Washington Systems Center - Storage

Spectrum Virtualize 8.3.1 Update

Accelerate with IBM Storage:

Byron Grossnickle – WSC Spectrum Virtualize SME
Matt Key – WSC FlashSystem SME

Accelerate with IBM Storage Webinars

The Free IBM Storage Technical Webinar Series Continues in 2020...

Washington Systems Center – Storage 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.

Located on the Accelerate with IBM Storage Site:
https://www.ibm.com/support/pages/node/1125513

Also, check out the WSC YouTube Channel here:
https://www.youtube.com/channel/UCNuks0go01_ZrVVF1jgOD6Q

2020 Webinars:

January 7 – What's New in IBM Spectrum Discover V2.0.2
January 14 – Brocade / IBM b-type SAN Modernization
January 21 – Cisco / IBM c-type SAN Analytics
February 20 – Spectrum Virtualize 8.3.1 Enhancements

April 9– GDPS V4.3 and SPEs Update
Register Here: Coming soon.....

Agenda

- Product Positioning/Hardware – Byron Grossnickle
- New Drives (NVMe, FCM2, SCM) – Matt Key
- STAT Tool in GUI – Matt Key
- Spectrum Virtualize for Public Cloud Update - Byron
- Performance
- 3 Site Coordinated Metro/Global Mirror
- DRAID Expansion
- Volume Protection
- HyperSwap Volume Expansion
- Miscellaneous Enhancements
  - SNMPv3 - 8.3.0.1
  - Enhanced Audit Logging – 8.3.0.1
  - NVMe-oF/RDMA Enhancements
  - Secure Data Deletion
  - Licensing/SCU Changes

Product Family Positioning

More is not better… but actually worse

Storage vendors use different platforms for different requirements

High-end platform	→ *Next gen* High-end platform
Midrange platform	→ *Next gen* Midrange platform
Entry platform	→ *Next gen* Entry platform

"Sustaining Innovation"

…but each platform adds complexity

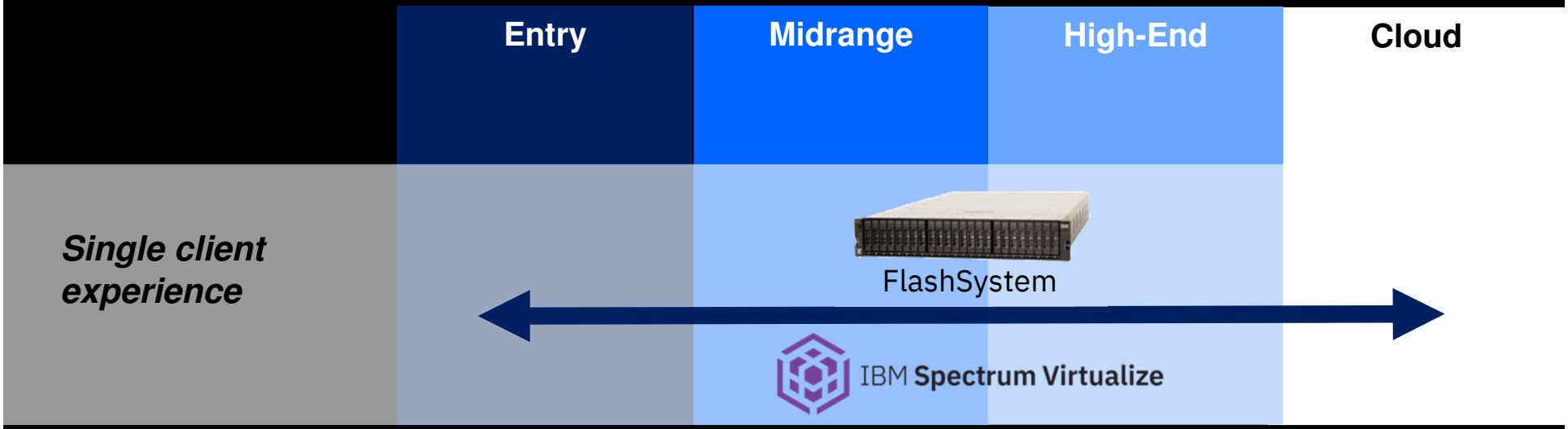
- *Different management and troubleshooting*
- *Different APIs and automation*
- *Different paths to cloud*

↑ **Innovation**
↑ **Complexity**
↑ **Cost**

Platform Complexity

	Entry	Midrange	High-End
Dell-EMC ***Five*** **product families**	 PowerVault  SC Series	 Unity XT  XtremIO	 PowerMax
HPE **Four product families**	 MSA	 Nimble  3Par	 Primera
NetApp **Four product families**	 EF	 SolidFire  FAS/AFF	 AFF A-Series
Pure **Two product families**	 C Series	 X Series	
Hitachi Vantara **Three product families**	 G Series  F Series	 	 5000 

Storage made Simple



↑ **Innovation** Leadership in all-flash arrays at each price and performance target

↓ **Complexity** Common data management across portfolio and to the cloud

↓ **Cost** Extend the benefits of IBM Storage to your entire storage estate

One Platform

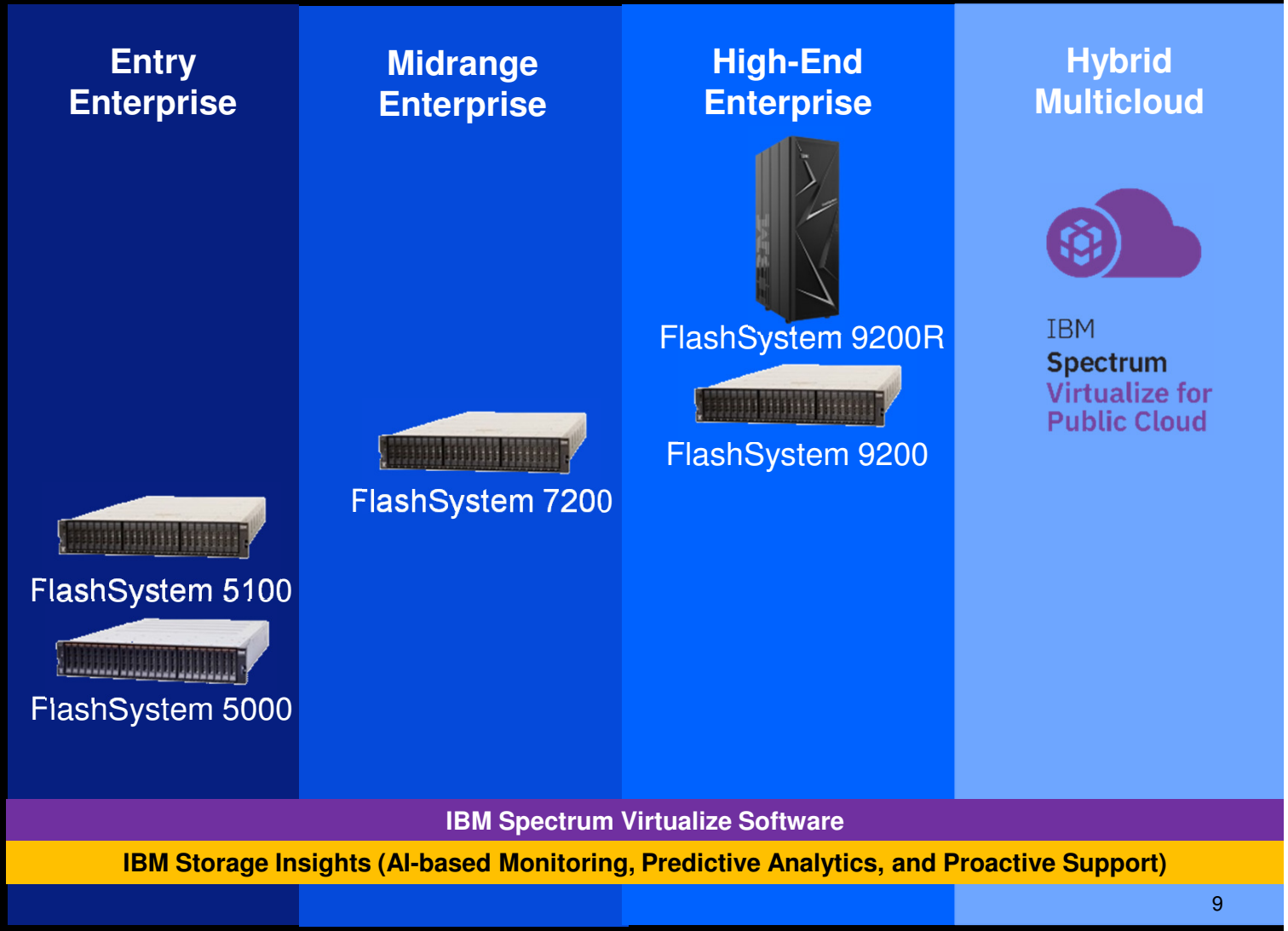
Operations made simple with IBM Storage Insights

Storage made simple with IBM Spectrum Virtualize

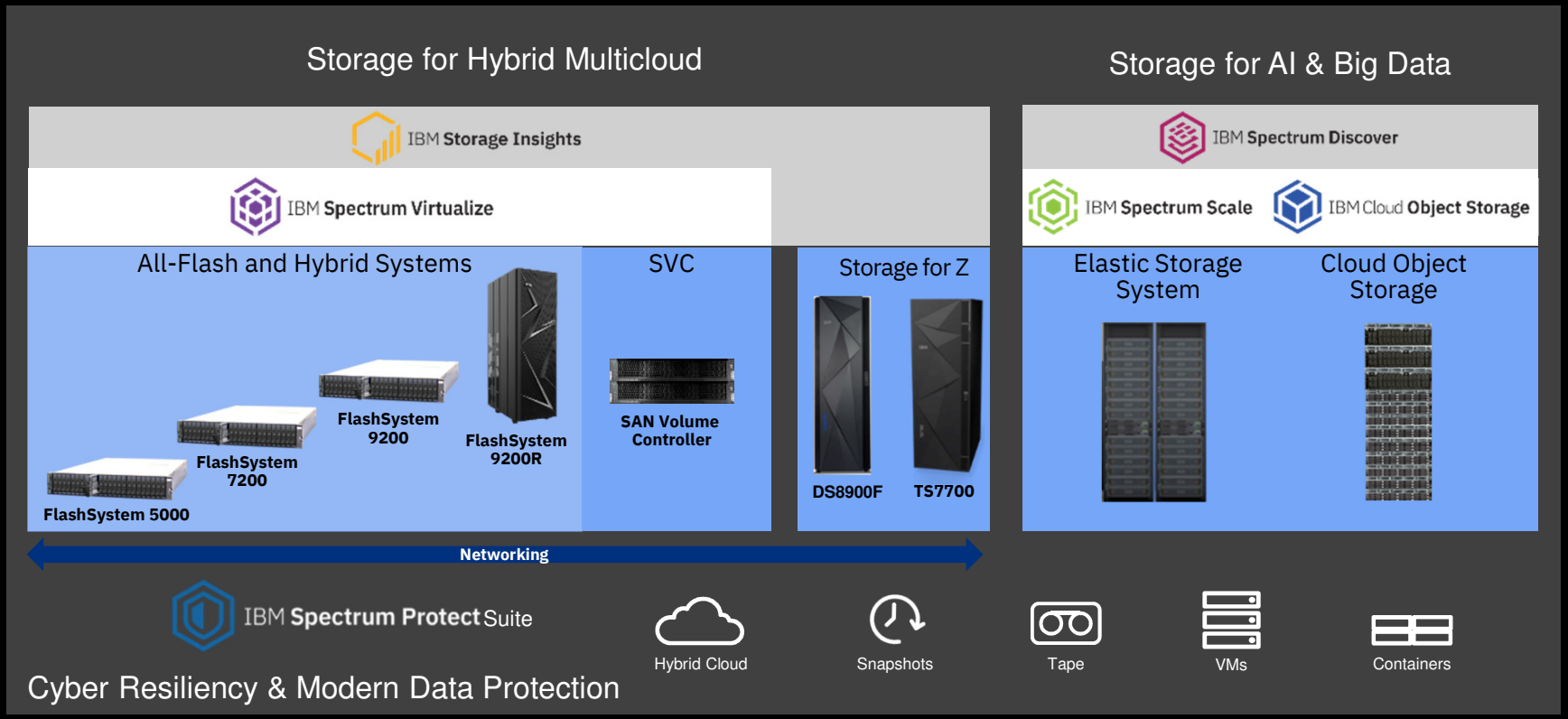
Entry Enterprise storage	Midrange Enterprise storage	High-end Enterprise storage	Storage Virtualization	Storage software for Public Cloud
				public

IBM FlashSystem Family

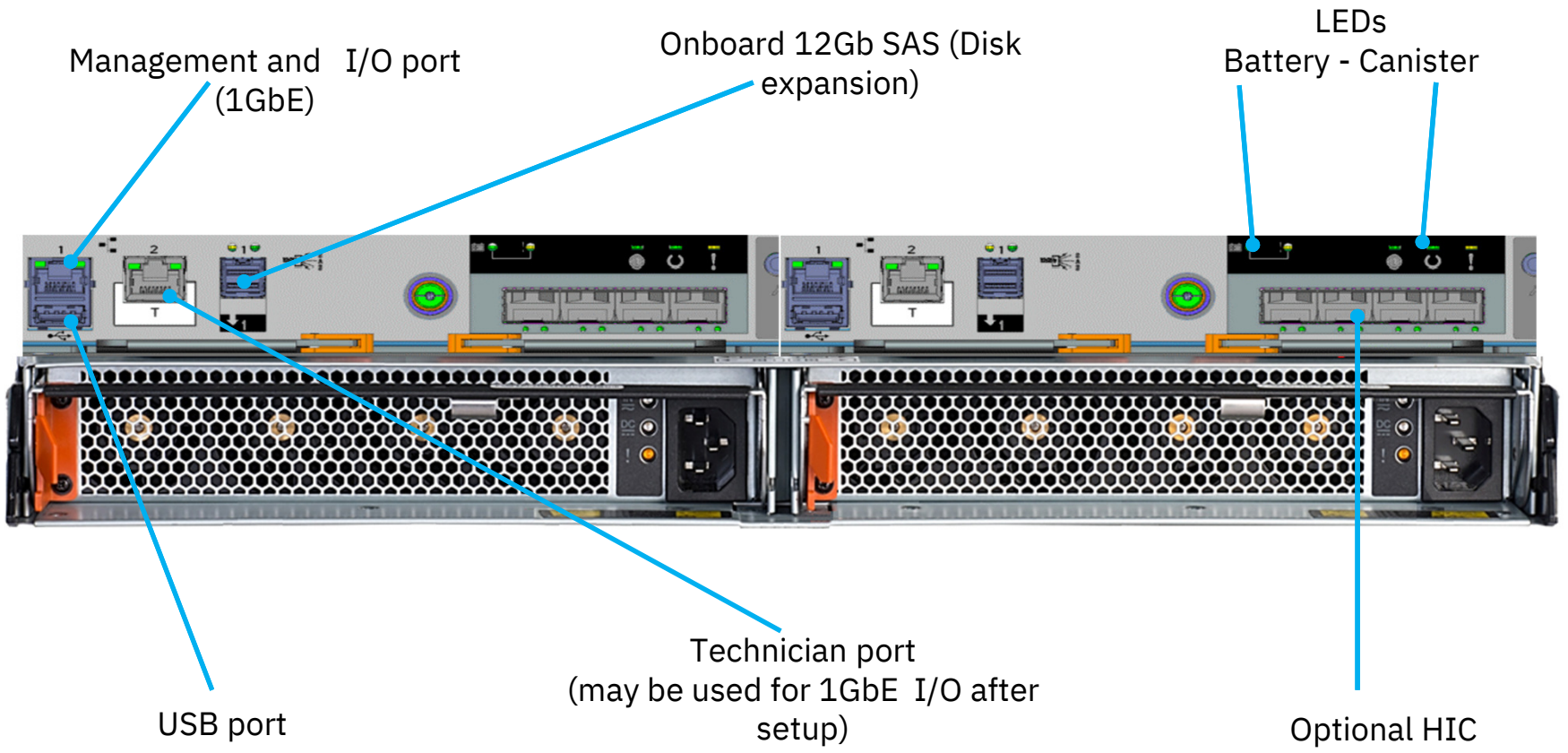
Storage made simple for hybrid multicloud



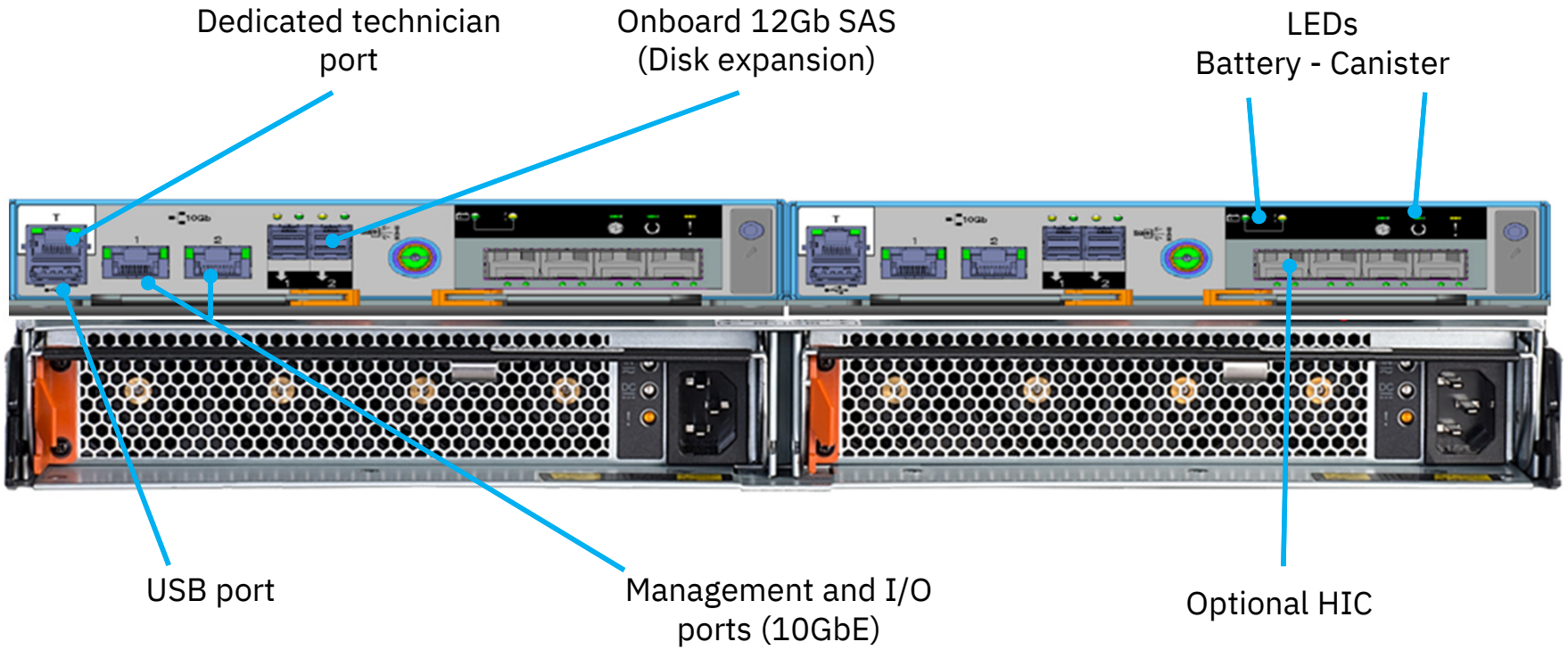
IBM Award winning Storage Portfolio



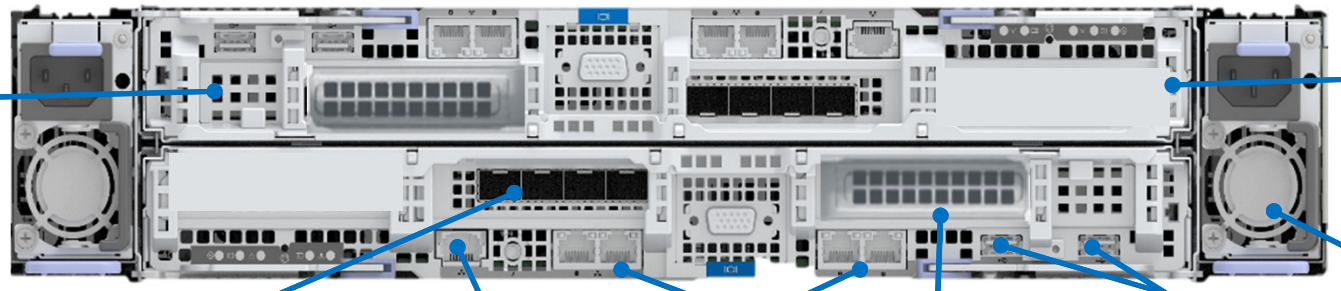
IBM FlashSystem 5010 (2072-2H2,2H4)



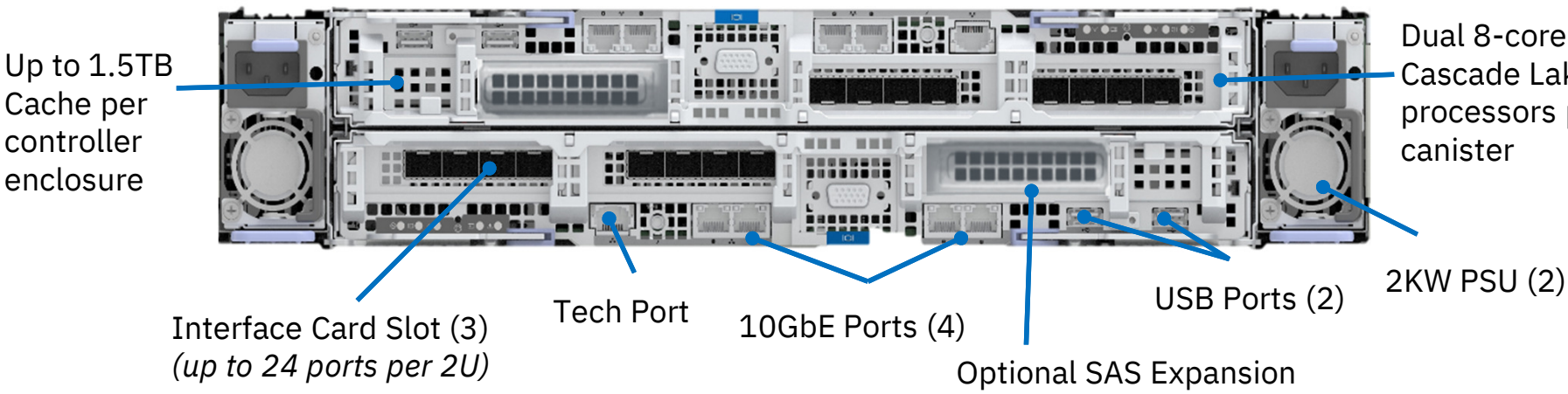
IBM FlashSystem 5030 (2072-3H2,3H4)



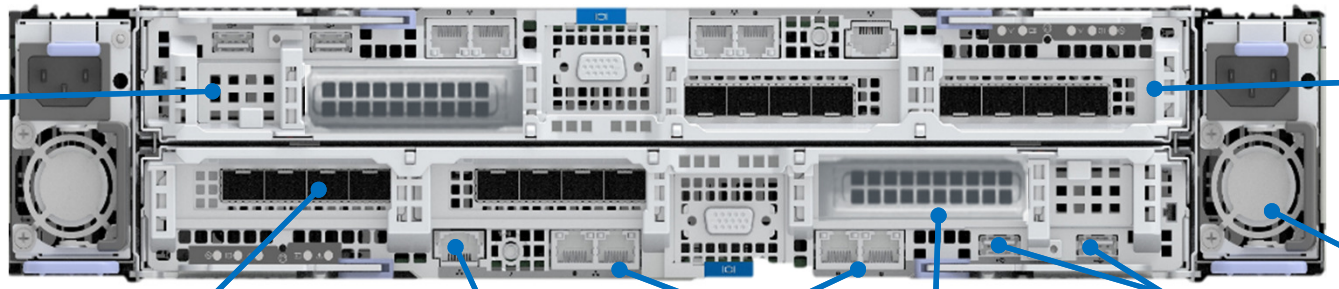
IBM FlashSystem 5100 (2077/2078-4H4,UHB)



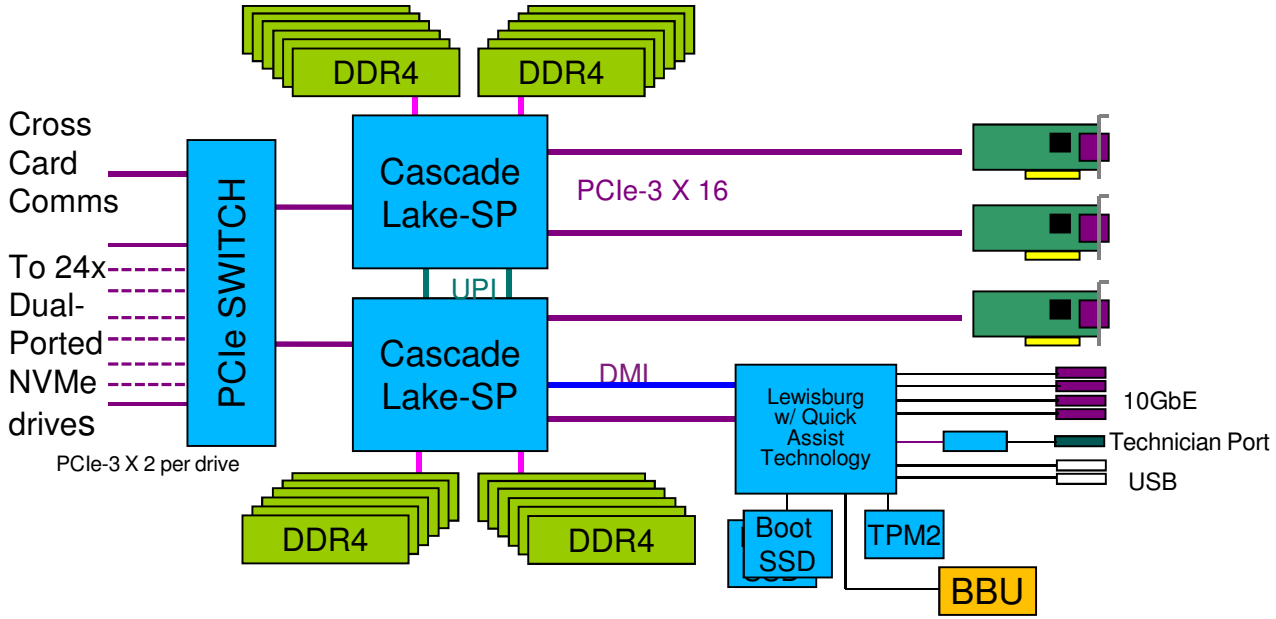
IBM FlashSystem 7200 (2076-824,U7C)



IBM FlashSystem 9200 (9846/9848-AG8,UG8)



Internal Diagram



Note: FlashSystem x1xx models are Skylake processors. FlashSystem models with x2xx are Cascade Lake processors.
Note: The FlashSystem 5100 only has a single processor

SAS Expansions

2U 3.5” x 12 SAS enclosure
(Hybrid only)

2U 2.5” x 24 SAS enclosure

5U 3.5” x 92 SAS enclosure

Adapter Options

	9150, 7200, 9200
Slot 1	4x16Gb FC 4x32Gb FC 2x25Gb Ethernet (RocE) 2x25Gb Ethernet (WARP)
Slot 2	4x16Gb FC 4x32Gb FC 2x25Gb Ethernet (RocE) 2x25Gb Ethernet (WARP)
Slot 3	**4x16Gb FC** **4x32Gb FC** **2x25Gb Ethernet (RocE)** **2x25Gb Ethernet (WARP)** **2x12Gb SAS Expansion**
Port Maximums Per Enclosure	24 Fibre Channel 12 Ethernet 4 SAS Expansion

Processor Comparison

5010	5030	5100	7200	9200	SVC SA2	SVC SV2
2 core 2.2GHz (Broadwell DE)	6 core 1.9GHz (Broadwell DE)	8 core 1.7GHz (Skylake)	8 core 2.1GHz x4 (Cascade Lake)	16 core 2.3GHz x4 (Cascade Lake)	8 core 2.1GHz x2 (Cascade Lake)	16 core 2.3GHz x2 (Cascade Lake)

	FlashSystem 5010	FlashSystem 5030	FlashSystem 5100	FlashSystem 7200	FlashSystem 9200	FlashSystem 9200R
All-Flash and/or Hybrid	AF ✓ - H ✓	AF ✓ - H ✓	AF ✓ - H ✓	AF ✓ - H ✓	AF ✓ - H ✗	AF ✓ - H ✗
Max cache per control enclosure	64GB	64GB	576GB	1.5TB	1.5TB	1.5TB x 4
Host adapter slots per control enc.	2	2	2	6	6	6 x 4
Storage Class Memory support	No	No	Yes	Yes	Yes	Yes
NVMe SSDs and IBM FCMs	No	No	Yes	Yes	Yes	Yes
NVMe-oF support	No	No	Yes	Yes	Yes	Yes
SAS SSDs and SAS HDDs	✓ and ✓	✓ and ✓	✓ and ✓	✓ and ✓	✓ and ✗	✓ and ✗
Support for SAS devices	Yes – Control & Exp.	Yes – Control & Exp.	Yes – Expansion	Yes – Expansion	Yes – Expansion	Yes – Expansion
Max physical capacity raw in 2U control enclosure	720TB	720TB	921.6TB	921.6TB	921.6TB	921.6TB x 4
Max usable effective capacity raw in 2U control enclosure	720TB	Up to 3.6PB*	1.8PB** Or up to 4.6PB*	1.8PB** Or up to 4.6PB*	1.8PB** Or up to 4.6PB*	1.8PB** x 4 Or up to 4.6PB* x 4
Maximum capacity with clustering	NA	32PB (2-way)	32PB (2-way)	32PB (4-way)	32PB (4-way)	32PB (4-way)
Data Reduction	None	Software DRP	FCM (no impact) DRP (Hdw assist)	FCM (no impact) DRP (Hdw assist)	FCM (no impact) DRP (Hdw assist)	FCM (no impact) DRP (Hdw assist)
Installation and support	Customer set-up	Customer set-up	Customer set-up	Customer set-up ECS optional	IBM install, technical advisor and ECS	IBM install, technical advisor and ECS
Support for CSI and other multicloud	Yes	Yes	Yes	Yes	Yes	Yes

← All-Flash and Hybrid | All-Flash Only →

IBM SAN Volume Controller

SV2

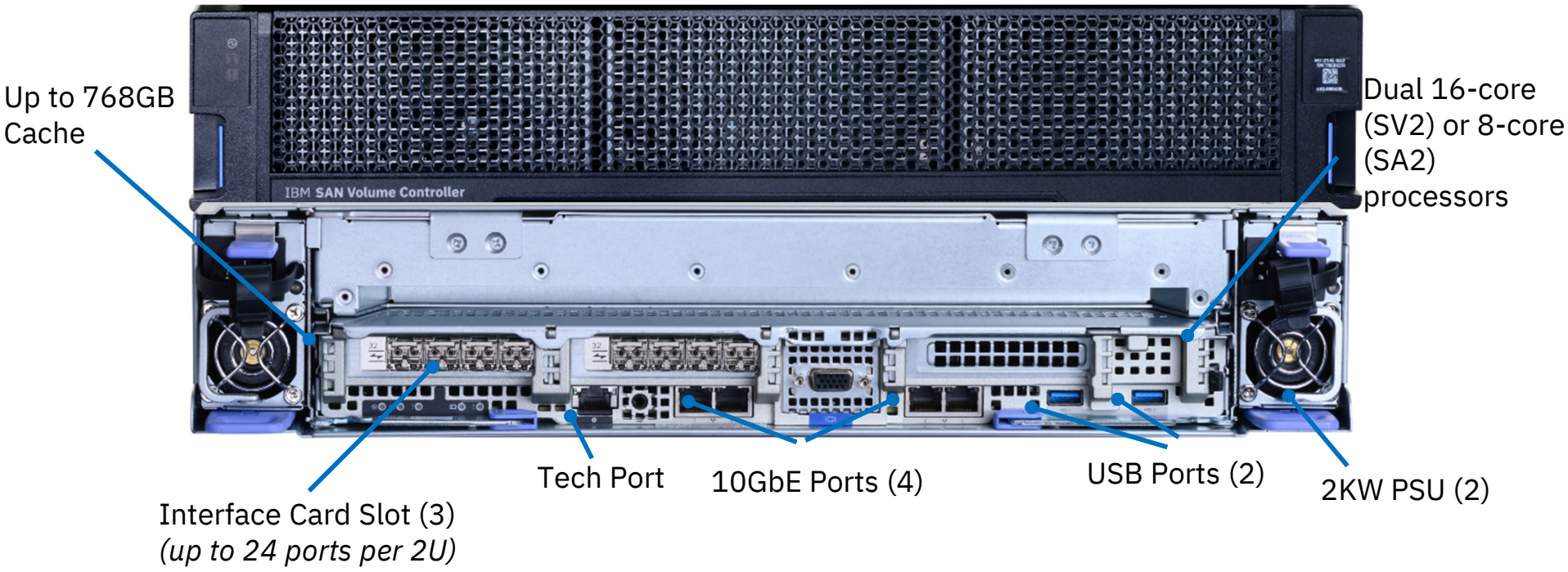
SV1

SA2

IBM Spectrum Virtualize

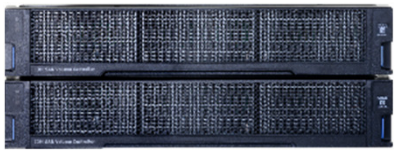
IBM Storage Insights

IBM SAN Volume Controller (SV2/SA2)

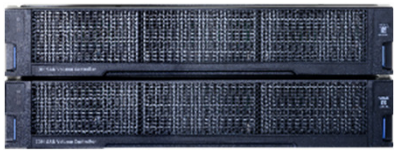


SAN Volume Controller Model Comparison

SV1



SV2



SA2

	Per Node	Per Node	Per Node
CPU	8 core 3.2GHz x2 (Broadwell EP)	16 core 2.3GHz x2 (Cascade Lake)	8 core 2.1GHz x2 (Cascade Lake)
Mem (min/max)	64GiB / 256GiB	128GiB / 768GiB	128GiB / 768GiB
Boot Drives	2 x external	2 x internal	2 x internal
Battery	2 x external	1 x internal	1 x internal
Compression	0-2 compression cards RACE or DRP	Built in (equivalent to 5 SV1 cards) **DRP only**	Built in (equivalent to 2 SV1 cards) **DRP only**
FC attach	Quad port 16Gbs	Quad port 32Gbs Quad port 16Gbs	Quad port 32Gbs Quad port 16Gbs
HIC Slots	4	3	3

Increased FC bandwidth + double PCIe bandwidth + huge cache increase = improved latency and IOPs

IBM SVC Memory

Per Node	SV2		SA2	
	Memory	Memory Bandwidth	Memory	Memory Bandwidth
Minimum Memory	128 GB	33%	128 GB	33%
Upgrade 1	384 GB	100%	384 GB	100%
Upgrade 2	768 GB	100%	768 GB	100%
Upgrade 3	n/a		n/a	

SAN Volume Controller SV2/SA2 Adapter Options

	SV2/SA2
Slot 1	4x16Gb FC 4x32Gb FC 2x25 Gb Ethernet (RocE or iWARP)*
Slot 2	
Slot 3	
Port Maximums	12 Fibre Channel Ports 6 Ethernet Ports

* iWARP and RocE are 2 different Interface Cards

Clustering Rules

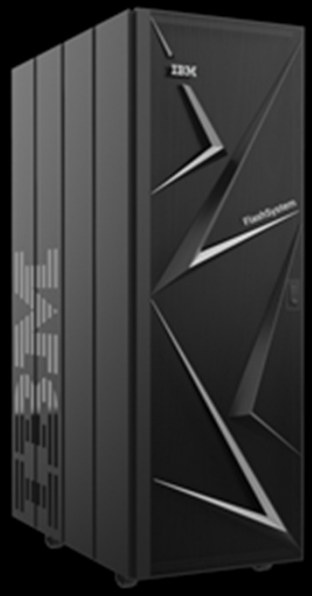
- SV2/SA2 can cluster with any SV1 and DH8 SVC clusters running the minimum level of software (8.3.1) up to 4 IO Groups

- Storwize V7000 (model 524, 624 and 724) can cluster with FlashSystem 7200, 9100, and 9200 running the appropriate minimum level of software up to 4 IO Groups
- FlashSystem 5100 can cluster with Storwize V5100 up to 2 IO groups
- FlashSystem 5030 can cluster with Storwize V5030 up to 2 IO Groups
- FlashSystem 5010 supports only a single IO Group

When Clustered, the name of the system changes in the following priority:

- Storwize -> FlashSystem
- V7000->7200 -> 9100 -> 9200

FlashSystem 9200R



FlashSystem 9200R

IBM FlashSystem 9200R

A Clustered FlashSystem 9200 Packaged in a Rack

2, 3 or 4 FlashSystem 9200s, clustered
9848 AG8 (no UG8)
Branded as 9202R, 9203R, 9204R, per number of 9200s

Optional expansions
9848 AFF 2U 24 drive and 9848 A9F 5U 92 drive options, SSD only

Dedicated fibre channel backbone
Isolated from host and copy service traffic
Choice of 2 each of 32Gb 8960-F24 Brocade or 8977-T32 Cisco switches

Packaged in 7965-S42 rack
FlashSystem branding
With dual PDUs appropriate for selected system

Assembled and delivered by IBM
Manufacturing places components in rack and ships to customer
IBM Lab Service engagement performs final signal cabling and integration into customer SAN if required

New Drive Technology
New Drives
STAT Tool in GUI

Matt Key

Storage Class Memory (SCM)

- A non-volatile media that sits between DRAM & commercial NAND Flash in performance (latency), density & endurance

- Spectrum Virtualize brings support for SCM in 8.3.1 for a use as a storage tier
- FlashSystem 5100/7200/9250, with Cascade Lake, set stage for SCM as extended caching capacity

- SV 8.3.1 Use Cases for SCM
  - EasyTier (tier_scm >> tier0_ssd >> tier1_ssd)
  - Storage Pool (exclusively SCM for highest perf)
  - DRP metadata
    DRP will automatically pull extents for exclusive use by metadata, typically index and reverse lookup volumes

SCM Offerings

Intel Optane (3D XPoint technology)

IBM Part #	01CM517	01CM518
Capacity	375 GB	750 GB
Endurance	30 DWPD	
Latency	~20us	
Base Supply Model	Intel Optane D4800X	

Samsung Z-SSD (Z-NAND or Low-latency NAND)

IBM Part #	01CM522	01CM523
Capacity	800 GB	1600 GB
Endurance	30 DWPD	
Latency	~20us	
Base Supply Model	Samsung SZ1735	

All SCM drives supported by IBM are:

Dual-ported PCIe 3.0 NVMe
Self-encrypting (standard TCG Opal specifications)
Very Fast

SCM Nuances

Due to restrictions in firmware, SCM drives may only be populated in the four (4) right-most slots (#21-24) of an NVMe enclosure

- Up to 16 drives per cluster following the above rule
- Error will alert if drive inserted into a slot below #21

Array restrictions apply to TRAID10 and DRAID5

- Smaller SSDs can be used as forced spare to SCM drive **Arrays perform as fast as slowest member **

No UNMAP support (it's not traditional NAND, little value) so host calls write 0's instead

Intel Optane drives may take 15min to format to become 'candidate' drives

FCM Update

Up to

38.4 TB Usable Capacity

Endurance

DWPD 2

Performance

Up to 50% Better Performance

over First Gen FCM

FlashCore

Gzip Compression

Consistency

Longevity

Adaptive

FCM Options

	Usable Capacity (TBu)	Max Effective Capacity (TBe)
Small	4.8	22
Medium	9.6	22
Large	19.2	44
XLarge	38.4	88

FCMs still require DRAID6 → Do not override this

Rebuilds are still 2.5TB/hr against TBe capacity

Small (4.8 TB) has only two dies per package

Medium, Large, Xlarge share same performance profile

Inside the new FCM

96 Layer 3D NAND

Mix of pSLC, QLC, and FlashCore

Dynamic SLC cache

Smart Data Placement

Same 2” U.2 Dual-Port NVMe form factor

Same 25W Power Max

Same Gzip Compression

Same Encryption

Runs a 2.x microcode

New FCMs can be a field replacement for First-Gen FCM

FCMs can be intermixed between generaions within a DRAID config

New FCM can exist in a 8.3.0 or lower code level**

**Largest 38.4 TB FCM needs SV 8.3.1

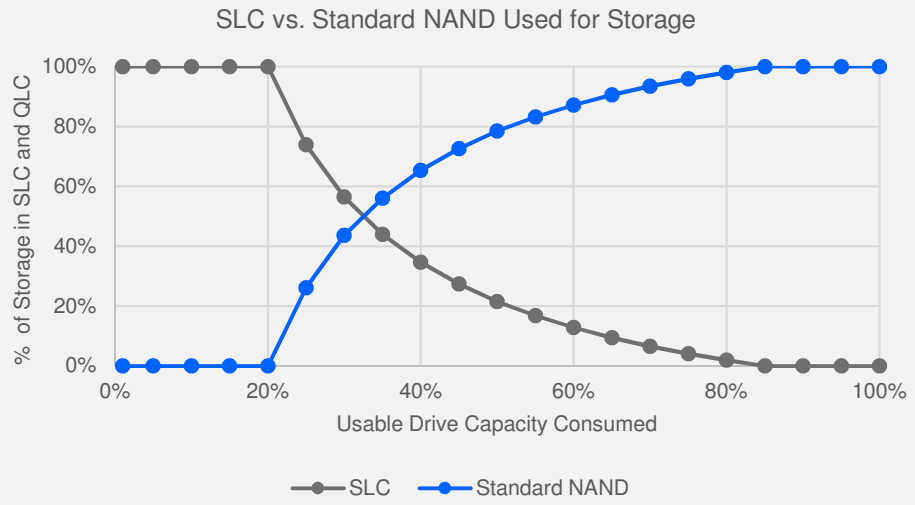
All drives look to DRAID6 for Variable Stripe RAID protection

Dynamic SLC Cache

Blocks within drive get reprogrammed to run as Single-Layer-Cell mode flash

- Highest Resiliency
- Highest Performance

- All writes land on SLC (after ACK by MRAM)
- Write-heavy workloads STAY on SLC tier
- SLC cache will ‘float’ across blocks to wear evenly



When 20% or less, then all data is on SLC

SLC will concede blocks to capacity storage upon demand

Following 80% guidance for capacity, performance will stay nominal

Above 80% Usage and has potential to show capacity-biased performance

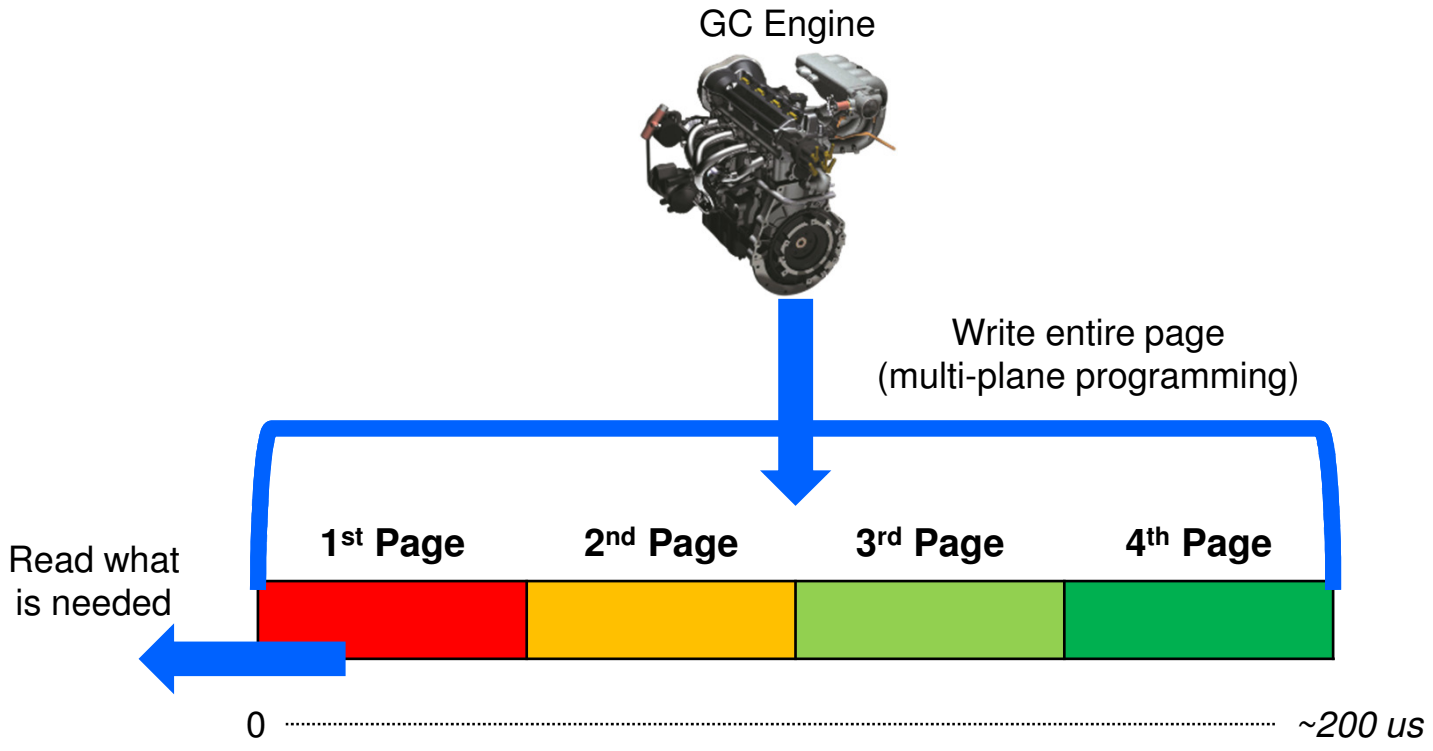
Smart Data Placement

During routine garbage collection, data is re-ordered onto capacity flash based upon read intensity

Read-intensive data stays on lowest page for lowest latency

Stages four different tiers of reads on capacity flash

Recently written data gets read from SLC



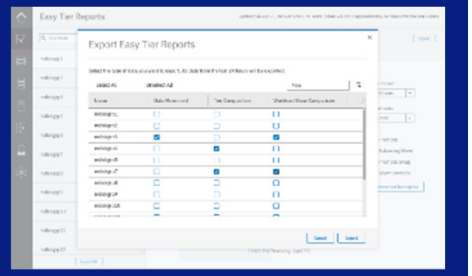
STAT Tool

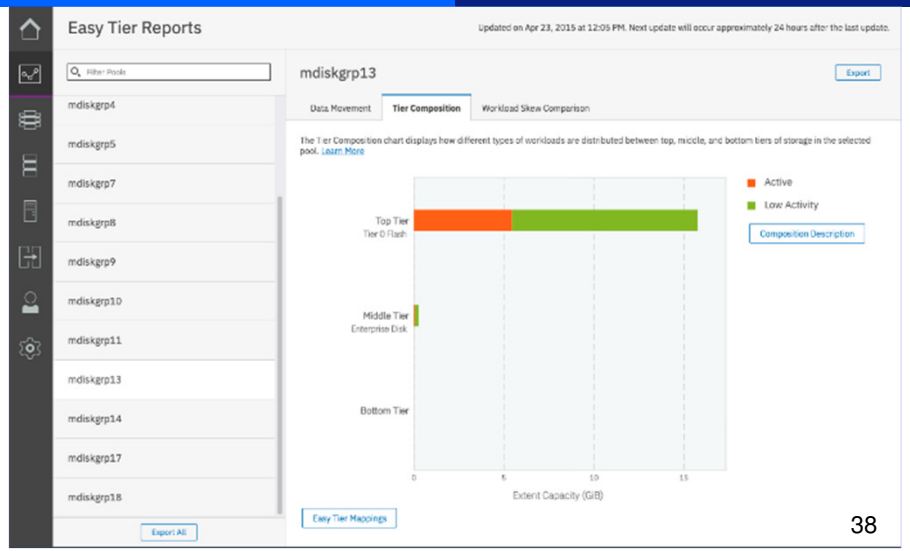
View & Export EasyTier data right from the GUI

No more dpa_heat file exports, parse utility, then STAT excel macro

Select individual pools to view & export

Currently for Standard Pools only
(inaccurate for DRP for volume or data view)





Easy Tier Storage Class Memory Support

Configuration	ET Top Tier	ET Middle Tier	ET Bottom Tier	Comments
Tier_SCM	Tier_SCM			
Tier_SCM + Tier0_Flash	Tier_SCM	Tier0_Flash		
Tier_SCM + Tier0_Flash + Tier1_Flash	Tier_SCM	Tier0_Flash	Tier1_Flash	
Tier_SCM + Tier0_Flash+Tier_Enterprise + Tier_Nearline				Not supported: ET will go into measure mode
Tier_SCM + Tier0_Flash + Tier_Enterprise / Tier_Nearline	Tier_SCM	Tier0_Flash	Tier_Enterprise/Tier_Nearline	
Tier_SCM + Tier0_Flash + Tier1_Flash + Tier_Enterprise /+ Tier_Nearline				Not supported: ET will go into measure mode
Tier_SCM + Tier1_Flash	Tier_SCM	Tier1_Flash		
Tier_SCM + Tier1_Flash + Tier_Enterprise/Tier_Nearline	Tier_SCM	Tier1_Flash	Tier_Enterprise / Tier_Nearline	
Tier_SCM + Tier1_Flash + Tier_Enterprise + Tier_Nearline	Tier_SCM	Tier1_Flash + Tier_Enterprise	Tier_Nearline	
Tier_SCM + Tier_Enterprise/Tier_Nearline	Tier_SCM	Tier_Enterprise / Tier_Nearline		
Tier_SCM + Tier_Enterprise + Tier_Nearline	Tier_SCM	Tier_Enterprise	Tier_Nearline	

+ = and / = or +/ = and/or

Spectrum Virtualize for Public Cloud

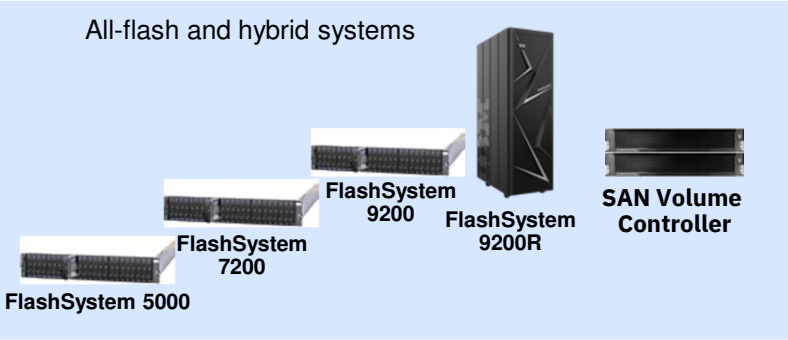
Byron Grossnickle

Positioning

Differences between Spectrum Virtualize & Spectrum Virtualize for Public Cloud

IBM Spectrum Virtualize

Spectrum Virtualize is the software that runs on premises in the IBM Storwize family, FlashSystem family, SVC (SAN Volume Controller).



* New option in SVC available March 6, 2020

IBM Spectrum Virtualize for Public Cloud

Spectrum Virtualize for Public Cloud is a separate product, with separate licensing terms and runs on supported public cloud infrastructure - IBM Cloud, and AWS today.

SVC/VSC SCU Capacity can be used for SV4PC with a simple one time registration license purchased. Use Category 1 SCUs

Technically, Spectrum Virtualize for Public Cloud is largely based on Spectrum Virtualize, but adapted to run on public cloud infrastructure, such as IBM Cloud and AWS EC2. Key differences:

- Clustering / Failover
- Shared Network Infrastructure
- Back end virtualization
- Deployment

8.3.1 Improvements in AWS

25% Capacity improvement per cluster

- Increased EBS Volumes per node, from 16 EBS to 20
- Increasing the total capacity of a cluster to 320TB

Twice the Nodes to Improve Performance

- Deploy up to 4-node cluster on AWS for improved scalability and performance
- Expand from 2 to 4 nodes or shrink from 4 to 2 node non-disruptively

Code changes to gain 20% IOPs improvement and 60% bandwidth improvement

- Configure Ethernet MTU as 9000 by default
- Multi-threaded I/O to EBS
- Improve CPU resource assignment

Data Reduction Pools (DRP) in Cloud

In Spectrum Virtualize 8.3.1 we enabled DRP which is an increased storage efficiency functionalities that will allow customers to:

- Automatically de-allocate and reclaim capacity of thin provisioned volumes that had data deleted.
- Reclaimed capacity can be reused by other volumes.
- Use the deduplication process to identify unique chunks of data, or byte patterns, and stores a signature of the chunk for reference when writing new data chunks.
- Depending on the data and workload, this can provide customers with the ability to reduce their monthly costs of EBS or IBM Cloud block storage.

Machine Type	DRP	DEDUP	Compress (DRP based)	Host unmap	Backend unmap
IBM Cloud bare metal	Yes	Yes	**No**	Yes	No
AWS EC2 c5.4xlarge	Yes	**No**	Yes	Yes	No
AWS EC2 c5.9xlarge	Yes	Yes	Yes	Yes	No
AWS EC2 c5.18xlarge	Yes	Yes	Yes	Yes	No

- Compression is CPU based
- Compression in only enabled in DRP
  - *RTC based compression is not available*

IBM Cloud Currency Update in 8.3.1

- Previously we only supported Intel Xeon E5-2620 or E5-2650 bare metal servers.
- In this release, you will be able to provision newer bare metal servers running Intel Xeon 5120 on the IBM Cloud for Spectrum Virtualize for Public Cloud.
- Previously Spectrum Virtualize for IBM Cloud only used 24GB RAM out of 64GB available RAM
- In this release, Spectrum Virtualize for IBM Cloud uses 58GB RAM out of 64GB available RAM
- The newer servers can be purchased through the IBM Cloud portal.

Spectrum Virtualize 8.3.1 Performance

Technical Update

February 2020

Olga Yiparaki, Chief Storage Performance Engineer
Brian Twichell, Lead for Distributed Storage Performance
Lee Sanders, Storwize Performance Team Lead

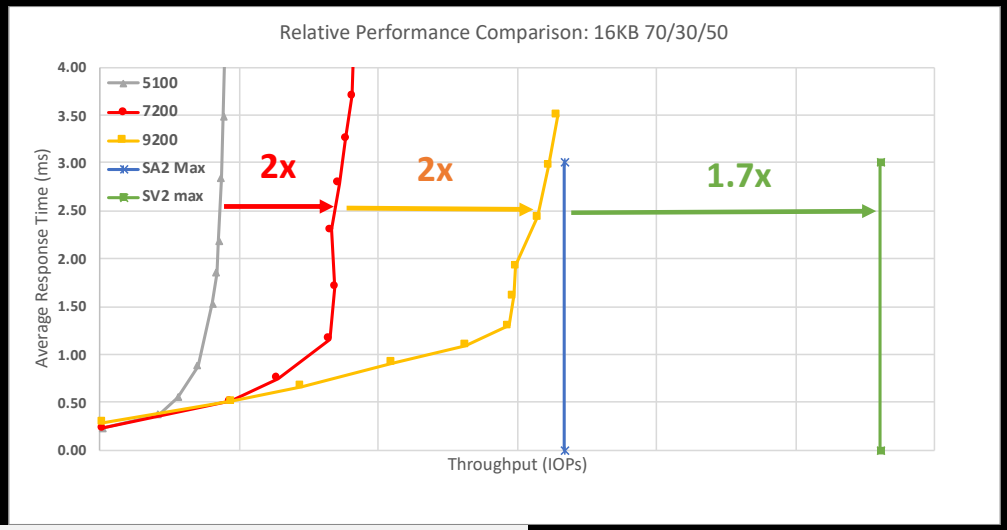
8.3.0 vs 8.3.1 DRP Performance

- Code improvements that have aided performance
  - Multi-chunk write support
  - Multi-threaded lower cache
    - Immense improvement in single volume performance
  - Internode messaging improvements

- Working Set: FS9150
  - 24 – 4.8 TB FCM modules
  - DRAID6
  - 768GB Cache (cache 3% of working set)
  - 25TB Working set
  - 67% capacity utilization AFTER data reduction
  - Compression 2.85:1, deduplication 2:1

Max Performance: New Platform At a Glance

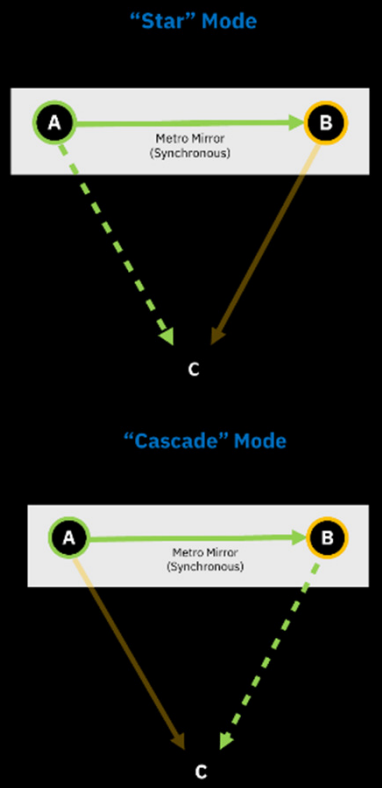
	Max IOPS	*Min Latency*	*Max Bandwidth*
	4K Read Hit	min latency for 4K Read Hit	Read BW, 256 KB reads from flash
FS 9200	**4.5 M**	**under 70**	**45 GB/s**
FS 7200	**2.4 M**	**under 70**	**35 GB/s**
FS 5100	**900 K IOPS**	**under 70**	**15 GB/s**
SV2	**4M**	**under 70**	**49 GB/s**
SA2	**2.4 M**	**under 70**	**33 GB/s**



Fully Allocated volumes, max performance shown on this page.
Performance shown is achievable and sustainable long-term (not a burst)
Capacity: Physical capacity utilization on the flash drives around 80%.
With FCM drives and drive compression ratio 2.85:1.

3 Site Coordinated Metro/Global Mirror

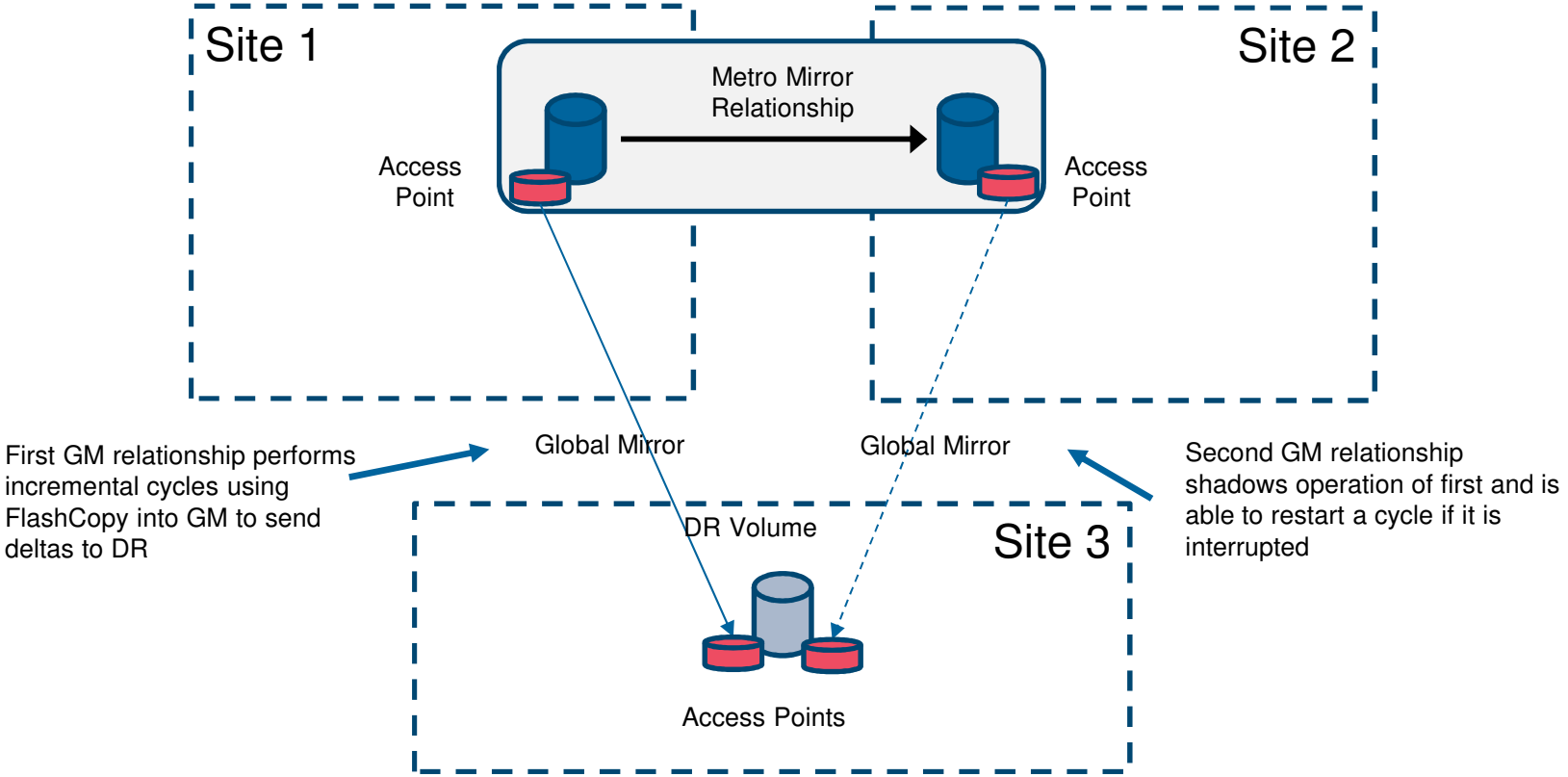
Byron Grossnickle



3 Site Metro/Global Mirror Timeline

- Three site coordinated Metro/Global Mirror will be available by SCORE(RPQ) 3/20/2020 with version 8.3.1.2
- Version 8.3.1.3, which will happen sometime in the 2nd or 3rd qtr will provide the function for GA

Metro/Global Mirror
How Does it Work?

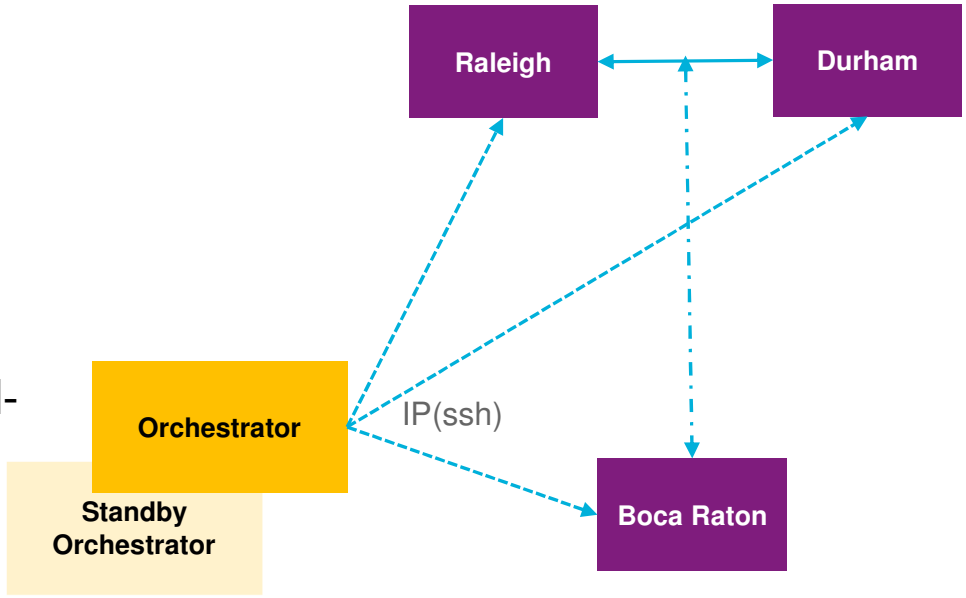


3 Site Metro/Global Mirror Overview

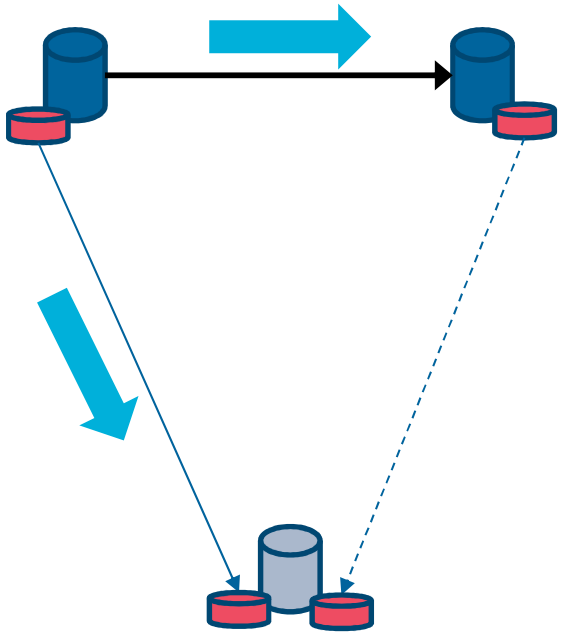
- You may hear this referred to a 3-Site Replication in other presentations
  - While “technically” true this is NOT 3 site replication, meaning any means of replication between 3 sites
- Provides a coordinated Metro/Global Mirror function to Spectrum Virtualize
  - Existing MM implementations can be converted non-disruptively
- Coordinated Metro/Global Mirror is orchestrated by a small RHEL 7.5 (or above) instance running on a standalone server or VM
- Synchronous mirror between 2 near sites with a 3rd far site
  - Standard MM restrictions apply for near sites
  - Standard GM restrictions apply for far site
  - Uses a methodology similar to GMCV
- Far site copy can be maintained from either the main site or the near site without the need to restart the mirror from scratch
- Will function between all Spectrum Virtualize products with a remote copy license
  - Layer between all systems must be the same (replication/storage)

How is the solution actually deployed?

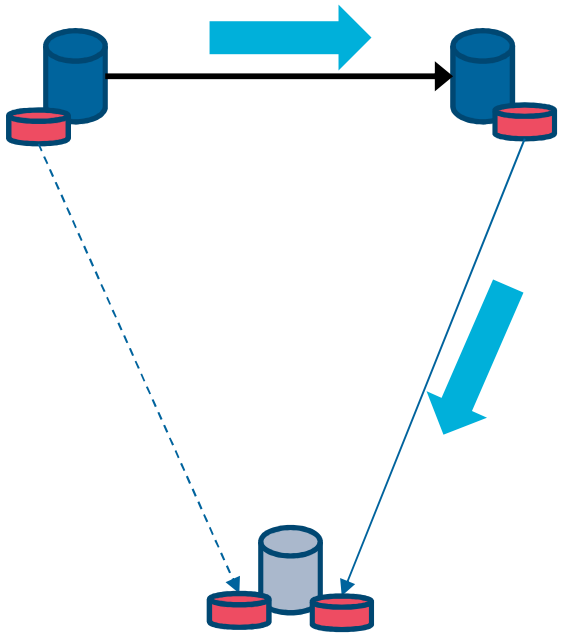
- Three systems with FC partnerships between them
- Orchestrator hosted on RHEL installation
- Provides overall monitoring plus end-user command line to monitor
- Alerts sent via existing systems



Star or Cascade



Star



Cascade

3-Site Metro/GM Prerequisites

- **Spectrum Virtualize systems**
- 3 Spectrum Virtualize systems with version 8.3.1.2 acting as 3 sites
  - Storage layer must be the same between systems (replication/storage)
- Fibre Channel remote copy partnerships established between all 3 sites
- MM remote copy consistency groups setup between near sites
- 3-site user with role ‘3SiteAdmin’

- **3-Site Orchestrator host system**
- A single, standalone Linux host running RHEL 7.5 (or higher)
- IP connectivity from orchestrator to all 3 sites
- Public/private keypair SSH established between orchestrator host and all 3 sites for 3-site user
- Firewall configured to allow orchestrator ports to establish ssh tunnel with 3 sites
- Orchestrator host should be preferably setup on a location different from 3 sites or at Auxiliary Far site

- **3-Site Orchestrator packaging and installation**
- This will be made available as a Linux RPM on IBM Fix central
  - When GA, until then it will be provided as part of the SCORE/RPQ
- Administrator will install this package on Linux RHEL host. This installs 3-Site services on orchestrator host
- A detailed 3-site configuration and CLI usage guide will be available

Metro/GM Limitations

- Initially Metro/GM replication is available as a SCORE/RPQ
- Spectrum Virtualize GUI does not support setting up, managing and monitoring 3-Site replication
  - Orchestrator is CLI only
- Auxfar/Tertiary site can not be used as primary/production site and/or as source for periodic replication in 3-Site configuration
- **Does not support HyperSwap between two near sites**
  - **In roadmap plan**
- Following operations are not supported on volumes which are part of 3-Site replication : restore and incremental fcmap.
- 3-Site replication is not supported over IP partnership
- Only a single 3-site orchestrator host instance can be active at a time
- Standalone relationships are not supported in 3-site
- MetroMirror Consistency Protection is not supported
- 16 consistency groups supported with up to 256 relationships in each group
- Minimum cycle period for far site replication is 5 minutes and RPO is 10 minutes

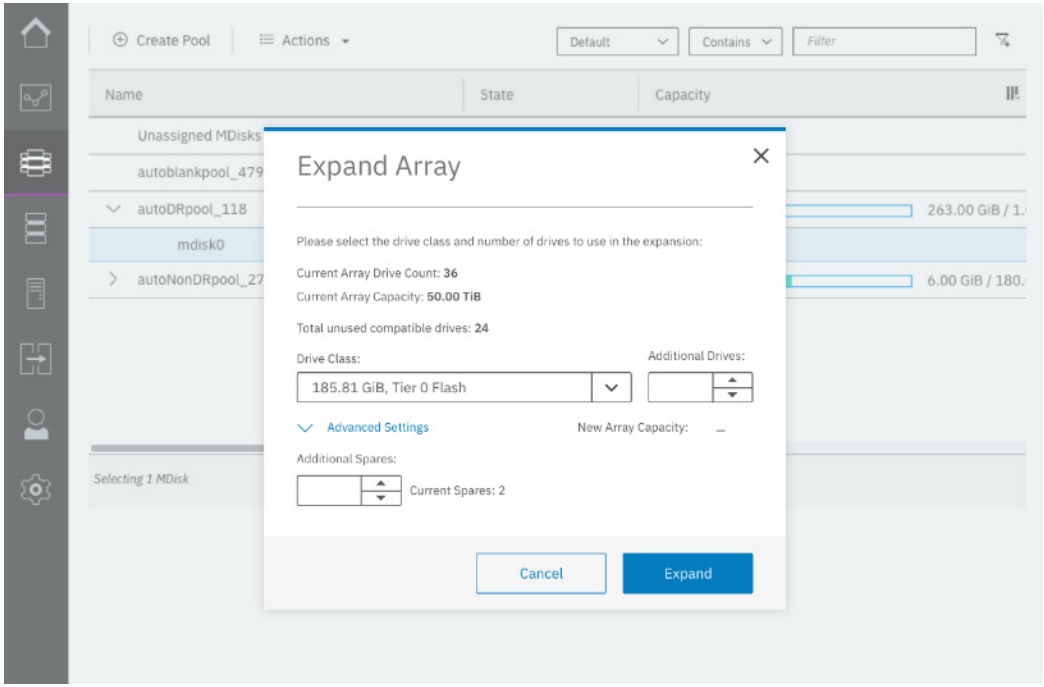
Distributed RAID (DRAID) Expansion

Byron Grossnickle

Distributed RAID Array Expansion – What is it?

- Ability to add one or more drives to an existing DRAID array non-disruptively
- New drives are integrated into array and existing data is re-striped to maintain DRAID efficiencies
  - Final result is the same as if that array had been created from scratch
  - Stripe size does not change

Array Expansion



What Drives Can Be Added?

- Ideally: Same as existing drives!
  - This means no performance bottlenecks
- There is allowance for cases where that is not possible
  - Similar to using the allow superior flag when creating arrays
  - Must be same or larger capacity
  - Extra capacity is not used
  - Compressing arrays must use drives of the same capacity
- Must be same or better performance
  - Though in reality we expect same performance class is preferred

How Many Drives/ How Long?

- Between 1 and 12 in any one expansion process
  - Adding >12 drives can be done with a series of expansions
  - When adding lots of drives, it might be quicker to create a new array, migrate the data and then expand the new array to incorporate the old drives
- A drive can either add extra capacity or extra rebuild areas
  - Recommended to stick with 1 rebuild area for every 24 drives in the array
  - Does not apply to FCMs
- Length of expansion depends on a number of factors
  - Expect between many hours (esp. NVMe drives) and many days (NL-SAS)
- Goal is to run in the background without impacting host IO performance
  - Affected by:
  - Drive capacity – but also how much is filled with data
  - Host IO workload – and its distribution
- Estimated completion time in CLI output
  - lsarrayexpansionprogress
  - Seen as long running task in the GUI

Stripe Size Will Not Change

- This means that your overall percentage of usable space will stay the same
  - Can be a challenge for smaller customers that purchased 6 drives.
    - In this case the usable capacity is 50%. After expansion (of any number of drives) the overall usable capacity will still be 50% of the total
- If purchasing a larger number of drives, consider making a new DRAID array with a larger stripe size. Then take the old DRAID array out of the pool and expand the DRAID array with the larger stripe size

Other Points of Consideration

- Array rebuild takes priority over expansion
  - If a drive fails mid-expansion, expansion is suspended to allow rebuild to run unimpeded
  - Expansion will resume after rebuild process is complete
- Expansion takes priority over copy-back process
  - If a failed drive is replaced mid-expansion, the expansion continues
  - Copy-back process will be started after expansion
  - Copy-back can only be performed after drive initialization is complete
    - Same serialization takes place on array creation
- DRAID-6 is now default array type recommended by GUI for 6 drives or higher
- Expansion cannot be cancelled or reversed
  - Deleting an array suspends expansion while data migrated off
- Work area will consume 5 extents in most cases and this area will be reserved until the expansion is finished

Volume Protection/HyperSwap
Volume Expansion

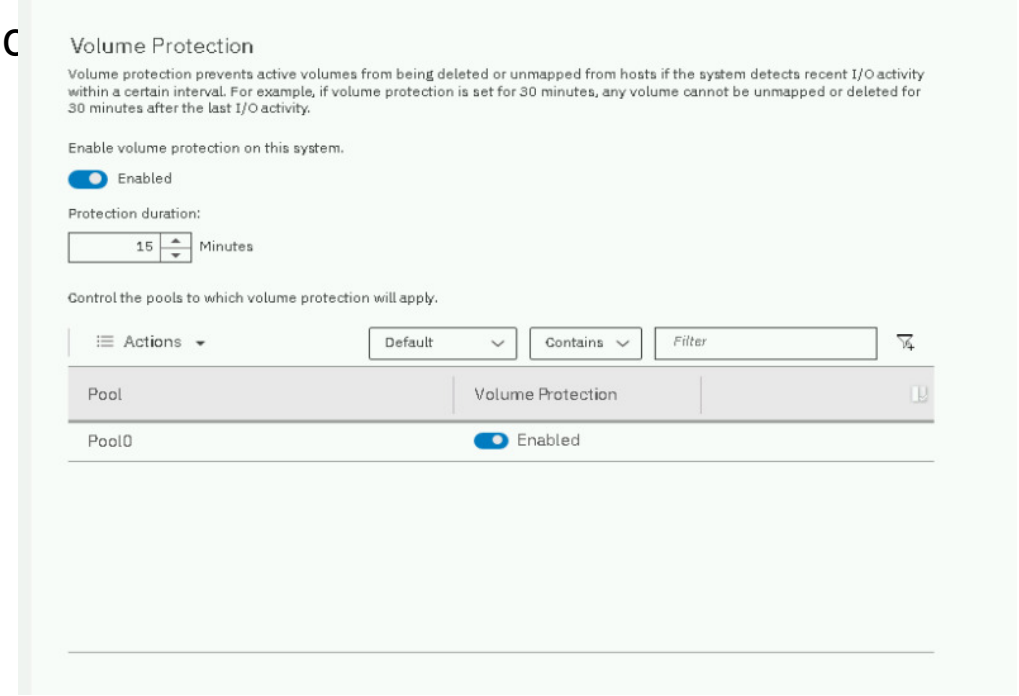
Byron Grossnickle

Volume Protection

- Volume protection is a feature that has been available since 7.4 code and by default is set to off.
- In 8.3.1 on new systems, volume protection will set to on by default
  - Existing systems that are upgraded will remain the same unless upgrading from a system prior to 7.4, in which case it will be on after the upgrade
  - Volume protection can be enable/disabled at a system level or at an individual pool/child pool level
  - When creating a new remote copy relationship, the target volume now falls under the protection of volume protection
- Feature is now available in the GUI

Volume Protection

- System/Settings/Volume Protectic



Volume Protection

vdisk_protection_enabled (system)	vdisk_protection_enabled (pool)	vdisk_protection_status (pool)
yes	yes	active
yes	no	inactive
no	yes	inactive
no	no	inactive

Other Changes to Enhance System/Data Protection

- The system will enforce SCSI LUN ID parity between all I/O groups
  - Previously this had to be done manually and if the SCSI LUN ID was different between 2 I/O groups it could confuse certain OS's
  - A volume assignment will fail if the system cannot give access to that LUN on the same SCSI ID to a server/cluster across all I/O groups
  - Policing to ensure the owning I/O group is always part of the access I/O groups
    - Previously, a user could accidently shut off the owning I/O group from access. This would cause undo traffic across I/O groups and at times cause problems
    - The Upgrade Test Utility will also police for this and not let an upgrade continue if the condition exists

Expanding HyperSwap Volumes

- Support for expanding HyperSwap volumes using GUI and CLI.
- Volumes can be expanded without stopping host I/O and the new capacity is available immediately.
- You can expand the size of a HyperSwap volume provided:
  - All copies of the volume are synchronized
  - All copies of the volume are thin or compressed
  - There are no mirrored copies
    - This one will likely be lifted later
  - The volume is not in a remote copy consistency group
  - The volume is not in any user-owned FlashCopy maps
- New CLI command: **`expandvolume`**
- An example to increase the capacity of volume5 by 10 GB:

```
expandvolume -size 10 -unit gb volume5
```

Miscellaneous Enhancements

Byron Grossnickle

SNMPv3

Outbound support for SNMP (Simple Network Management Protocol) in Spectrum Virtualize has been updated from version 2 to include both version 2 and version 3.

Version 3 uses the same base protocol as earlier versions but introduces encryption and much improved authentication mechanisms. Depending on how you authorize with the SNMP agent on a device, you may be granted different levels of access. The security level depends on what credentials you must provide to authenticate successfully

SNMP V3 Security Options

Security Level	Description	Security Name and Engine id	Authentication Protocol	Authentication Password	Privacy Protocol	Privacy Password
NoAuthNoPriv	No authentication nor encryption	✓	X	X	X	X
AuthNoPriv	Messages are authenticated but not encrypted	✓	✓	✓	X	X
AuthPriv	Messages are authenticated and encrypted	✓	✓	✓	✓	✓

Enhanced Audit Logging

- Audit logging has been enhanced to cover
  - Auditing of satask and svctask commands at the CLI level
  - Auditing of login attempts to the cluster or an individual node (GUI)
  - SSH login attempts
- Captured by default at the config node
  - Authentication attempts do not show up in the GUI audit log
    - CLI logs are located in: var/log/cli_audit
    - Authentication attempt are located in: var/log/secure
  - These files are part of a snap capture
- By default, on upgrades these messages are off for syslog servers
  - chsyslogserver –audit on <server_id>
    - CLI log messages use the "NOTICE" level of the syslog selector, whose format is usually <facility>.<level>
    - *.notice /var/log/svc_commands (syslog rule to write to a unique file)
  - chsyslogserver –login on <server_id>
    - Authentication messages are sent to "authpriv" facility of a syslog selector
    - authpriv.* /var/log/svc_auth (syslog rule to write to a unique file)

Enhanced Audit Logging

- Now supports both UDP and TCP with the ability to designate the port
- Syslog command now supports DNS name (If DNS server designated) rather than just IP address

NVMe-oF Big Host Support

- Improving performance for hosts with high number of CPU cores
  - The more CPU cores the host has, the more IO queues it can use, thus more concurrent IO requests towards NVMe controller
- The actual change:
  - Increasing the number of IO queues from 8 to 16
  - Increasing the size of IO queues from 10 to 32
- The results
  - More throughput
  - Lower latency

Secure Data Deletion 8.3.1.1

- Allows data to be securely erased from internal drives as well as boot drives
- 3 levels
  - Cryptographic Erase – Instant
    - Always available on FCM and industry standards NVMe self encrypting drives (SED)
    - Only available on SAS drives if encryption has been enabled on the system
  - Block Erase – Fast
    - Raises and lowers voltage level of the storage element. Specific elements are modified with vendor specific value
  - Data Overwrite – Slow
    - Replaces existing data with random data
- chdrive –task erase
- Boot drive erase
  - satask rescuenode –secureerase

1H2020 Storage Capacity Unit (SCU) Changes

IBM Spectrum Virtualize for SVC and External Virtualization for 9100, 9200 and 7200

Current

Storage Category	SCU Category	SCU
Flash and SSD	Category 1	1 TiB to 1 SCU
10K, 15K and systems using Category 3 drives with advanced architectures	Category 2	1 TiB to .847 SCU
NL-SAS and SATA	Category 3	1 TiB to .25 SCU

New

Storage Category	SCU Category	SCU
Storage Class Memory and SV4PC	Category 1	1 TiB to 1 SCU
Flash and SSD	**Category 2**	**1 TiB to .85 SCU**
10K, 15K and systems using Category 3 drives with advanced architectures	**Category 3**	**1 TiB to .50 SCU**
NL-SAS and SATA	**Category 4**	1 TiB to .25 SCU

Note: SVC4PC is Category 1 because it is all in licensing, meaning all functions are included

Accelerate with IBM Storage Webinars

The Free IBM Storage Technical Webinar Series Continues in 2020...

Washington Systems Center – Storage 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.

Located on the Accelerate with IBM Storage Site:
https://www.ibm.com/support/pages/node/1125513

Also, check out the WSC YouTube Channel here:
https://www.youtube.com/channel/UCNuks0go01_ZrVVF1jgOD6Q

2020 Webinars:

January 7 – What's New in IBM Spectrum Discover V2.0.2
January 14 – Brocade / IBM b-type SAN Modernization
January 21 – Cisco / IBM c-type SAN Analytics
February 20 – Spectrum Virtualize 8.3.1 Enhancements

April 9– GDPS V4.3 and SPEs Update
Register Here: Coming soon.....

Thank You!