

Software-Defined Data Infrastructure Essentials

Cloud, Converged, and Virtual Fundamental Server Storage I/O Tradecraft

Greg Schulz
Server StorageIO
@StorageIO

Contents

Preface	xv
Who Should Read This Book	xviii
How This Book Is Organized	xix
Acknowledgments	xxiii
About the Author	xxv
Part One	
Server Storage I/O, Software-Defined and Data Infrastructures	1
Chapter 1: Server Storage I/O and Data Infrastructure Fundamentals	3
What You Will Learn in This Chapter	3
1.1 Getting Started	4
1.2 What's the Buzz in and around Servers, Storage, and I/O?	6
1.2.1 Data Infrastructures—How Server Storage I/O Resources Are Used	9
1.2.2 Why Servers Storage and I/O Are Important (Demand Drivers)	13
1.2.3 Data Value	14
1.3 Server Storage I/O Past, Present, and Future	15
1.3.1 Where Are We Today? (Balancing Legacy with Emerging)	15
1.3.2 Where Are We Going? (Future Planning, Leveraging Lessons Learned)	17
1.4 Server and Storage I/O Tradecraft	18
1.5 Fundamental Server and Storage I/O Terminology (Context Matters)	20
1.6 What's in Your Fundamental Toolbox?	21
1.7 Common Questions and Tips	22
1.8 Learning Experience	23
1.9 Chapter Summary	24

vi Software-Defined Data Infrastructure Essentials

Chapter 2: Application and IT Environments	25
What You Will Learn in This Chapter	25
2.1 Getting Started	25
2.1.1 Tradecraft and Context for the Chapter	26
2.2 Everything Is Not the Same with Servers Storage and I/O	26
2.2.1 Various Types of Environments (Big and Small)	27
2.2.2 Gaining Data and Application Insight	28
2.2.3 Various Types of Applications	29
2.2.4 Various Types of Data	32
2.3 Common Applications Characteristics	41
2.3.1 Performance and Activity (How Resources Get Used)	42
2.3.2 Availability (Accessibility, Durability, Consistency)	44
2.3.3 Capacity and Space (What Gets Consumed and Occupied)	45
2.3.4 Economics (People, Budgets, Energy and other Constraints)	46
2.4 Where Applications and Data Get Processed and Reside	46
2.4.1 The Past, Today, and Tomorrow	47
2.5 What's in Your Application Toolbox?	48
2.6 Common Questions and Tips	48
2.7 Learning Experience	49
2.8 Chapter Summary	49
Part Two	
Server I/O and Networking Deep Dive	51
Chapter 3: Bits, Bytes, Blobs, and Software-Defined Building Blocks	53
What You Will Learn in This Chapter	53
3.1 Getting Started	53
3.1.1 Tradecraft and Context for the Chapter	55
3.1.2 Removing Object Obstacles: Context Matters	57
3.2 Server and Storage I/O Basics	57
3.2.1 From Bits to Bytes, Blocks to Blobs (Server Storage I/O Counting)	58
3.2.2 Where Are My Missing Bytes?	61
3.2.3 Server Memory and Storage Hierarchy	62
3.3 How Servers Organize, Access, and Use Storage	64
3.3.1 Resource Tenancy (Sharing vs. Dedicated)	67
3.3.2 Basic Storage Organization (Partitions, LUNs, and Volumes)	69
3.3.3 How Data Gets Written to and Read from a Storage Device	74
3.4 Volumes and Volume Managers	77
3.5 Files and File Systems	78

3.5.1	Metadata Matters and Management	81
3.5.2	File System and Storage Allocation	82
3.6	Cloud, Object, and API-Accessible Storage	83
3.7	Cloud Services (Public, Private, Hybrid)	87
3.7.1	Cloud and Object Access	87
3.8	What's in Your Server Storage I/O Toolbox?	88
3.9	Common Questions and Tips	89
3.10	Learning Experience	90
3.11	Chapter Summary	90
Chapter 4: Servers: Physical, Virtual, Cloud, and Containers		93
	What You Will Learn in This Chapter	93
4.1	Getting Started	93
4.1.1	Tradecraft and Context for the Chapter	95
4.2	Server Fundamentals	96
4.2.1	Server and I/O Architectures	97
4.2.2	Applications PACE and Defining Your Server	99
4.3	Bare-Metal, Physical Machines, and Hosts	100
4.3.1	Fundamental Server Components	100
4.3.2	Server Reliability, Availability, and Serviceability	102
4.3.3	Processors: Sockets, Cores, and Threads	102
4.3.4	Memory Matters	104
4.3.5	I/O and Expansion (Internal and External)	107
4.3.6	PCIe, Including Mini-PCIe, U.2, M.2, and GPU	107
4.3.7	LAN and Storage Ports and Internal Storage	113
4.4	Server Packaging and Enclosures	114
4.4.1	Appliances: Converged, Hyper-Converged, and CiB	115
4.5	Scaling Servers (Up, Down, and Out)	117
4.6	Management Tools and Security	118
4.7	Operating Systems	119
4.8	Hypervisors and Virtual Server Infrastructures (VSI)	121
4.8.1	Virtual Desktop Infrastructure	124
4.9	Containers and Microservices	125
4.10	Cloud Machines and Instances	128
4.11	Putting the Server Together	129
4.12	What's in Your Server Toolbox?	130
4.13	Common Questions and Tips	131
4.14	Learning Experience	133
4.15	Chapter Summary	133

viii Software-Defined Data Infrastructure Essentials

Chapter 5: Server I/O and Networking	135
What You Will Learn in This Chapter	135
5.1 Getting Started	135
5.1.1 Tradecraft and Context for the Chapter	137
5.1.2 Server I/O PACE and Performance Fundamentals	139
5.1.3 Server I/O and Networking Fundamentals	143
5.1.4 Inter- vs. Intra-Server I/O Activity	144
5.2 Server I/O Basic Building Blocks	145
5.2.1 Cache and Locality of Reference	148
5.3 Access Methods and Personalities	150
5.3.1 Logical Server and Storage I/O	150
5.3.2 Namespace and Endpoints	152
5.3.3 Connection-Oriented and Connectionless Transport Modes	155
5.3.4 RDMA (Remote Direct Memory Access)	155
5.3.5 RoCE (RDMA over Converged Ethernet)	156
5.3.6 SATA, SCSI, and NVMe Protocols	156
5.3.7 TCP/IP and TCP/UDP	156
5.3.8 Protocol Droop and Deterministic Behavior	158
5.4 Physical Server and Storage I/O	159
5.5 Server I/O Topologies	159
5.6 Server I/O Management Topics	162
5.7 What's in Your Server Storage I/O Toolbox?	163
5.8 Common Questions and Tips	164
5.9 Learning Experience	165
5.10 Chapter Summary	165
Chapter 6: Servers and Storage-Defined Networking	167
What You Will Learn in This Chapter	167
6.1 Getting Started	167
6.1.1 Tradecraft and Context for the Chapter	168
6.2 Server and Storage I/O Networking Devices	170
6.2.1 Cabling and Connectors	171
6.2.2 Host Bus Adapters and Network Interface Cards	172
6.2.3 Switches and Directors	176
6.2.4 Bridges and Gateways	178
6.3 Local Server I/O and Storage Access	178
6.3.1 PCIe, M.2, and U.2 (SFF-8639) Drive Connectors	181
6.3.2 Serial ATA (SATA)	182
6.3.3 Serial Attached SCSI (SAS)	182
6.3.4 NVM Express (NVMe)	184

6.3.5	Ethernet (IEEE 802)	186
6.3.6	Fibre Channel over Ethernet (FCoE)	187
6.3.7	iSCSI (SCSI on TCP/IP)	187
6.3.8	Fibre Channel (FC)	188
6.3.9	InfiniBand (IBA)	189
6.4	Metropolitan and Wide-Area Networking	191
6.4.1	Enabling Server I/O and Storage over Distance	192
6.5	Software-Defined Networks (SDN) and Network Function Virtualization (NFV)	194
6.6	Accessing Cloud Storage and Services	197
6.7	What's in Your Server Storage I/O Networking Toolbox?	201
6.8	Common Questions and Tips	202
6.9	Learning Experience	204
6.10	Chapter Summary	204
Part Three		
Storage Deep Dive and Data Services		207
Chapter 7: Storage Mediums and Component Devices		209
What You Will Learn in This Chapter		209
7.1	Getting Started	210
7.1.1	Tradecraft and Context for the Chapter	212
7.2	Common Storage Device and Media Matters	213
7.2.1	Storage Device Media PACE and Metrics that Matter	215
7.2.2	Packaging, Form Factors, and Interfaces	217
7.2.3	Enclosures	218
7.3	Aligning Technology to Application Needs	221
7.4	Volatile Memory (DRAM) and Non-Persistent Storage	223
7.5	Non-Volatile Memory (NVM) and SSD	223
7.5.1	NVM, Flash, and SSD Fundamentals	225
7.5.2	Flash SSD TRIM and UNMAP Garbage Collection	226
7.5.3	Different Types and Classes of SSD	227
7.5.4	NVM and SSD Considerations	227
7.6	Magnetic Hard Disk Drives	228
7.6.1	Types and Classes of HDD	230
7.6.2	HDD Solid-State Hybrid Disk Considerations and Trends	232
7.7	Magnetic Tape, Optical, and Removable Media	233
7.8	What's in Your Storage Device Toolbox?	236
7.9	Common Questions and Tips	236
7.10	Learning Experience	238
7.11	Chapter Summary	240

x Software-Defined Data Infrastructure Essentials

Chapter 8: Data Infrastructure Services: Access and Performance	241
What You Will Learn in This Chapter	241
8.1 Getting Started: What’s in Your Server Storage I/O Toolbox?	242
8.1.1 Tradecraft and Context for the Chapter	244
8.2 Virtual Disks, Endpoints, Namespaces, and Access	246
8.2.1 Access, Personality, and Protocols	247
8.2.2 Migrating Server Storage	255
8.2.3 Data Movement and Migration	256
8.2.4 Volume Managers and Virtual Disks	257
8.2.5 Multi-tenancy and Virtualization	261
8.3 Performance (Productivity and Effectiveness) Services	262
8.3.1 Server Storage I/O Acceleration (Cache and Micro-tiering)	266
8.4 Economics, Analytics, and Reporting (Insight and Awareness)	271
8.4.1 Metrics That Matter and Insight Awareness	271
8.4.2 Metrics and Data Infrastructure Cost Considerations	272
8.4.3 Analytics, Discovery, and Search	274
8.5 Management Services	275
8.5.1 Management Interfaces	275
8.5.2 Management Functionalities	280
8.5.3 Third-Party Storage and Cloud Support	283
8.6 What’s in Your Toolbox?	284
8.7 Common Questions and Tips	284
8.8 Chapter Summary	285
 Chapter 9: Data Infrastructure Services: Availability, RAS, and RAID	 287
What You Will Learn in This Chapter	287
9.1 Getting Started	287
9.1.1 Tradecraft and Tools	288
9.2 Copy Data Management	289
9.3 Availability (Resiliency and Data Protection) Services	291
9.3.1 Revisiting 4 3 2 1—The Golden Rule of Data Protection	293
9.3.2 Availability, FTT, and FTM Fundamentals	295
9.3.3 Common Availability Characteristics and Functionalities	297
9.3.4 Reliability, Availability, Analytics, and Data Protection Management	300
9.3.5 Enabling Availability, Resiliency, Accessibility, and RTO	302
9.4 What’s in Your Toolbox?	321
9.5 Common Questions and Tips	322
9.6 Learning Experience	323
9.7 Chapter Summary	324

Chapter 10: Data Infrastructure Services: Availability, Recovery-Point Objective, and Security	325
What You Will Learn in This Chapter	325
10.1 Getting Started	325
10.1.1 Tradecraft and Tools	326
10.2 Enabling RPO (Archive, Backup, CDP, Snapshots, Versions)	326
10.3 Backup and Archiving	332
10.3.1 Local and Remote Copies	332
10.3.2 Walking the Protection Talk	333
10.3.3 Point-in-Time Protection for Different Points of Interest	334
10.3.4 Point-in-Time Protection and Backup/Restore Considerations	335
10.4 Snapshots, CDP, Versioning, Consistency, and Checkpoints	340
10.5 Data Infrastructure Security (Logical and Physical)	344
10.5.1 Data Infrastructure Security Implementation	345
10.5.2 Physical Security and Facilities	346
10.5.3 Logical and Software-Defined Security	346
10.5.4 Checksum, Hash, and SHA Cipher Encryption Codes	348
10.5.5 Identity Access Management and Control	351
10.5.6 Server, Storage, and I/O Networking	352
10.5.7 General Data Infrastructure Security-Related Topics	353
10.6 What's In Your Toolbox?	354
10.7 Common Questions and Tips	355
10.8 Learning Experience	358
10.9 Chapter Summary	358
Chapter 11: Data Infrastructure Services: Capacity and Data Reduction	361
What You Will Learn in This Chapter	361
11.1 Getting Started	361
11.1.1 Tradecraft and Tools	362
11.2 Capacity (Space Optimization and Efficiency) Services	362
11.2.1 General Capacity Items	363
11.2.2 Data Footprint Reduction	366
11.3 What's in Your Toolbox?	394
11.4 Common Questions and Tips	394
11.5 Learning Experience	396
11.6 Chapter Summary	396
Chapter 12: Storage Systems and Solutions (Products and Cloud)	399
What You Will Learn in This Chapter	399

xii Software-Defined Data Infrastructure Essentials

12.1	Getting Started	399
12.1.1	Tradecraft, and Context for the Chapter	403
12.2	Why So Many Types of Storage Systems?	404
12.2.1	Different IT Environments	407
12.2.2	Application Workloads and Usage Scenarios	407
12.2.3	Storage System, Solution, and Service Options	408
12.2.4	Storage Systems and Solutions—Yesterday, Today, and Tomorrow	409
12.3	Common Storage System Characteristics	410
12.3.1	Storage System PACE Considerations	411
12.3.2	Additional Functional and Physical Attributes	416
12.4	Storage Systems Architectures	416
12.4.1	Scaling Up, Down, Out, or Out and Up	418
12.4.2	Storage Front-End Matters	418
12.4.3	Internal Processing and Functionality	421
12.4.4	Storage Back-End Media Matters	428
12.4.5	Common Storage Architecture Considerations	429
12.5	Storage Solution Packaging Approaches	431
12.5.1	Tin-Wrapped and Hardware Defined	432
12.5.2	Software-Defined Storage	434
12.5.3	Virtual Storage and Storage Virtualization	438
12.5.4	Hybrid, All-NVM, All-Flash, and All-SCM Arrays	439
12.6	Converged Infrastructure (CI) and Hyper-CI (HCI)	440
12.7	Cloud Instances and Services	442
12.8	Gateways, Appliances, Adapters, and Accessories	444
12.9	Storage Management Software	445
12.10	Resiliency Inside and Outside Storage Solutions	445
12.11	What’s in Your Storage Toolbox?	447
12.12	Common Questions and Tips	447
12.13	Learning Experience	449
12.14	Chapter Summary	449
Part Four		
Putting Software-Defined Data Infrastructures Together		451
Chapter 13: Data Infrastructure and Software-Defined Management		453
	What You Will Learn in This Chapter	453
13.1	Getting Started	453
13.1.1	Management Tradecraft and Tools	454
13.1.2	Data Infrastructure Habitats and Facilities	457
13.2	Data Infrastructure Management	458

13.2.1	Troubleshooting, Problem Solving, Remediation, and Repairs	463
13.2.2	Availability, Data Protection, and Security	473
13.2.3	Analytics, Insight, and Awareness (Monitoring and Reporting)	474
13.3	Data Infrastructure Decision Making	482
13.3.1	Comparing Data Infrastructure Components and Services	485
13.3.2	Analysis, Benchmark, Comparison, Simulation, and Tests	488
13.4	Data Infrastructure Design Considerations	493
13.5	Common Questions and Tips	494
13.6	Chapter Summary	495
Chapter 14: Data Infrastructure Deployment Considerations		497
	What You Will Learn in This Chapter	497
14.1	Getting Started	497
14.1.1	Deployment Tradecraft and Tools	498
14.2	Applications, Tips, and Learning Experiences	498
14.2.1	Software-Defined, Virtual, Containers, and Clouds	499
14.2.2	Microsoft Azure, Hyper-V, Windows, and Other Tools	517
14.2.3	VMware vSphere, vSAN, NSX, and Cloud Foundation	521
14.2.4	Data Databases: Little Data SQL and NoSQL	525
14.2.5	Big Data, Data Ponds, Pools, and Bulk-Content Data Stores	534
14.3	Legacy vs. Converged vs. Hyper-Converged vs. Cloud and Containers	539
14.4	Common Questions and Tips	542
14.5	Chapter Summary	545
Chapter 15: Software-Defined Data Infrastructure Futures, Wrap-up, and Summary		547
	What You Will Learn in This Chapter	547
15.1	Getting Started on the Wrap-up	547
15.1.1	What's in Your Server Storage I/O Toolbox?	549
15.1.2	Tradecraft and Context Topics	549
15.2	People, Process, and Best Practices	551
15.2.1	Skills Development	552
15.3	Emerging Topics, Trends, and Predictions	553
15.4	Chapter and Book Summary	557
Appendix A: Learning Experiences		561
Appendix B: Additional Learning, Tools, and Tradecraft Tricks		567

xiv Software-Defined Data Infrastructure Essentials

Appendix C: Frequently Asked Questions	579
Appendix D: Book Shelf and Recommended Reading	583
Appendix E: Tools and Technologies Used in Support of This Book	587
Appendix F: How to Use This Book for Various Audiences	589
Appendix G: Companion Website and Where to Learn More	591
Glossary	593
Index	615