

DATA CENTER NETWORK DESIGN AND TECHNOLOGIES

MAHESH SUBRAMANIAM
MICHAL STYSZYNSKI
HIMANSHU TAMBAKUWALA

Foreword by JEFF DOYLE



FREE SAMPLE CHAPTER |



AI DATA CENTER NETWORK DESIGN AND TECHNOLOGIES

This page intentionally left blank

AI DATA CENTER NETWORK DESIGN AND TECHNOLOGIES

Mahesh Subramaniam, Michal Styszynski,
Himanshu Tambakuwala

◆◆ Addison-Wesley

Hoboken, New Jersey

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and the publisher was aware of a trademark claim, the designations have been printed with initial capital letters or in all capitals.

The authors and publisher have taken care in the preparation of this book, but make no expressed or implied warranty of any kind and assume no responsibility for errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of the use of the information or programs contained herein.

Visit us on the Web: informit.com/aw

Library of Congress Control Number: 2025949447

Copyright © 2026 Pearson Education, Inc.

All rights reserved. This publication is protected by copyright, and permission must be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. For information regarding permissions, request forms, and the appropriate contacts within the Pearson Education Global Rights & Permissions Department, please visit www.pearson.com/global-permission-granting.html.

No patent liability is assumed with respect to the use of the information contained herein. Although every precaution has been taken in the preparation of this book, the publisher and author assume no responsibility for errors or omissions. Nor is any liability assumed for damages resulting from the use of the information contained herein.

Please contact us with concerns about any potential bias at <https://www.pearson.com/report-bias.html>.

ISBN-13: 978-0-13-543628-8

ISBN-10: 0-13-543628-1

\$PrintCode

**Head of Enterprise Content
and Training, Enterprise
Learning and Skills**

Julie Phifer

Executive Editor

James Manly

Development Editor

Ellie C. Bru

Managing Editor

Sandra Schroeder

Senior Project Editor

Mandie Frank

Copy Editor

Kitty Wilson

Indexer

Timothy Wright

Proofreader

Donna E. Mulder

Technical Reviewer

Tanveer Dandiana

Designer

Chuti Prasertsith

Compositor

codeMantra

To my brother, Rajesh, who taught me the meaning of connection—long before I ever studied networks. “Your absence echoes in every page, a quiet hum between the lines... A bright star in my sky, the legacy I live and build for... You are the architecture of my soul—my reason to dream, my need to create, my home in every connection.”

I also want to dedicate this book to my parents, Sarasu and Subramaniam, my wife, Ramya, and my sons, Rithvik and Rithish. They are my steady roots beneath every life storm, the strength that holds the ground I walk upon. Their love is my unwavering light; that light is the one and only thing which makes me stand.

—Mahesh

I want to dedicate this book to my wife, Kasia, and my sons, Ernest and Marcel, for their love, patience, and understanding. Their encouragement, and joy for life inspire me every day and have helped me complete this book.

—Michal

To the ones I owe it all: Neha, Rachit, Aanvi, and my parents, Prakash and Indu. Thank you for every late night and every quiet sacrifice you made.

—Himanshu

This page intentionally left blank

Contents

Foreword	xv
Preface	xvii
Acknowledgments	xix
About the Authors	xxi
1 Wonders in the Workload	1
What's New in AI Data Center Workloads	1
The Life Cycle of an AI Model	2
Training an AI Model	3
Parallelism	4
Data Parallelism	4
Model Parallelism	5
Job Completion Time (JCT)	6
Tail Latency	7
Understanding Traffic and RDMA	8
RDMA Transport Types	9
RoCEv2	11
Summary	16
Test Your Knowledge	17
2 “The Common-Man View” of AI Data Center Fabrics	19
Training vs. Inference AI Data Centers	19
AI Training Data Centers	19
AI Inference Data Centers	20
InfiniBand vs. Ethernet for AI Training Data Centers	21
Advantages of InfiniBand in AI Training Clusters	21
Advantages of Ethernet in AI Training Clusters	22
Ethernet Hardware Switches and Advanced Software Features	22
Handling Elephant Flows	24

	Load-Balancing Techniques	25
	Congestion Management and Mitigation Techniques	26
	Summary	28
	Test Your Knowledge	29
3	Network Design Considerations	31
	Background Introduction	31
	Training Data Center Architecture	33
	Rail-Optimized Design (ROD)	34
	Rail-Unified Design (RUD)	42
	Rack Design	45
	Top-of-Rack	45
	Middle-of-Row	47
	End-of-Row	47
	Design Comparison	48
	Scheduled Fabric	49
	Topologies	50
	Dragonfly Topology	50
	Torus Topology	53
	Inference Data Center Architecture	56
	Multi-Planar Scale-Out Architectures	56
	Summary	63
	Test Your Knowledge	64
	References	66
4	Optics and Cable Management	67
	Scaling Optics for AI Clusters	67
	Challenges in Optical Innovation	70
	Packet Flow	70
	Demultiplexers and Multiplexers	71
	Digital Signal Processors (DSPs)	71
	Transmission Modes	73

Multi-Mode Fiber (MMF)	74
Single-Mode Fiber (SMF)	75
Dense Wavelength Division Multiplexing (DWDM)	75
AI Server Connectivity Options	75
Transceiver Types	76
Cable and Connector Types	78
Standards	79
High-Bandwidth Optics	80
Further Innovations in Optics	82
Summary	83
Test Your Knowledge	85
References	86
5 Thermal and Power Efficiency Considerations	87
Thermal Footprints in AI Data Centers	87
Airflow Options	88
Front-to-Back Airflow	88
Back-to-Front Airflow	88
Bidirectional Fans	89
Choosing the Fan Direction	89
Liquid Cooling	89
Immersion Liquid Cooling	90
Cold Plate Liquid Cooling	91
Rear-Door Heat Exchanger Liquid Cooling	92
Sprayed Liquid Cooling	93
Summary	93
Test Your Knowledge	94
References	95
6 Efficient Load Balancing	97
Per-Flow Load Balancing	99
Static Load Balancing	100

	Dynamic Load Balancing	103
	Global Load Balancing	106
	Traffic Engineering–Based Load Balancing	110
	Per-Packet Load Balancing	115
	Random Spray	116
	DLB Per-Packet Mode	116
	Selective Packet Spraying	116
	Load-Balancing Mechanism Comparison	117
	Summary	118
	Test Your Knowledge	119
7	RoCEv2 Transport and Congestion Management	123
	Congestion Points	123
	Explicit Congestion Notification (ECN)	127
	Priority Flow Control (PFC)	130
	Data Center Quantized Congestion Notification (DCQCN)	134
	Source Flow Control (SFC)	136
	Congestion Signaling	137
	Summary	139
	Test Your Knowledge	140
8	IP Routing for AI/ML Fabrics	143
	Dynamic IP Routing Options	144
	eBGP Underlay for Three-Stage/Five-Stage Fabric for an AI Data Center	145
	eBGP Underlay and BGP Unnumbered	145
	BGP ASN Allocation	149
	BGP Advanced Capabilities	151
	BGP DPF for AI Data Center Path Diversity	155
	Traffic-to-Fabric Color Mapping	161
	Interior Gateway Protocols for AI Data Center Fabric	162
	IS-IS for IP Fabric: Flood Reduction and Flex Algo	168

Multi-tenancy for an AI/ML Cluster Data Center Network	171
Network-Level Multi-tenancy	171
Server-Level Multi-tenancy	173
Combining Server- and Network-Level Multi-tenancy for AI/ML Clusters	175
Microsegmentation and Multi-tenancy for an AI/ML Data Center	177
Extending IP Routing to the Server	177
Traffic Engineering in the AI Data Center Fabric	178
Segment Routing and SRv6 for AI/ML Fabrics	179
SRv6-Based Traffic Engineering	182
SRv6 Compressed SID	183
Segment Routing for AI/ML Clusters	183
Summary	184
Test Your Knowledge	185
References	187
9 Storage Network Design and Technologies	189
The AI Data Center Life Cycle and Storage Networks	191
Storage Network Design Types	193
Block, Object, and File Storage Systems	198
NVMe-oF for Block-Level Access	199
NVMe-o-TCP State Machine	201
NVMe-o-RDMA/RoCEv2 State Machine	206
High-Performance File Systems	208
GPUDirect Storage	211
InfiniBand for Storage	213
Typical InfiniBand Designs	214
LID: InfiniBand Addressing and Frame Formats	215
InfiniBand QoS Support	216
Summary	217
Test Your Knowledge	218
References	219

10	AI Network Performance KPIs	221
	Significance of Performance Benchmarking	221
	MLCommons for AI Data Centers	223
	MLCommons Initiatives	224
	MLCommons Benchmarking Suites	224
	Benchmarking a Data Center for Machine Learning	225
	Summary	226
	Test Your Knowledge	227
	References	228
11	Monitoring and Telemetry	229
	Exploring Monitoring Options	229
	Network Monitoring in an AI/ML Data Center Network	231
	In-Band Flow Analyzer (IFA)	234
	Corrective Actions	237
	Summary	238
	Reference	238
12	Ultra Ethernet Consortium (UEC)	239
	UEC Developments and Working Groups	241
	Link Layer Working Group	242
	Transport Working Group	242
	Software Working Group	243
	Other Working Groups	243
	UEC Key Terminology	244
	The UEC and Network Architectures	246
	UEC and the Existing Ethernet	246
	A New Protocol Stack	247
	Data Plan: Packet Forwarding Options	252
	UDP-Based Encapsulation for UET	252
	UET over IP: The New Entropy Field	255
	Packet Delivery Modes	257

Reliable Ordered Delivery (ROD)	259
Reliable Unordered Delivery (RUD)	259
Reliable Unordered Delivery Idempotent (RUDI)	259
Unreliable Unordered Delivery (UUD)	260
UEC Delivery Modes Comparison	260
Congestion Management (CM) in the UEC Specification	261
Network Signal Congestion Control (NSCC)	261
Receiver Credit Congestion Control (RCCC)	262
Credit-Based Flow Control (CBFC)	263
Packet Trimming and Fast Retransmissions	264
Link Layer Reliability (LLR) Mechanism	265
In-Network Collectives (INC) and xCCL	266
Management and Orchestration	268
Interoperability and Backward Compatibility	269
Compliance and Certification	269
UEC Challenges and Future Directions	269
Comparing UEC to InfiniBand and RoCEv2	270
Summary	271
Test Your Knowledge	272
References	273
13 Scale-Up Systems	275
Key Building Blocks of Scale-Up Systems	278
Scale-Up Ethernet Transport (SUE-T)	281
Ultra Accelerator Link (UALink)	286
Memory Coherence in Scale-Up Systems	291
Scale-Up Systems: Key Differences and Similarities	292
Summary	294
Test Your Knowledge	295
References	297

14	Conclusion	299
	DC Network Role for AI	299
	Caveats and Challenges	300
	Future Developments	302
	Final Remarks	304
	References	305
	Appendix A Questions and Answers	307
	Appendix B Acronyms	329
	Index	335

Foreword

If you've picked up this book, you're likely not someone who has bought into the media hysteria around AI or who thinks that it's a danger to humanity. Those folks have watched too many *Matrix* or *Terminator* movies. Don't get me wrong, as there are reasons to be concerned about AI and reasons there should be intelligent regulations around some of it. But AI is a tool, nothing more. Like any tool it can be used or misused.

Like it or not, AI is integrated into our everyday lives, often in ways that we don't even notice. Arguably, the first widespread use of AI was in spam filters. When you buy spam detection software, you buy a *database* of patterns common to most spam email and an algorithm (*model*) that watches for characteristics in your email that match those patterns. When it thinks it recognizes (*infers*) spam, it moves the email to an isolated mailbox. You actively *train* your spam filter by periodically reviewing your spam mailbox, identifying email that is not spam, so that the app can better infer spam from new data.

There are some fundamental AI/ML concepts in what I just described: a curated reference database, a pattern-matching model, training, and inference based on new data.

Want another example? The first "smart" doorbell I purchased was little more than a camera, a button, and some motion detection capabilities. The one I have now can differentiate a person from random motion. It can tell me if that person dropped off or picked up a package. With a little training (there's that word again), it can identify friends and family members who come to the door. Through my Wi-Fi network, it connects to a database and a pattern recognition model.

And, of course, there's your smartphone with AI apps from face ID and voice commands, to music and entertainment apps that learn your preferences and make recommendations, to targeted advertisements in your social media, to the ability to "Google" (the company name has become a verb) an answer to almost any question you might have. Your new television likely has AI and maybe your new refrigerator as well. Your new car is swimming with it, giving you directions and keeping you in your lane and telling you if you're following too closely. And as I type this foreword, Microsoft Word follows what I'm typing, suggests sentence completion, and catches spelling and grammatical errors.

Then there are the AI apps that are not quietly hiding in your refrigerator or word processor: generative AI that can create images and music, write academic papers or code, or generate designs; predictive AI that can make forecasts based on analysis of historical data (note that predictive AI can make some embarrassingly bad predictions when analyzing particularly chaotic systems, such as human behavior); and natural language models that allow you to communicate naturally with AI systems and receive human-sounding answers.

And, of course, self-driving automobiles utilize a variety of AI models to operate safely on the road. I personally believe self-driving automobiles are not quite ready for prime time, but I also believe they are our future. There could well be a not-so-distant future in which manually driving a car is considered illegally dangerous.

Finally, as network nerds we've all heard the term "autonomous networks." Networks tend to be highly predictable systems, and with machine learning and analytics we are moving closer to network autonomy and further away from the one factor that accounts for most network outages: humans touching the network.

AI, in its many subgroups, can be simple enough for you to create and run programs right on your laptop. But processing massive databases can be hugely expensive in infrastructure, power, and cooling costs. Massive databases require massive, quickly accessible storage infrastructure. Processing that data within reasonable job completion times requires powerful servers running GPU/TPU/DPU clusters in parallel, sometimes requiring tens of thousands of processors running the same model. In the largest-scale cases the storage, processing backends, and inference frontends can involve not just separate types of data centers but also physically separate data centers. Reliably distributing and collecting data among these processors calls for new kinds of data center networking because storage, processing (training), and inference all have different requirements.

That's where this first-of-its-kind book comes in. My friends Mahesh, Himanshu, and Michal have among them a vast level of experience designing AI data centers. They explain in great detail the unique requirements of AI data centers and the technologies created to address these problems. Many of these technologies are already finding their way into high-performance compute (HPC) data centers and will eventually find their way into normal cloud data centers. So even if you are not an AIDC specialist, it behooves you to become familiar with the technologies and design principles covered here.

–Jeff Doyle

Preface

AI Data Center Network Design and Technologies is a comprehensive, practical guide for network engineers, architects, and technology leaders who are building, scaling, or optimizing the infrastructure that supports artificial intelligence. This book bridges the gap between theory and practice, providing an in-depth look at the design, deployment, and management of high-performance AI data centers—where scale-out and scale-up networking, xPU-based computing, storage, and operations work together.

Register your copy of *AI Data Center Network Design and Technologies* on the InformIT site for convenient access to updates and/or corrections as they become available. To start the registration process, go to informit.com/register and log in or create an account. Enter the product ISBN (9780135436288) and click Submit. Look on the Registered Products tab for an Access Bonus Content link next to this product, and follow that link to access any available bonus materials. If you would like to be notified of exclusive offers on new editions and updates, please check the box to receive email from us.

This page intentionally left blank

Acknowledgments

We take this opportunity to sincerely acknowledge those individuals who have provided insightful meets and encouragement from the very start of our careers, helping us lay a strong foundation for the advancements in networking technologies we see today:

Manish Prakash, Shivam Srivastava, Raghu Subramanian, Gouthaman Guna, Tim McCarthy, Rajesh Sunkara, Ansuha P, John Steele, Jim Lowe, Ken D, Varun Arya, Rob Nath, Jim Hurd, Mathias Kokot, Sanjeev Ugalkar, Kaushal Agrawal, Karthik Vugane, Vinod Subramaniam (Vinod), Praful Lalchandani, Mike Bushong, Praveen Jain, A.E. Natarajan, Mansour Karam, Murali Vemula, Kiran K.L., Raj Yavatkar, T. Sridhar, Jay Ramalingam, Arun Viswanathan, Jasmeet Sawhney, Yogesh Kumar, Chirag Kachalia, Rajesh Dhople, Dharshan Ravichandran, Ashwin Balaji Sundaresan, Kumuthini Ratnasingham, Francis Guiller, Jeff Doyle, Amit Sanyal, Arun Gandhi, Ben Baker, Himansu Sahu, Mallikarjun Tallapragada, Suryanarayana MNV, Suraj Kumar, Harisankar Ramalingam, Rajashekhar Reddy, Parthipan T.S., Balaji Palanisami, Rajesh Venmanad, Manoj Kulandaivel, Sureshkumar Chinnappan, Pavan Kurapati, Dilip Sundarraj, Murugan Kanniappan, Russ White, Amogh Vijayakumar, Vince Loschiavo, Gilbert Montanez, Gustavo Castellanos, Nishal Singh, Kevin Oriet, Vinay Kakkar, Praveen Shetty, Sivakumar Gnanasundram, Shankar Nagaraj, Murali Chandran, Santha Kumar, Shahid Ali Khan, Abey Rajan, and Senthilnathan Ramasubbu.

We also extend our sincere appreciation to:

Nipun Chawla, Rahul Mehta, Aroshi Handa, Rajat Setia, Shaleen Vashisht, Mike Albano, Ramesh Kandula, Manish Gupta, Tamas Mondal, Kevin Wang, Antoni Przygienda, Wen Lin, Ajay Gaonkar, Selvakumar Sivaraj, Aquin Mathai, Vrishab Sikand, Kishore Tiruveedhula, Thorbjørn Zieger, Zacharias El Banna, Lars Axelund, Elisabeth Rodrigues, Washid Lootfun, Adrien Desportes, Krzysztof Szarkowicz, Riccardo Belli, Zuhair Makawa, Krishnamachari KR, Ridha Hamidi, Grzegorz Koscian, Piotr Czechowicz, Fathi Ben Zaied, Vikram Nagarajan, Jonathan Scherman, Rene Triana, Danny He, Chris Hackett, Muzamil Khan, Jeff Haas, Andy McCarthy, Pawel Rabiej, Pawel Kocimowski, Ahmed Bilal Khan, Sławomir Karas, Soumyodeep Joarder, and Jeffrey Zhang.

This page intentionally left blank

About the Authors

Mahesh Subramaniam is a proven leader in AI data centers and next-generation networking technologies. He played a key role in defining the advanced software roadmap for AI fabrics, which are now deployed in production networks across various AI data centers worldwide. As the Senior Director of Product Management for AI Data Centers at HPE Juniper Networks, he leads cutting-edge innovations in AI infrastructure and cloud-scale solutions, optimized for both scale-up and scale-out architectures. Mahesh is also an inventor with several technology patents and a recognized speaker at global forums, including the UEC Summit, OCP, and Tokyo MPLS forum. His work has earned him accolades, including the CEO Excellence Award, the Record High Business Award, and the Star Award for the Cloud DC Reference Architecture. With a remarkable history in the networking industry, Mahesh has a strong track record of leading products and managing technical and business strategies across cross-functional teams.

Michal Styszynski is a Product Management Director in the Data Center Networks Business Unit (DC BU) at HPE Juniper Networking. Michal has been with Juniper Networks for more than 13 years. Before his current role, he was a Technical Marketing Engineer (TME) in the DC BU and a Technical Solution Consultant at Juniper. In these roles, he handled data center projects for large-scale enterprises and federal networks and worked closely with Tier 2 cloud and telco-cloud service providers. Before joining Juniper, he spent around 10 years working at Orange, FT R&D, and TPSA Polpak engineering. Michal graduated from the Electronics & Telecommunications department at Wroclaw University of Science & Technology with a master's degree in engineering. He also holds an MBA from Paris Sorbonne Business School and is a JNCIE-DC#523, as well as PEC, PLC, and PMC certified from the Product School in San Francisco.

Himanshu Tambakuwala is a highly accomplished networking expert and certified technical architect whose experience spans the entire product lifecycle—from hands-on engineering to product strategy. He is a JNCIE holder in Data Center and Service Provider technologies and an inventor with four granted technology patents and two additional patents currently filed. As a Product Manager at Juniper Networks, Himanshu was instrumental in defining the feature roadmap for network fabrics that power cutting-edge AI/ML data centers.

What You'll Find Inside

Chapter 1: Wonders in the Workload

Start your journey by examining the complete lifecycle of AI/ML workloads. This chapter explains how raw data is collected, labeled, and preprocessed before being routed into distributed training pipelines. You will understand the fundamentals of forward and backward propagation, gradient descent, and iterative optimization—and see how these processes scale across thousands of GPUs through data, pipeline, and tensor parallelism. The chapter introduces job completion time (JCT) and tail latency as key metrics and discusses how RDMA (in RoCEv2) facilitates the low-latency, high-throughput transfers needed for modern AI. The technical groundwork clarifies why lossless, high-radix, dynamically balanced fabrics are crucial.

Chapter 2: “The Common-Man View” of AI Data Center Fabrics

This chapter examines the different requirements for AI training and inference data centers. It provides a technical comparison of InfiniBand and Ethernet fabrics, emphasizing their trade-offs in latency, scalability, cost, and ecosystem support. The section discusses AI-specific traffic patterns, such as low-entropy flows, elephant bursts, and synchronization-induced congestion, and explains how advanced load-balancing techniques (static, dynamic, and global) and congestion management methods (ECN, PFC, and DCQCN) are used to maintain fabric efficiency and resilience.

Chapter 3: Network Design Considerations

In this chapter, the book becomes a practical blueprint for building scalable AI clusters. It introduces rail-optimized design and rail-unified design (aka non-rail-optimized design), explaining how each approach affects latency, scalability, and fault tolerance. You'll get hands-on guidance for rack layouts (ToR, MoR, and EoR), cable management, and power/cooling planning. The text then expands to advanced topologies—Clos, Dragonfly, and Torus—showing how to scale from a few racks to tens of thousands of GPUs and how to adapt these principles for inference-centric data centers.

Chapter 4: Optics and Cable Management

Dive into the physical layer with a technical tour of optics and cabling. This chapter covers the evolution from 10 Gbps to 400 Gbps/800 Gbps/1.6 Tbps, the role of DSPs, advanced modulation (PAM4 and QAM), and FEC. It compares transceiver types (QSFP, OSFP, and CFP), connectors (MTP/MPO and LC), and cable options (DAC, AEC, AOC, MMF, SMF, and DWDM). The chapter also explores the latest in pluggable, co-packaged, and linear-drive optics, emphasizing the importance of power and thermal management in high-density environments. We have not addressed the upcoming CPO and LRO/LPO optics technologies in this book as they are subject to ongoing architectural development and discussions with multiple vendors regarding power-efficient rack solutions.

Chapter 5: Thermal and Power Efficiency Considerations

AI clusters push power and cooling to their limits. This chapter addresses the engineering required to keep them running. You'll learn about airflow management (front-to-back, back-to-front, and bidirectional), advanced liquid cooling (immersion, cold plate, rear-door heat exchangers, and spray cooling), and the impact of high-density servers and optics on rack design. Real-world examples from hyperscalers illustrate how next-generation clusters are redefining what's possible in data center thermal management.

Chapter 6: Efficient Load Balancing

This chapter is the technical heart of the networking discussion. It starts with the limitations of traditional ECMP and hash-based load balancing and then introduces a suite of modern technologies: static load balancing, dynamic load balancing (DLB), global load balancing (GLB), flowlet-based rebalancing, per-packet spraying, and selective spraying for RDMA. You'll see how each method works, where it excels, and how to tune it for AI/ML traffic patterns, ensuring optimal utilization and minimal congestion.

Chapter 7: RoCEv2 Transport and Congestion Management

This chapter provides an in-depth technical overview of RoCEv2, the primary transport protocol for AI clusters. The chapter details every potential congestion point and explains the mechanisms—including ECN, PFC, DCQCN, SFC, and CSIG—that can be used to manage them. You'll learn how to tune these controls for different environments, how to prevent PFC storms, and how to use emerging techniques for even more precise congestion management.

Chapter 8: IP Routing for AI/ML Fabrics

Routing is the silent backbone of AI performance. This chapter explores BGP, OSPF, IS-IS, and RIFT in the context of AI data centers, detailing advanced features such as BGP unnumbered, add-path, bandwidth communities, and deterministic path forwarding. You'll learn how to design for multi-tenancy, overlay networks, and server-level routing, as well as how to integrate telemetry and controllers for adaptive, performance-aware routing.

Chapter 9: Storage Network Design and Technologies

Training is fundamentally an I/O challenge. This chapter reviews storage technologies—block, file, and object storage; NVMe-oF; parallel file systems; and GPUDirect Storage—and explains the protocols, state machines, and design patterns that enable high-performance, scalable, and resilient storage networks. The chapter provides practical advice on integrating on-premises and cloud storage, as well as on designing for both hot and cold data paths.

Chapter 10: AI Network Performance KPIs

This chapter outlines the key KPIs—throughput, latency, accuracy, power, efficiency, and scalability—and presents the MLCommons/MLPerf benchmarking suite. You'll learn how to create and analyze benchmarks for both training and inference, ensuring that your performance claims are meaningful, reproducible, and actionable.

Chapter 11: Monitoring and Telemetry

Operations is where design meets reality. This chapter covers the tools and techniques for monitoring and telemetry, from SNMP and syslog to streaming telemetry and in-band flow analyzers. You'll see how to combine real-time and historical data, leverage AI-driven analytics, and automate corrective actions to keep your fabric healthy and performant.

Chapter 12: Ultra Ethernet Consortium (UEC)

This chapter introduces the Ultra Ethernet Consortium and its new protocol stack, which is designed for million-node AI clusters. You'll learn about technical innovations in transport (UET), congestion management (NSCC, RCCC, and CBFC), and packet delivery (ROD, RUD, RUDI, and UUD) and see how UEC compares to InfiniBand and RoCEv2. The chapter offers practical guidance for transitioning to UEC-ready fabrics and highlights the challenges and opportunities ahead.

Chapter 13: Scale-Up Systems

The emergence of scale-up systems represents a significant development in AI data center infrastructure. This chapter details advanced integrated high-performance computing platforms for AI data centers, referred to as super-accelerators. These platforms employ more than eight XPU (accelerators) within a single system rack, interconnected via next-generation technologies such as UALink, which is a viable alternative to Nvidia NVLink. The chapter presents an overview of encapsulation formats and their key characteristics. It also discusses the placement of these new scale-up systems within the data center network and their integration with scale-out backend solutions, including ESUN (Ethernet for Scale-Up Networking), as proposed at a high level by the Open Compute Project Foundation (OCP).

Credits

Cover: BAIVECTOR/Shutterstock

Figures 4.2.1 Dell'Oro Group

This page intentionally left blank

4

Optics and Cable Management

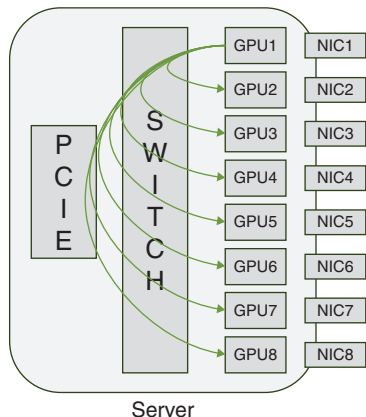
Scaling Optics for AI Clusters

In Chapter 3, “Network Design Considerations,” we discussed design concepts involved in building an AI/ML data center. Optics and cables are important components of any data center, so requirements of an AI/ML data center.

With AI/ML clusters, the ports on a server currently support 200 Gbps and 400 Gbps throughput and are moving toward 800 Gbps and 1.6 Tbps throughput. Nvidia is the top vendor for the GPU server supplies for AI/ML clusters. It has multiple generations of GPUs—Volta, Ampere, Hopper, and Blackwell. Nvidia currently has A100 (Ampere) and H100 (Hopper) on the market and is moving toward H200 and further versions of GPU based on Blackwell.

From the GPU, Nvidia uses NVLink to connect to the NVSwitch, which is used within the server to communicate between the GPUs that are internal to the server. Each NVLink is 300 Gbps for the Volta generation, 600 Gbps for the Ampere generation, and 900 Gbps for the Hopper generation, moving toward 1800 Gbps for the Blackwell generation. Other vendors have similar solutions; for example, AMD has Infinity Fabric, and Intel has CXL (Compute Express Link), UCIe (Universal Chiplet Interconnect Express), and PCIe (Peripheral Component Interconnect Express) switches.

Figure 4-1 illustrates the system topology of a server with 8 GPUs, where an internal switch helps with communication across the GPUs in the server.

**Figure 4-1***System topology*

In addition to the NVSwitch for internal communication, there are different NICs that are used for external connectivity between the GPUs across multiple servers. In either case of NVIDIA or AMD-based GPU servers for the scale-out DC use-case, each GPU is connected to a dedicated NIC card, and then each NIC card from the same server connects to different top-of-rack Ethernet or InfiniBand switches. It means for local intra-server communication, the NV switch will be used, and for any server-to-server communication, the external switch is used. GPU-to-NIC card connections are 400 Gbps or 800 Gbps Ethernet or IB. This means a top-of-rack switch is also typically a high-port-density 400Gbps/800Gbps and 1.6Tbps switch. From the AI DC server perspective, besides the GPU-connected NIC cards, there are also storage NIC cards (NVMe-o-F, for example) and out-of-band connections. They are all interconnected via PCIe generation 5 and newer to offer even higher local server interconnects to memory blocks or the CPU.

Figure 4-2 shows a chart from a Dell'Oro report on market adoption of optics for AI clusters from 2020 to 2027. This chart indicates that optics adoption is going to move toward 1.6 Tbps. With the requirement of high-bandwidth optics, there is also a need for high-radix switches that can support a large number of ports per rack unit (RU). In addition, optics need to be power-efficient to reduce the power and thermal budget of a rack. These needs are driving the enhancement of small-form-factor optics, modulation, connectors, and cables.

Figure 4-2.1 shows a 2025 study from Dell'Oro showcasing the adoption of 800 Gbps in 2025 and growing demand for 3.2 Tbps optics by 2029.

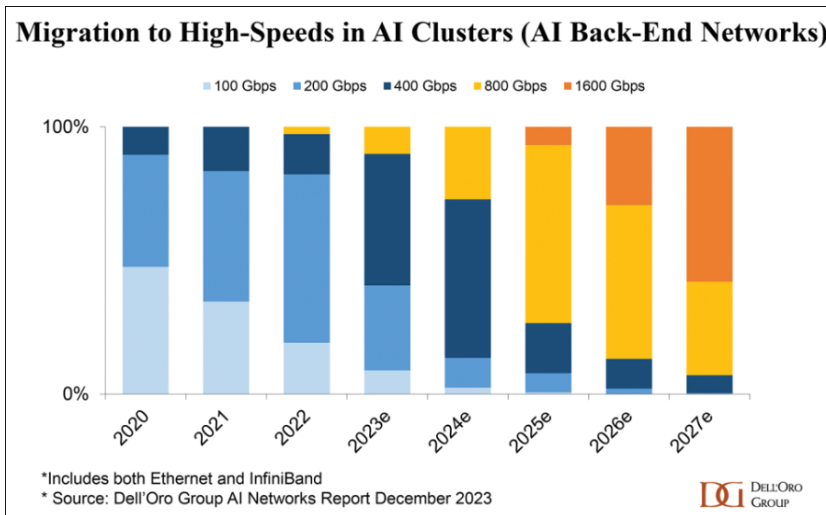


Figure 4-2

Dell'Oro chart on optics evolution

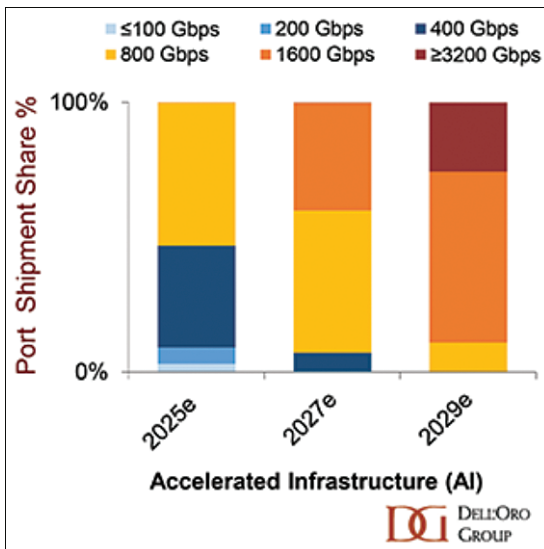


Figure 4-2.1

Dell'Oro chart report on optics evolution 2025

The throughput requirement from server to leaf is expanding to 200 Gbps/400 Gbps/800 Gbps/1.6 Tbps, which is leading vendors to focus on optics to support high-bandwidth connections.

Challenges in Optical Innovation

The optics industry has rapidly advanced from 10 Gbps to 100 Gbps and now to 800 Gbps and beyond. This growth is outpacing traditional models, driven by the need for more bandwidth and faster innovation. Although optical technology doesn't directly follow Moore's Law, it is propelled by data center demands, powerful processors, and expanding data volumes.

These are some of the challenges related to optics technology:

- **Signal quality:** Maintaining signal integrity gets harder as data rates rise. Higher speeds mean more attenuation, dispersion, and noise from crosstalk, all of which can harm signal quality. To tackle these challenges, advanced modulation schemes and error correction are required.
- **Signal conditioning:** At elevated speeds, a signal may experience more severe impairments. To preserve signal integrity, sophisticated methods like digital signal processing and equalization are required.
- **Power consumption and cooling requirements:** Higher-speed optics usually require more power, which can create significant challenges in data centers, where energy efficiency is vital. In addition, increased power usage results in high heat dissipation and requires cooling.
- **Availability and cost:** As AI/ML data centers are demanding higher-speed optics, different organizations are working toward coming up with standards for high-speed optics. Optics technology is lagging, and challenges include availability and cost of high-speed optics units.

Packet Flow

Whenever a network device receives data, it is in the form of either electrical signals or optical signals. A signal goes through multiple stages before it reaches the packet-forwarding engine. Similarly, a packet goes through multiple stages before it is transmitted out. Figure 4-3 illustrates the major components related to the optics in a switch.

In the following sections, we'll discuss packet flow for the newer 400 Gbps and onward optics.

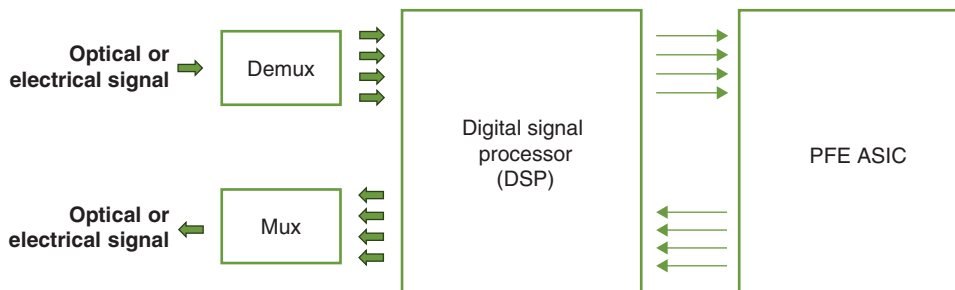


Figure 4-3

Packet flow from optics toward the PFE ASIC

Demultiplexers and Multiplexers

As shown in Figure 4-4, when a signal—either in electrical form when using copper cables or optical form when using optical fiber—reaches the pluggable optics, it goes through a demultiplexer (demux) function that splits the signals.



Figure 4-4
Demultiplexer

The PFE ASIC has multiple SerDes (serializer/deserializers), which form a high-speed interface used to convert data streams received between serial and parallel forms. As the PFE ASIC processes packets serially, the SerDes performs the job of serializing the data received from multiple SerDes links in parallel. If the PFE ASIC supports ~50 Gbps SerDes, a demux could perform the demultiplexing to 8×50 Gbps from 400 Gbps optics. If the PFE ASIC supports ~100 Gbps SerDes, a demux could perform the demultiplexing to 4×100 Gbps from 400 Gbps optics or 8×100 Gbps from 800 Gbps optics.

After the PFE ASIC processes the data, the SerDes again parallelizes the data into separate streams. Either single or multiple SerDes can be mapped to a single optics unit to achieve the required rate. This process is used for creating N:1 conversion of signals, and it is achieved with the help of multiplexing, as illustrated in Figure 4-5.



Figure 4-5
Multiplexer

With AI/ML requirements of 400 Gbps, 800 Gbps, and beyond, the SerDes links are now at 200 Gbps and moving toward higher speeds. Further improvements in the mux and demux are required to be able to achieve higher speeds.

Digital Signal Processors (DSPs)

The signals from the demultiplexer are passed to the digital signal processor (DSP), which carries out several functions. Let's investigate each of them in detail.

Modulation and Demodulation

DSPs are responsible for dealing with sophisticated modulation methods in high-speed optical communications, as illustrated in Figure 4-6. They encode digital data onto the optical carrier wave by converting it into diverse amplitude and phase states. On the receiving side, DSPs decode the optical signal by interpreting the detected changes in amplitude and phase to retrieve the original data.

Modulation is the process of converting data or information to electrical or optical signals. Modulation is required from the ASIC to the electrical signals and then later from electrical signals to optical signals. The modulation at each layer may be different. NRZ, also known as PAM-2, is a traditional modulation technique that does not support higher-bandwidth requirements. PAM-4 and above are being used for newer optics to support the higher-bandwidth requirements of AI/ML data centers.

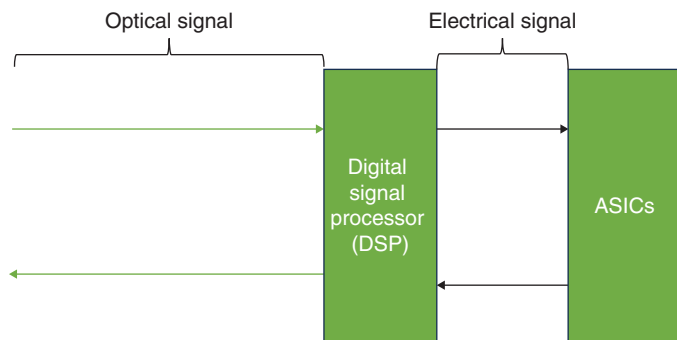


Figure 4-6

Conversion of optical to electrical signals and vice versa

The following modulations are used in different types of network connectivity:

- **NRZ:** Non-Return-to-Zero (NRZ), which used to be a widely used scheme, has two voltage levels to represent 0 and 1. It is commonly used in the 28 GHz range and for some 56 GHz channels.
- **PAM-4:** Pulse Amplitude Modulation with four levels (PAM-4) is a signal-encoding technique that uses four distinct signal levels (that is, voltage levels) to represent four combinations of two-bit logic (00, 01, 10, and 11). PAM-4 modules double the bandwidth of a connection, with each signal level representing 2 bits of logic information. PAM-4 is used for some 56 GHz channels and all 112 GHz channels.
- **Higher-order modulation:** For greater speeds, modulation schemes such as PAM-8 or QAM (quadrature amplitude modulation) are considered. These methods boost data rates by using additional signal levels or combining various amplitude and phase states.

- **DWDM (dense wavelength division multiplexing):** Coherent DWDM technology uses amplitude, phase, and polarization of light. It supports much higher bit rates on a single wavelength when DSPs are used. DWDM is used for transmitting multiple data channels over a single fiber, expanding data capacity in high-speed connections between data centers. 400G ZR modules have ushered in a new era of DWDM technology marked by open, standards-based, and pluggable DWDM optics, enabling true IP-over-DWDM. 400G ZR modules are used for connectivity between data centers (up to 80 km).

To achieve higher data transmission rates, the industry is trying out different modulation methods that make it possible to transmit more data at a time. We can compare it to adding more seats to carry a larger number of passengers on a flight, where the downside is that privacy is reduced. Similarly, transmitting more data at a time results in more noise to the signals and creates a need for more sophisticated mechanisms to deal with it.

Error Detection and Correction

DSPs are responsible for error detection and handling retransmission or correction of received packets to maintain data accuracy. They use forward error correction (FEC) algorithms to rectify errors that occur during transmission. Methods like low-density parity-check (LDPC) codes and Bose-Chaudhuri-Hocquenghem (BCH) codes are often used to enhance data integrity and reliability. At 400 Gbps and above, FEC is needed for reliability, although it introduces latency.

Clock Data Recovery

DSPs play a critical role in synchronizing the transmitter and receiver to maintain data integrity and minimize errors. A DSP extracts clock signals from the incoming data, which is vital for precise data sampling and decoding. This function becomes especially important at high data rates, where accurate timing is essential.

Equalization

DSPs use equalization to improve the signal-to-noise ratio (SNR). Equalization is a signal processing technique that restores the shape of a signal waveform in optics. Equalization algorithms include feed-forward equalization (FFE) and decision-feedback equalization (DFE).

Transmission Modes

Beyond the DSP and mux/demux, we move more toward optics. Two decades ago, most network gear, including switches and routers, relied on copper cables for data transmission. But these cables are limited in terms of how fast data can move. As data speeds picked up, these cables couldn't keep up, as they only reached about 3 meters for 100 Gbps Ethernet links. They were good for connections in the same rack, though, such as connecting servers to the top of rack (TOR) switch.

Today, we have high-performance switches, routers, and SmartNICs that can handle speeds of 200 Gbps, 400 Gbps, and even 800 Gbps. To handle these speeds, fiber-optic cables have taken over from copper ones. Instead of using electrical signals, optical fiber uses light to send data.

The advantage in fiber-optic cables is that they can go much farther—up to 80 km or even 120 km—without losing signal strength. They're also much more reliable than copper because they use total internal reflection to carry light and therefore aren't affected by electromagnetic interference. In addition, fiber-optic cables handle changes in temperature and pressure well.

The following sections detail the different cable options available for connectivity in data center fabrics.

Multi-Mode Fiber (MMF)

MMF is optical fiber that is designed for the transmission of multiple rays of light at a time (see Figure 4-7). Normally, the core diameter of MMF is 50 μm (micrometers) and 62.5 μm . It propagates 850 nm (nanometers) and 1300 nm wavelengths from low-cost light sources like LEDs or VCSELs and has more attenuation. MMF is used for short-to medium-range connections, as well as for connecting devices within the same rack or in nearby racks.

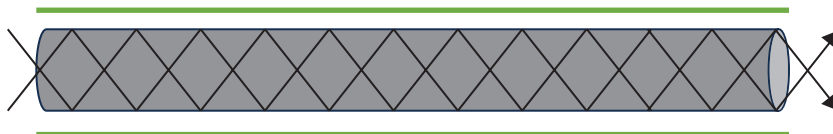


Figure 4-7

Multi-mode fiber

There are five different grades of MMF:

- **OM1:** OM1 supports 1 Gbps speeds up to a distance of 300 m and 10 Gbps speeds up to 33 m.
- **OM2:** OM2 supports 1 Gbps speeds up to a distance of 550 m and 10 Gbps speeds up to 82 m.
- **OM3:** OM3 supports 40 Gbps speeds up to a distance of 240 m and 100 Gbps to 400 Gbps speeds up to a distance of 100 m.
- **OM4:** OM4 supports speeds of 100 Gbps to 400 Gbps up to a distance of 150 m.
- **OM5:** OM5 supports speeds of 100 Gbps to 400 Gbps up to a distance of 150 m. It can accomplish 400 Gbps with just four transmit and receive cables and wavelength division multiplexing (WDM) technology. OM5 uses the wavelength 953 nm along with the wavelength of 850 (OM4).

OM1 has a glass core diameter of 62.5 μm . The other grades of MMF have a glass core diameter of 50 μm .

Single-Mode Fiber (SMF)

SMF is an optical fiber that is designed for the transmission of a single ray of light at a time, eliminating distortion from overlapping light pulses (see Figure 4-8). Normally, SMF has a core diameter of 8 μm to 10 μm , which can propagate higher wavelengths of 1310 nm and 1550 nm from the laser. It is an expensive type of cable that is used for long-distance data transmission between different buildings or data centers.



Figure 4-8
Single-Mode Fiber

Dense Wavelength Division Multiplexing (DWDM)

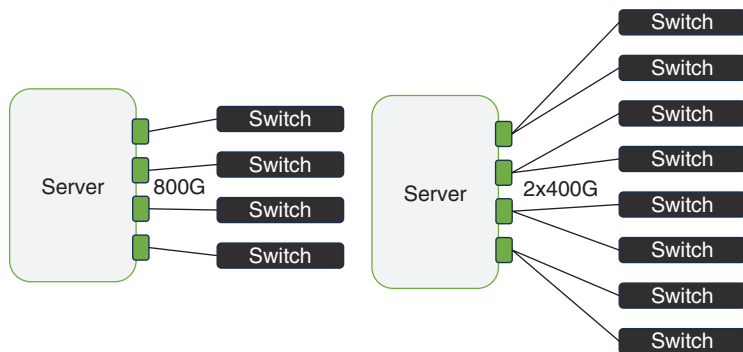
DWDM enables dense wavelength multiplexing by multiplexing multiple electrical signals to a single optical lane, as illustrated in Figure 4-9. DWDM technology helps in transmitting data over long distances, as described earlier in this chapter. This is specially required in case of data center interconnects.



Figure 4-9
DWDM

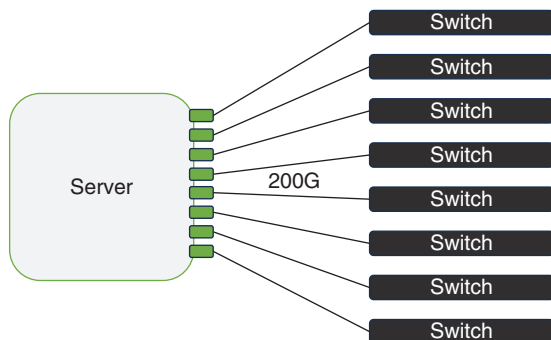
AI Server Connectivity Options

On the switch side, the connection could be either single-mode or multi-mode fiber and optics, depending on the distance. Since the server ports an OSFP transceiver with 800 Gbps, the switch side must be a 800 Gbps port or two 400 Gbps ports with the use of breakout cables. Figure 4-10 illustrates both of these connectivity options. For rail-optimized design (ROD), another option is able to connect 8 GPUs to 8 different switches.

**Figure 4-10**

Connection options with the Nvidia H100

Figure 4-11 illustrates the connectivity option with the A100 server. There are 8 ports on the server that can connect to 8 different switches in ROD.

**Figure 4-11**

Connection option with the Nvidia A100

Transceiver Types

The choice of optics is based on the connectivity distance between two nodes. In general, optics meant for smaller distances are cheaper than those meant for longer distances. Each type of optical fiber has a suffix to denote the reach and optical lanes. For example, in 400G-SR8, SR stands for *short reach*, which is up to 100 m, and 8 denotes the number of optical lanes; 400G-SR8 can have 8 optical lanes of 53 Gbps each, which are multiplexed to support 400 Gbps bandwidth.

Table 4-1 lists the transceiver types and the reach and mode of optical fiber with which it can be used.

Index

Numerics

400G ZR, 73
400G-SR8, 76

A

- A100 server, 68
- AALC (air-assisted liquid chromatography), 92
 - See also* cooling systems
- ACC (accelerator), 276–277, 279, 280. *See also* GPU/s; scale-up; xPU
- accuracy, 223
- acknowledgement packet, RoCEv2, 14–15
- ACLs (access lists), 175
- Adjacency SID, 180
- AEC (active electrical cable), 79
- AETH header, 14–15
- AFH (AI Fabric Header), 281–282
- agent-based network monitoring, 231
- AI. *See also* models
 - models, 1
 - training, 1
- AI data center/s. *See also* fabric; inference; training
 - benchmarking for ML, 225–226
 - burstiness, 23
- congestion management, 24–25, 26–27, 156
 - CSIG (Congestion Signaling), 137–139
 - DCQCN (Data Center Quantized Congestion Notification), 27–28, 134–136
 - ECN (Explicit Congestion Notification), 127–129
 - orchestration, 268
 - PFC (Priority Flow Control), 130–134
 - SFC (Source Flow Control), 136–137
 - super spine-to-spine link, 126
 - in the UEC specification, 261–265
 - UET (Ultra Ethernet Transport), 242–243
- cooling systems, 88
 - back-to-front airflow, 88–89
 - bidirectional fans, 89
 - choosing airflow direction, 89
 - cold plate liquid cooling, 91–92
 - front-to-back airflow, 88
 - immersion liquid cooling, 90
 - liquid cooling, 89–90
 - rear-door heat exchanger liquid cooling, 92–93
 - single-phase immersion cooling, 90
 - sprayed liquid cooling, 93
 - two-phase immersion cooling, 91

- elephant flows, 24–25
- Ethernet, 22. *See also* Ethernet
- flowlets, 23
- GPUs, 23
- high-bandwidth flows, 23
- inference, 3, 20, 33, 226
- Infiniband, 21. *See also* IB (InfiniBand)
- JCT (job completion time), 6–7, 24
- KPIs (key performance indicators), 223
- load balancing, 25–26. *See also* load
 - balancing
 - dynamic, 25
 - global, 25
 - IP ECMP, 24
 - selective, 26
 - static, 25
- monitoring, 229, 231
 - agent-based, 231
 - collector software, 231–232
 - corrective actions, 237–238
 - IFA (In-band Flow Analyzer), 230, 234–237
 - IPFIX, 231
 - network utilization, 232–233
 - real-time, 231, 233
 - RPC query, 230
 - sFlow, 231
 - SNMP polling, 230
 - SNMP traps, 230
 - syslog, 230
 - telemetry, 230, 232
- performance benchmarking, 222
- power consumption, 87, 301
- rack design, 45, 48–49
 - end-of-row, 47–48, 77
 - middle-of-row, 47, 77
 - thermal management, 88
 - ToR (Top of Rack), 45–47
- RDMA (remote direct memory access), 8. *See also* RDMA (remote direct memory access)
 - data transfer, 9
 - transport types, 9–10
 - WQE (work queue element), 9
- RoCEv2 (RDMA over Converged Ethernet version 2), 11. *See also* RoCEv2 (RDMA over Converged Ethernet version 2)
 - acknowledgement packet, 14–15
 - header, 11
 - low entropy, 22–23
 - messages, 11–14
 - sessions, 11
- ROD (rail-optimized design), 34–40
- storage, 32
- tail latency, 6–8
- training, 19–20, 33. *See also* training
- vendor interoperability, 300
- AI training lifecycle, 209
 - data preparation, 191
 - data storage positioning, 192
 - training and tuning phase, 192
- algorithm/s, 1, 20
 - Dijkstra, 144
 - FEC (Forward Error Correction), 73
 - job, 6
 - link quality band, 103
- all-gather/all-reduce, 4–5
- AMD, Infinity Fabric, 67
- AOC (active optical cable), 79
- API, 243
 - LibFabric, 241
 - Verbs, 9
- architecture, 76–77. *See also* scale-out; scale-up
 - inference DC, 56–60
 - rack design, 48–49
 - end-of-row, 47–48
 - middle-of-row, 47
 - ToR (Top of Rack), 45–47
 - scheduled fabric, 49–50
 - training DC, 33
 - GPU-to-leaf connectivity, 33–34
 - ROD (rail-optimized design), 34–40, 300–301
 - RUD (rail-unified design), 42
 - servers, 33–34
- ASIC, 98, 99, 101, 104–105
- INC agents, 268

PFE, 70

 SerDes (serializer/deserializer), 71

ASNs (autonomous system numbers), 145–146,
149–151

assigned flow mode, DLB (dynamic load
balancing), 104

AXI interface, SUE-T, 285

B

backend network, storage, 192

back-to-front airflow, 88–89

backward propagation, 3

bandwidth, 280

 data center, 68

 DLB, 99, 103–104

 ECMP, 97

 optics, 68–69

 UEC scaling, 260

BCH (Bose-Chaudhuri-Hocquenghem) codes, 73

benchmarking, 221–222

 for ML, 225–226

 suites, MLCommons, 224–225

BGP, 111, 143–144, 145

 ADD-PATH, 151–152

 ASNs (autonomous system numbers),
 145–146, 149–151

 DPF (Deterministic Path Forwarding),
 155–162

 Link Bandwidth Extended Community,
 153–155

 minimum peers per IP prefix, 155

 NNHN (next-next-hop nodes), 108–110

 AS-PATH strip and replace, 152–153

 session establishment, 146–148

 unnumbered implementation, 145–149

bidirectional fans, 89

block, 38–39

block storage, 198, 199–201

BMSaaS (Bare Metal Server as a Service), 171

on-box agent, 162

brick, 38–39

BTH header, InfiniBand, 11

burstiness, 23

C

cabling

 fiber-optic. *See also* optics

 AOC (active optical cable), 79

 LC (Lucent Connector), 78

 MMF (multi-mode fiber), 74–75

 MPO, 78

 SMF (single-mode fiber), 75

 transceiver types, 76–77

 multiplanar architecture, 60

cache coherence, 291

CBFC (Credit-Based Flow Control), 242, 263–264

CCL (collective communication library), 3–4, 243,
246, 266–268

central controller, 162

certification, Ultra Ethernet Specification v1.0,
269

CFP2, 80

CFP4, 80, 81

clock data recovery, 73

Clos topology, 50, 52, 56. *See also* topology

clusters, 240. *See also* GPUs; server/s
 combining server- and network-level
 multi-tenancy, 175

 GPUaaS, 171

 GPUs, 36–40

 SR (Segment Routing), 183–184

 training, 33

CNP (congestion notification packet), 128

cold plate liquid cooling, 91–92

collapsed storage design, 196–197

compliance, Ultra Ethernet Specification v1.0,
269

Compliance working group, 243–244

congestion

 leaf-to-server link, 125

 leaf-to-spine link, 124

 local leaf link, 123–124

 spine-to-leaf link, 124–125

 spine-to-super spine link, 126

congestion management, 24–25, 26–27, 156

 CSIG (Congestion Signaling), 137–139

 DCQCN (Data Center Quantized Congestion
 Notification), 27–28, 134–136

- ECN (Explicit Congestion Notification), 127
 - CNPs, 128
 - data transfer, 129
 - ECN bits, 127–128
 - ECN threshold, 129
 - lossy queue, 129
 - message flow, 128
 - packet flow between servers, 128
 - orchestration, 268
 - PFC (Priority Flow Control)
 - frames, 130
 - packet flow, 131–132
 - pause frames, 130
 - storm, 133
 - watchdog, 133–134
 - XOFF/XON thresholds, 130–132
 - SFC (Source Flow Control), 136–137
 - super spine-to-spine link, 126
 - in the UEC specification, 261
 - CBFC (Credit-Based Flow Control), 263–264
 - NSCC (Network Signal Congestion Control), 261–262
 - packet trimming, 264–265
 - RCCC (Receiver Credit Congestion Control), 262
 - UET (Ultra Ethernet Transport), 242–243
 - control plane, SR (Segment Routing), 180
 - control verbs, 292
 - controller-based fabric, 178–179
 - controller-based load balancing, 179–180
 - cooling systems, 88
 - back-to-front airflow, 88–89
 - bidirectional fans, 89
 - choosing airflow direction, 89
 - front-to-back airflow, 88
 - liquid cooling, 89–90
 - cold plate liquid cooling, 91–92
 - immersion liquid cooling, 90
 - rear-door heat exchanger liquid cooling, 92–93
 - single-phase immersion cooling, 90
 - sprayed liquid cooling, 93
 - two-phase immersion cooling, 91
 - co-packaged optics, 82–83
 - CPAK, 80
 - CPO (co-packaged optics), 82
 - CPU, cold plate liquid cooling, 91–92
 - CRC (cyclic redundancy checking), 253–254
 - CSIG (Congestion Signaling), 137–139
 - CtIOS, 248
 - CUDA (Compute Unified Device Architecture), 213, 304
 - CXL (Compute Express Link), 67, 275
- ## D
- DAC (Direct Attach Copper), 78–79
 - data centers. *See* AI data centers
 - data gathering, 2
 - data link layer, UALink, 287
 - data parallelism, 4–5
 - data plane, SR (Segment Routing), 180
 - data preparation, 191
 - data tagging, 2
 - DCQCN (Data Center Quantized Congestion Notification), 27–28, 40, 108, 134–136, 240
 - DD-QSFP. *See* QSFP-DD (QSFP Double Density)
 - Dell’Oro report, 68–69
 - demultiplexers, 71
 - Dijkstra algorithm, 144
 - direct-to-chip liquid cooling, 91. *See also* liquid cooling
 - DLB (dynamic load balancing), 25, 99, 103–104, 156
 - assigned flow mode, 104
 - flowlet mode, 104–106
 - per-packet mode, 116
 - DPF (Deterministic Path Forwarding). *See* BGP, DPF (Deterministic Path Forwarding)
 - Dragonfly, 50–52, 164–167, 246–247
 - DSP (Digital Signal Processor), 60, 71–72
 - clock data recovery, 73
 - equalization, 73
 - error detection and correction, 73
 - FEC (Forward Error Correction), 73
 - modulation, 72–73

DWDM (dense wavelength division multiplexing), 73, 75
 dynamic IP routing protocols. *See* IP routing protocols

E

eBGP underlay, 145
 ECMP (equal cost multipathing), 97–99, 102, 153–155. *See also* load balancing
 ECN (Explicit Congestion Notification), 24, 27–28, 127. *See also* DCQCN(Data Center Quantized Congestion Notification)
 CNPs, 128
 data transfer, 129
 ECN bits, 127–128
 ECN threshold, 129
 lossy queue, 129
 message flow, 128
 packet flow between servers, 128
 EDR (Enhanced Data Rate), 15
 efficiency, 223
 elephant flows, 24–25, 106
 encapsulation, 252
 UDP-based, 252–255
 UET over IP, 255–256
 end-of-row design, 47–48, 77
 entanglement, 304
 equalization, 73
 error detection and correction, 73
 ESUN (Ethernet Scale-Up Networking), 276
 Ethernet. *See also* RoCEv2 (RDMA over Converged Ethernet version 2); SUE-T (Scale-Up Ethernet Transport)
 advantages in AI training clusters, 22
 comparison with InfiniBand, 15–16
 and UEC, 246–247
 extension header, SRv6, 182–183

F

FA (fabric address), 244–245
 fabric, 32, 301. *See also* Ethernet; IB (InfiniBand); IP routing protocols; NVMe-oF (NVME

overFabrics); scale-up; UEC (Ultra Ethernet Consortium)
 burstiness, 23
 controller-based, 178–179
 eBGP underlay, 145
 extending IP routing to the server, 177–178
 IGPs (interior gateway protocols), 162–163
 IP Clos, 143
 IS-IS, 168–171
 RIFT (Routing in Fat Trees), 163–168
 JCT (job completion time), 6–7
 scale-up, 276
 UEC IP, 246
 underlay, 156
 fans
 back-to-front airflow, 88–89
 bidirectional, 89
 choosing airflow direction, 89
 front-to-back airflow, 88
 FBF (filter-based forwarding), 161
 FCoIP (Fibre Channel over IP), 194
 FEC (Forward Error Correction), 73
 FEP (Fabric Endpoint), 244–245, 249, 250–252
 FFE (feed-forward equalization), 73
 fiber-optic cables. *See* optics
 FIFO interface, SUE-T, 285
 file storage, 198
 file systems, high-performance, 208–209
 VAST, 211
 WekaFS, 209–211
 Flex Algo, 169–171
 flit packing, 286–289
 flood reduction, IS-IS, 168–169
 flow pinning, 26
 flow-based load balancing, 99–100
 DLB (dynamic load balancing), 99, 103–104, 156
 assigned flow mode, 104
 DLB flowlet mode, 104–106
 GLB (global load balancing), 99, 106–108
 NNHN (next-next-hop nodes), 108–110
 SLB (static load balancing), 99, 100–102
 TELB (traffic engineering load balancing), 99, 110–111

flowlet mode, DLB (dynamic load balancing),
104–106
flowlets, 23
forward propagation, 3
four-pod scale-up design, 278–279
frames, PFC, 130
front-to-back airflow, 88

G

GBP (group-based policy), 177
generative AI, 1
GLB (global load balancing), 25, 99, 106–110,
156, 300–301
GPT model, 302
GPUDirect Storage, 211–213
GPU/s, 23, 97–98, 104, 111, 280
cluster, 35–37
cold plate liquid cooling, 91
job scheduler, 179
-to-leaf connectivity, 33–34
Multi-Instance, 173
multiplanar architecture, 57
NICs, 33, 68
Nvidia, 67
NVLink, 67
optical interconnects, 304
parallel processing, 3–4
rails, 34
-server, 24
vendors, 302
gradients, 3

H

H100 server, 68
hardware coherence, 291
hash, 99
hash key, 101–102
HBM (high-bandwidth memory), 279
HCA (host channel adapter), 22
HDD
network-connected, 190
PCIe-connected, 190–191

HDR (Hundred Gigabit Data Rate), 15
header
IFA (In-band Flow Analyzer), 235–237
RoCEv2 (RDMA over Converged Ethernet
version 2), 11
high-bandwidth flows, 23
higher-order modulation, 72
high-performance file systems
VAST, 211
WekaFS, 209–211
historical network monitoring, 231
home agents, 291
HPC (high-performance computing)
systems, 19, 241
hybrid fabric storage, 196–197

I

IB (InfiniBand), 9, 15–16, 21
advantages in AI training clusters, 21
AETH header, 14–15
BTH header, 11
comparison with UEC, 270–271
HCA (host channel adapter), 22
for storage, 213–214
LID (local identifier), 215–216
typical designs, 214–215
transport services, 21
idempotent, 259–260
IEEE 802.3bm, 79
IEEE 802.3bs, 79
IFA (In-band Flow Analyzer), 230, 234–235, 237
header, 235–237
latency, 237
IGPs (interior gateway protocols), 143, 162–168.
See also IP routing protocols
IM (INC Manager), 246
immersion liquid cooling, 90
single-phase, 90
two-phase, 91
inactivity timer, DLB (dynamic load balancing),
104–105
INC (In-Network Collectives), 243, 246,
266–268

inference, 3, 20, 33, 226
 DC architecture, 56–60
 fabric, 32
 multi-node, 56
 quality of data, 31
 requirements, 20

Intel, CXL (Compute Express Link), 67

interconnect points, 281–282

interoperability and backward compatibility,
 Ultra Ethernet Specification v1.0, 269

inter-rail communication, 36, 40

intra-rail communication, 36

IP ECMP load balancing, 24

IP routing protocols, 143, 144
 BGP, 143–144, 145
 ADD-PATH, 151–152
 ASNs (autonomous system numbers),
 145–146, 149–151
 DPF (Deterministic Path Forwarding),
 155–161
 Link Bandwidth Extended Community,
 153–155
 minimum peers per IP prefix, 155
 AS-PATH strip and replace, 152–153
 session establishment, 146–148
 traffic-to-fabric color mapping, 161–162
 unnumbered implementation, 145–149
 IS-IS, 144
 Flex Algo, 169–171
 flood reduction, 168–169
 OSPF, 144

IPFIX, 231

IS-IS, 143
 Flex Algo, 169–171
 flood reduction, 168–169
 TLV (Type-Length-Value), 170–171

iteration, training, 4

ITU-T G.959.1, 79

iWARP (Internet Wide Area RDMA Protocol), 10

J-K

JCT (job completion time), 6–8, 24, 222

job, 6

job scheduler, 179

JobID, 245–246, 254

KPIs (key performance indicators), 221
 AI data center, 223
 model, 222

L

labeled data, 19

latency, 223
 IFA (In-band Flow Analyzer), 237
 NVLink, 275
 PCIe, 191, 275
 residence time, 237
 RUD (rail-unified design), 42
 storage network, 190
 tail, 120

LC (Lucent Connector), 78

LDPC (low-density parity-check), 73

leaf switch/node, 34–35, 146, 149, 151, 152,
 153–155, 161, 170, 173
 link congestion, 124
 top-of-rack, 45–47

leaf-to-server link congestion, 125

leaf-to-spine link congestion, 124

LibFabric, 241, 243

LID (local identifier), 215–216

life cycle, AI model, 2–3

Link Layer Working Group, 242

link-state protocols
 IS-IS, 144, 168–169
 OSFP, 144

liquid cooling, 89–90
 cold plate, 91–92
 immersion, 90
 single-phase, 90
 two-phase, 91
 rear-door heat exchanger, 92–93
 sprayed, 93

LLMs (large language models), 1, 189

LLR (Link-Layer Reliability), 242, 265–266

load balancing, 25, 97, 98–99
 controller-based, 179–180
 dynamic, 25

- flow-based, 99–100
 - DLB (dynamic load balancing), 99, 103–106, 156
 - GLB (global load balancing), 99, 106–108
 - SLB (static load balancing), 99, 100–102
 - TE-LB (traffic engineering load balancing), 99
 - TELB (traffic engineering load balancing), 110–111
 - global, 25, 300–301
 - hash key, 101–102
 - IP ECMP, 24
 - mechanism comparison, 117–118
 - per-packet, 115–116
 - DLB, 116
 - random spray, 116
 - resilient hashing, 101–102
 - selective, 26
 - selective packet spraying, 116–117
 - static, 25
 - LoadGen, 225
 - local leaf link congestion, 123–124
 - logical separation, storage network, 195–196
 - lossless queue, 130
 - lossy queue, 129
 - low entropy, RoCEv2 (RDMA over Converged Ethernet version 2), 22–23
 - LPO (Linear Pluggable Optics), 60, 82–83
 - LRO (Linear Receive Optics), 60, 82–83
- ## M
-
- MAC-VRF/RT5-VRF, 162
 - Management working group, 243–244
 - MCP (Model Context Protocol), 3
 - memory coherence, 291–292
 - micro-segment SRv6, 183
 - middle-of-row design, 47, 77
 - MIG (Multi-Instance GPU), 173
 - MLCommons, 223
 - benchmarking suites, 224–225
 - initiatives, 224
 - objectives, 224
 - MLCube, 224
 - MLPerf, 224
 - MMF (multi-mode fiber), 74–75
 - model/s, 1, 33, 97. *See also* algorithms; training
 - deployment, 31
 - GPT, 302
 - inference, 3
 - KPIs (key performance indicators), 222
 - life cycle, 2–3
 - monitoring, 3
 - optimization, 302–303
 - parallelism, 5–6
 - parameters, 302–303
 - selection, 31
 - training, 3
 - backward propagation, 3
 - collective communication, 4
 - data parallelism, 4–5
 - forward propagation, 3
 - gradients, 3
 - iteration, 4
 - JCT (job completion time), 6–7
 - model parallelism, 5–6
 - parallelism, 4
 - modulation, 72–73
 - higher-order, 72
 - NRZ (Non-Return-to-Zero), 81
 - PAM4, 72, 81
 - quadrature amplitude, 72
 - MoE (Mixture of Experts), 277–278
 - monitoring, 229, 231
 - agent-based, 231
 - collector software, 231–232
 - corrective actions, 237–238
 - historical network, 231
 - IFA (In-band Flow Analyzer), 230, 234–235, 237
 - header, 235–237
 - latency, 237
 - nodes, 234
 - IPFIX, 231
 - real-time, 231, 233
 - RPC query, 230
 - sFlow, 231
 - SNMP polling, 230

- SNMP traps, 230
- syslog, 230
- telemetry, 230, 232
- MPLS (Multiprotocol Label Switching), 180
- MPO, 78
- multi-node inference, 56
- multiplanar architecture, 56–60, 281
- multiplexors, 71
- multi-tenancy
 - combining server- and network-level, 175
 - network-level, 171–173
 - server-level, 173

N

- NCCL (NVIDIA Collective Communication Library), 10, 266–267. *See also* RCCL (ROCm Communication Collectives Library)
- NDR (Next Data Rate), 15
- network-level multi-tenancy, 171–173, 175
- NICs, 33, 68, 240
 - in multiplanar architecture, 57
 - storage network, 189–190
- NNHN (next-next-hop nodes), 107, 108–110
- Node SID, 180
- NRZ (Non-Return-to-Zero), 72, 81
- NSSC (Network Signal Congestion Control), 242, 247, 261–262
- Nvidia
 - A100 server, 68
 - B300, 280
 - CUDA (Compute Unified Device Architecture), 304
 - GPUDirect Storage, 211–213
 - GPUs, 67
 - H100 server, 68
 - H100/H200 servers, 155
 - NVSwitch, 42
- NVLink, 67, 275, 292–293
- NVMe-oF (NVME over Fabrics), 190, 191, 199–201
- NVMe-o-FC (NVMe over Fibre Channel), 200
- NVMe-o-RDMA (NVMe over RDMA), 192, 200

- NVMe-o-RDMA/RoCEv2 state machine, 206–208
- NVMe-o-TCP (NVME over TCP), 123, 194, 200–201
- NVMe-o-TCP state machine, 201–206
- NVSwitch, 67

O

- OAM (Ops/Admin/Maintenance), 107
- object storage, 198
- off-premises storage, 196–197
- OIF-OSPF-1.0, 79
- optics, 73–74
 - 400G-SR8, 76
 - AOC (active optical cable), 79
 - bandwidth, 68–69
 - CFP4, 81
 - co-packaged, 82–83
 - DWDM (dense wavelength division multiplexing), 75
 - industry challenges, 70
 - LC (Lucent Connector), 78
 - Linear Receive, 82–83
 - MMF (multi-mode fiber), 74–75
 - MPO, 78
 - OSFP, 81–82
 - QSFP (Quad Small Form-factor Pluggable), 80–82
 - SMF (single-mode fiber), 75
 - transceiver types, 76–77
- orchestration, 268
- OSFP (Octal Small-Form Factor Pluggable), 79, 81–82
- OSPF (Open Shortest Path First), 144

P

- packet delivery modes (UEC), 257–259, 260–261
 - ROD (reliable ordered delivery), 259
 - RUD (reliable unordered delivery), 259
 - RUDI (reliable unordered delivery idempotent), 259–260
 - UUD (unreliable unordered delivery), 260

packet flow

- demultiplexers and multiplexers, 71
- DSP (Digital Signal Processor), 71–72
 - clock data recovery, 73
 - equalization, 73
 - error detection and correction, 73
 - modulation, 72–73

packet spraying, selective, 116–117

packet trimming, 264–265

PAIR (People+AI Research), 224

PAM4, 72, 81

parallelism, 3–4

- data, 4–5
- model, 5–6
- pipeline, 5
- tensor, 5–6

parameters

- model, 302–303
- training, 2

path attributes, ADD-PATH, 151–152

AS-PATH strip and replace, 152–153

path-vector protocol, BGP, 145

pause frames, 130

PBR (policy-based routing), 111

PCIe (Peripheral Component Interconnect Express), 67, 189

- connected HDD/SDD, 190–191
- latency, 275

PDC (Packet Delivery Context), 245

PDS (Packet Delivery Sublayer), 245, 248, 250–251

per-flow load balancing, 99–100

performance

- benchmarking, 221–222
- JCT (job completion time), 222
- KPIs (key performance indicators), 221
- network, monitoring, 232–233
- tail latency, 6–8

Performance and Debug working group, 243–244

per-packet load balancing, 115–116

- DLB (dynamic load balancing), 116
- random spray, 116

PFC (Priority Flow Control), 24, 27–28. *See also*

- DCQCN (Data Center Quantized Congestion Notification)
 - frames, 130
 - lossless queue, 130
 - packet flow, 131–132
 - pause frames, 130
 - storm, 133
 - watchdog, 133–134
 - XOFF/XON thresholds, 130–132

PFE (packet forwarding engine), 99

PFE ASIC, 70, 71

physical layer (PHY), 247–248

Physical Layer working group, 243–244

physical separation, storage network, 194

PIDonFEP, 245–246, 254

pipeline parallelism, 5

ports, server, 34

POSIX (Portable Operating System Interface), 209–210

power consumption, data centers, 87, 301

preprocessing, workload, 31

PRI (Packet Rate Improvement), 242

profiles, UEC (Ultra Ethernet Consortium), 241

protocol stack, UEC, 247–248

- PDS (Packet Delivery Sublayer), 250–251
- physical layer (PHY), 247–248
- session establishment, 248–250

PSN (packet sequence number), 11, 258–259, 260, 265

PUE (power usage effectiveness), 88

PyTorch, 210–211

Q

QAM (quadrature amplitude modulation), 72

Qiskit SDK, 304

QML (quantum machine learning), 304

QoS, 216

QP (queue pair), 11

QSFP (Quad Small Form-factor Pluggable), 80–82

QSFP28, 79

QSFP56, 79, 80–81, 87
 QSFP-DD (QSPF Double Density), 79, 80–82
 quantum computing, 303–304

R

rack design, 48–49. *See also* cooling systems
 end-of-row, 47–48, 77
 middle-of-row, 47, 77
 scale-up, 276
 thermal management, 88
 ToR (Top of Rack), 45–47
 RAG (retrieval-augmented generation), 3, 189
 rail-only design, 33–34
 rails, 34
 random spray per-packet load balancing, 116
 RCCC (Receiver Credit Congestion Control),
 242, 262
 RCCL (ROCM Communication Collectives
 Library), 10, 247
 RDMA (remote direct memory access), 8, 301
 data transfer, 9
 message exchange, 11–14
 NVMe over, 200
 over Converged Ethernet (RoCE), 9–16.
 See also RoCEv2 (RDMA over Converged
 Ethernet version 2)
 PSN (packet sequence number), 11
 QP (queue pair), 11
 transport types, 9–10
 Verbs API, 9
 WQE (work queue element), 9
 real-time network monitoring, 231, 233
 rear-door heat exchanger liquid cooling, 92–93
 redundancy, storage network, 193–194
 resilient hashing, 101–102
 Resource Index, 245–246, 254
 RH (Reliability Header), 282–283
 RIFT (Routing in Fat Trees), 143, 163–168
 RoCEv2 (RDMA over Converged Ethernet version
 2), 11, 22, 99, 104, 117, 123, 155, 194, 240,
 247, 301
 acknowledgement packet, 14–15
 comparison with UEC, 270–271

DCQCN (Data Center Quantized Congestion
 Notification), 40
 header, 11
 messages, 11–14
 sessions, 11
 ROD (rail-optimized design), 33–40, 300–301
 BGP DPF, 156–159
 block, 38–39
 GPU scaling, 38–40
 inter-rail communication, 36, 40
 leaf switch, 34–35
 spine switch, 36–37
 super spine switch, 40
 ROD (reliable ordered delivery), 245, 259
 RPC query, 230
 RUD (rail-unified design), 33–34, 42, 300–301
 RUD (reliable unordered delivery), 245, 259
 RUDI (reliable unordered delivery idempotent),
 245, 259–260

S

scalability, 223
 scale-out, 275, 276
 scale-up, 1, 275, 293. *See also* SUE-T (Scale-Up
 Ethernet Transport)
 chassis, 278
 fabric interconnect points, 276
 four-pod design, 278–279
 memory coherence, 291–292
 multi-planar design, 281
 physical constraints, 276
 rack systems, 276, 280
 server, 276–277
 SUE-T (Scale-Up Ethernet Transport), 276,
 280, 292–293
 AFH (AI Fabric Header), 281–282
 AXI interface, 285
 on-die memory, 285
 end-to-end protocol stack, 283–284
 FIFO interface, 285
 interconnect points, 281–282
 RH (Reliability Header), 282–283
 shared memory model, 284

- systems comparison, 292–294
- UALink, 286
 - control verbs, 292
 - data link layer, 287
 - flit packing, 286–289
 - PGAS (Partitioned Global Address Space), 292
 - physical layer, 286–287
 - shared memory semantics, 286–287
 - transaction layer, 287, 288, 292
 - UPLI (UALink Protocol Level Interface), 289–290
 - virtual pod partitioning, 289
 - xPU, 286
- use cases, 277–278
- xPU, 276–277, 279
 - ACC (accelerator), 280
 - bandwidth, 280
 - HBM, 279
- scaling, GPU cluster, 36–40
- “Scaling Distributed Machine Learning with Gigabit Ethernet”, 68
- scheduled fabric, 49–50
- sDLB (selective load balancing), 26
- sectors, 288
- selective packet spraying, 116–117
- semi-supervised training, 2
- SerDes (serializer/deserializer), 71, 286–287
- server/s, 299
 - A100, 68
 - clusters, 239
 - connectivity options, 75–76
 - extending IP routing to the, 177–178
 - H100, 68
 - internal switch, 34
 - level multi-tenancy, 173
 - multi-tenancy, 175
 - combining server- and network-level, 175
 - network-level, 171–173
 - server-level, 173
 - PCIe connection, 190
 - ports, 34
 - quantum, 304
 - scale-up, 276–277
 - topology, 67–68
 - training DC, 33–34
- SES (Semantic Sublayer), 242–243, 245, 248, 250–251
- session establishment, UEC, 248–250
- SFC (Source Flow Control), 136–137
- sFlow, 231
- SID (SR identifier), 180
- single-phase immersion cooling, 90
- SLB (static load balancing), 25, 99, 100–102
- SMF (single-mode fiber), 75
- SNMP polling, 230
- SNMP traps, 230
- SNR (signal-to-noise ratio), 73
- software coherence, 291–292
- Software working group, 243
- Source PFC. *See* SFC (Source Flow Control)
- spine switch/node, 36–37, 146, 149, 152, 153–155, 170, 173
- spine-to-leaf link congestion, 124–125
- spine-to-super spine link congestion, 126
- sprayed liquid cooling, 93
- SR (Segment Routing), 180
 - for AI/ML clusters, 183–184
 - control plane, 180
 - data plane, 180
 - SID (SR identifier), 180
 - SR domain, 180
- SRv6
 - based traffic engineering, 182–183
 - compressed SID, 183
 - extension header, 182–183
 - micro-segment, 183
- SSD (solid-state drive), 189
 - network-connected, 190–191
 - PCIe-connected, 190–191
- state machine
 - NVMe-o-RDMA/RoCEv2, 206–208
 - NVMe-o-TCP, 201–206
- storage, 189
 - block, 198
 - file, 198
 - GPUDirect Storage, 211–213

- object, 198
 - off-premises, 196–197
 - storage data center, 32
 - storage networks, 217
 - AI training server connection, 192–193
 - backend, 192
 - IB (InfiniBand), 213–214
 - LID (local identifier), 215–216
 - typical designs, 214–215
 - VLs (virtual lanes), 216
 - latency, 190
 - NICs, 189–190
 - NVMe-oF for block-level access, 199–201
 - speed, 192
 - SSD/HDD storage, 189–191
 - training
 - collapsed storage design, 196–197
 - hybrid fabric storage, 196–197
 - inter-site redundancy, 194
 - logical separation, 195–196
 - node- and link-level redundancy, 193
 - physical separation, 194
 - Storage working group, 243–244
 - structured data, 191
 - SUE-T (Scale-Up Ethernet Transport), 276, 280, 292–293
 - AFH (AI Fabric Header), 281–282
 - AXI interface, 285
 - on-die memory, 285
 - end-to-end protocol stack, 283–284
 - FIFO interface, 285
 - interconnect points, 281–282
 - RH (Reliability Header), 282–283
 - shared memory model, 284
 - super spine switch, 40
 - super spine-to-spine link congestion, 126
 - supervised training, 2
 - switches
 - high-radix, 68
 - internal, 34
 - leaf, 34
 - PFE (packet forwarding engine), 99
 - rack design, 48–49
 - end-of-row, 47–48
 - middle-of-row, 47
 - ToR (top-of-rack), 45–47
 - server connectivity options, 75–76
 - spine, 36–37
 - super spine, 40
 - VOQ (Virtual Output Queues), 50
 - syslog, 230
-
- ## T
-
- tagging, 2
 - tail latency, 6–8, 120
 - TCP, 126
 - TE-LB (traffic engineering load balancing), 99
 - TELB (traffic engineering load balancing), 110–111
 - telemetry, 230
 - tensor parallelism, 5–6
 - thermal management, rack systems, 88. *See also* cooling systems
 - thread execution, 6
 - throughput, 223
 - TLV (Type-Length-Value), 170–171
 - topologies, 50
 - Clos, 50, 52, 56
 - Dragonfly, 50–52
 - middle-of-row, 47
 - Torus, 53–55
 - topology. *See also* architecture; rack design
 - rail-optimized, 34–40
 - ROD (rail-optimized design)
 - block, 38–39
 - GPU scaling, 38–40
 - inter-rail communication, 36, 40
 - spine switches, 36–37
 - super spine switch, 40
 - RUD (rail-unified design), 42
 - server, 67–68
 - ToR (Top of Rack) design, 45–47
 - Torus, 53–55, 246–247
 - traffic engineering, 110–111, 179, 182–183. *See also* TELB (traffic engineering load balancing)
 - traffic-to-fabric color mapping, 161–162
 - training, 2, 31. *See also* AI training lifecycle

- architecture, 33. *See also* architecture; topology
 - GPU-to-leaf connectivity, 33–34
 - ROD (rail-optimized design), 34–40, 300–301
 - RUD (rail-unified design), 42
 - servers, 33–34
 - backward propagation, 3
 - clusters, 33, 240
 - data centers, 19–20
 - distributed, 226
 - fabric, 32
 - forward propagation, 3
 - gradients, 3
 - iteration, 4
 - JCT (job completion time), 6–7
 - labeled data, 19
 - parallelism, 3–4
 - data, 4–5
 - model, 5–6
 - semi-supervised, 2
 - storage, 189
 - storage network
 - collapsed storage design, 196–197
 - hybrid fabric storage, 196–197
 - inter-site redundancy, 194
 - logical separation, 195–196
 - node- and link-level redundancy, 193
 - physical separation, 194
 - supervised, 2
 - time requirements, 226, 302
 - unsupervised, 3
 - Transport working group, 242–243, 257–258
 - TSS (UET Transport Security Sublayer), 254, 256
 - two-phase immersion cooling, 91
- U**
-
- UALink, 276, 280, 285, 286, 288
 - control verbs, 292
 - data link layer, 287
 - flit packing, 286–289
 - PGAS (Partitioned Global Address Space), 292
 - physical layer, 286–287
 - shared memory semantics, 286–287
 - transaction layer, 287, 288, 292
 - UPLI (UALink Protocol Level Interface), 289–290
 - virtual pod partitioning, 289
 - xPU, 286
 - UCle (Universal Chiplet Interconnect Express), 67
 - UDP, 123
 - UDP-based encapsulation, 252–255
 - UEC (Ultra Ethernet Consortium)
 - challenges and future directions, 269–270
 - comparison with InfiniBand and RoCEv2, 270–271
 - Compliance WG, 243–244
 - congestion management, 261
 - CBFC (Credit-Based Flow Control), 263–264
 - NSSC (Network Signal Congestion Control), 261–262
 - packet trimming, 264–265
 - RCCC (Receiver Credit Congestion Control), 262
 - existing Ethernet and, 246–247
 - INC (In-Network Collectives), 266–268
 - interoperability and backward compatibility, 269
 - Link Layer WG, 242
 - LLR (Link-Layer Reliability), 265–266
 - Management WG, 243–244
 - motivations for redesigning AI/ML fabrics, 240–241
 - network architectures, 246
 - orchestration, 268
 - packet delivery modes, 257–259, 260–261
 - ROD (reliable ordered delivery), 259
 - RUD (reliable unordered delivery), 259
 - RUDI (reliable unordered delivery idempotent), 259–260
 - UUD (unreliable unordered delivery), 260
 - Performance and Debug WG, 243–244
 - Physical Layer WG, 243–244
 - profiles, 241
 - protocol stack, 247–248

PDS (Packet Delivery Sublayer),
 250–251
 physical layer (PHY), 247–248
 session establishment, 248–250
 Software WG, 243
 Storage WG, 243–244
 Transport WG, 242–243, 257–258
 working groups, 241

UET (Ultra Ethernet Transport), 241, 242, 243,
 246
 over IP, 255–256
 UDP-based encapsulation, 252–255

Ultra Ethernet, 239

underlay, eBGP, 145–149

unstructured data, 191

unsupervised training, 3

UPLI (UALink Protocol Level Interface),
 289–290

use cases
 scale-up systems, 277–278
 UEC (Ultra Ethernet Consortium), 241

UUD (unreliable unordered delivery), 245, 260

V

VAST, 211

vendor interoperability, 300

Verbs API, 9

virtual pod partitioning, 289

VLs (virtual lanes), 216

VOQ (Virtual Output Queues), 50

VRF, 162

VXLAN, 145, 162, 171
 GBP (group-based policy), 177
 RT5, 173

W

watchdog, PFC, 133–134

WDM (wavelength division multiplexing), 74

weighted ECMP, 102, 153–154

WekaFS, 209–211

working groups, 241
 Compliance, 243–244
 Link Layer, 242
 Management, 243–244
 Performance and Debug, 243–244
 Physical Layer, 243–244
 Software, 243
 Storage, 243–244
 Transport, 242–243, 257–258

WQE (work queue element), 9

X-Y-Z

XOFF/XON thresholds, 130–132

xPU, 275, 276–277, 279. *See also* GPU/s
 ACC (accelerator), 280
 bandwidth, 280
 HBM (high-bandwidth memory), 279
 SPA (System Physical Address), 292
 UALink, 286

ZR (extended reach), 77