



ADVANCED TECHNOLOGY GROUP (ATG)



Accelerate with ATG Webinar: Introducing the Storage Scale System 6000

Shu Mookerjee

ATG Senior Storage Technical Specialist

Shu.Mookerjee@ibm.com



Accelerate with ATG Technical Webinar Series

Advanced Technology Group experts cover a variety of technical topics.

Audience: Clients who have or are considering acquiring IBM Storage solutions. Business Partners and IBMers are also welcome.

To automatically receive announcements of upcoming Accelerate with IBM Storage webinars, Clients, Business Partners and IBMers are welcome to send an email request to accelerate-join@hursley.ibm.com.



2024 Upcoming Webinars – Register Here!

[IBM c-Type SAN Update - January 16th, 2024](#)

[Exploring the IBM Storage Scale GUI - January 23rd, 2024](#)

[IBM Copy Services Manager V6.3.9.0 & DS8900F R9.4 Updates - January 30th, 2024](#)

[IBM Storage Ceph Updates for Versions 6 and 7 - February 6th, 2024](#)

Important Links to bookmark:



ATG Accelerate Site: <https://ibm.biz/BdSUFN>

ATG MediaCenter Channel: <https://ibm.biz/BdfEgQ>

Offerings

Client Technical Workshops

- **IBM DS8900F Advanced Functions:** Feb. 7-8 in Dallas, TX
- **IBM Cyber Resiliency with IBM Storage Defender:** Feb. 14-15 in Dallas, TX
- **IBM Storage Fusion:** Feb. 21-22 in Dallas, TX
- **IBM Storage Scale & Storage Scale System:** Feb. 28-29 in Costa Mesa, CA
- **IBM FlashSystem Deep Dive & Advanced Functions:** March 6-7 in Dallas, TX
- **IBM Storage Ceph:** March 20-21 in Costa Mesa, CA
- IBM FlashSystem and Storage Virtualize
- IBM Cloud Object Storage

TechZone Test Drive / Demo's

- IBM Storage Scale and Storage Scale System GUI
- IBM Storage Virtualize Test Drive
- IBM DS8900F Storage Management Test Drive
- Managing Copy Services on the DS8000 Using IBM Copy Services Manager Test Drive
- IBM DS8900F Safeguarded Copy (SGC) Test Drive
- IBM Cloud Object Storage Test Drive - (Appliance based)
- IBM Cloud Object Storage Test Drive - (VMware based)
- IBM Storage Protect Live Test Drive
- IBM Storage Ceph Test Drive - (VMware based)

Please reach out to your IBM Representative or Business Partner for more information.

***IMPORTANT* The ATG team serves clients and Business Partners in the Americas, concentrating on North America.**

Save the date

Storage @ IBM TechXchange Conference 2024

October 21-24, 2024

Mandalay Bay | Las Vegas
#IBMTechXchange

Key Learnings

- Practical how-to advice
- Patterns and best practices
- Success stories, IBM PoV, proven techniques

Featured Products

IBM Storage Defender

IBM Storage Fusion

IBM Storage Scale + IBM Storage Ceph

IBM Tape + IBM SAN

IBM Storage FlashSystem + IBM Storage DS8000

Collaborate. Learn. Play.

Community

IBM Champions

User Groups

Tech Peers

Business Partners



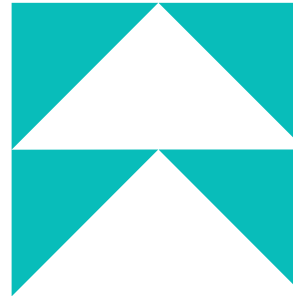
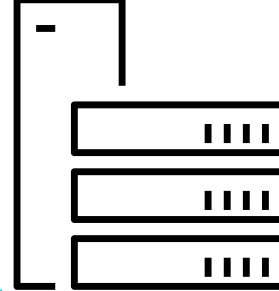
Sandbox

Network

Learn

Collaborate

Play



Accelerate your Career

Labs (Instructor-Led, Self-paced)

IBM Certification Testing

Earn up to 25 hours in CPE credits

Breakout Sessions

Trends and Directions

User Groups

Product Deep Dives

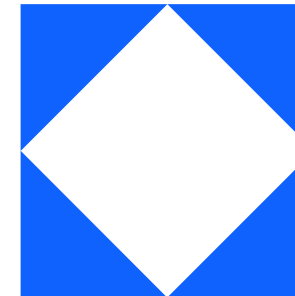
Meet the Expert

Professional Development

Show the Code

Birds of a Feather

Academic/Research



Roadmaps

Go deep with people in the know and set the stage for where IBM is going in the future



<https://www.ibm.com/community/ibm-techxchange-conference/>

Game On!



Accelerate with ATG Survey

Please take a moment to share your feedback with our team!

You can access this 6-question survey via [Menti.com](https://www.menti.com/join/17086924) with code 1708 6924 or

Direct link <https://www.menti.com/alwhyze7z1gz>

Or

QR Code





ADVANCED TECHNOLOGY GROUP (ATG)



Accelerate with ATG Webinar: Introducing the Storage Scale System 6000

Shu Mookerjee

ATG Senior Storage Technical Specialist

Shu.Mookerjee@ibm.com



Meet the Speaker



Shu Mookerjee is a Level 2 Certified Technical Specialist with over twenty years at IBM, working in a variety of roles including sales, management and technology. For the last decade, he has focused exclusively on storage and has been the co-author of four (4) Redbooks. Currently, Shu is part of the Advanced Technology Group where he provides education, technical guidance and Proofs of Technology to IBMers, business partners and clients. His focus solutions are IBM Storage Fusion, IBM Storage Scale and ESS/SSS.

Agenda

- Goals and Objectives
- An Overview of IBM Storage Scale
- A Brief Detour into Storage Scale RAID
- The Current ESS/SSS Family
- Introducing the SSS 6000
 - General Overview
 - Hardware
 - Software Stack
- Utility Nodes – EMS and Protocols
- Integration with an Existing/Legacy ESS Deployment
- Performance

Goals and Objectives

Objective:

Provide an overview of the new SSS 6000

We WILL:

- Provide a *mid-level* review hardware and software architecture
- Discuss the coexistence and interoperability with some legacy ESS/SSS deployments
- Cover relative performance data

We WILL Not:

- Review product positioning
- Cover deep technical architecture and functions
- Discuss the implementation and installation process

IBM Storage Scale – Overview

What is Storage Scale?

- Highly scalable distributed parallel POSIX file system
- Runs on any Open System OS (AIX, Windows, Linux) environment
- Utilizes any block storage device (internal media or SAN storage)

Under the file system

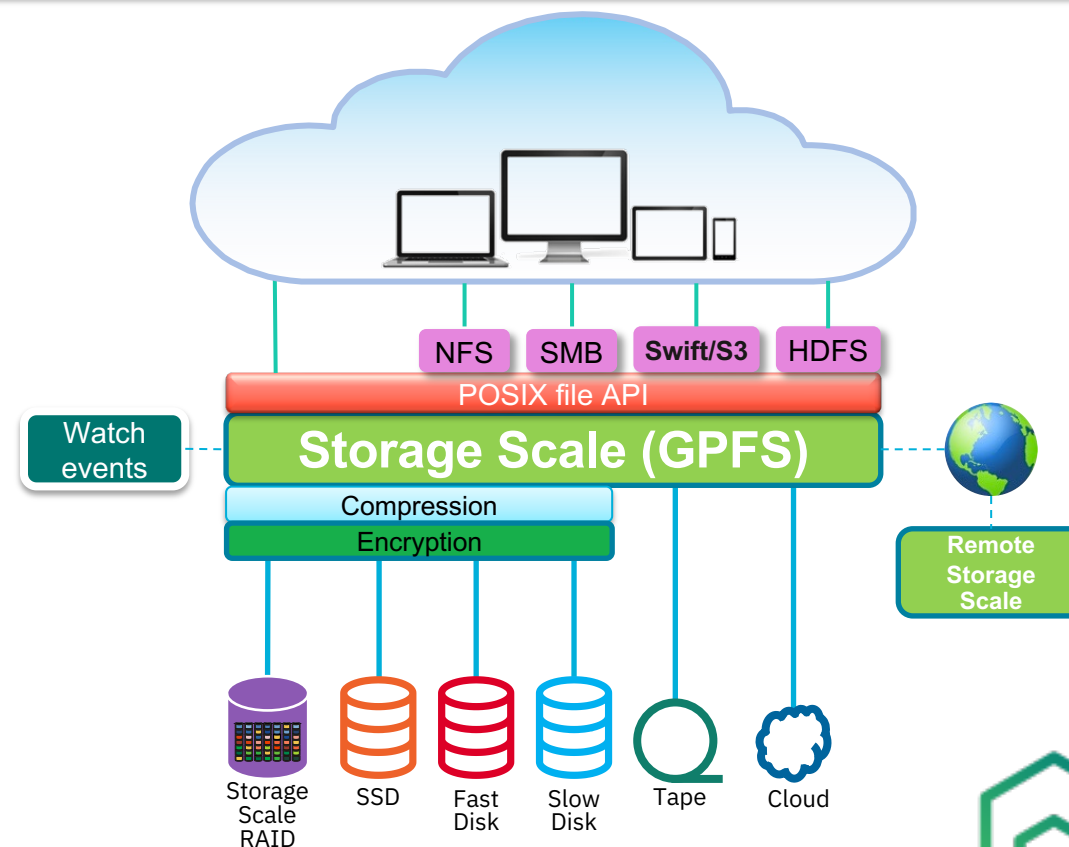
- Use any block storage usable by the OS.
- Information Lifecycle Management features enable tiering
- Built-in encryption and compression
- Tape and cloud-based storage usable as tiers.
- Storage Scale RAID gives unrivaled data integrity.

Above the file system – since not everything will run Scale directly...

- Export the data through multiple protocols (SMB, NFS, Object, HDFS)
- Container Storage Interface (CSI) extends storage access to containers.

Beside the file system

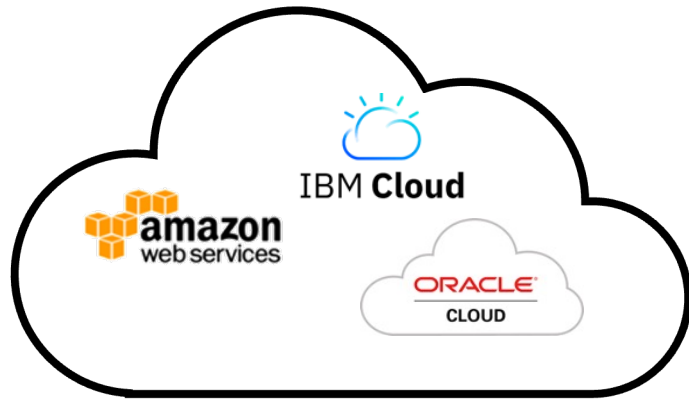
- Peer with remote Storage Scale file systems for efficient remote access.
- Replication capabilities for robust Disaster-Recovery capabilities.
- Watch events in the file system with Watch Folders and File Audit Logging.



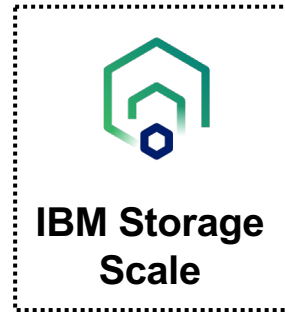
In a single name space, vastly simplify administration and reduce costs by eliminating storage islands.

IBM Storage Scale – Deployment Options

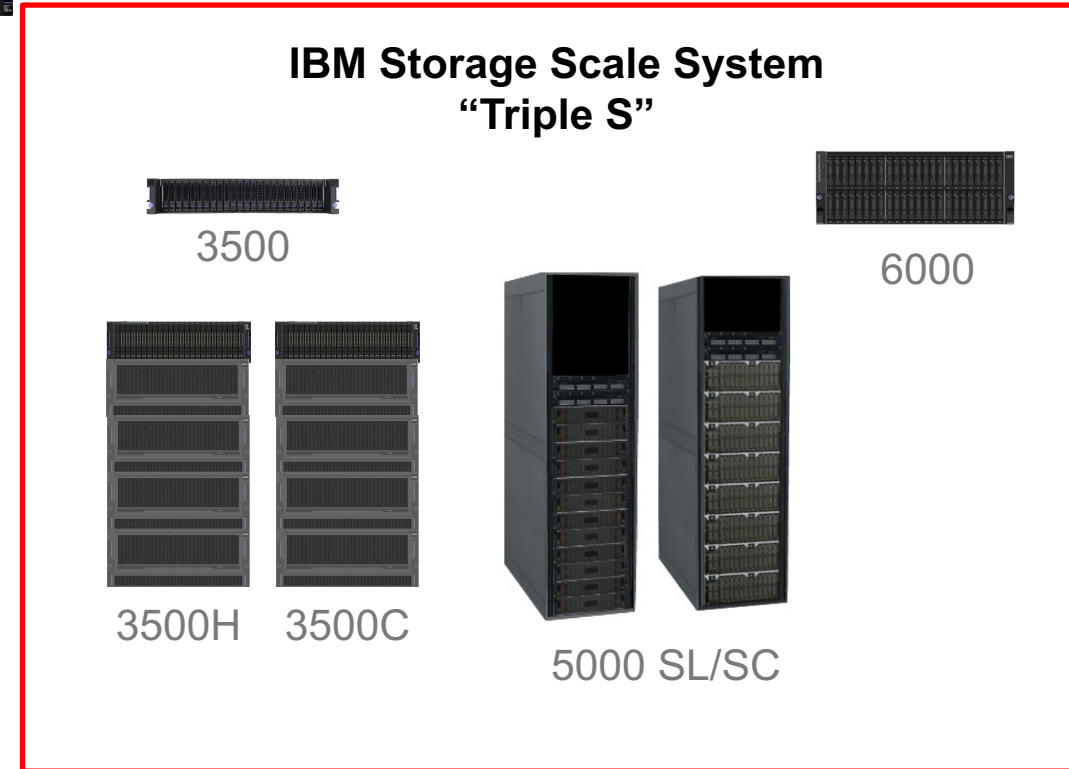
Cloud



Software



Hardware



IBM Storage Scale Server

What is The Storage Scale System? It's a "Scale Building Block"

Okay, here's a more technical definition:

The Triple S is a pair of cross-connected Scale Servers deployed in a fully integrated enclosure running declustered RAID

Scalable, reliable high performance enterprise storage

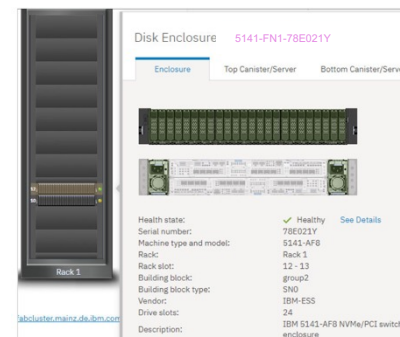
- Faster data access with the choice to scale-up or scale-out
- Easy to deploy clusters with unified system GUI
- Simplified storage administration with IBM Spectrum Control integration

One solution for all Storage Scale data needs

- Single repository of data with unified file and object support
- Anywhere access with multi-protocol support using protocol nodes: NFS 3/4.0/4.1, SMB, Object, and HDFS
- Ideal for big data analytics including full Hadoop transparency

Ready for business-critical data

- Disaster recovery with synchronous or asynchronous replication
- Ensure reliability and fast rebuild times using Storage Scale RAID's dispersed data and erasure code
- Five 9s (99.999%) of availability and online scalability and upgrades

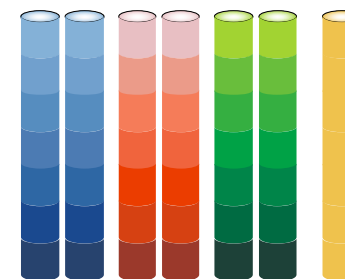


IBM Storage Scale Server – A Brief Detour on Storage Scale RAID

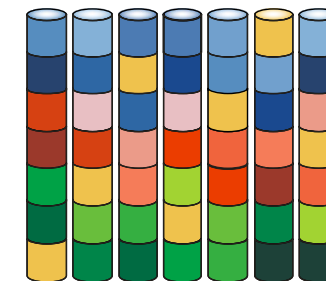
Storage Scale RAID is a *software* implementation of “declustered RAID” that is unique to the ESS/SSS

- A Standard Scale deployment can generally leave the drive media management (integrity, performance, block placement, etc) to the server or storage array
- Since the drives exist in the same enclosure, the Scale software must assume responsibility
- Storage Scale RAID uses Declustered RAID and Erasure Coding to provide:
 - Extremely fast rebuild times with minimal impact on performance
 - Very strong data integrity checks
 - Error detection codes enable detecting track errors and dropped writes
 - Consistent performance from 0 – 99% utilization or 1 to many jobs in parallel

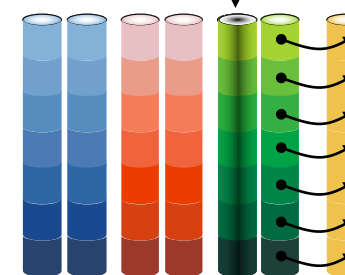
Conventional



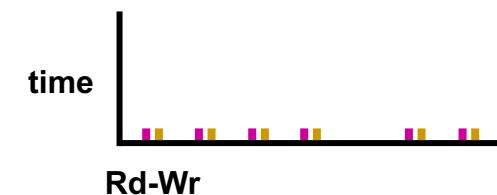
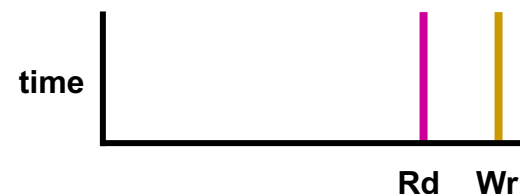
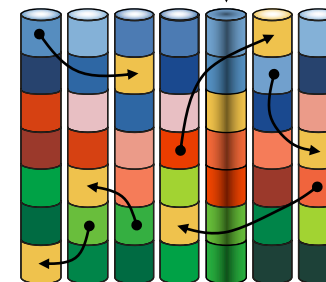
De-clustered



failed disk



failed disk



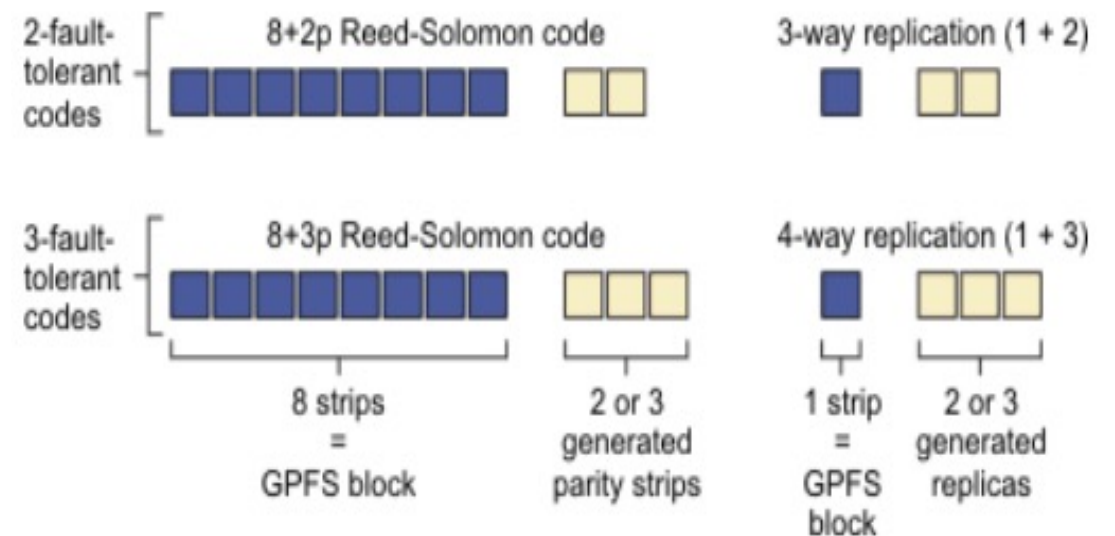
IBM Storage Scale Server – A Brief Detour on Storage Scale RAID

Reed-Solomon Encoding

- 8 Data Strips + 2 or 3 parity strips
- Stripe width 10 or 11 strips
- Storage efficiency 80% or 73% respectively*

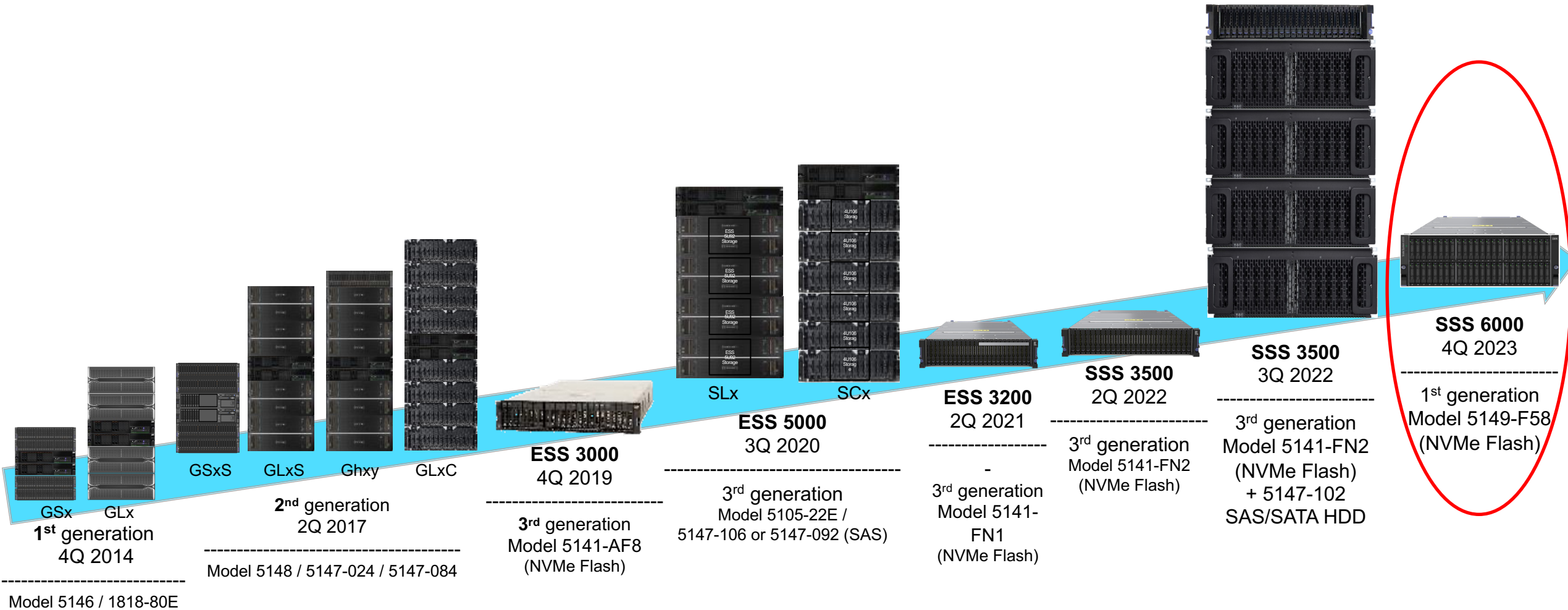
3-way or 4-way replication

- Strip size is file system data block size
- Storage efficiency 33% or 25% respectively



*Excluding user-configurable spare space for rebuilds

IBM Storage Scale Server - Family



Introducing...The SSS 6000 - Overview

Introducing the new IBM Storage Scale System 6000!

System Features

- 4U Enclosure
- Up to 48 x 2.5" drives
- Dual AMD EPYC Genoa 48 cores per controller
- 768 GB – 1.5 TB Memory
- 24 or 48 x NVMe / FCM (PCIe Gen 4 drives)
- From 96 TB to 2.6 PBe Usable Capacity

Performance and Sustainability

- 2x throughput improvement
- Hybrid/converged performance and capacity support
- Containerized protocol support on IO nodes (as with the ESS3500)

Supports 24 or 48 NVMe SSDs

- 3.84 TB
- 7.68 TB
- 15.36 TB
- 30.74 TB

Supports 24 or 48 NVMe FCMs

- 19 TB
- 38 TB



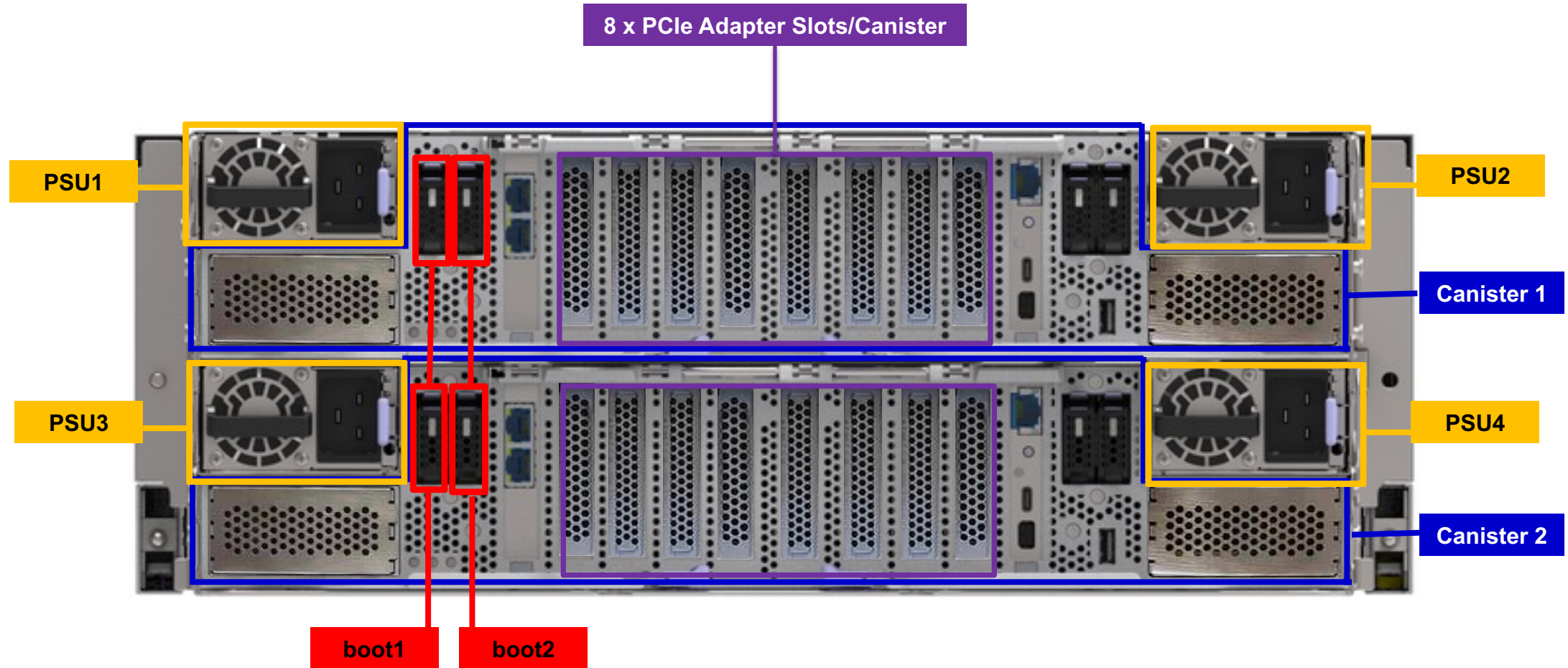
Introducing...The SSS 6000 – Hardware Overview



Front View – Fully Populated (48 Drives)

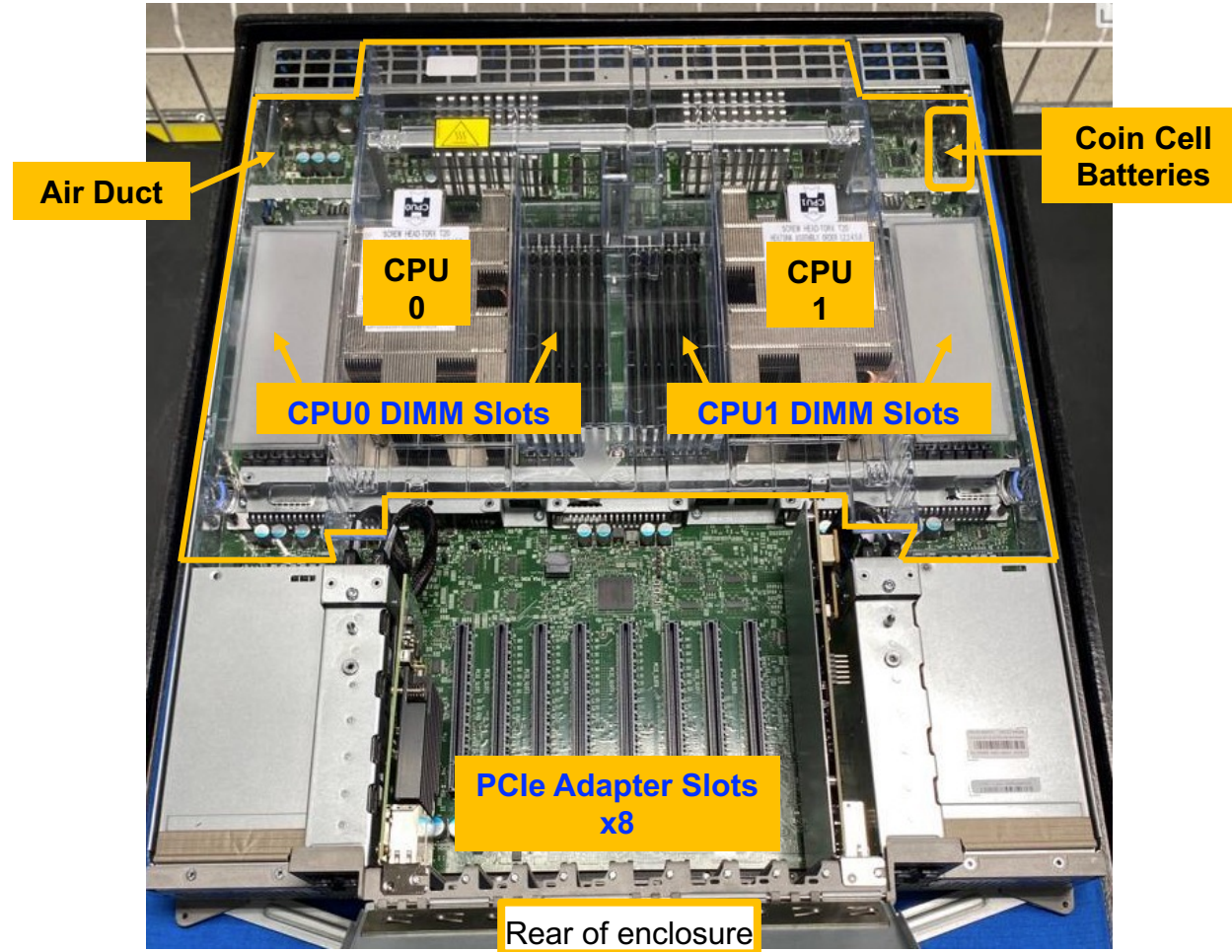
Note: Half Populated systems (24 Drives) will fill all the EVEN slots for Canister 1 and the ODD slots for Canister 2

Introducing...The SSS 6000 – Hardware Overview



Rear View

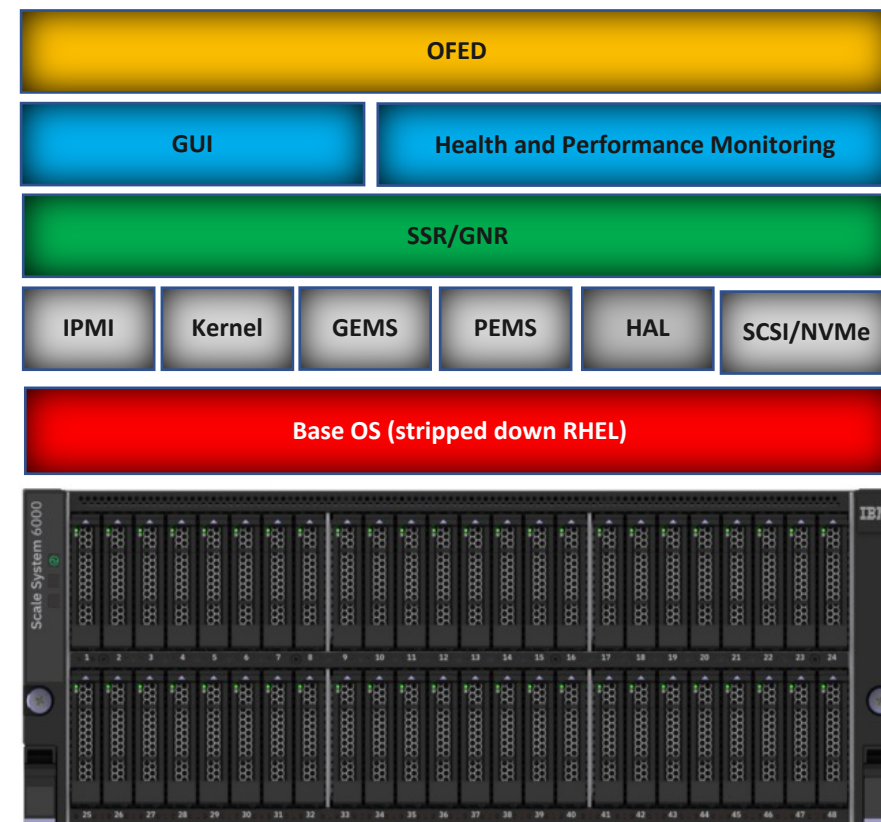
Introducing...The SSS 6000 – Hardware Overview



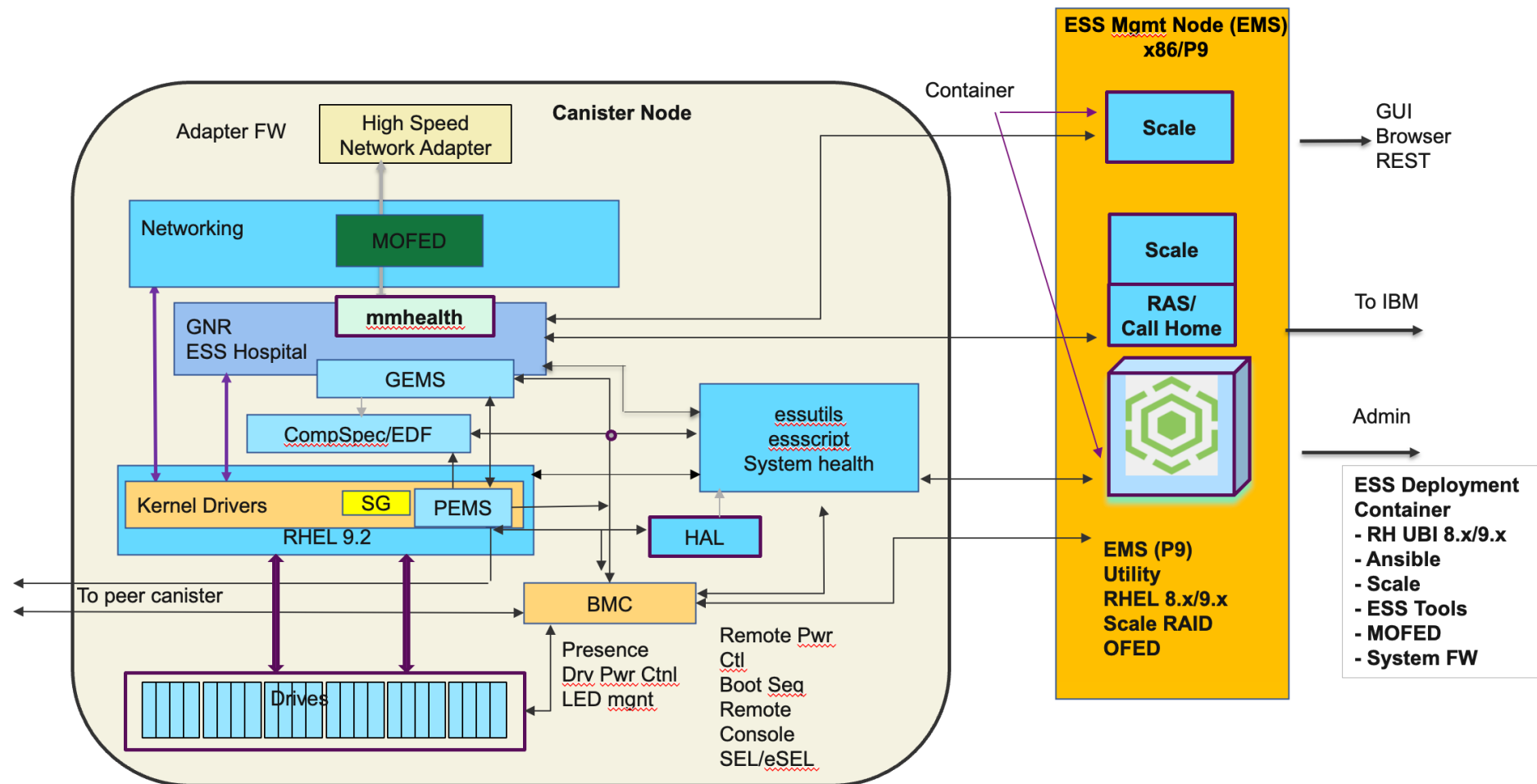
Under the Covers

Introducing...The SSS 6000 – Software Stack

- Base Operating System – RHEL 9.X
- HAL (Hardware Abstraction Layer)
 - Hardware monitoring
 - Call home
- GEMS/PEMS (GNR Enclosure Management Services/Platform Enclosure Management Services)
 - End point inventory
 - State management (presence, power control, LED control)
 - SES emulation (HBA, Encl)
- Storage Scale RAID/GPFS Native RAID (SSR/GNR)
 - Error recovery and reporting
 - Checksum and Disk Hospital
 - Topology and Configuration
 - mmvdisk
 - Component Specification
- Health monitoring, data gathering, reporting and analysis
 - Disk and Fabric Hospital
 - HAL
 - mmhealth
 - Call home
- GUI
 - Setup wizard
 - Inventory and health monitoring
- Networking
 - Management
 - High-speed OFED (OpenFabric Enterprise Distribution) drivers

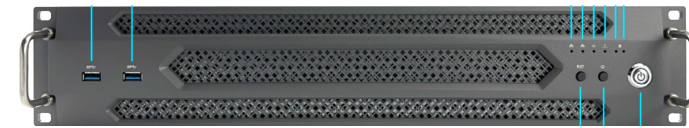


Introducing...The SSS 6000 – Software Architecture and Logical Layout



Utility Nodes – General Specifications

- General purpose, multifunctional x86 server
- 2U (25" Depth) Rack Server, CRPS Redundant P/S
- Processor
 - Dual Socket SP3 (LGA4094)
 - Support AMD® EPYCTM 7003 and 7002
- Memory
 - Up to 8TB
 - 16 DIMM slots per CPU (8 channels per CPU)
 - Supports RDIMM, LRDIMM, RDIMM/LRDIMM-3DS
- PCIe Adapter Slot
 - 4 x PCIe 4.0 x16 + 1 PCIe 4.0 x8 RISER Slots (not used)
 - 1 x OCP NIC 3.0 Slot (PCIe 4.0 x16) (not used)
- Boot Drive
 - 2 x M.2 Micron 7450 PCIe 4.0 x4 for OS
- System Management and BMC
 - 1 x 1Gb Enet for BMC (IPMI 2.0) and SOL
 - 2 x 10Gbit LAN for management (Mgmt/SSR/Others)



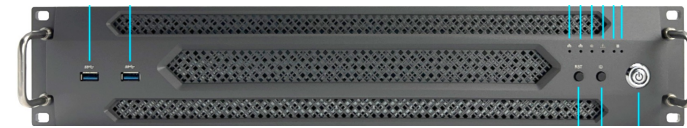
Front View



Rear View
and internal Components

Utility Nodes – General Information

- Currently only supports EMS and Protocol Node functionality on separate servers
- Same Machine Type/Model for both protocol and EMS: 5149-23E
 - However: Protocol node has a different **feature code**
 - Protocol Node default is sold/ordered in pairs
- Utility Node EMS is a requisite for ordering Utility Protocol nodes
- Cannot mix Utility Node Protocol Node with any POWER based Protocol nodes
- SSR set up. Note that CES set up is client responsibility
- GUI, Call Home, etc is still done via EMS guest

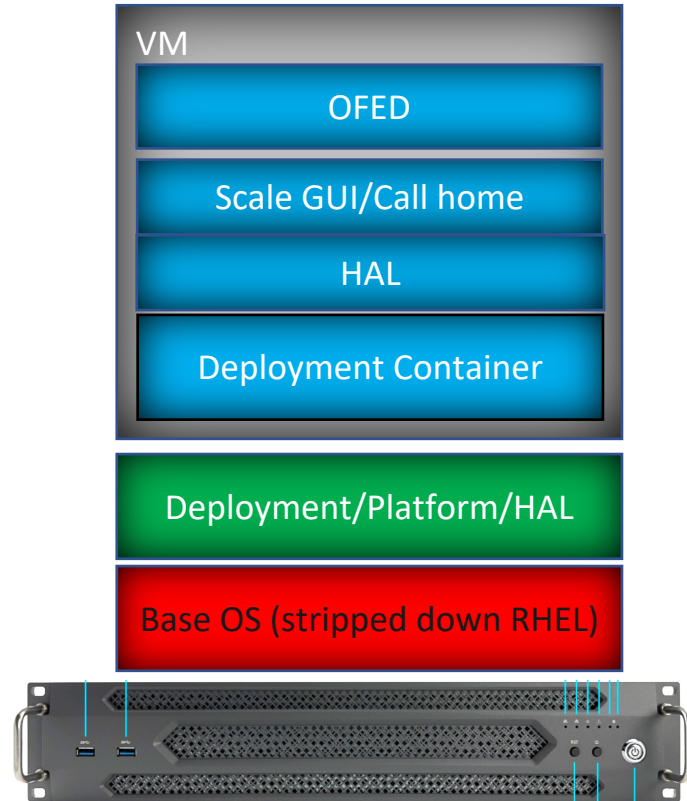


Front View

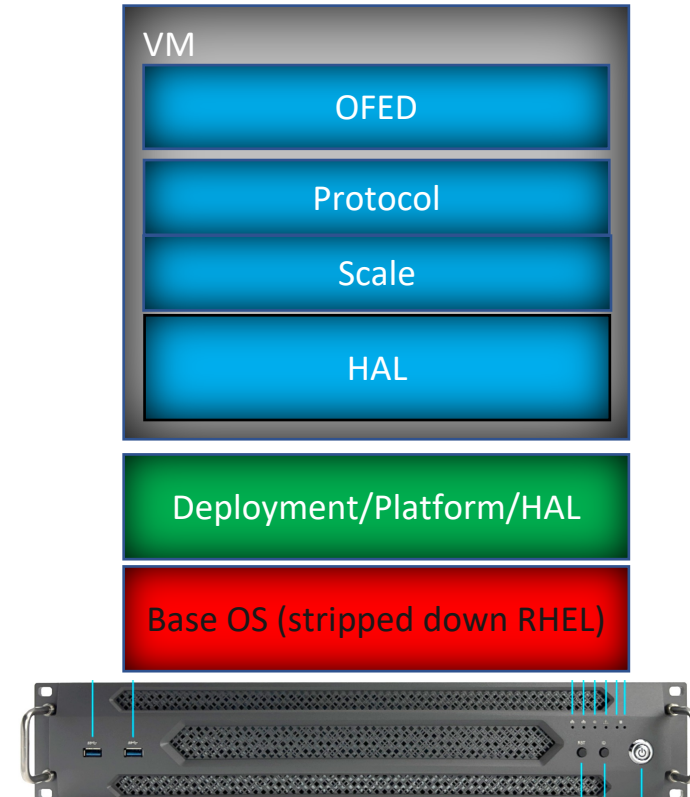


Rear View
and internal Components

Utility Nodes – Software Stack

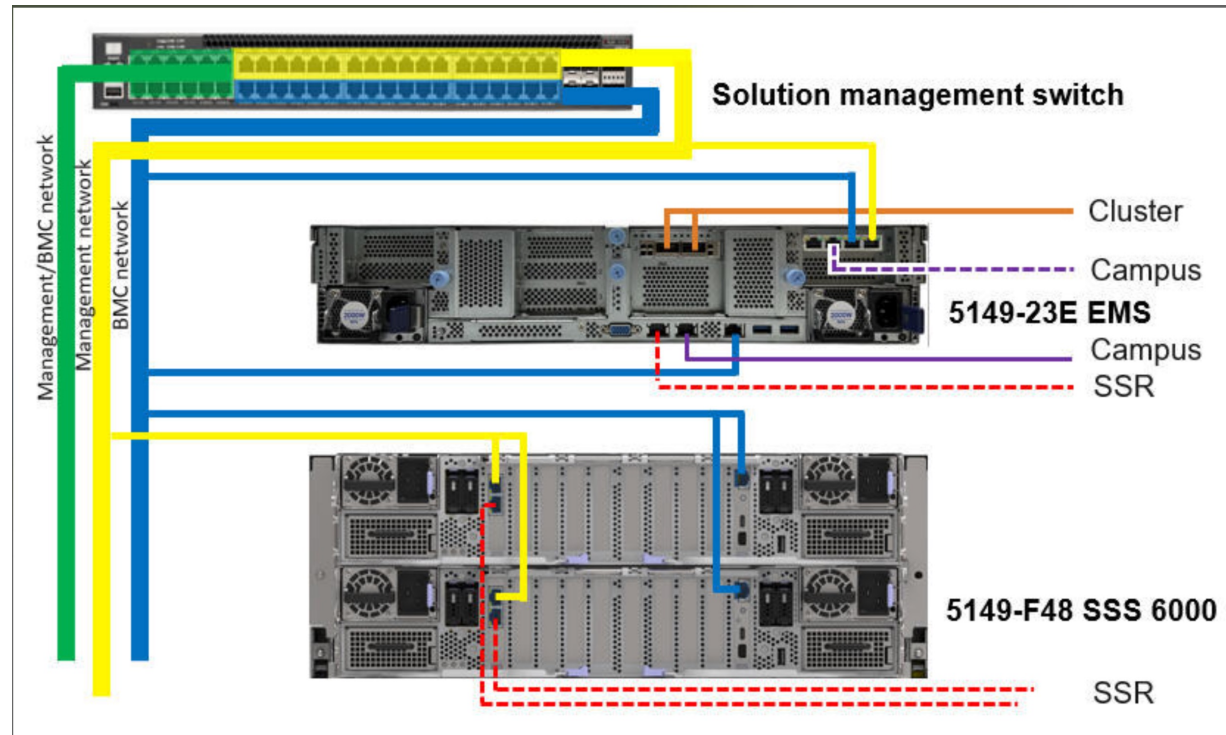


Utility Node - EMS

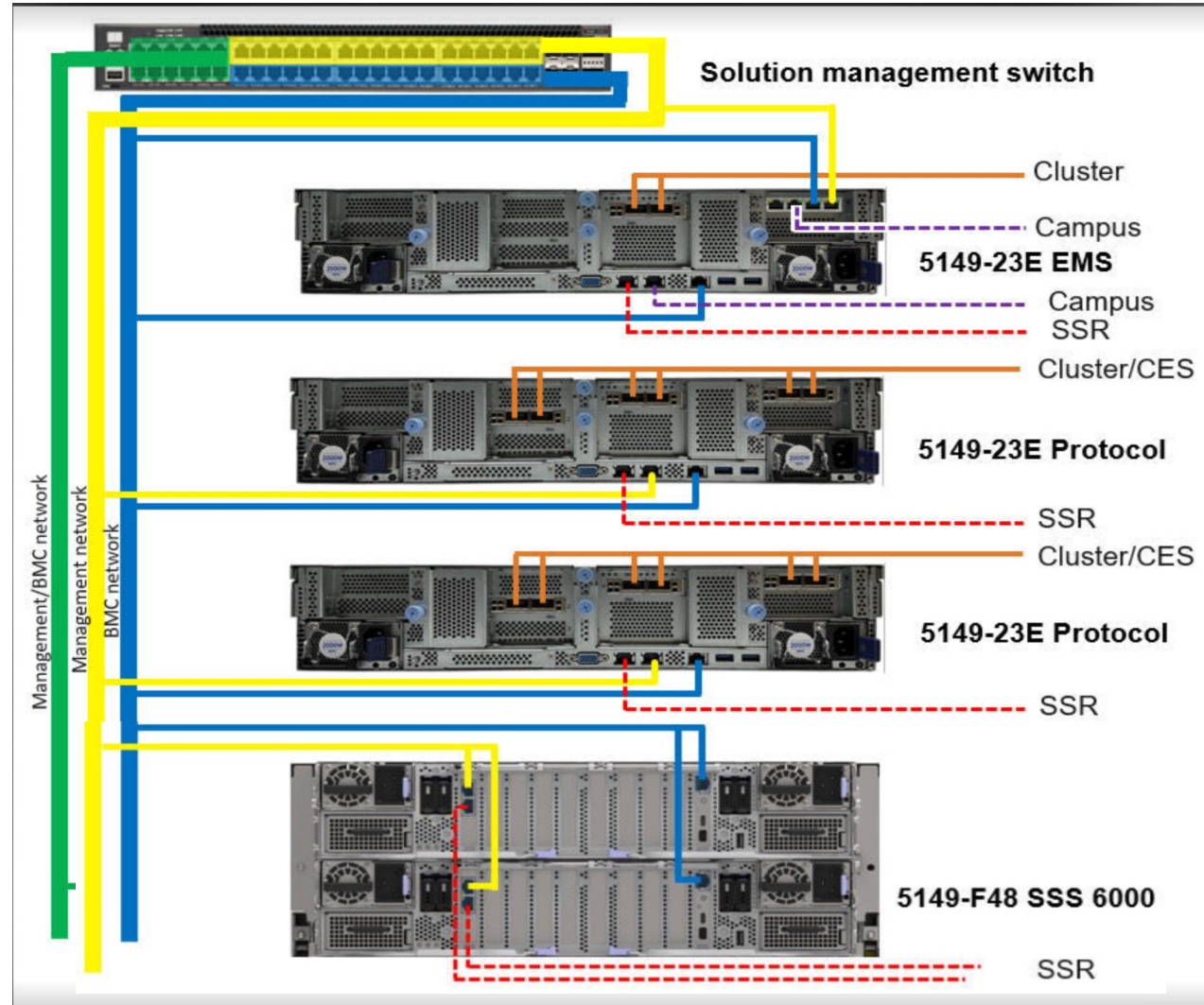


Utility Node – Protocol

Utility Nodes – Networking - EMS

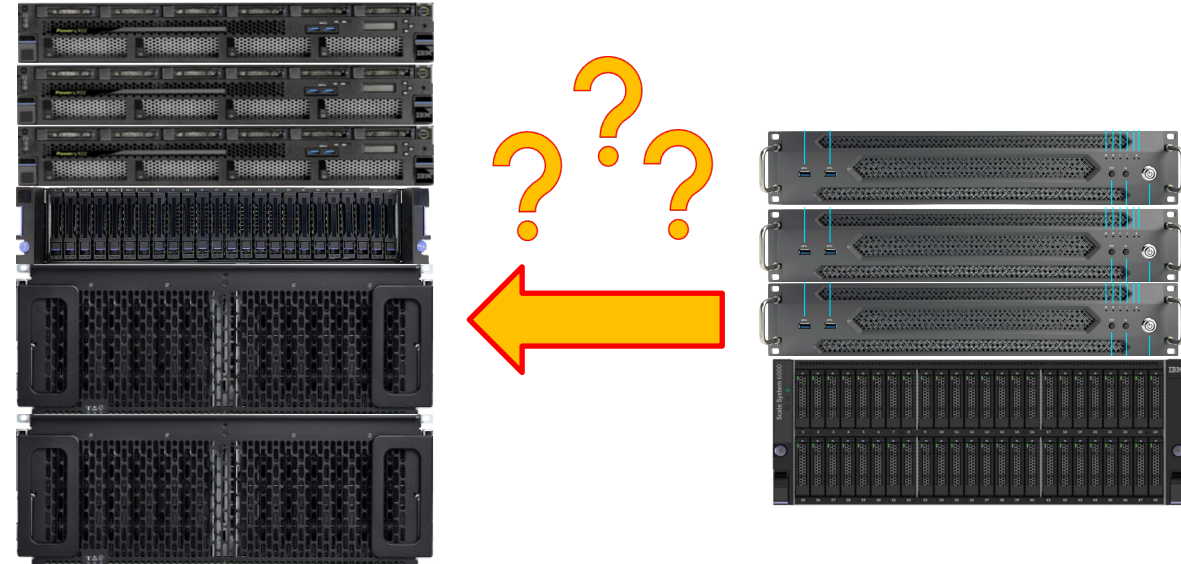


Utility Nodes – Networking – EMS + Protocol Node



Integrating SSS6000 with Legacy Deployments

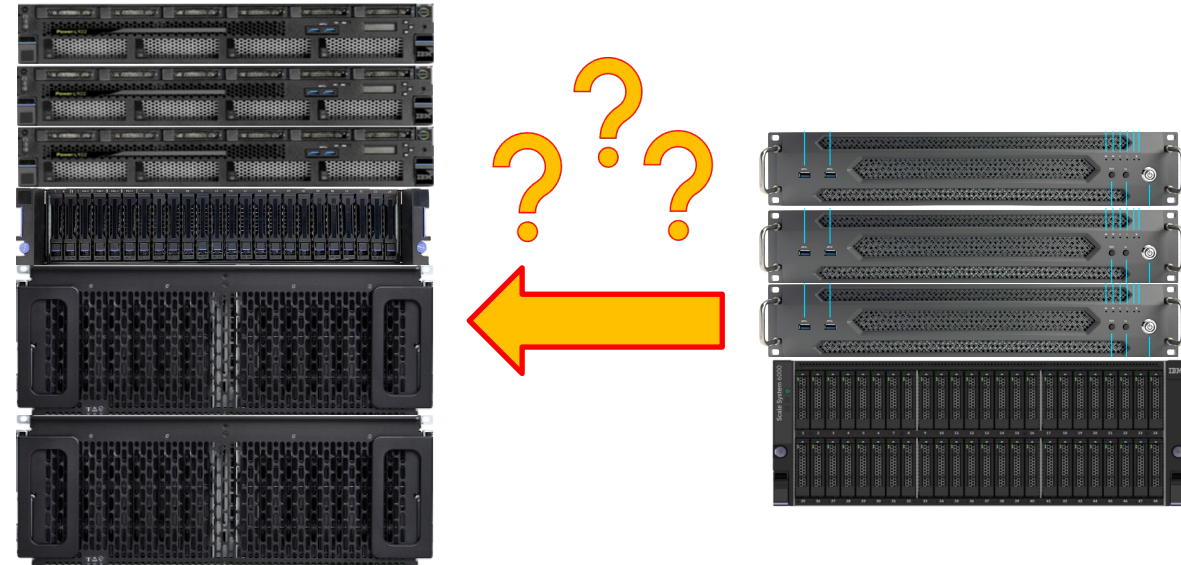
- Commonly asked question: “How do I add a 6000 to an existing ESS/SSS deployment?”
- Especially since:
 - SSS 6000 (and dual socket protocol) can only be deployed with a single socket Utility EMS
 - Utility Node (EMS) cannot deploy / manage any legacy hardware (Besides ESS 3500)
 - P9 EMS cannot deploy / manage ESS 6000 (or dual socket protocol)



Integrating SSS6000 with Legacy Deployments – Short Term

So what's the solution? In the short-term:

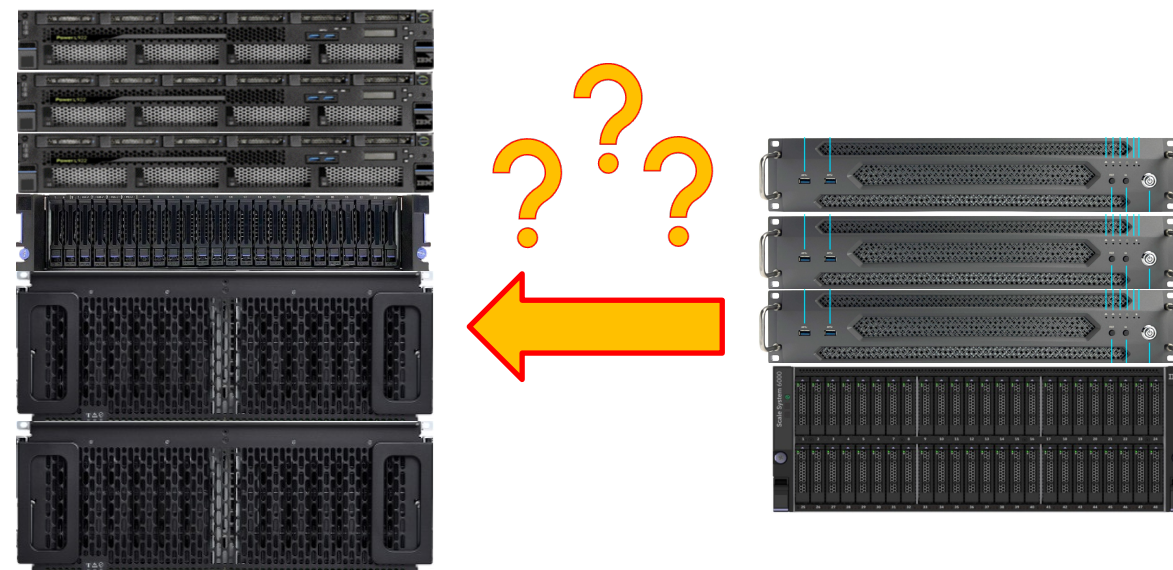
- Add SSS 6000 + Utility node to the existing p9 EMS cluster with BOTH EMS in the same cluster
- This means:
 - Storage Scale GUI running on both EMS
 - Utility Node will manage the 6000 “side” of the cluster while the P9 will manage the Legacy “side”
 - Define health monitoring and collectors on two nodes (one per EMS)
 - Reasonable restrictions regarding call home (ESA on p9 EMS/holistic on x86) and GUI functionality
 - Tiers/Pools/policies up to the user. Normal scale-based operations.



Integrating SSS6000 with Legacy Deployments – Long Term

The long-term solution will enable:

- Conversion of p9 EMS to a single EMS Utility node, after which the p9 EMS would be decommissioned or repurposed
- Single management node (Utility), Single GUI, Collector
- This support to be extended to all supported ESS types to the Utility Node, including:
 - p8 IO, p8 protocol, 5K,
 - p9 protocol, 3200, 3500, 5000
- Call home fully holistic (no ESA)



Integrating SSS6000 with Legacy Deployments - Restrictions

- Existing p9 EMS environments must be upgraded to 6.1.9.1 before the Utility Nodes and SSS6000 can be added
- P8 EMS environments not in plan for coexistence until May 2024
- No current BYOE coexistence support
- Dual EMS (p9) under consideration for coexistence support
- Initial support for conversion will only support single EMS
- BYOE to only support ESS 3500/VM for foreseeable future.




Comparative Performance Numbers

The SSS 6000 demonstrates significant performance improvements over the ESS3500

- Faster processor/more cores
- More memory
- More SSD/NVMe drives
- Next gen network adapters/more ports

Note:



- These are not official benchmark testing numbers!! 
- They are merely designed to showcase the performance improvement within the brand
- Evaluation system and parameters are shown below

SSS 6000 Configuration Evaluated	
NVMe Drives	48x 3.84TB NVMe Tier-1 Flash
Memory per canister	1536GB (24x 64GB dimms)
IB adapters per canister	4x CX7 adapters, each with 2x 200Gb ports
Clients	10x clients each with 2 cpu chips/48 cores and 2x CX7 400Gb/s IB 1-port adapters
Switch	NDR 400Gib MQM9700 - Infiniband

Comparative Performance Numbers – Overall Performance

Using:

- A “maxed out” Config
- Sequential Bandwidth
- 8+2p RAID
- 16MB filesystem block size

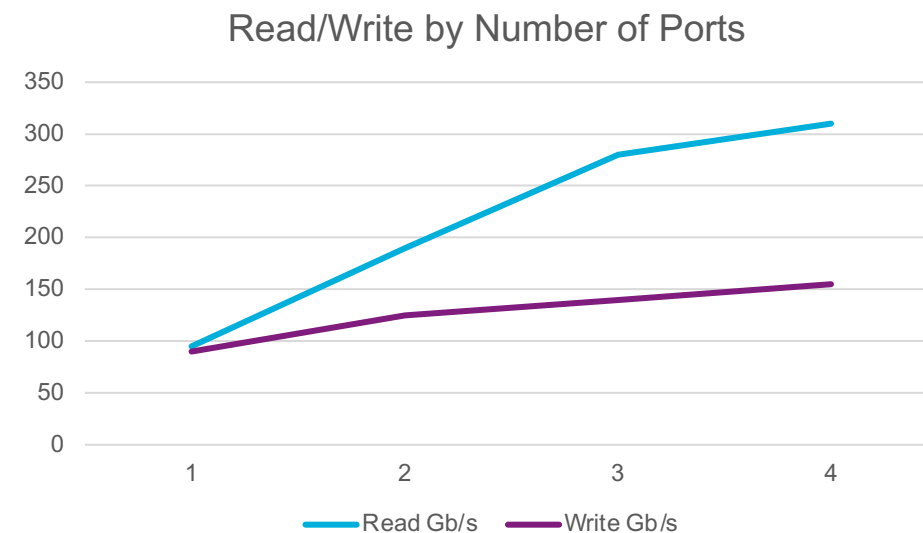
	ESS 3500	SSS 6000
NVMe Drives	24	48
Raw Drive Read BW (Read: 6.8 GB/s per drive) – fio benchmark	160 GB/s	320 GB/s
IB Ports per canister	4x 200Gb	8x 200Gb
IB RDMA Network Bandwidth NSDPerf benchmark	200 GB/s	400 GB/s
Memory Bandwidth AMD Stream benchmark	282 GB/s	1,400 GB/s
Measured Bandwidth from Clients IOR Sequential IO 16M 8+2p File System	Read 126 GB/s Write 60 GB/s	Read 310 GB/s Write 155 GB/s

Comparative Performance Numbers – Varying Bandwidth

Using:

- Scaling out the Infiniband Adapters
- Sequential Bandwidth
- 8+2p RAID
- 16MB filesystem block size

SSS 6000 IB adapters (2x 200Gb ports each)	Read GB/s	Write GB/s
1	95	90
2	190	125
3	280	140
4	310	155



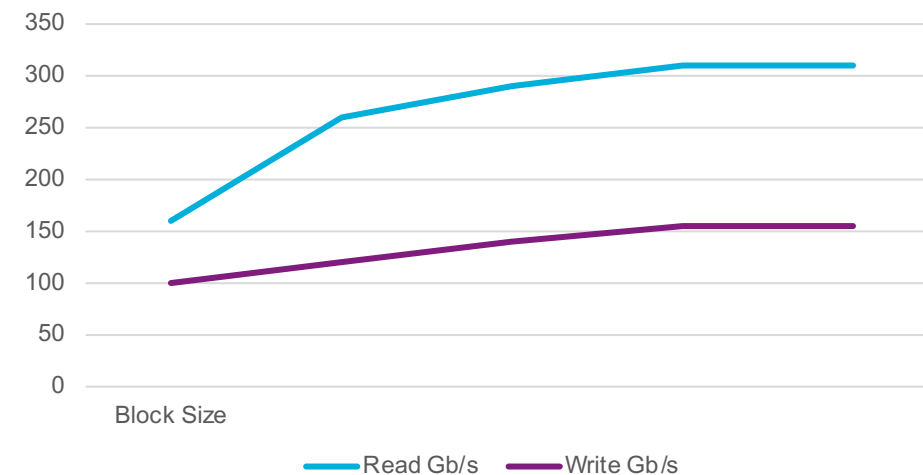
Comparative Performance Numbers – Various Block Sizes

Using:

- Varying filesystem block sizes
- Sequential Bandwidth
- A “maxed out” Config
- 8+2p RAID
- 16MB filesystem block size

SSS 6000 Filesystem Block Size(MB)	Read GB/s	Write GB/s
1	160	100
2	260	120
4	290	140
8	310	155
16	310	155

Read/Write by File Block Size



Thank you!

Accelerate with ATG Survey

Please take a moment to share your feedback with our team!

You can access this 6-question survey via [Menti.com](https://www.menti.com/join/17086924) with code 1708 6924 or

Direct link <https://www.menti.com/alwhyze7z1gz>

Or

QR Code

