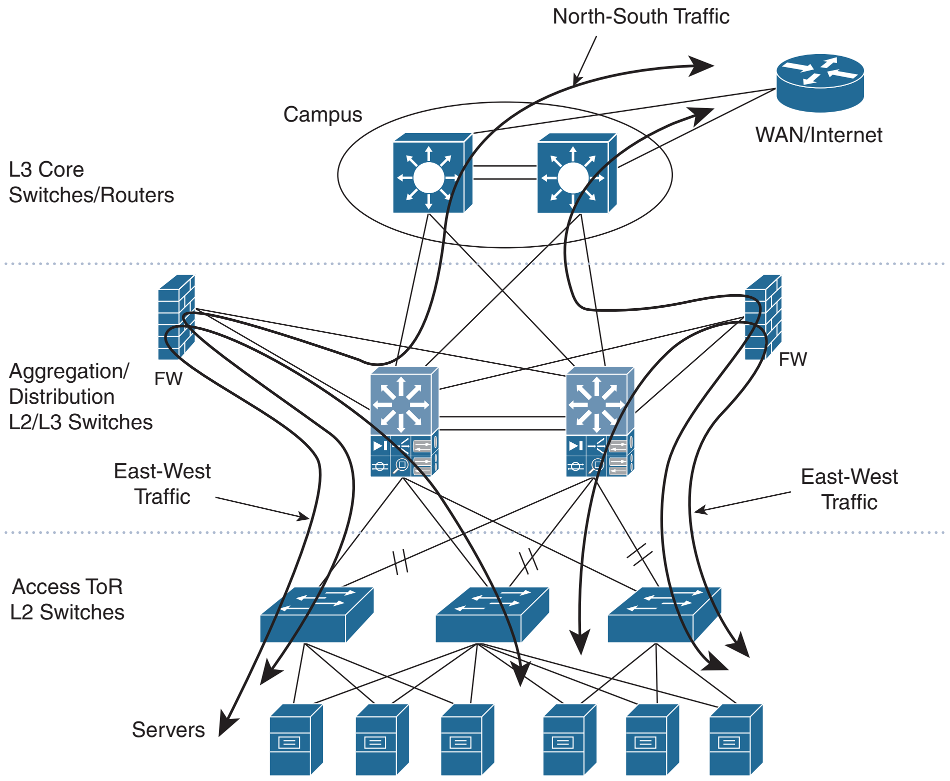


Figure 1-9 *Firewall Overload*

With so many issues in legacy data centers, newer data center designs are making modifications, including these:

- Moving away from the three-tier architecture to a two-tier leaf and spine architecture
- Moving L3 into the access layer and adopting L2 tunneling over L3 networks
- Implementing link aggregation technologies such as virtual port channel (vPC) and multichassis link aggregation (MLAG) to offer greater efficiency by making all links in the network active; traffic is load-balanced between different switches
- Moving firewall policies for east-west traffic from the aggregation layer to the access layer or directly attaching them to the application

The preceding enhancements and more are discussed in detail in Part VI of this book, "Hyperconverged Networking."

Looking Ahead

If you look at the combined picture of data networking and storage networks, it is obvious that something must change. Storage traffic constitutes different protocols, such as Internet Small Computer System Interface (iSCSI), Network File System (NFS), Server Message Block (SMB), and others. Merging such protocols with data traffic over an oversubscribed infrastructure can have huge implications on the stability of storage access and the performance of applications. Also, networking in general needs to move from complex manual configuration to a more automated and dynamic way of network deployment. The advances in switching and routing hardware and software allow you to dramatically simplify the network and minimize the tiers. Advances in server virtualization have changed the way servers and compute are deployed in the data center, which challenges the existing storage networking designs. Hyperconvergence collapses the compute, networking, and storage tiers, offering a pay-as-you-grow approach to storage.

However, the basic needs of the enterprises do not change. Simplifying network deployments should not come at the expense of having highly available and robust networks. To understand the benefits of hyperconvergence, it helps to understand the current storage deployments and their challenges. You cannot solve a problem if you do not understand what the problem is.

Chapter 2, "Storage Networks: Existing Designs," discusses some storage basics that might be familiar to a storage administrator but challenging for a virtualization or a networking engineer. It covers the storage multitier architecture; block, file, and object file systems; storage connectivity; storage protocols and protection methods; and advanced storage functionality. Part II of this book, "Evolution in Host Hardware and Software," discusses the evolution in compute, storage, and networking that allows you to build next-generation hyperconverged data centers.

Index

A

- **ABS (Acropolis Block Services), 272–273**
- **acceleration, hardware, 125**
- **access time, 24**
- **Accolade Technology Inc., 311**
- **ACI (application-centric infrastructure), 252, 312–313**
  - ACI Anywhere, 386
  - APIC (Application Policy Infrastructure Controller), 360–361
  - domains
    - *external, 364–365*
    - *fabric connectivity, 367–369*
    - *physical, 364–365*
    - *switching and routing, 365–367, 369–371*
    - *VMM (Virtual Machine Manager), 362–364*
  - microsegmentation constructs
    - *ANPs (application network profiles), 355–358*
    - *EPGs (endpoint groups), 353–355*
    - *service graphs, 358–359*
    - *tetration model, 360*
  - multicast, 383
  - Multi-PoD (point of delivery), 383–384
  - Multi-Site, 384–386
  - NSX (network virtualization and security) compared to
    - *automation and visibility, 390–391*
    - *control plane, 389*
    - *data forwarding, 390*
    - *networking learning curve, 391*
    - *overview of, 386–387*
    - *policy setting and enforcement, 387–388*
    - *VXLAN performance, 388–389*
  - overlay network
    - *L2 switching in, 378–380*
    - *L3 switching/routing, 380–383*
    - *overview of, 373–376*
    - *VNIDs (VXLAN Instance IDs), 376–378*

overview of, 352–353
underlay network, 371–373
Acropolis Block Services, 272–273
Acropolis File Services, 273–274
Acropolis Hypervisor (AHV), 256
Acropolis OS (AOS), 274
ACS (atomic compare and swap), 263
active-active data centers, 148
address resolution protocol (ARP), 15, 315
addressing
FC (fibre channel), 42
IP (Internet Protocol), 12–13
iSCSI (Internet Small Computer Interface), 47–48
MAC (media access control), 108
WWNN (World Wide Node Name), 108
advanced catalogs, 431
Advanced Host Controller Interface (AHCI), 71–72
Advanced Technology Attachment over Ethernet (AoE), 40, 50
agents, Neutron, 321–322
aggregation switches, 7
AHCI (Advanced Host Controller Interface), 71–72
AHV (Acropolis Hypervisor), 256
Alarms menu (Cisco Intersight), 209
all-flash arrays, 73–75
all-flash nodes, 156
allocation, memory, 85
alternative HCI implementations. *See* Nutanix ECP (Enterprise Cloud Platform); OpenStack; vSAN
ALU (arithmetic logic unit), 66–67
Amazon Machine Images (AMI), 410
Amazon Web Services. *See* AWS (Amazon Web Services)
AMI (Amazon Machine Images), 410
ANPs (application network profiles), 355–358
Aodh, 293
AoE (Advanced Technology Attachment over Ethernet), 40, 50
AOS (Acropolis OS), 274
APIC (Application Policy Infrastructure Controller), 93, 360–361
APIs (application programming interfaces)
Cisco CloudCenter, 446
Cisco UCSD (USC Director) integration, 434–437
Neutron, 321
RESTful, 36, 215, 295, 297, 434–436
VMware vSAN, 244
AppDynamics, 217–219, 446
Appfog, 397
Application Mobility Fabric, 256–257
application network profiles (ANPs), 355–358
Application Policy Infrastructure Controller (APIC), 93, 360–361
application programming interfaces. *See* APIs (application programming interfaces)
Application Virtual Switch (AVS), 93
application-centric infrastructure (ACI), 252, 312–313
application-specific integrated circuits (ASICs), 311
arithmetic logic unit (ALU), 66–67
ARP (Address Resolution Protocol), 15, 315
arrays
caching storage, 60–61
flash, 73–75
storage, 102

ASICs (application-specific integrated circuits), 311

asynchronous native replication for DR (NRDR), 189–190

atomic compare and swap (ACS), 263

automation

ACI (application-centric infrastructure), 390–391

AWS (Amazon Web Services), 416

infrastructure as a code, 417–420

SDKs (software development kits), 420–421

networking, 126–127

NSX (network virtualization and security platform), 390–391

private cloud, 425–427, 431–432

availability. ***See also*** **HA (high availability)**

AWS (Amazon Web Services), 399–401

EBS (Elastic Block Storage), 409

Nutanix ECP (Enterprise Cloud Platform), 263–264

Avamar, 241–243

AVS (Application Virtual Switch), 93

AWS (Amazon Web Services)

AMI (Amazon Machine Images), 410

availability zones, 399–401

cloud automation, 416

infrastructure as a code, 417–420

SDKs (software development kits), 420–421

cloud monitoring, 416

CloudFormation, 417–420

compute instances, 409

EC2 instances

auto scaling, 414–416

launching, 411–412

volumes, adding, 414

Glacier, 441–442

global infrastructure, 399–401

IAM (identity and access management), 411–412

networking, 401–404

overview of, 398–399

regions, 399–401

S3 (Simple Storage Service), 36

security groups, 410–411

storage

EBS (Elastic Block Storage), 406–409

Glacier, 405

overview of, 404–405

RRS (Reduced Redundancy Storage), 405

S3 (Simple Storage Service), 405–406

Azure, 397

B

B200 blade server, 159

backup, 129–130

Nutanix ECP (Enterprise Cloud Platform), 270–271

third-party backup tools, 191–192

VMware vSAN, 241–243

balancing. ***See*** **load balancing**

Barbican, 292

bare-metal servers, 78–79, 396

BDs (bridge domains), 366

Berkeley Software Distribution (BSD), 312

BGP (Border Gateway Protocol), 8, 308

block-level deduplication, 56–57

block-level storage, 35

blocks (Nutanix), 263–264

boot policy, 165

Border Gateway Protocol (BGP), 8, 308

bridge domains (BDs), 366

Broadcom Ltd, 311

BSD (Berkeley Software Distribution), 312

buckets (S3), 405–406

BUM (broadcast, unknown unicast, and multicast) traffic, 15–16

business benefits

deployment speed, 136

HA (high availability), 139

infrastructure scaling, 136–137

IT operational model, 137–138

low-entry cost structure, 139

overview of, 135–136

private cloud agility/flexibility, 138–139

system management, 138

TCO (total cost of ownership), 140

BW (Business Warehouse), 147

C

cache vNode, 182

caching

caching storage arrays, 60–61

host-based, 75

VMware vSAN, 232

Calm, 278–279

CAM (content-addressable memory), 308

Capacity Optimization Engine (COE), 265

CAPEX (capital expenditures), 127, 140, 212, 216, 399

Cassandra, 258, 260

catalogs (UCSD), 431, 432

Cavium Inc., 311

CBT (changed block tracking), 192, 275

CCM (CloudCenter Manager), 448

CCO (CloudCenter orchestrator), 448

CEE (Converged Enhanced Ethernet), 49

Ceilometer, 293

CentOS KVM hypervisor, 256

central processing units. *See* CPUs (central processing units)

centralized firewalls, 319

Ceph, 300–303

Cerebro, 259, 270

certifications, security, 194

CGs (consistency groups), 270

changed block tracking (CBT), 192, 275

Chef OpenStack, 292

CI (converged infrastructure), 108–116

Cinder Block Storage, 296–297

CIOs (chief information officers), 399, 442

Cisco 5108 UCS blade chassis, 159–160

Cisco ACI (application-centric infrastructure), 312–313

Cisco AppDynamics, 217–219

Cisco application-centric infrastructure. *See* ACI (application-centric infrastructure)

Cisco AVS (Application Virtual Switch), 93

Cisco C220/C240 M4/M5 rack servers, 158–159

Cisco CloudCenter, 446–449

Cisco FlashStack, 112

Cisco FlexPod, 112

Cisco HyperFlex. *See* HyperFlex (HX)

Cisco HyperFlex Connect HTML5 Management, 200–202

Cisco HyperFlex Data Platform. *See* HXDP (HX Data Platform)

Cisco Intersight. *See* Intersight

Cisco Nexus 1000v, 93, 321

Cisco TAC, Cisco Intersight integration with, 205

Cisco Tetration. *See* Tetration

Cisco UCS (Unified Computing System). *See* UCS (Unified Computing System)

Cisco Unified Computing System (UCS), 223

Cisco Validated Design (CVD), 331

Cisco VesraStack, 112

Cisco VIC MLOM interface cards, 159

Cisco VxBlock, 112

Cisco WebEx, 443

Cisco Workload Optimizer (CWOM), 216–217, 449

Cisco/Unified Computing System (UCS) Nexus, 321

classless interdomain routing (CIDR). *See* SVI (switch virtual interface)

CLI (command-line interface), 244, 308

CliQr. *See* CloudCenter

cloning, 55

- HyperFlex (HX), 189
- Nutanix ECP (Enterprise Cloud Platform), 268–269
- VMware vSAN, 236–237

cloud bursting, 441

cloud creep, 212

cloud service providers (CSPs), 4

cloud technology. *See* multicloud; private cloud; public cloud

CloudCenter, 446–449

CloudCenter Manager (CCM), 448

CloudCenter orchestrator (CCO), 448

CloudFormation, 417–420

Cloudkitty, 293

Cloud-Scale Acceleration Architecture, 311

CloudWatch, 416

clusters

- definition of, 86–87
- Nutanix ECP (Enterprise Cloud Platform), 258–259, 271–272
- resiliency, 179–180
- VMware vSAN, 238–241

CMVP (Cryptographic Module Validation Program), 235

CNA (converged network adapter), 49

COE (Capacity Optimization Engine), 265

command-line interface (CLI), 244, 308

community switch ports, 14

compression, 58–59, 128

- HyperFlex (HX), 187
- Nutanix ECP (Enterprise Cloud Platform), 265–266
- VMware vSAN, 235

compute-only nodes, 154–155, 157, 250–251

Config-VVol-Metadata, 101

Congress, 292

consistency groups (CGs), 270

containers
Docker, 80–82
HyperFlex support for, 172–173
Nutanix support for, 275
Open Virtuozzo, 80–81
persistent storage, 244
Swift, 299
content-addressable memory (CAM), 308
Contiv project, 353
control plane, 308, 389
control VM (CVM), 119, 257, 264
Converged Enhanced Ethernet (CEE), 49
converged infrastructure. *See* CI (converged infrastructure)
converged network adapter (CNA), 49
converged nodes. *See* nodes
converged systems, 112–116
COOP, 371
copy on write (COW) logical snapshots, 52–54
core switches/routers, 7
CoreOS rkt, 80–81
COW (copy on write) logical snapshots, 52–54
CPUs (central processing units)
dual-core, 66–67
logical cores, 67–68
Moore's Law, 65–66
multi-core, 66–67
pCPUs (physical CPUs), 67
physical cores, 67–68
single-core, 66–67
vCPUs (virtual CPUs), 68–70
virtual cores, 67–68
x86 standard architecture, 66
CRC-32 (32-bit cyclic redundancy check), 234
CRE (generic routing encapsulation), 214
CRM (customer relationship management), 4, 146
CRUSH algorithm, 303
Cryptographic Module Validation Program (CMVP), 235
CSPs (cloud service providers), 4
Curator, 259, 261
CVD (Cisco Validated Design), 331
CVM (control VM), 119, 257, 264
CWOM (Cisco Workload Optimizer), 216–217, 449

D

daemons, Ovsdb-vswitchd, 325–326
DAS (direct-attached storage), 38–39, 118–127
Dashboard view
Cisco Intersight, 207
HyperFlex Connect HTML5 Management, 200–201
Data Center Bridging (DCB), 49
data compression. *See* compression
data distribution
cluster resiliency, 179–180
data protection via replication, 178–179
data striping, 176–177
overview of, 174–176
data encryption. *See* encryption
data encryption keys (DEKs), 192–193
data forwarding, 390
data locality, 175, 285–286
data networks, 4

design challenges

BUM traffic flooding, 15–16

firewall overload, 17–18

flat L2 networks with stretched VLANs, 10–11

interVLAN routing via SVI, 12

IP4 address scarcity, 12–13

loop prevention via spanning tree, 16

oversubscription between tiers, 9–10

private VLANs, 13–15

traffic hopping between tiers, 11–12

logical server grouping, 8–9

multitier architecture, 6–8

network equipment, 4

security, 5–6

traffic redirection and optimization, 5

types of, 4

data optimization, 185–186

data plane, 308

Data Plane Development Kit (DPDK), 312

data platform. ***See*** **HXDP (HX Data Platform)**

data platform controller (DPC), 119, 169–170

Data Protection Manager (DPM), 275

data storage

backup, 129–130

deduplication and compression, 128

EC (erasure coding), 128–129

overview of, 127–128

replication, 129–130

data striping, 176–177

data tiering, 282

data virtualization, 182–183

data vNode, 182, 185

database transaction units (DTUs), 216

databases, 4, 146

datastores, 82–83, 84, 228

DATA-VVol-VMDK, 101

DCB (Data Center Bridging), 49

DDA (Discrete Device Assignment), 274

deduplication, 128

block-level, 56–57

definition of, 55–56

file-level, 56

global, 58

HyperFlex (HX), 187

inline, 58

local, 58

Nutanix ECP (Enterprise Cloud Platform), 264–265

Nutanix ECP versus HyperFlex, 283–284

object-level, 56

post-process, 58

source-based, 57–58

target-based, 57–58

VMware vSAN, 235

Defense Information Systems Agency (DISA), 235

DEKs (data encryption keys), 192–193

DELETE method, 435

Dell EMC, 246, 255

delta disks, 237

denial of service (DoS) attacks, 6

deployment

HyperFlex (HX)

Cisco HyperFlex Connect HTML5 Management, 200–202

Cisco Intersight, 204
installation phase, 197–199
sizer, 199
VMware vSphere management plug-in, 203–204
workload profiler, 199
iSCSI (Internet Small Computer Interface) SANs, 46–47
speed of, 136
Designate, 292
desktops, VDI (virtual desktop infrastructure), 141–144
destaging, 185–186
Device menu (Cisco Intersight), 209
DevOps, 141
DHCP (Dynamic Host Configuration Protocol), 321
dies, 66–67
Digital Ocean, 397
direct-attached storage. *See* DAS (direct-attached storage)
DISA (Defense Information Systems Agency), 235
Disable Object Checksum, 231
disaster recovery, 148–149
backup, 129–130
HCI use case, 148–149
Nutanix ECP (Enterprise Cloud Platform), 270–271
replication, 129–130
VMware vSAN, 238–243
Discrete Device Assignment (DDA), 274
distributed architecture
DAS (direct-attached storage), 118–127
HyperFlex (HX), 281–282
Nutanix ECP (Enterprise Cloud Platform), 281–282
distributed controllers, 119
distributed file systems, 249–250
distributed logical router (DLR), 343–345
distributed objects, 183
distributed resource scheduling (DRS), 82
distributed storage fabric (DSF), 257–259
distributed virtual port groups (dvPG), 92
distributed virtual ports (dvPorts), 92
distributed virtual uplinks (dvUplink), 92
distribution switches, 7
DLR (distributed logical router), 343–345
Docker containers, 80–82
HyperFlex support for, 172–173
Nutanix support for, 275
domains
external, 364–365
fault, 229–230
identifiers, 44
physical, 364–365
VMM (Virtual Machine Manager), 362–364
DoS (denial of service) attacks, 6
DPC (data platform controller), 119, 169–170
DPDK (Data Plane Development Kit), 312
DPM (Data Protection Manager), 275
DR. *See* disaster recovery
Dragonflow, 292
DRAM (dynamic random-access memory), 147, 232
drives
delta disks, 237

HDDs (hard disk drives), 22–23, 155, 406–408
Nutanix ECP (Enterprise Cloud Platform), 260–261
performance
access time, 24
IOPS (input/output operations per second), 24–25
latency, 24–25
throughput, 21–24
RAID (Redundant Array of Independent Disks), 26–30
SEDs (self-encrypting drives), 192–193
SSDs (solid-state drives), 23, 155
DRS (distributed resource scheduling), 82
DSF (distributed storage fabric), 257–259
DTUs (database transaction units), 216
dual-core CPUs (central processing units), 66–67
dvPG (distributed virtual port groups), 92
dvPorts (distributed virtual ports), 92
dvUplink (distributed virtual uplinks), 92
Dynamic Host Configuration Protocol (DHCP), 321
dynamic random-access memory (DRAM), 147, 232

E

EBS (Elastic Block Storage), 406–409, 425–426
EC. *See* erasure coding
EC2 instances, 411–416
ECMP (Equal Cost Multi-path), 315
ECP (Enterprise Cloud Platform). ***See*** **Nutanix ECP (Enterprise Cloud Platform)**
edge computing, 146
edge nodes, 157
Elastic Block Storage (EBS), 406–409, 425–426
email servers, 4
EMC
Avamar, 129–130, 241–243
Vipr controller, 99
VxRail, 223
Encapsulated Remote Switched Port Analyzer (ERSPAN) agents, 214
encryption, 59
disk, 59
EBS (Elastic Block Storage), 408
Nutanix ECP (Enterprise Cloud Platform), 272
SEDs (self-encrypting drives), 23, 59, 129–130, 192–193, 199, 235
VMware vSAN, 234–235
endpoint groups. ***See*** **EPGs (endpoint groups)**
Engine Yard, 397
enterprise class applications, 146–147
Enterprise Cloud Platform. ***See*** **ECP (Enterprise Cloud Platform)**
enterprise resource management (ERM), 146
enterprise resource planning (ERP), 4
enterprise workload issues, 211–212
EPGs (endpoint groups), 353–355, 366
Equal Cost Multi-path (ECMP), 315
equipment, network, 4

Era database services, 269–270

erasure coding, 128–129

Nutanix ECP (Enterprise Cloud Platform), 266

VMware vSAN, 236

ERM (enterprise resource management), 146

ERP (enterprise resource planning), 4

ERSPAN (Encapsulated Remote Switched Port Analyzer) agents, 214

ESXi, 79, 136, 170–171

Ethernet, Fibre Channel over, 49–50, 82, 109

EUI (Extended Unique Identifier), 47

Extended Unique Identifier (EUI), 47

Extensible Markup Language (XML), 215, 434

Extent store (Nutanix), 261

F

fabric, 369

ACI (application-centric infrastructure)

connectivity, 367–369

switching and routing, 365–367, 369–371

definition of, 44

DSF (distributed storage fabric), 257–259

FIs (fabric interconnects), 108

Nutanix Application Mobility Fabric, 256–257

NVMe-oF (NVMe over fabrics), 72

switching fabric, 313–315

Fabric Interconnect menu (Cisco Intersight), 207

Fabric Login process (FLOGI), 44

Fabric loopback TEP (FTEP), 370

failure tolerance (vSAN), 231

Fast Ethernet, 4

fault domains, 229–230

fault tolerance, 89

FC (fibre channel)

addressing, 42

drives, 23

FCNS (fibre channel network service), 44

FCoE (Fibre Channel over Ethernet), 49–50, 82, 109

interfaces, 4, 82, 109

LUN masking, 43

multipathing, 44–46

overview of, 41–42

switching, 44–45

zoning, 43

Federal Information Processing Standards (FIPS), 235

field programmable gate arrays (FPGAs), 311

file server VMs (FSVMs), 273–274

file systems

CephFS (Ceph File System), 301

LFS (log-structured file system), 122–124, 181–182

NFS (Network File System), 36, 82, 244

Nutanix ECP (Enterprise Cloud Platform), 282

object file system, 226–227

overview of, 122–124

union file system, 173

VMFS (Virtual Machine File System), 82, 226

vSAN, 228, 249–250

File Transfer Protocol. *See* **FTP (File Transfer Protocol)**

File Transfer Protocol (FTP), 5
file-level deduplication, 56
file-level storage, 35–36
FIPS (Federal Information Processing Standards), 235
firewalls, 6
  centralized, 319
  NGFWs (next generation firewalls), 314
  overload, 17–18
firmware policy, 165
FIs (fabric interconnects), 108
flash arrays, 73–75
Flash Read Cache Reservation (vSAN), 231
flash wear, 73
flash-based products, 72–75
FlashStack, 112
flat L2 networks, 10–11
FlexPod, 112
FLOGI (Fabric Login process), 44
flooding, 15–16, 342
Flow, 331
force provisioning, 231
forks, 183
FPGAs (field programmable gate arrays), 311
Freezer, 292
FSVMs (file server VMs), 273–274
FTEP (Fabric loopback TEP), 370
FTP (File Transfer Protocol), 5
fully converged systems, 112–116

G

GCP (Google Cloud Platform), 397, 446
GE (Gigabit Ethernet), 4, 20
generic routing encapsulation (GRE), 214
GENEVE (Generic Network Virtualization Encapsulation), 126, 315–316
GET method, 435
GFS (global file system), 40
GIPo (Group IP outer), 383
Glacier, 405, 441–442
Glance, 291, 293
global deduplication, 58
Google App Platform, 397
Google Apps, 443
Google Cloud Platform (GCP), 397, 446
GPUs (graphics processing units), 155
granular volumes, 100–101
Group IP outer (GIPo), 383
groups
  endpoint, 353–355, 366
  security, 410–411
  VMware NSX, 336–337
GST (global station table), 371
guest OS assignment, 85
GUIs (graphical user interfaces), 138, 244, 308

H

HA (high availability)
  datastores, 82
  HCI business benefits to, 139
  Nutanix ECP (Enterprise Cloud Platform), 264
  VMware vSAN, 239–240
hard disk drives. *See* HDDs (hard disk drives)
hardware. *See* host hardware

Hardware Compatibility List (HCL), 245

hashing

CRUSH algorithm, 303

MD5 algorithm, 57, 299

SHA-1 algorithm, 57

HBAs (host bus adapters)

Cisco UCS (Unified Computing System), 108

overview of, 30–31

vHBAs (virtual host bus adapters), 109, 166

HCL (Hardware Compatibility List), 245

HDDs (hard disk drives), 22–23, 155, 406–408

Heat, 293

Help menu (Cisco Intersight), 209

Heroku, 397

high availability. ***See*** **HA (high availability)**

Horizon, 292, 293

host bus adapters. ***See*** **HBAs (host bus adapters)**

host bus interconnect, 70–72

host hardware, 65

CPUs (central processing units)

dual-core, 66–67

logical cores, 67–68

Moore's Law, 65–66

multi-core, 66–67

pCPUs (physical CPUs), 67

physical cores, 67–68

single-core, 66–67

vCPUs (virtual CPUs), 68–70

virtual cores, 67–68

x86 standard architecture, 66

flash-based products, 72–75

hardware acceleration, 125

host bus interconnect, 70–72

host-based caching, 75

host-based networking, 310–312

HP Enterprise Converged Infrastructure, 397

HT (hyperthreading), 67

HTTP (Hypertext Transfer Protocol), 5, 338, 434–435

HX. ***See*** **HyperFlex (HX)**

HXDP (HX Data Platform)

data distribution

cluster resiliency, 179–180

data protection via replication, 178–179

data striping, 176–177

overview of, 174–176

Docker containers support, 172–173

DPC (data platform controller), 169–170

Hyper-V environment, 171–172

installation, 197–199

management platforms

Cisco HyperFlex Connect HTML5 Management, 200–202

Cisco Intersight, 204

VMware vSphere management plug-in, 203–204

overview of, 168

read/write operations, 181–187

data virtualization, 182–183

destaging and data optimization, 185–186

LFS (log-structured file system), 181–182

read operation lifecycle, 186–187

write operation lifecycle, 183–185

sizer, 199

VMware ESXi environment, 170–171

workload profiler, 199

hybrid arrays, 74

hybrid cloud, 439–442

hybrid nodes, 156

hyperconverged networking

SDN (software-defined networking)

host-based networking, 310–312

microsegmentation, 309–310, 319

open source initiatives, 320–324

overlay network, 309–310, 315–319

overview of, 308–309

OVN (Open Virtual Network), 326–331

switch-based networking, 312–313

switching fabric, 313–315

underlay network, 315

vendors, 331

VMware NSX (network virtualization and security platform)

flooding avoidance, 342

IP routing, 343–345

L2 switching, 343

multidestination traffic, 347–349

NSX controller cluster, 340–341

NSX manager, 339–340

overview of, 335–336

policies, 336–338

vDS (virtual Distributed Switch) enhancements, 341–342

hyperconvergence software

Nutanix ECP (Enterprise Cloud Platform)

availability domains, 263–264

cluster components, 258–259

data path resiliency, 264

data protection, 262

DSF (distributed storage fabric), 257–259

I/O path, 261

metadata, 263

physical drives, 260–261

read I/O, 262

write I/O, 261–262

VMware vSAN

caching, 232

datastore, 228

high-level software architecture, 225–226

I/O operation details, 232–233

object file system, 226–227

storage policies, 228–231

HyperFlex (HX), 138, 164–165. ***See also*** **HXDP (HX Data Platform)**

Cisco UCSD (USC Director) integration

APIs (application programming interfaces), 434–437

automation and orchestration, 431–432

infrastructure as a code, 434–437

UCSD interfaces, 433–434

cloning, 189

compression, 187

deduplication, 187

deployment
  Cisco HyperFlex Connect HTML5 Management, 200–202
  Cisco Intersight, 204
  installation phase, 197–199
  sizer, 199
  VMware vSphere management plug-in, 203–204
  workload profiler, 199
HXDP (HX Data Platform)
  data distribution, 174–182
  Docker containers support, 172–173
  DPC (data platform controller), 169–170
  Hyper-V environment, 171–172
  overview of, 168
  read/write operations, 181–187
  VMware ESXi environment, 170–171
hyperconverged product portfolio, 108
NRDR (native replication for DR), 189–191
Nutanix ECP versus
  data locality, 285–286
  data tiering, 282
  deduplication, 283–284
  distributed architecture, 281–282
  file systems, 282
  hardware comparison, 280–281
  overview of, 279–280
overview of, 153–154
performance benchmarks, 160–162
physical components, 154–156
  all-flash nodes, 156
  Cisco 5108 UCS blade chassis, 159–160
  Cisco C220/C240 M4/M5 rack servers, 158–159
  Cisco VIC MLOM interface cards, 159
  compute-only nodes, 157
  edge nodes, 157
  hybrid nodes, 156
  UCS 6200 and 6300 fabric interconnect, 158
security, 192
  data at rest, 192–193
  management security, 194
  operational security, 193–194
  regulation and certification compliance, 194
snapshots, 188–189
third-party backup tool integration, 191–192
UCS (Unified Computing System) integration, 162
  external storage, 167–168
  logical network design, 162–164
  service templates and profiles, 164–165
  vNIC templates, 166–167
VMware vSAN versus, 246
  advanced data services, 251–252
  compute-only nodes, 250–251
  file systems, 249–250
  hardware comparison, 247
  management software, 252
  networking, 252–253
  one-to-one versus many-to-many rebuild, 250

scaling up, 247–248

vSAN in-kernel versus controller-based solutions, 248–249

workload optimization and efficiency

Cisco AppDynamics, 217–219

CWOM (Cisco Workload Optimizer), 216–217

enterprise workload issues, 211–212

HyperFlex with Cisco Tetration, 212–215

overview of, 211

HyperFlex cluster menu (Cisco Intersight), 207

HyperFlex Connect HTML5 Management, 200–202

Hypertext Transfer Protocol (HTTP), 5, 338, 434–435

hyperthreading (HT), 67

Hyper-V, 136

HyperFlex in, 171–172

Nutanix support for, 274–275

hypervisors, 79–80

I

IA (Infrequent Access) S3 storage, 405

IaaS (infrastructure as a service), 396–397, 423

IAM (identity and access management), 411–412, 425–426

IANA (Internet Assigned Numbers Authority), 13

IB (InfiniBand), 40

IBM DB2, 272–273

IBM Smartcloud Enterprise, 397

identity and access management (IAM), 411–412, 425–426

IDSs (intrusion detection systems), 6

IEEE (Institute of Electrical and Electronics Engineers), 49

IETF (Internet Engineering Task Force), 117

IGMP (Internet Group Management Protocol), 315, 347

InfiniBand (IB), 40

infrastructure as a code, 417–420, 434–437

infrastructure as a service (IaaS), 396–397, 423

infrastructure loopback IP addresses, 370

Infrequent Access (IA) S3 storage, 405

inline deduplication, 58

input/output operations per second (IOPS), 24–25, 73, 138, 156, 199, 216, 247, 296

Institute of Electrical and Electronics Engineers (IEEE), 49

Intel, 155, 311

Intermediate System to Intermediate System (IS-IS), 8, 126, 308, 371

Internet Assigned Numbers Authority (IANA), 13

Internet Engineering Task Force (IETF), 117

Internet Group Management Protocol (IGMP), 315, 347

Internet Protocol. *See* IP (Internet Protocol)

Internet Small Computer System Interface. *See* iSCSI (Internet Small Computer Interface)

Intersight, 204, 446

advantages of, 205

claiming nodes in, 205–209

conceptual view of, 204–205

menus, 207–209

interVLAN routing, 12

intrusion detection systems (IDSs), 6

intrusion prevention systems (IPSs), 6

IOPS (input/output operations per second), 24–25, 73, 138, 156, 199, 216, 247, 296

IOVisor, 170

IP (Internet Protocol)

addressing, 12–13, 370–371

routing, 343–345

IP address management (IPAM), 214

IPAM (IP address management), 214

IPSs (intrusion prevention systems), 6

IQN (iSCSI Qualified Name), 47

Ironic, 291

iSCSI (Internet Small Computer Interface), 82

addressing, 47–48

deployments, 46–47

IQN (iSCSI Qualified Name), 47

iSER (iSCSI extensions for Remote Direct Memory Access), 40

LUN masking, 48

multipathing, 48–49

network segmentation, 48

vSAN iSCSI target service, 243

iSER (iSCSI extensions for Remote Direct Memory Access), 40

IS-IS (Intermediate System to Intermediate System), 8, 126, 308, 371

IT as a service (ITaaS), 423

J

Java Eclipse, 437

JavaScript Object Notation. ***See*** **JSON (JavaScript Object Notation)**

JBOD (just a bunch of disks), 98

JSON (JavaScript Object Notation), 434

CloudFormation templates, 417–420

exporting Tetration data to, 215

IAM (identity and access management), 412

K

Karbor, 292

KEKs (key encryption keys), 59, 192–193

Kerberos, 275

Keystone, 292, 293

KMIP (Key Management Interoperability Protocol), 131, 234

KMSs (key management servers), 411

Kolla, 292

Kuryr, 292

KVMs (kernel-based virtual machines), 136, 256

L

LAN connectivity policy, 165

large enterprises (LEs), 136–137

latency, 11–12, 24–25

Layer 2 (L2), 4, 343, 378–380

Layer 3 (L3), 4, 380–383

LAZ (logical availability zones), 181

leaf layer, 313

leaf node loopback IP addresses, 370
least recently used (LRU) counters, 262
legacy installations
SDS (software-defined storage), 97–99
VDI (virtual desktop infrastructure) with, 143–144
Lenovo, 255
LEs (large enterprises), 136–137
LFS (log-structured file system), 122–124, 181–182
LIBRADOS, 300–301
lightweight snapshots (LWSs), 272
linked clones, 237
Linux containers (LXC), 80–81
load balancing, 5, 266–267, 314
ACI (application-centric infrastructure), 373
computing, 89
storage, 90
local data protection, 148
local deduplication, 58
local snapshots, 148
local station tables (LSTs), 371
Log Insight, vRealize, 246
logical availability zones (LAZ), 181
logical cores, 67–68
logical network design, 162–164
logical server grouping, 8–9
logical snapshots, 52–54
logical switches (LSWs), 341–342
logical unit numbers. *See* LUNs (logical unit numbers)
logical volume manager (LVM), 33–35
log-structured file system (LFS), 122–124, 181–182
LPM (longest prefix match), 373
LRU (least recently used) counters, 262
LSTs (local station tables), 371
LSWs (logical switches), 341–342
LUNs (logical unit numbers), 31–33, 82, 124, 136–137
configuration, 84
masking, 43, 48
LVM (logical volume manager), 33–35
LWSs (lightweight snapshots), 272
LXC (Linux container), 80–81
LZ4 algorithm, 235

M

MAC (media access control) addresses, 108
Magnum, 291
management. *See also* monitoring
HXDP (HX Data Platform), 200–204
management plane, 308
NSX manager, 339–340
Nutanix ECP (Enterprise Cloud Platform), 275–279
overview of, 131–133
security, 194
server virtualization, 90
VMware vSAN, 244–246
vSAN versus HyperFlex, 252
Manila, 291
many-to-many rebuild, 250
mapping
between clouds, 447
databases, 371
MapReduce, 258
masking LUNs, 43, 48
MD5 hashing algorithm, 57, 299

media access control (MAC) addresses, 108

MEKs (media encryption keys), 59

Mellanox Technologies Ltd., 311

memory. ***See also*** **caching**

allocation of, 85

CAM (content-addressable memory), 308

DRAM (dynamic random-access memory), 147

NVMe (Non-Volatile Memory Express), 71–72, 199, 311

objects, 227

MEM-VVol-Snapshots, 101

Message Digest 5 (MD5) algorithm, 57

metadata, 182, 263

metadata vNode, 185

methods (HTTP), 435

MG-BGP (Multi-Protocol Border Gateway Protocol), 126, 371

microsegmentation constructs, 309–310, 319

ANPs (application network profiles), 355–358

EPGs (endpoint groups), 353–355

service graphs, 358–359

tetration model, 360

Microsoft Cloud-Scale Acceleration Architecture, 311

migration

storage, 90

VMs (virtual machines), 87

Mistral, 292

ML2 (Modular Layer 2), 321

MLAG (multichassis link aggregation), 18

MLC (multilevel cell) NAND, 73

MLOM (modular LAN-on-motherboard), 159

Monasca, 293

monitoring

AWS (Amazon Web Services), 416

HXDP (HX Data Platform), 200–204

Nutanix ECP (Enterprise Cloud Platform), 275–279

overview of, 131–133

Moore's Law, 65–66

MP-BGP (Multi-Protocol BGP), 126, 308, 371

MPIO (multipathing I/O), 273

MPIT (multiple points in time), 240

multicast, 383

multichassis link aggregation (MLAG), 18

multicloud, 442–446. ***See also*** **hybrid cloud**

multi-core CPUs (central processing units), 66–67

multidestination traffic

multicast model with no optimization, 347

overview of, 346–347

unicast mode, 348–349

multilevel cell (MLC) NAND, 73

multipathing, 44–46, 48–49

multipathing I/O (MPIO), 273

multiple points in time (MPIT), 240

Multi-PoD (point of delivery), 383–384

Multi-Protocol Border Gateway Protocol (MP-BGP), 126, 308, 371

Multi-Site, 384–386

multitier applications, launching in AWS (Amazon Web Services)

AMI (Amazon Machine Images), 410

compute instances, 409

EC2 instances, 412–416

IAM (identity and access management), 411–412
security groups, 410–411
multitier design, 6–8, 20–21, 41
BUM traffic flooding, 15–16
firewall overload, 17–18
flat L2 networks with stretched VLANs, 10–11
interVLAN routing via SVI, 12
IP4 address scarcity, 12–13
loop prevention via spanning tree, 16
oversubscription between tiers, 9–10
private VLANs, 13–15
traffic hopping between tiers, 11–12
Murano, 293

N

NAND flash, 73
NAS (network-attached storage), 39–40, 118, 243–244
NAT (Network Address Translation), 6, 13, 322–323
National Institute of Standards and Technology (NIST), 194, 235
native command queueing (NCQ), 71–72
native replication for DR (NRDR), 189–191
NCQ (native command queueing), 71–72
NearSync replication, 272
Network Address Translation (NAT), 6, 13, 322–323
network equipment, 4
Network File System. *See* NFS (Network File System)
network interface cards (NICs), 7, 85
Network Management Systems (NMSs), 107–108
Network Processors Units (NPUs), 312
network virtualization and security platform. *See* NSX (network virtualization and security platform)
Network Virtualization using Generic Routing Encapsulation (NVGRE), 126, 315–316
network-attached storage (NAS), 39–40, 118
networking and security software product. *See* NSX (network virtualization and security platform)
networks. *See* data networks; storage networks
Neutron, 292
agents, 321–322
APIs (application programming interfaces), 321
overview of, 320
plug-ins, 321
server, 320
Nexus 1000v, 93, 321
NFS (Network File System), 36, 82, 102, 244
NGFWs (next generation firewalls), 6, 314
NICs (network interface cards), 7, 85
Cisco UCS (Unified Computing System), 108
smart NICs, 310–312
vNICs (virtual NICs), 109
NIST (National Institute of Standards and Technology), 194, 235
NMSs (Network Management Systems), 107–108

No Structured Query Language (NoSQL), 258

nodes, 154–155

all-flash, 156

claiming, 205–209

clusters, 86–87

compute-only, 157, 250–251

edge, 157

hybrid, 156

vNodes, 182, 185

nonpersistent desktops, 142

Non-Volatile Memory Express (NVMe), 71–72, 199, 311

NoSQL (No Structured Query Language), 258

Nova, 294–296

NPUs (Network Processors Units), 312

NRDR (native replication for DR), 189–191

NSX (network virtualization and security platform)

ACI (application-centric infrastructure) compared to

automation and visibility, 390–391

control plane, 389

data forwarding, 390

networking learning curve, 391

overview of, 386–387

policy setting and enforcement, 387–388

VXLAN performance, 388–389

controller clusters, 340–341

flooding avoidance, 342

IP routing, 343–345

L2 switching, 343

multidestination traffic

multicast model with no optimization, 347

overview of, 346–347

unicast mode, 348–349

NSX manager, 339–340

overview of, 15, 252, 335–336

policies

enforcement of, 338

overview of, 335–336

sample policy rules, 337–338

security groups, 336–337

vDS (virtual Distributed Switch) enhancements, 341–342

NSX (networking and security software), 310

Nutanix ECP (Enterprise Cloud Platform)

Acropolis Block Services, 272–273

Acropolis File Services, 273–274

availability domains, 263–264

backup and restore, 270–271

clones, 268–269

cluster components, 258–259

data at rest encryption, 272

data compression, 265–266

data path resiliency, 264

data protection, 262

deduplication, 264–265

disk balancing, 266–267

Docker containers, 275

DSF (distributed storage fabric), 257–259

Era database services, 269–270

erasure coding, 266

HyperFlex versus

data locality, 285–286

data tiering, 282

deduplication, 283–284
distributed architecture, 281–282
file systems, 282
hardware comparison, 280–281
overview of, 279–280
Hyper-V support, 274–275
I/O path, 261
metadata, 263
overview of, 255–257
persistent storage, 275
physical drives, 260–261
Prism management system, 275–277
read I/O, 262
replication, 270–271
snapshots, 268–269
storage tiering, 267
stretched clusters, 271–272
write I/O, 261–262
Nutanix Flow, 331
NVGRE (Network Virtualization using Generic Routing Encapsulation), 126, 315–316
NVMe (Non-Volatile Memory Express), 71–72, 199, 311
NVMe-oF (NVMe over fabrics), 72

O

object file system, 226–227
Object Space Reservation, 231
object store daemons (OSDs), 301
objects
distributed, 183
object-based storage, 36–37
object-level deduplication, 56
S3 (Simple Storage Service), 405–406
Swift, 299
OCP (Open Container Project), 80–81
Octavia, 292
ODL (OpenDaylight), 324
ODX (Offload Data Transfer), 172
Office 365, 443
Offload Data Transfer (ODX), 172
one-to-one rebuild, 250
Open Container Project (OCP), 80–81
Open Shortest Path First (OSPF), 8, 126, 308
open source, 289–290
Ceph, 300–303
networking initiatives
Neutron, 320–322
OVS (Open vSwitch) architecture, 322–324
OpenStack, 99, 331, 353
Cinder Block Storage, 296–297
current projects, 290–293
Nova, 294–296
OpenStack Foundation, 290
Swift, 296–297
vendors, 331
Open Virtual Appliance (OVA), 198
Open Virtual Network. *See* OVN (Open Virtual Network)
Open Virtuozzo containers, 80–81
Open vSwitch Database (OVSDB), 359
Open vSwitch (OVS). *See* OVS (Open vSwitch)
OpenCloud, 397
OpenDaylight (ODL), 324
OpenStack, 99, 331, 353
Charms, 292
Cinder Block Storage, 296–297

client, 292
current projects, 290–293
Nova, 294–296
OpenStack Foundation, 290
OpenStack-Ansible, 293
Swift, 290–293, 297–300
OpenVZ, 80–81
operating systems, 85–86
operational security, 193–194
OPEX (operational expenditures), 127, 140, 212, 216, 399
OpFlex, 359
Oplog, 260–262, 265, 272, 282
optimization
data, 185–186
multidestination traffic without, 347
traffic, 5
workload, 211–219
Oracle Real Application Clusters (RAC), 272–273
orchestration
Cisco UCSD (USC Director), 430
installation, 431–432
overview of, 425–427
OSDs (object store daemons), 301
OSPF (Open Shortest Path First), 8, 126, 308
OSs (operating systems), 85–86
Other VVols, 101
OVA (Open Virtual Appliance), 198
overlay network, 309–310, 315–319. ***See also*** **ACI (application-centric infrastructure)**
oversubscription between tiers, 9–10
OVN (Open Virtual Network), 127
control plane improvements, 328–331
data plane improvements, 329–331
functionalities, 326–327
high-level architecture, 327–328
overview of, 326
OVS (Open vSwitch), 93, 127
architecture, 322–324
Nutanix's use of, 256
OVN (Open Virtual Network) compared to, 328–331
OVSDB (Open vSwitch Database), 359
Ovsdb-server, 325
Ovsdb-vswitchd, 325–326

P

PaaS (platform as a service), 397, 423
packet-based firewalls, 6
Panko, 293
partial data locality, 175
partitions, Swift, 299
PCI (Peripheral Component Interconnect), 30–31
PCIe (Peripheral Computer Interconnect express), 30–31, 70–71, 311
PCIX (PCI eXtended), 70
pCPUs (physical CPUs), 67
PDs (protection domains), 270
performance
disk
access time, 24
IOPS (input/output operations per second), 24–25
latency, 24–25
throughput, 21–24
HCI (hyperconverged infrastructure), 120–121
HyperFlex (HX), 160–162

VXLAN (Virtual Extensible LAN), 388–389

persistent desktops, 142

persistent storage, 244, 275

PFTT (primary level of failures to tolerate), 228–230, 241

physical cores, 67–68

physical CPUs (pCPUs), 67

physical domains, 364–365

physical NICs (pNICs), 92–93

physical snapshots, 52

PIM (Protocol Independent Multicast), 315, 347

Planar NAND, 73

platform as a service (PaaS), 397, 423

PLOGI (port login) process, 44

pNICs (physical NICs), 92–93

policies

ACI (application-centric infrastructure), 387–388

boot, 165

firmware, 165

LAN connectivity, 165

networking, 126

NSX (network virtualization and security platform), 387–388

storage, 228–231

VMware NSX (network virtualization and security platform)

enforcement of, 338

overview of, 335–336

sample policy rules, 337–338

security groups, 336–337

Policy menu (Cisco Intersight), 209

Portable Operating System Interface (POSIX), 181

ports

community switch, 14

PLOGI (port login) process, 44

promiscuous, 14

virtual switching, 90–93

POSIX (Portable Operating System Interface), 181

POST method, 435

post-process deduplication, 58

Primar, 241

primary level of failures to tolerate (PFTT), 228–230, 241

principals, 411

Prism management system, 257, 275–277

private cloud. *See also* **hybrid cloud; public cloud**

automation, 425–427

Cisco UCSD (USC Director), 425–427

catalogs, 431

HyperFlex integration, 431–437

orchestration, 425–427, 430

overview of, 427

policies, 427–429

vDCs (virtual data centers), 429–430

consumption models, 423–425

convergence and hyperconvergence, 425

HCI business benefits to, 138–139

orchestration, 425–427

private VLANs, 13–15

PRLI (process login), 44

profiles

ANPs (application network profiles), 355–358

Cisco UCS (Unified Computing System), 111–112

service, 111–112, 164–165, 207–209

promiscuous ports, 14

protection domains (PDs), 270

Protocol Independent Multicast (PIM), 315, 347

provisioning

force, 231

HXDP (HX Data Platform)

Cisco HyperFlex Connect HTML5 Management, 200–202

Cisco Intersight, 204

VMware vSphere management plug-in, 203–204

Nutanix ECP (Enterprise Cloud Platform)

Calm, 278–279

infrastructure management, 276–277

operational insight, 277

overview of, 275–276

tools, 278

overprovisioning, 212

overview of, 124–125, 131–133

server virtualization, 90

thick, 50–51

thin, 50–51

proxy servers, Swift, 298–300

public cloud, 406–409. ***See also*** **hybrid cloud; private cloud**

AWS (Amazon Web Services)

AMI (Amazon Machine Images), 410

availability zones, 399–401

cloud automation, 416–421

cloud monitoring, 416

CloudFormation, 417–420

compute instances, 409

EC2 instances, 412–416

global infrastructure, 399–401

IAM (identity and access management), 411–412

networking, 401–404

overview of, 398–399

regions, 399–401

S3 (Simple Storage Service), 36

security groups, 410–411

storage, 404–409

IaaS (infrastructure as a service), 396–397

overview of, 395–396

PaaS (platform as a service), 397

SaaS (software as a service), 397

Puppet OpenStack, 293

PUT method, 435

Q-R

question mark (?), 436

RabbitMQ, 295

RAC (Real Application Clusters), 272–273

Rackspace, 290, 397. ***See also*** **OpenStack**

RADOS (Reliable Autonomous Distributed Object Store), 300–301

RAID (Redundant Array of Independent Disks), 26–30, 98, 127

Rally, 292

RAM (random access memory), 65

Raw Device Mapping (RDM), 82

RBAC (role-based access control), 200

RBD (RADOS block device), 300–301

RCT (resilient change tracking), 275

RDBs (relational databases), 4, 146

RDM (Raw Device Mapping), 82
RDMA (Remote Direct Memory Access), 72
read/write operations, 181–187
  data virtualization, 182–183
  destaging and data optimization, 185–186
  LFS (log-structured file system), 181–182
  Nutanix ECP (Enterprise Cloud Platform), 261–262
  read operation lifecycle, 186–187
  write operation lifecycle, 183–185
ReadyClones, 188
Real Application Clusters (RAC), 272–273
recovery. *See* disaster recovery
recovery point objectives (RPOs), 51, 129–130, 148, 271–272
recovery time objectives (RTOs), 51, 148, 271–272
Red Hat Enterprise Linux OpenStack Platform, 331
Red Hat OpenShift, 397
redirect on write (ROW) logical snapshots, 52–54
redirection of traffic, 5
Reduced Redundancy Storage (RRS), 405
Redundant Array of Independent Disks (RAID), 26–30, 98, 127
redundant array of independent disks (RAID), 127
ReFS (Resilient File System), 275
regions (AWS), 399–401
regulations
  multicloud adoption and, 443
  security, 194
relational databases (RDBs), 4, 146
Reliable Autonomous Distributed Object Store (RADOS), 300–301
Remote Direct Memory Access (RDMA), 72
remote procedure calls (RPCs), 295, 321
replication, 55, 129–130
  data protection via, 121–122, 178–179
  Nutanix ECP (Enterprise Cloud Platform), 270–271
  unicast mode, 348–349
  VMware vSAN, 238–241
replication, cloud, 148
Representation State Transfer. *See* RESTful (Representation State Transfer) APIs
resilient change tracking (RCT), 275
Resilient File System (ReFS), 275
RESTful (Representation State Transfer) APIs, 36, 215, 295, 297, 434–436
return on investment (ROI), 445
RGW (RADOS Gateway), 300
rings, Swift, 299
ROBO (remote office business office), 144–146
RoCE (RDMA over Converged Ethernet), 72
ROI (return on investment), 445
role-based access control (RBAC), 200
routing equipment, 4
ROW (redirect on write) logical snapshots, 52–54
RPCs (remote procedure calls), 295, 321
RPOs (recovery point objectives), 51, 129–130, 148, 271–272

RRS (Reduced Redundancy Storage), 405

RTOs (recovery time objectives), 51, 148, 271–272

S

S3 (Simple Storage Service), 36, 405–406, 425–426

SaaS (software as a service), 214, 397, 423

Sahara, 292

Salesforce, 443

SANs (storage area networks), 4, 78, 118, 135. *See also* **vSAN**

- converged systems, 113
- FC (fibre channel)
  - *addressing, 42*
  - *LUN masking, 43*
  - *multipathing, 44–46*
  - *overview of, 41–42*
  - *switching, 44–45*
  - *zoning, 43*
- FCoE (Fibre Channel over Ethernet), 49–50, 82, 109
- iSCSI (Internet Small Computer Interface)
  - *addressing, 47–48*
  - *deployments, 46–47*
  - *LUN masking, 48*
  - *multipathing, 48–49*
  - *network segmentation, 48*
- overview of, 40

SAP Business Warehouse (BW), 147

SAP HANA, 114, 147

SAS (serial attached SCSI), 23, 70

SATA (serial advanced technology attachment), 22, 70

scaling

- distributed DAS architecture, 120
- EC2 auto scaling, 414–416
- HCI business benefits, 136–137
- scalable hashing, 303
- vSAN versus HyperFlex, 247–248

SCSI controllers, 85

SDDC (vSAN software-defined data center), 245–246

SDKs (software development kits), 420–421, 437

SDN (software-defined networking)

- host-based networking, 310–312
- microsegmentation, 309–310, 319
- open source initiatives, 320–324
- overlay network, 309–310, 315–319
- overview of, 308–309
- OVN (Open Virtual Network)
  - *control plane improvements, 329*
  - *data plane improvements, 329–331*
  - *functionalities, 326–327*
  - *high-level architecture, 327–328*
  - *overview of, 326*
- switch-based networking, 312–313
- switching fabric, 313–315
- underlay network, 315

SDS (software-defined storage), 135

- legacy installations, 97–99
- objectives, 96–97
- open source vendors, 331
- overview of, 95–96, 135
- VASA (vSphere APIs for Storage Awareness), 99–104
- Vipr controller, 99
- VVols (Virtual Volumes), 99–104

search engine, Intersight, 209
Searchlight, 292
secondary level of failures to tolerate (SFTT), 231, 241
Secure Hash Algorithm 1 (SHA-1), 57
Secure Shell (SSH), 338
Secure Sockets Layer (SSL), 5
security. *See also* policies
  data networks, 5–6
  data protection, 121–122, 148–149, 262
  HyperFlex, 192–194
  overview of, 130–131
security groups
  AWS (Amazon Web Services), 410–411
  VMware NSX (network virtualization and security platform), 336–337
Security Technical Implementation Guide (STIG), 131, 235
SEDs (self-encrypting drives), 23, 59, 129–130, 192–193, 199, 235
segmentation, 48, 342, 346–347
self-encrypting drives (SEDs), 23, 59, 129–130, 192–193, 199, 235
Senlin, 292
serial advanced technology attachment (SATA), 22, 70
serial attached SCSI (SAS), 23, 70
server load balancing (SLB), 5
Server menu (Cisco Intersight), 207
Server Message Block (SMB), 36, 244
servers. *See also* virtualization, server
  Cisco C220/C240 M4/M5 rack servers, 158–159
  Cisco USC B200, 159
  clusters, 86–87
  logical, 8–9
  Neutron, 320
  Ovsdb-server, 325
  storage-to-server connectivity, 38–40
  Swift proxy servers, 298–300
service catalogs (UCSD), 431, 432
service graphs, 358–359
service profiles, 111–112, 164–165, 207–209
ServiceNow, 448
services, virtualization. *See* virtualization, server
services delivery policy (UCSD), 429
Settings menu (Cisco Intersight), 209
SFTT (secondary level of failures to tolerate), 231, 241
SHA-1 algorithm, 57
shadow clones, 269
Simple Network Management Protocol (SNMP), 434
Simple Storage Service. *See* S3 (Simple Storage Service)
single root input/output virtualization (SR-IOV), 312
single sign-on (SSO), 200
single-core CPUs, 66–67
Site Recovery Manager (SRM), 240
sizer (HX), 199
Sizing tool (Nutanix), 278
SLB (server load balancing), 5
Slicehost, 290
smart NICs (network interface cards), 310–312
SMB (Server Message Block), 36, 244
SMBs (small and medium-sized businesses), 136–137
snapshots
  EBS (Elastic Block Storage), 408

HyperFlex (HX), 188–189
local, 148
logical, 52–54
Nutanix ECP (Enterprise Cloud Platform), 268–269, 272
overview of, 51–52
traditional or physical, 52
VMware vSAN, 236

SNMP (Simple Network Management Protocol), 434

software as a service (SaaS), 214, 397, 423

software development kits (SDKs), 420–421, 437

software-defined networking. *See* SDN (software-defined networking)

software-defined storage. *See* SDS (software-defined storage)

Solarflare Communications Inc., 311

solid-state drives. *See* SSDs (solid-state drives)

Solum, 293

source-based deduplication, 57–58

spanning trees, 16

spine nodes, 314

spine-proxy IP, 370

SQL (Structured Query Language), 146, 272–273

SR-IOV (single root input/output virtualization), 312

SRM (Site Recovery Manager), 240

SSDs (solid-state drives), 23, 66–67, 70, 82–83, 118, 147, 155, 406–408

SSH (Secure Shell), 338

SSL (Secure Sockets Layer), 5

SSO (single sign-on), 200

standard catalogs, 431

Stargate, 258, 270

Stateless Transport Tunneling (STT), 126, 315–316

stHypervisorSvc, 171

STIG (Security Technical Implementation Guide), 131, 235

storage area networks. *See* SANs (storage area networks)

storage arrays, 102

storage controllers, 30–31

storage efficiency technologies
caching storage arrays, 60–61
cloning, 55
data compression, 58–59
deduplication
block-level, 56–57
definition of, 55–56
file-level, 56
global, 58
inline, 58
local, 58
object-level, 56
post-process, 58
source-based, 57–58
target-based, 57–58
disk encryption, 59
replication, 55
snapshots
logical, 52–54
overview of, 51–52
traditional or physical, 52
storage tiering, 59–60
thin provisioning, 50–51

storage load balancing, 90

storage networks. *See also* SANs (storage area networks)
block-level storage, 35
caching storage arrays, 60–61
cloning, 55

DAS (direct-attached storage), 38–39
data compression, 58–59
deduplication
block-level, 56–57
definition of, 55–56
file-level, 56
global, 58
inline, 58
local, 58
object-level, 56
post-process, 58
source-based, 57–58
target-based, 57–58
disk encryption, 59
disk performance
access time, 24
IOPS (input/output operations per second), 24–25
latency, 24–25
throughput, 21–24
file-level storage, 35–36
HDDs (hard disk drives), 22–23
LUNs (logical unit numbers), 31–33, 82, 124, 136–137
configuration, 84
masking, 43, 48
LVM (logical volume manager), 33–35
multitier design, 20–21
NAS (network-attached storage), 39–40
object-based storage, 36–37
RAID (Redundant Array of Independent Disks), 26–30
replication, 55
snapshots
logical, 52–54
overview of, 51–52
traditional or physical, 52
SSDs (solid-state drives), 23, 66–67, 70, 82–83, 118, 147, 155, 406–408
storage controllers, 30–31
storage tiering, 59–60
thin provisioning, 50–51
storage policies, 228–231, 428
storage tiering, 59–60, 267
storage-defined networking (SDN), 126
Storlets, 291
stretched clusters
Nutanix ECP (Enterprise Cloud Platform), 271–272
VMware vSAN, 238–241
stretched VLANs, 10–11
strict data locality, 175
striping, 176–177
Structured Query Language (SQL), 146, 272–273
STT (Stateless Transport Tunneling), 126, 315–316
subnetting, 12–13
Support Insight, 245
SVI (switch virtual interface), 12
swap objects, 227
SWAP-VVol-Swap, 101
Swift, 297–300
switch virtual interface. *See* SVI (switch virtual interface)
switches, 4
logical, 341–342
NSX L2 switching, 343
OVS (Open vSwitch), 322–324, 328–331
switch-based networking, 312–313
ToR (top-of-rack), 7, 92–93, 115, 310, 313
virtual, 90–93

switching fabric, 313–315, 367–371
- ACI (application-centric infrastructure), 365–367
- FC (fibre channel), 44–45
- multitier architecture, 6–8
- SDN (software-defined networking), 313–315

synchronous native replication for DR (NRDR), 190–191

T

Tacker, 292

target initiator groups (TIGs), 243

target-based deduplication, 57–58

TBW (Total Bytes Written), 73

TCAM (ternary CAM), 308

TCO (total cost of ownership), 140, 444–445

TCP (Transmission Control Protocol), 5, 46–47, 109

telemetry, Cisco Intersight, 205

templates
- CloudFormation, 417–420
- service, 164–165
- vNIC, 166–167

tenants, 366

ternary CAM (TCAM), 308

Tetration
- HyperFlex with
  - *data collection, 214*
  - *data output, 215*
  - *open access, 215*
  - *overview of, 212–214*
  - *Tetration analytics cluster, 214*
- model for, 360
- Tetration-V, 214

thick provisioning, 50–51

thin provisioning, 50–51

third-party backup tools, 191–192

three-tier architecture. *See* **multitier design**

throughput, 21–24

tier-1 enterprise class applications, 146–147

tiers, storage, 59–60, 267

TIGs (target initiator groups), 243

time to live (TTL), 330

TOE (TCP Offload Engine), 46–47

ToR (top-of-rack) switches, 7, 92–93, 115, 310, 313

Total Bytes Written (TBW), 73

total cost of ownership (TCO), 140, 444–445

transfer speed, 21–24

Transmission Control Protocol (TCP), 5, 46–47, 109

Tricircle, 292

Tripelo, 293

Trove, 292

TTL (time to live), 330

U

UCS (Unified Computing System), 108–112, 138, 223
- HyperFlex integration, 162
  - *external storage, 167–168*
  - *logical network design, 162–164*
  - *service templates and profiles, 164–165*
  - *vNIC templates, 166–167*

UCS 6200 and 6300 fabric interconnect, 158

UCS Integrated Infrastructure for Red Hat Enterprise Linux OpenStack, 331

UCSD (USC Director)

catalogs, 431

HyperFlex integration, 431–437

orchestration, 430

overview of, 427

policies, 427–429

vDCs (virtual data centers), 429–430

UCSM (UCS Manager), 198

UDP (User Datagram Protocol), 5, 342

UKB (Universal Kernel Bypass), 312

underlay network, 315, 371–373

unicast mode, 348–349

Unified Cache, 261

Unified Computing System. *See* UCS (Unified Computing System)

unified threat management (UTM), 6

uniform resource identifiers (URIs), 435–436

union file system, 173

Universal Kernel Bypass (UKB), 312

Universal Unique Identifiers (UUIDs), 111

Uplink Subinterface IP addresses, 370–371

URIs (uniform resource identifiers), 435–436

use cases

data protection and disaster recovery, 148–149

DevOps, 141

edge computing, 146

overview of, 140

ROBO (remote office business office), 144–146

server virtualization, 140–141

tier-1 enterprise class applications, 146–147

VDI (virtual desktop infrastructure), 141–144

User Datagram Protocol. *See* UDP (User Datagram Protocol)

User Datagram Protocol (UDP), 5, 342

User menu (Cisco Intersight), 209

UTM (unified threat management), 6

UUIDs (Universal Unique Identifiers), 111

V

VAAI (vStorage API for Array Integration), 170

VASA (vSphere APIs for Storage Awareness), 99–104

vCenter Server Appliance (VCSA), 245

vCPUs (virtual CPUs), 68–70

VCSA (vCenter Server Appliance), 245

vDCs (virtual data centers), 341–342, 429–430

VDI (virtual desktop infrastructure), 55–56, 114, 141–144, 236, 269

VDP (vSphere Data Protection), 241–243

vDS (virtual Distributed Switch), 93

Veeam, 129–130

vendors, open source, 331

VesraStac, 112

VGs (volume groups), 273

vHBAs (virtual host bus adapters), 109, 166

VICs (virtual interface cards), 109, 159

Vipr, 99

virtual cores, 67–68

virtual CPUs (vCPUs), 68–70

virtual data centers (vDC), 429–430

virtual desktop infrastructure (VDI), 55–56, 114, 141–144, 236, 269

virtual Distributed Switch (vDS), 93

Virtual Extensible LANs. *See* VXLANs (Virtual Extensible LANs)

virtual host bus adapters (vHBAs), 109, 166

virtual interface cards (VICs), 109, 159

virtual local area networks (VLANs), 10–11, 13–15

virtual MAC addresses (vMACs), 92

virtual machine disks (VMDKs), 82, 100

Virtual Machine File System (VMFS), 82, 102, 226

Virtual Machine Manager (VMM) domain, 362–364

virtual machines. *See* VMs (virtual machines)

virtual NICs (vNICs), 79, 92–93, 109, 166–167

virtual port channels (vPCs), 18, 368–369

virtual private cloud (VPC), 401–404

virtual routing and forwarding (VRF), 323, 366

virtual storage appliance (VSA), 119

virtual switching, 90–93

Virtual Tunnel Endpoint (VTEP), 369–370

Virtual Volumes. *See* VVols (Virtual Volumes)

Virtual Volumes (VVols), 83, 99–104, 243–244

virtualization, server

datastores, 82–83

HCI use case, 140–141

overview of, 77–78

virtual switching, 90–93

virtualization layer

bare-metal versus virtualized servers, 78–79, 396

Docker containers, 80–82

type 1 hypervisors, 79

type 2 hypervisors, 80

virtualization services

clusters, 86–87

computing load balancing, 89

fault tolerance, 89

overview of, 86

provisioning and management, 90

storage load balancing, 90

storage migration, 90

VM migration, 87

VM (virtual machine) creation, 84–86

datastore configuration, 84

guest OS assignment, 85

LUN configuration, 84

memory allocation, 85

new VM creation, 84

NICs (network interface cards), 85

OS installation, 86

SCSI controllers, 85

virtual disk assignment, 85

Vitrage, 292

VLANs (virtual local area networks), 10–11, 13–15

vMACs (virtual MAC addresses), 92

VMDKs (virtual machine disks), 82, 100, 226–227

VMFS (Virtual Machine File System), 82, 102, 226

VMK (vmkernel) port groups, 92

VMM (Virtual Machine Manager) domain, 362–364

vmnics, 92–93

vMotion, 87

VMs (virtual machines). ***See also*** **vSAN**

- creation of, 84–86
  - *datastore configuration, 84*
  - *guest OS assignment, 85*
  - *LUN configuration, 84*
  - *memory allocation, 85*
  - *new VM creation, 84*
  - *NICs (network interface cards), 85*
  - *OS installation, 86*
  - *SCSI controllers, 85*
- migration, 87
- vCPUs (virtual CPUs), 68–70

VMware. ***See also*** **NSX (network virtualization and security platform); vSAN**

- ESXi, 79, 170–171
- HCL (Hardware Compatibility List), 245
- VASA (vSphere APIs for Storage Awareness), 99–104
- vSphere DRS, 89
- vSphere management plug-in, 203–204
- vSphere Storage vMotion, 90
- VSS (vNetwork Standard Switch), 93
- VUM (VMware Update Manager), 245
- VVols (Virtual Volumes), 99–104, 243–244

vNetwork Standard Switch (VSS), 93

vNICs (virtual NICs), 79, 92–93, 109, 166–167

VNIDs (VXLAN Instance IDs), 376–378

vNodes, 182, 185

volumes

- adding to EC2 instances, 414
- EBS (Elastic Block Storage), 407–409
- granular, 100–101
- groups, 273
- VVols (Virtual Volumes), 83, 99–104, 243–244

voting, 230

VPC (virtual private cloud), 401–404

vPCs (virtual port channels), 18, 368–369

vRealize, 245–246

VRF (virtual routing and forwarding), 323, 366

vROps (vSAN vRealize Operations), 245–246

VSA (virtual storage appliance), 119

vSAN

- backup, 241–243
- caching, 232
- cloning, 236–237
- Compatibility Guide, 224
- compression, 235
- data encryption, 234–235
- data integrity, 234
- datastore, 228
- deduplication, 235
- erasure coding, 236
- high-level software architecture, 225–226

- HyperFlex versus, 246
  - *advanced data services, 251–252*
  - *compute-only nodes, 250–251*
  - *file systems, 249–250*
  - *hardware comparison, 247*
  - *management software, 252*
  - *networking, 252–253*
  - *one-to-one versus many-to-many rebuild, 250*
  - *scaling up, 247–248*
  - *vSAN in-kernel versus controller-based solutions, 248–249*
- integration with legacy SAN and NAS, 243–244
- I/O operation details, 232–233
- management, 244–246
- object file system, 226–227
- overview of, 223–224
- persistent storage for containers, 244
- physical components, 224–225
- replication, 238–241
- SDDC (vSAN software-defined data center), 245–246
- snapshots, 236
- storage policies, 228–231
- Support Insight, 245
- vROps (vSAN vRealize Operations), 245–246

vSphere APIs for Storage Awareness (VASA), 99–104

vSphere Data Protection (VDP), 241–243

vSphere DRS, 89

vSphere management plug-in, 203–204

vSphere Storage vMotion, 90

VSS (vNetwork Standard Switch), 93

vStorage API for Array Integration (VAAI), 170

VTEP (VXLAN Tunnel Endpoint), 342, 346–347, 369–370

VUM (VMware Update Manager), 245

VVols (Virtual Volumes), 83, 99–104, 243–244

VxBlock, 112, 246

VXLANs (Virtual Extensible LANs), 93, 126, 315–319

- control plane, 389
- overview of, 373–376
- performance, 388–389
- VNIDs (VXLAN Instance IDs), 376–378
- VTEP (VXLAN Tunnel Endpoint), 342, 346–347

VxRail, 223

W

wafers, 66–67

WAN optimizers, 5, 314

Watcher, 292

WebEx, 443

Windows Server 2016, 275

Windows Server Failover Cluster (WSFC), 243

workload optimization

- Cisco AppDynamics, 217–219
- CWOM (Cisco Workload Optimizer), 216–217
- enterprise workload issues, 211–212
- HyperFlex with Cisco Tetration, 212–215
- overview of, 211

workload profiler, 199
write operations. *See* read/write operations
WSFC (Windows Server Failover Cluster), 243
WWNN (World Wide Node Name) addresses, 108
WWNs (World Wide Names), 42–43
WWPNs (World Wide Port Names), 42–43

X

x86 standard architecture, 66
XenServer, 256
XML (Extensible Markup Language), 215, 434
XP fabric extenders, 159–160
X-ray tool (Nutanix), 278
Xtract tool (Nutanix), 278

Y-Z

YAML Ain't Markup Language, 215
Zaqar, 293
zero-trust model, 319
Zerto, 129–130
zones
  AWS (Amazon Web Services), 299
  FC (fibre channel), 43
  LAZ (logical availability zones), 299
  Swift, 299
Zookeeper, 258
Zun, 291