



ADVANCED TECHNOLOGY GROUP (ATG)



Accelerate with ATG Webinar: Storage Virtualize 8.6.1 Update

Byron Grossnickle

IBM Advanced Technology Group - Storage Virtualize SME

byrongro@us.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 TS7700 Tape Solution Overview 101](#) - March 12th, 2024

[IBM Storage Virtualize CDR 8.6.2 Technical Update](#) – March 26th, 2024

[Leveraging Your Data Lake with IBM Storage Fusion Data Cataloging](#)

Part 1 - April 9th, 2024

Part 2 - April 16th, 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
- IBM Cyber Resiliency with IBM Storage Defender
- IBM Storage Fusion
- IBM Storage Scale & Storage Scale System
- **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

(Registrations open in March)

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



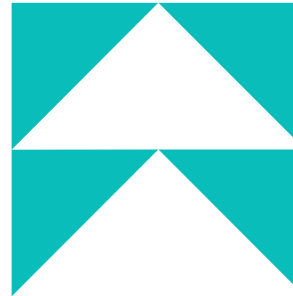
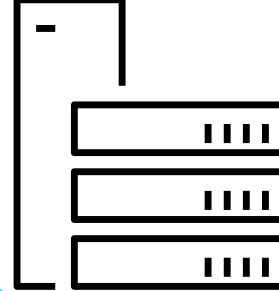
Sandbox

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

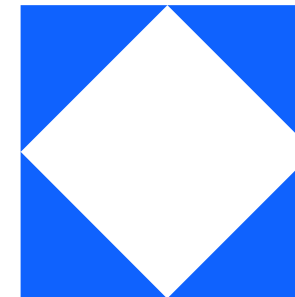
Meet the Expert

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

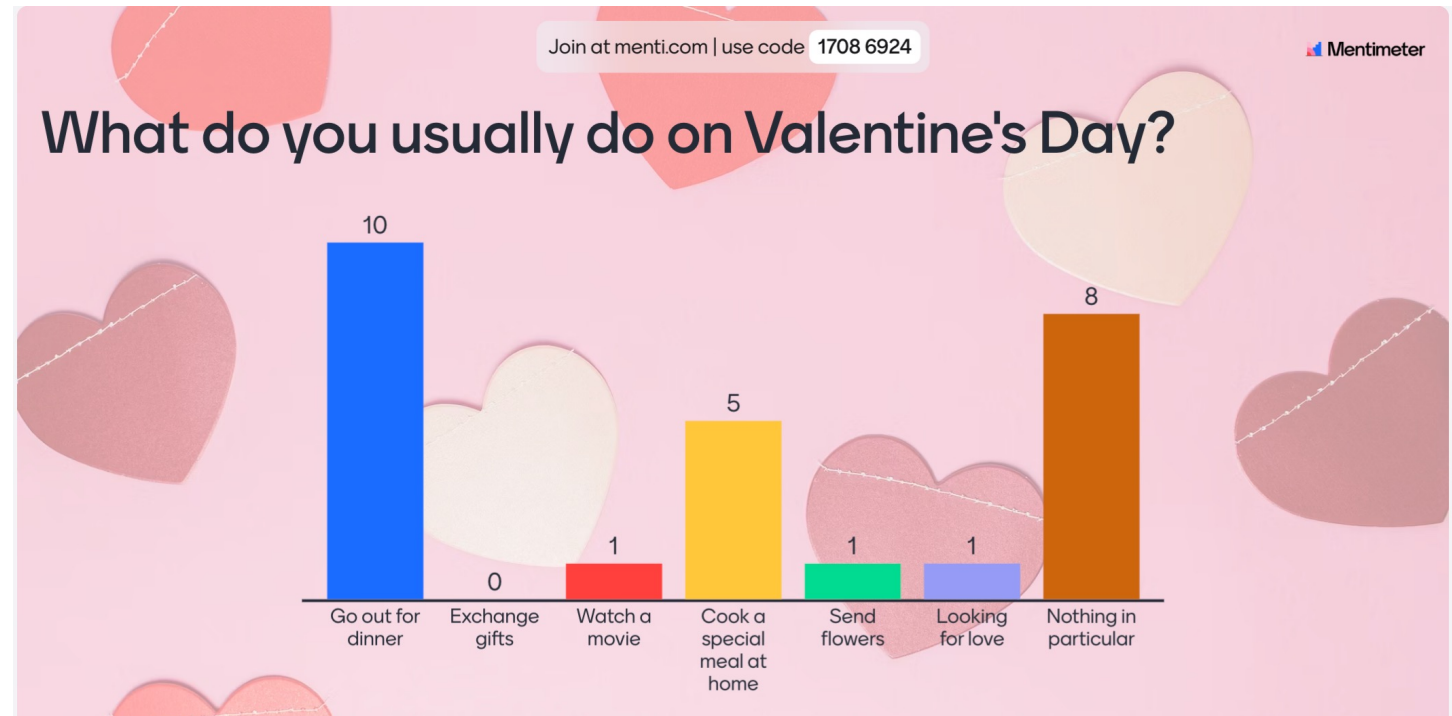
Please take a moment to share your feedback with our team!

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

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

Or

QR Code





ADVANCED TECHNOLOGY GROUP (ATG)



Accelerate with ATG Webinar: Storage Virtualize 8.6.1 Update

Byron Grossnickle

IBM Advanced Technology Group - Storage Virtualize SME

byrongro@us.ibm.com



Meet the Speakers



Byron Grossnickle is an IBM Storage Technical Specialist concentrating on Storage Virtualize software. This include FlashSystem, SVC, and Storage Virtualize for Public Cloud. Byron has been with IBM 18 years exclusively in storage. Prior to working for IBM, Byron spent 6 years engineering storage in the Telcom Industry. Prior to that he worked 8 years in healthcare IT. Byron lives in the Kansas City area and is available to travel to customer engagements.

Agenda



- Release Schedule
- FS9500/SV3 – 64 Gb FC Adapter Support
- 64 Gb FC Adapter Performance
- Policy-based High Availability
- Policy-based HA Performance
- Support Statements
 - GM/GMCMV
 - vVols with Spectrum Connect
 - FlashSystem/SVC NPIV for Host Ports

Release Schedule

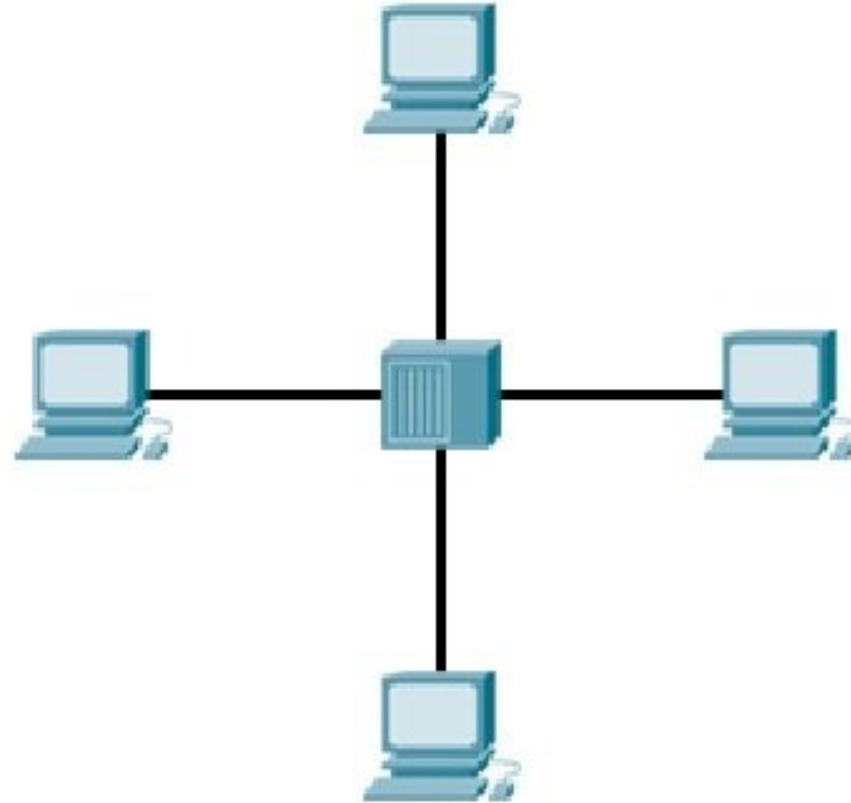
- RFA Announce – 8.6.1 – August 15, 2023
- pGA – 64 Gb FC on 8.6.0+ – August 26, 2023
- eGA – 8.6.1 – September 15, 2023

8.6.1 is a Continuous Development Release (CDR). This means that it will get no planned patches or updates until the next CDR (8.6.2) in 4Q2023

Spectrum Virtualize is now known as Storage Virtualize



FS9500/SV3 64 Gb FC Support



4 Port 64Gb FC Adapter Support

- Supported Products
 - FS9500
 - SV3
- Max of 1 – 64 Gb FC card per cage
 - Adjacent slot left blank
 - Not oversubscribed
 - 8.6.2+ allows you to fill cages. Filled cages are oversubscribed
- Speed Support is N-2
 - 64Gb
 - 32Gb
 - 16Gb
- FRU Part Number: 03JK911
- **Supported on 8.6.0+**

Hardware and Support Details

Platform	Adapter	Max adapter per node	Max ports per node	Max ports per enclosure	PCI slots
FS9500	4x64Gb FC	3	12	24	1, 5, 7
SV3	4x64Gb FC	3	12	12	1, 5, 7

Performance Summary

- SV3 and FS9500 with half of amounts of 64Gbps FC ports offer mostly similar performance as the same system with 32Gbps FC ports, except small transfer cache hits workloads due to lack of enough adapter cpu cores.

Policy-based High Availability



Agenda

Objectives

Feature overview

Quorum

HA behaviours

Host requirements

Limitations of first release

Useful information

Comparison vs HyperSwap and ESC

Design Hill

A storage architect can configure highly available storage between two independent systems using policy-based configuration with no single point of failure, with throughput four times greater than the current I/O group HyperSwap allowing continued application access during partial or complete failure of either system.

Design Hill

A storage architect can configure highly available storage between two independent systems using policy-based configuration with no single point of failure, with throughput four times greater than the current I/O group HyperSwap allowing continued application access during partial or complete failure of either system.

- HA is managed at a higher-level than day-to-day administrators
- Automatic failover capabilities to ensure host I/O does not fail
- ‘storage’ not just ‘volumes’ – the entire management plane is highly available

Design Hill

A storage architect can configure highly available storage **between two independent systems** using policy-based configuration with no single point of failure, **with throughput four times greater than the current I/O group HyperSwap allowing continued application access during partial or complete failure of either system.**

- HA between two systems partnered together
- No requirement of the same hardware in each location (e.g., FS9500 and SVC)
- [Future] No requirement of the same software version in each location (within reason) – systems upgrade independently

Design Hill

A storage architect can configure highly available storage between two independent systems using **policy-based configuration** with no single point of failure, **with throughput four times greater than the current I/O group HyperSwap allowing continued application access during partial or complete failure of either system.**

- HA configuration expressed using replication policies
- Extends the 8.5.2 async policy-based replication design
- Remote provisioning managed automatically based on pre-defined rules

Design Hill

A storage architect can configure highly available storage between two independent systems using policy-based configuration with **no single point of failure**, with throughput four times greater than the current I/O group HyperSwap allowing continued application access during partial or complete failure of either system.

Designed to ensure application I/O survives:

- Site failures
- Power failures
- Cluster recovery on either system
- Repeated warmstarts
- Simultaneous warmstarts in an I/O group
- Offline arrays/pools

To be continued...

Design Hill

A storage architect can configure highly available storage between two independent systems using policy-based configuration with **no single point of failure**, with throughput four times greater than the current I/O group HyperSwap **allowing continued application access during partial or complete failure of either system.**

- Quorum app failure/loss of connectivity
- ISL failures
- Significant ISL congestion
- Brief connectivity glitches

Additionally, non-HA storage will be unaffected by HA problems

Design Hill

A storage architect can configure highly available storage between two independent systems using policy-based configuration with no single point of failure, with throughput four times greater than the current I/O group HyperSwap allowing continued application access during partial or complete failure of either system.

- Negligible read performance impact (same as HyperSwap)
- Reduction in latency for writes vs HyperSwap
- Significant increase in IOPS/throughput vs HyperSwap
- Low overhead for HA to get the most out of hardware

Iterative Delivery Model!

- The first release is targeted at new installations and proof of concepts, not existing systems
- Support for adding HA to an existing non-HA system is planned to be included in the next 8.6.3+
- There will be functionality missing from the first release, but many enhancements are already being planned for upcoming releases

Agenda

Objectives

Feature overview

Quorum

HA behaviours

Host requirements

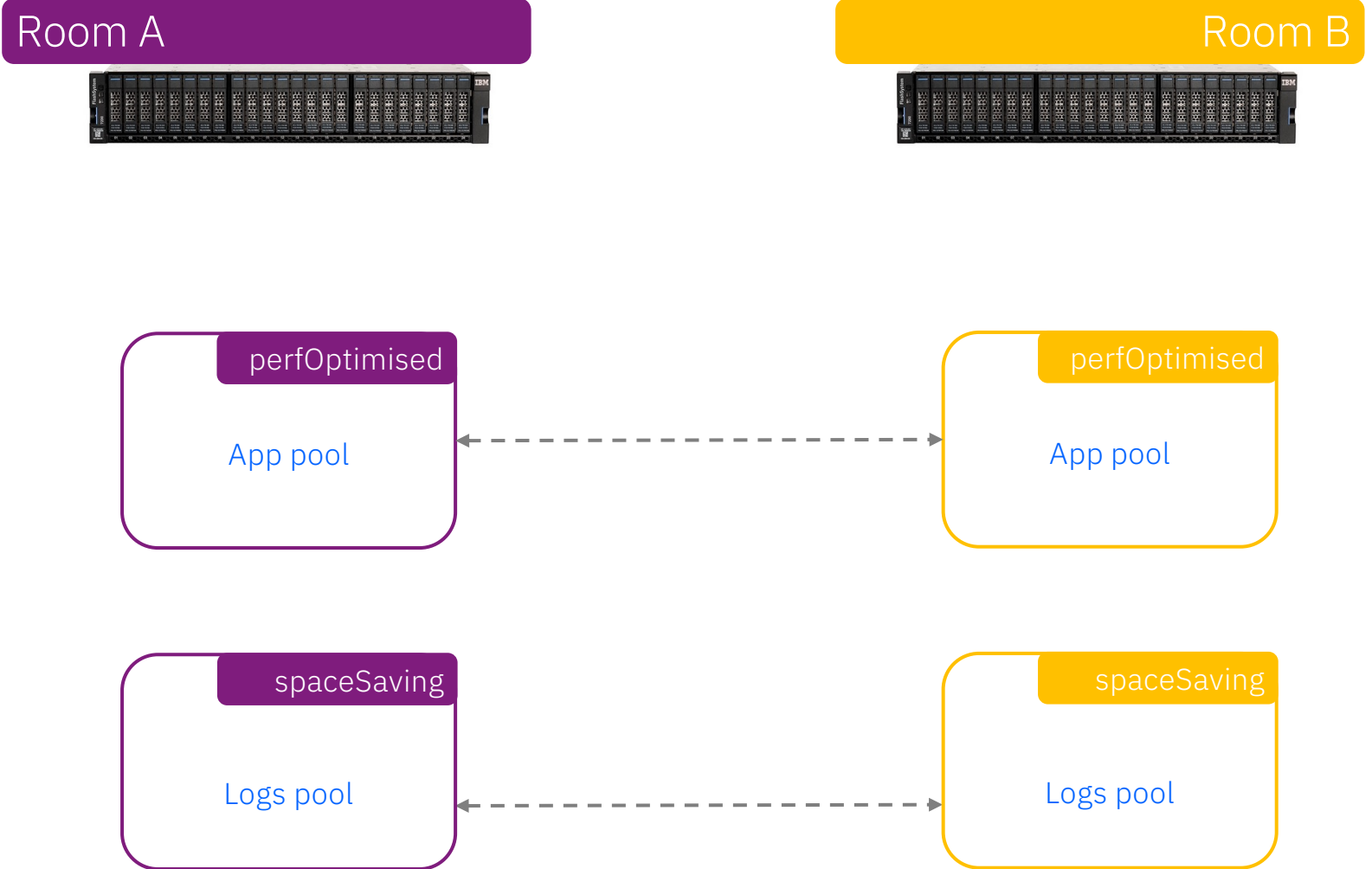
Limitations of first release

Useful information

Comparison vs HyperSwap and ESC

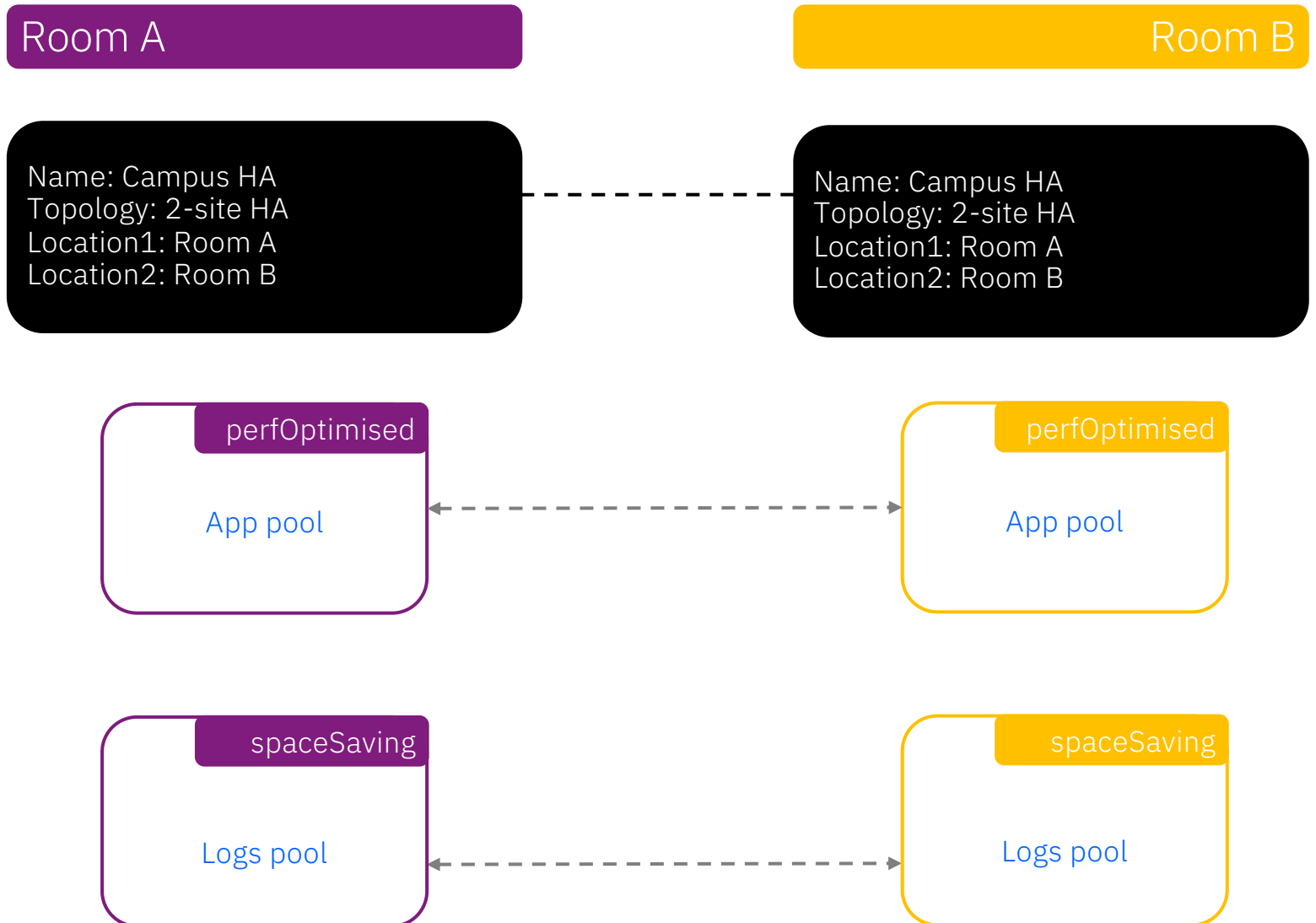
Storage architect performs the one-time configuration tasks...

- The storage architect performs the initial configuration tasks, similar to configuring asynchronous replication.
- Configures the partnership, pools, pool links and provisioning policies.



Storage architect creates a replication policy

- Much like with asynchronous replication, a replication policy defines how and where to replicate.
- 2-site-ha topology configures high availability between two independent systems.

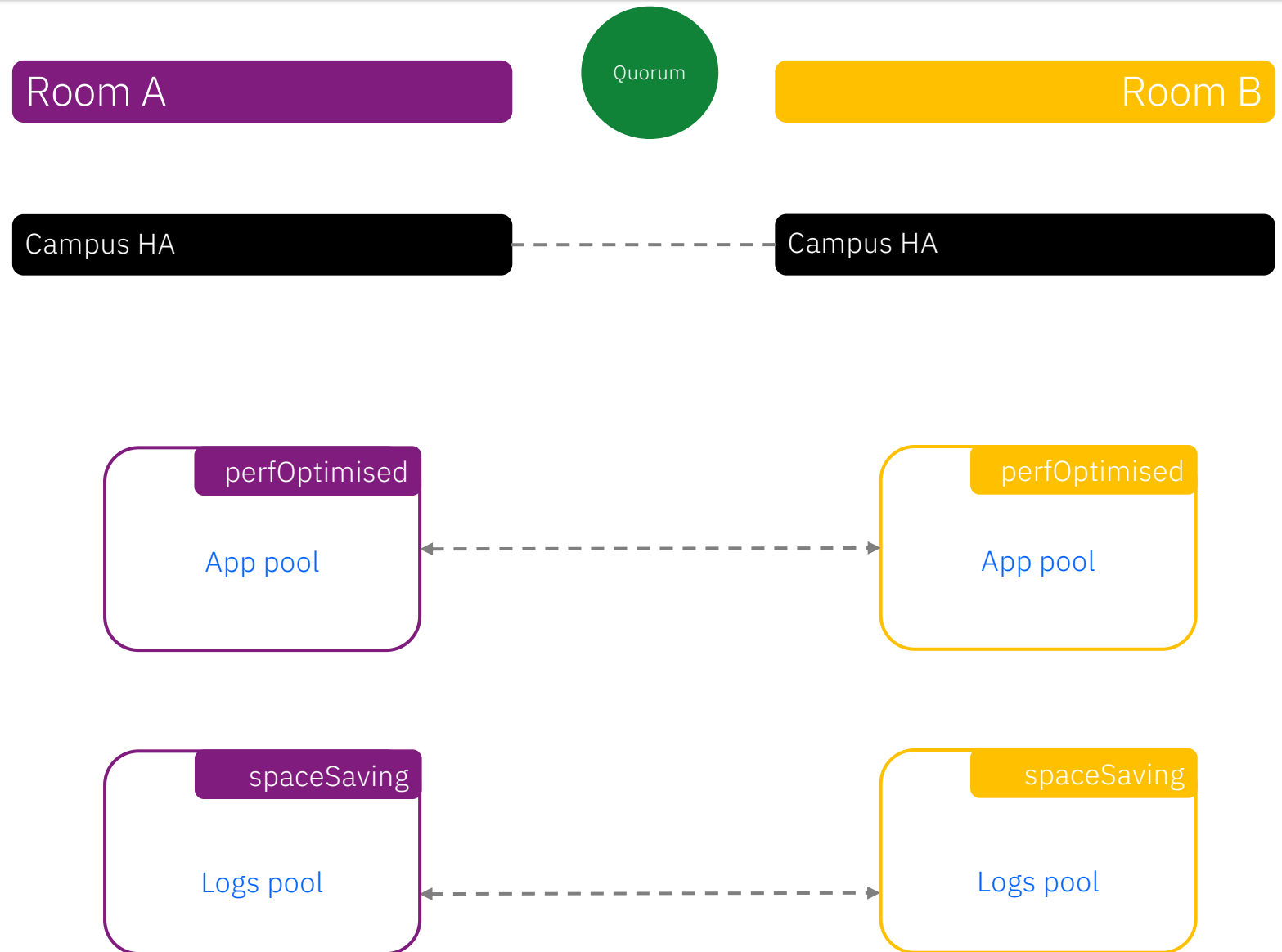


Storage architect deploys quorum for tie-breaking

- IP quorum can be configured to act as an independent witness in the event of loss of connectivity

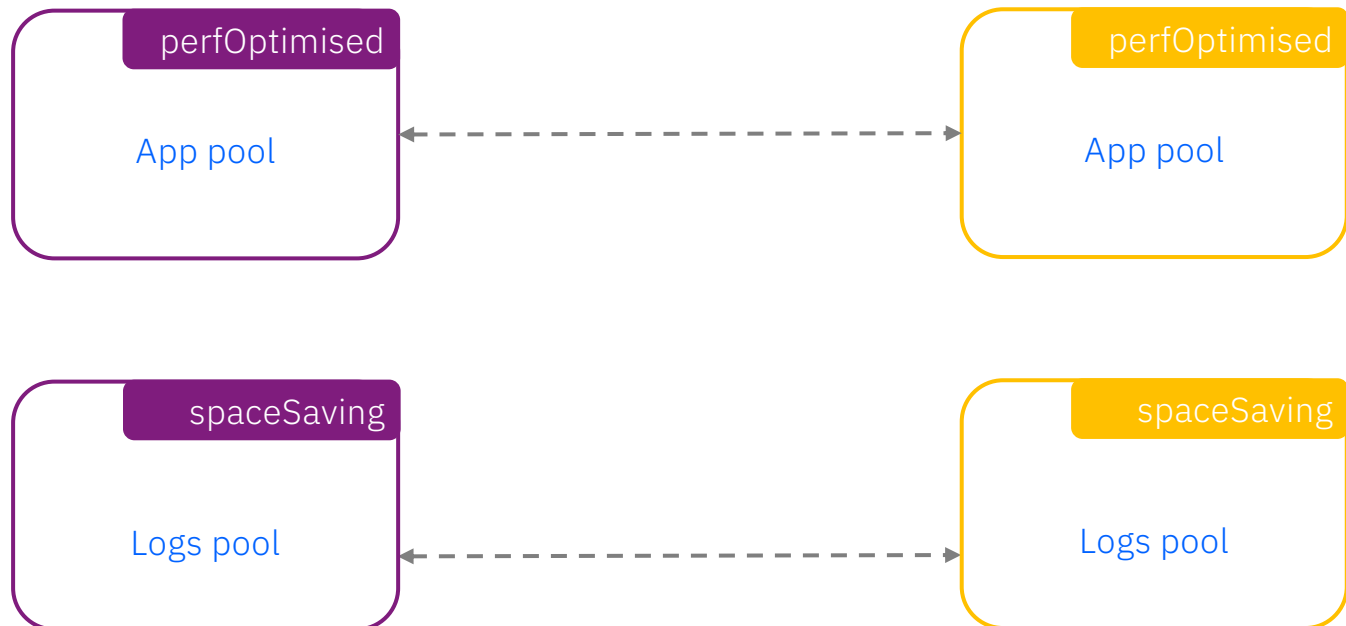
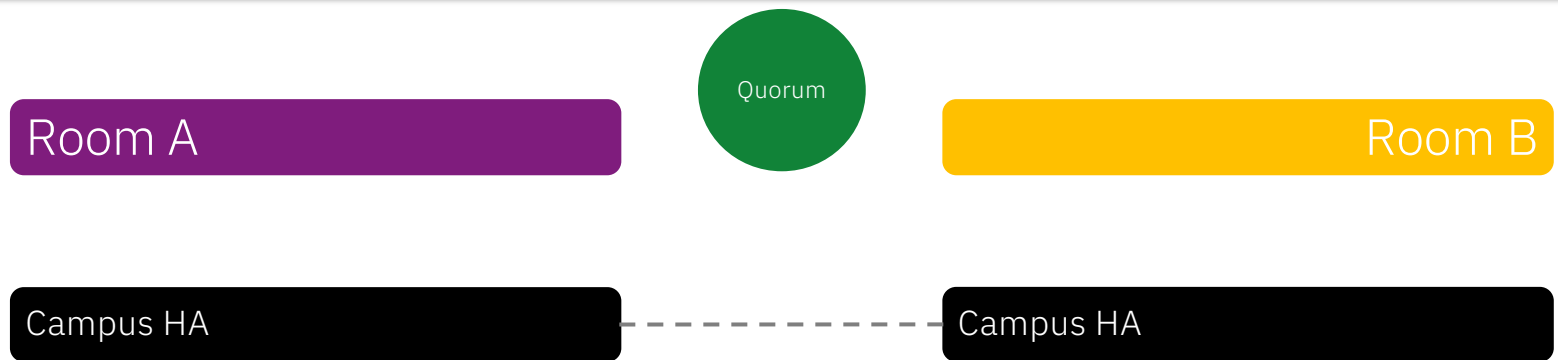
IP quorum is required, a FC quorum will not work

- Java app installed on a server that is not dependent on either system
- Multiple quorum applications can be deployed for redundancy

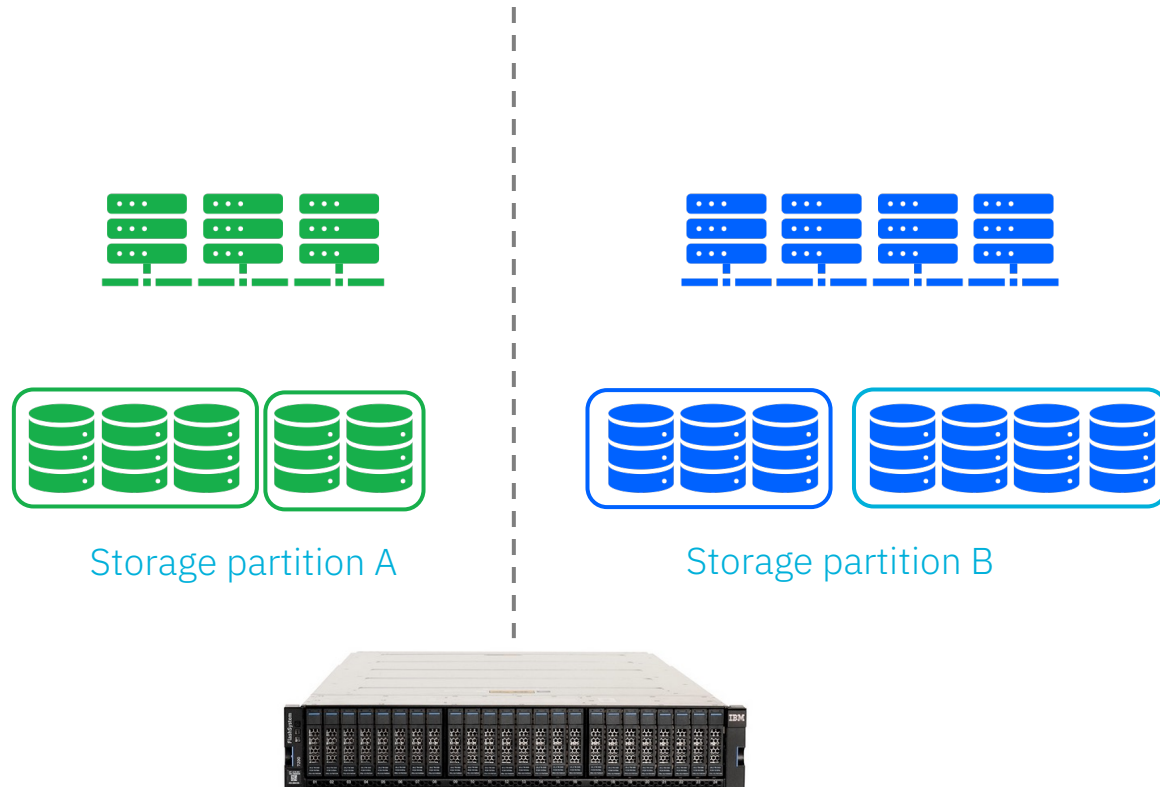


Storage architect setup is now complete

- One-time setup tasks have been completed.
- Ready for the storage admin to configure highly available storage.












Storage Partitions




- Partitions the system into disjoint sets of hosts, volume groups and volumes.
- Storage partitions can be configured to be highly available, automatically configuring every host and volume within it to be highly available.
- The management of hosts, volume groups and volumes automatically swaps between systems to provide highly available management access.
- Highly available storage partitions have an attribute that defines which system is preferred for management access when HA is established.

Partitions Panels

☰ IBM FlashSystem 5200

-  **Dashboard**
-  **Monitoring** ▼
-  **Pools** ▼
-  **Volumes** ▼
-  **Hosts** ▼
-  **Copy services** ▼
-  **Policies** ▼
-  **Access** ▼
-  **Settings** ▼

 **Storage partitions**

- myPartition >
- GoldTier >
- BronzeTier >

☰ IBM FlashSystem 5200 chestnut-c Storage partitions
🔔¹ 📄 ? 👤¹

All storage partitions Create partition

<div style="background-color: #fff; padding: 10px; border-radius: 5px;"> <p>myPartition ⋮</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; border-right: 1px solid #ccc; padding: 5px;">System 1 chestnut-c</td> <td style="width: 50%; padding: 5px;">System 2 peanut-c</td> </tr> <tr> <td style="border-right: 1px solid #ccc; padding: 5px;">High-availability ✔ Established</td> <td style="padding: 5px;">Replication policy HaRepPolicy</td> </tr> <tr> <td style="border-right: 1px solid #ccc; padding: 5px;">Hosts 2</td> <td style="padding: 5px;">Volumes 10</td> </tr> <tr> <td style="border-right: 1px solid #ccc; padding: 5px;">Volume groups 2</td> <td></td> </tr> </table> <p style="text-align: center; background-color: #007bff; color: white; padding: 5px; margin-top: 10px;">View partition</p> </div>	System 1 chestnut-c	System 2 peanut-c	High-availability ✔ Established	Replication policy HaRepPolicy	Hosts 2	Volumes 10	Volume groups 2		<div style="background-color: #fff; padding: 10px; border-radius: 5px;"> <p>GoldTier ⋮</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; border-right: 1px solid #ccc; padding: 5px;">System 1 chestnut-c</td> <td style="width: 50%; padding: 5px;">System 2 peanut-c</td> </tr> <tr> <td style="border-right: 1px solid #ccc; padding: 5px;">High-availability ❌ Problem</td> <td style="padding: 5px;">Replication policy HaRepPolicy</td> </tr> <tr> <td style="border-right: 1px solid #ccc; padding: 5px;">Hosts 1</td> <td style="padding: 5px;">Volumes 13</td> </tr> <tr> <td style="border-right: 1px solid #ccc; padding: 5px;">Volume groups 1</td> <td></td> </tr> </table> <p style="text-align: center; background-color: #007bff; color: white; padding: 5px; margin-top: 10px;">View partition</p> </div>	System 1 chestnut-c	System 2 peanut-c	High-availability ❌ Problem	Replication policy HaRepPolicy	Hosts 1	Volumes 13	Volume groups 1		<div style="background-color: #fff; padding: 10px; border-radius: 5px;"> <p>BronzeTier ⋮</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; padding: 5px;">Hosts 0</td> <td style="width: 50%; padding: 5px;">Volumes 100</td> </tr> <tr> <td style="padding: 5px;">Volume groups 9</td> <td></td> </tr> </table> <p style="text-align: center; background-color: #007bff; color: white; padding: 5px; margin-top: 10px;">View partition</p> </div>	Hosts 0	Volumes 100	Volume groups 9	
System 1 chestnut-c	System 2 peanut-c																					
High-availability ✔ Established	Replication policy HaRepPolicy																					
Hosts 2	Volumes 10																					
Volume groups 2																						
System 1 chestnut-c	System 2 peanut-c																					
High-availability ❌ Problem	Replication policy HaRepPolicy																					
Hosts 1	Volumes 13																					
Volume groups 1																						
Hosts 0	Volumes 100																					
Volume groups 9																						

Latency 0 ms Read 0 ms Write 0 ms Bandwidth 0 MBps Read 0 MBps Write 0 MBps IOPS 0 Read 0 Write 0

Partitions Panels

IBM FlashSystem 5200
chestnut-c
Storage partition: myPartition

🔔¹
📄
❓
👤¹

🏠 myPartition

Storage partition overview

Volumes ⌵

Hosts ⌵

Copy services ⌵

Policies ⌵

IP quorum

←| Exit myPartition

myPartition overview

Replication Policy

Policy name	Topology
HaRepPolicy	2 Site, High Availability

High availability status

📄 chestnut-c

Production system

Management through this system
 ✓ Preferred management is this system
 Logged in to this system

Go to system →

📄 peanut-c

Production system

Connectivity

Partnership (peanut-c) →

● Configured

IP quorum →

Storage components

Hosts (2) →

✓ Online

Volumes (10) →

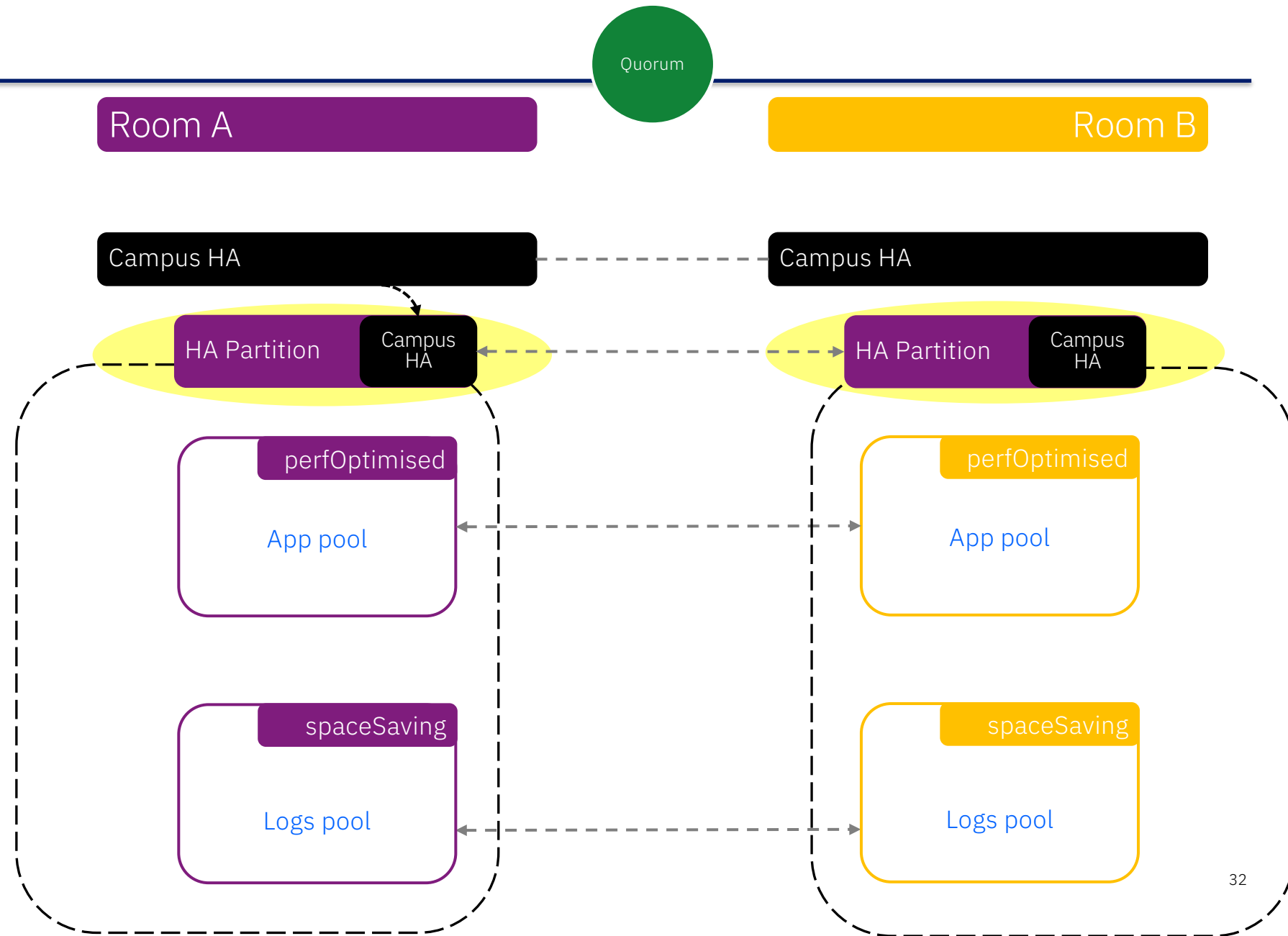
Volume groups (2) →

✓ Online

Latency 0 ms Read 0 ms Write 0 ms
Bandwidth 0 MBps Read 0 MBps Write 0 MBps
IOPS 0 Read 0 Write 0

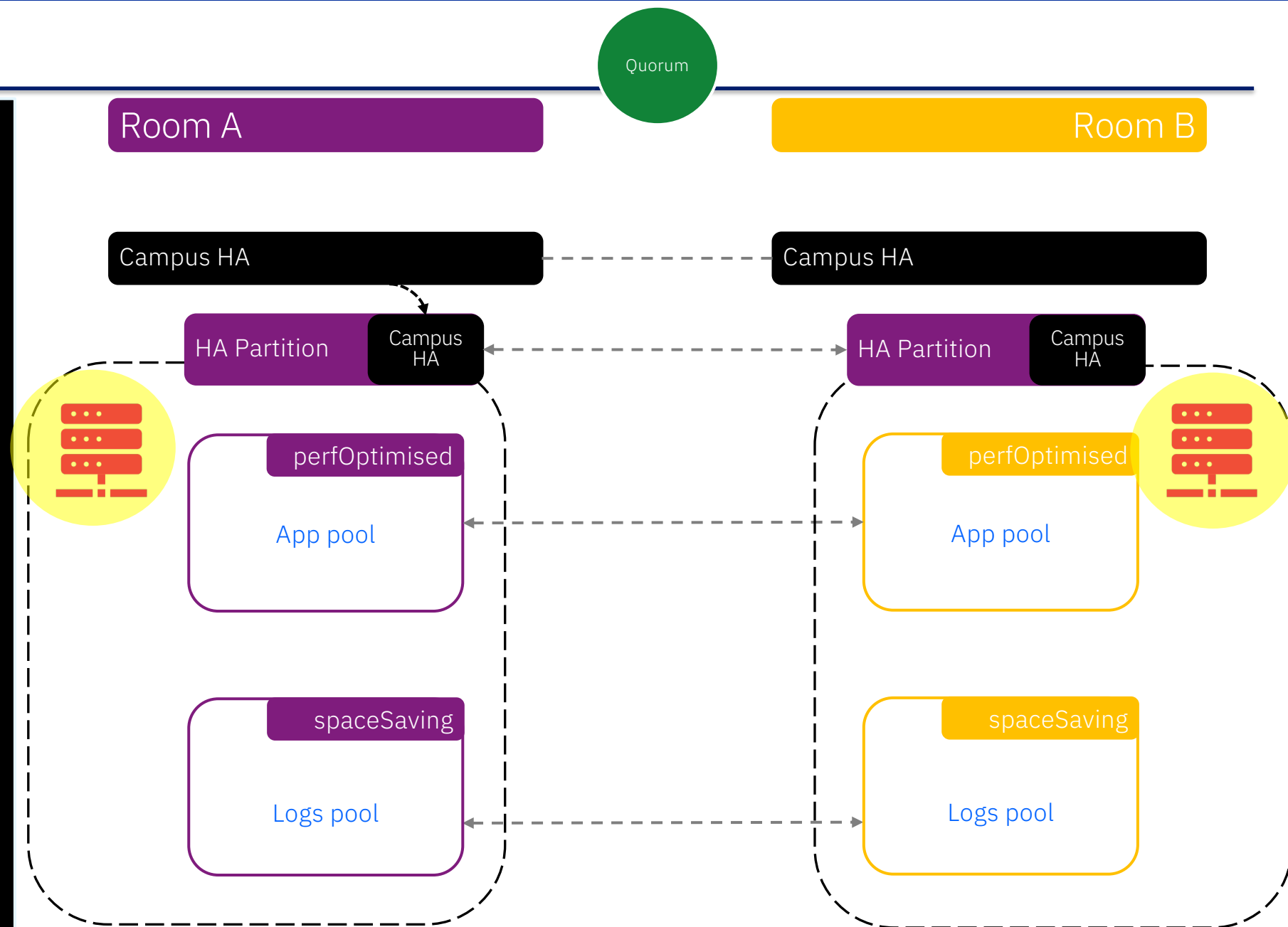
Storage admin creates a new storage partition

- Storage admin creates a new storage partition to act as a container for hosts, volume groups and volumes.
- Assigning a HA replication policy to the storage partition enables HA for all hosts, volume groups and volumes in it.
- Automatically configured on the partnered system.



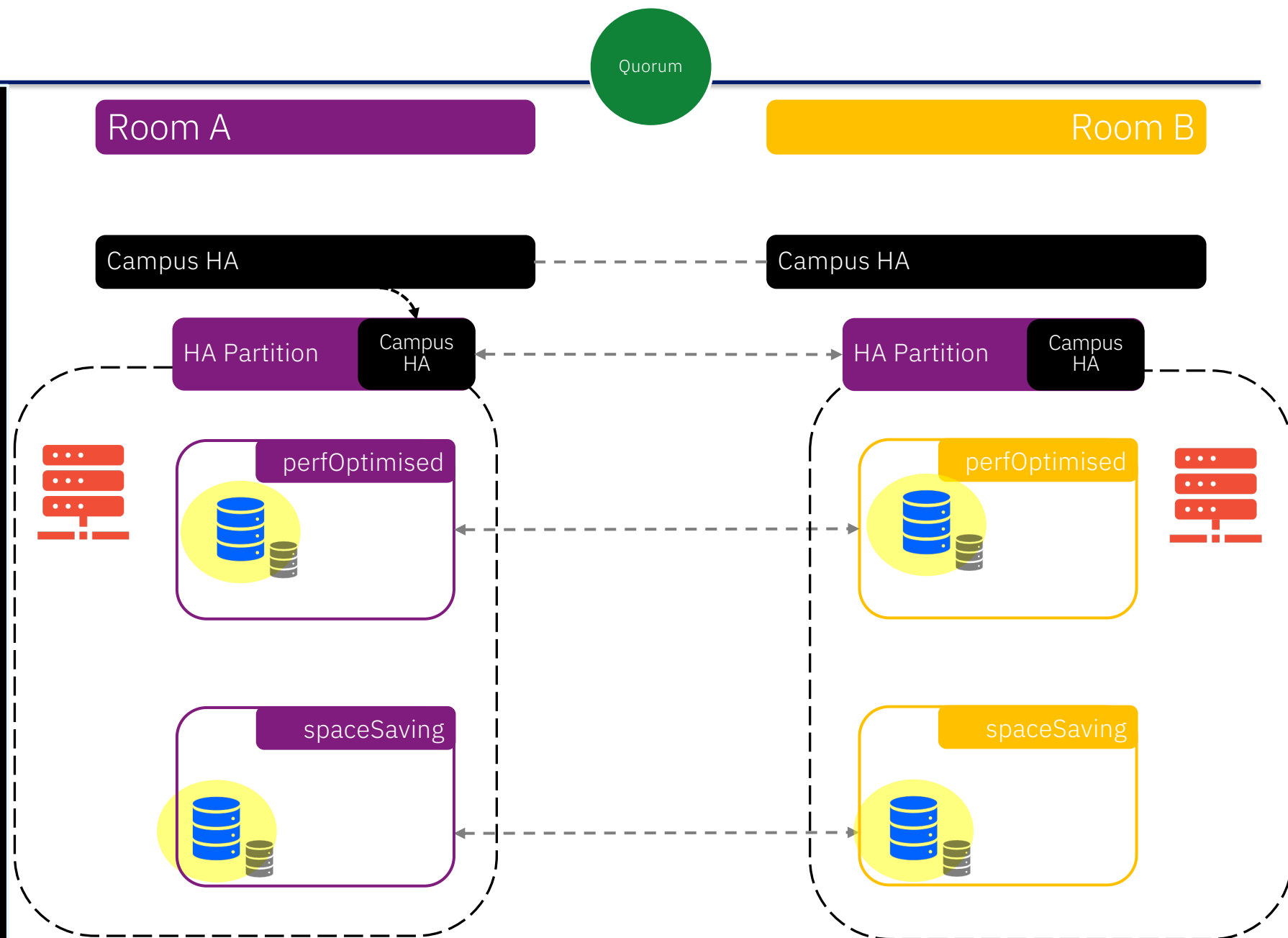
Storage admin creates a new host

- Specifying a highly available storage partition when creating a host automatically configures HA for it
- Hosts configured in a HA partition can only be mapped to volumes in the same partition
- Non-HA volumes to the same host require a second host definition (and ports) that is in a non-HA partition
- Configuration changes made to the host are automatically applied to both systems



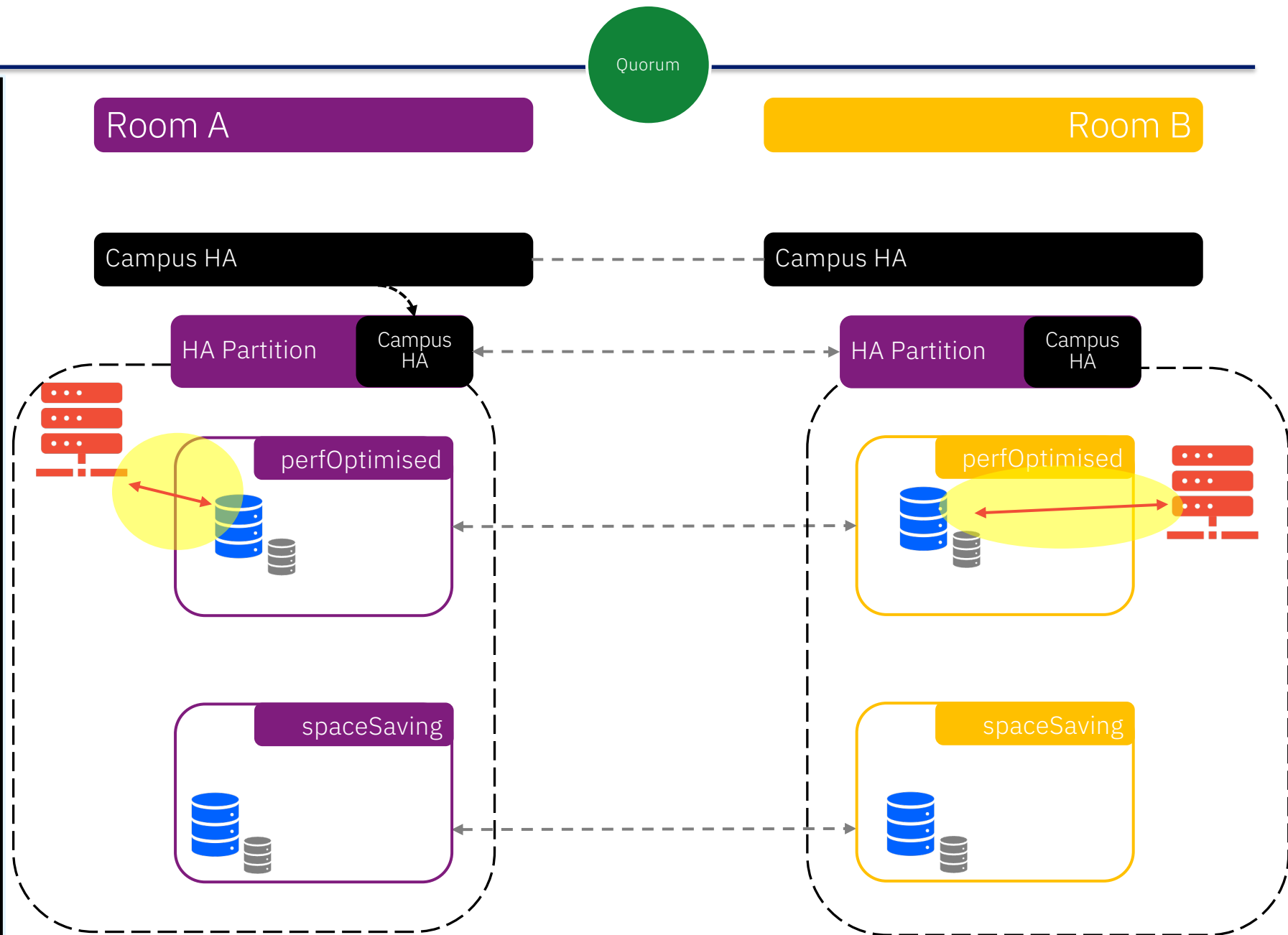
Storage admin creates a new volume group and volumes

- Storage admin creates a volume group in the storage partition
- Storage admin creates volumes, specifying a size, pool and volume group
- The volume will be created according to the provisioning policy for the linked storage pool
- The volume will have the identical UUID of the source volume similar to Volume Migration



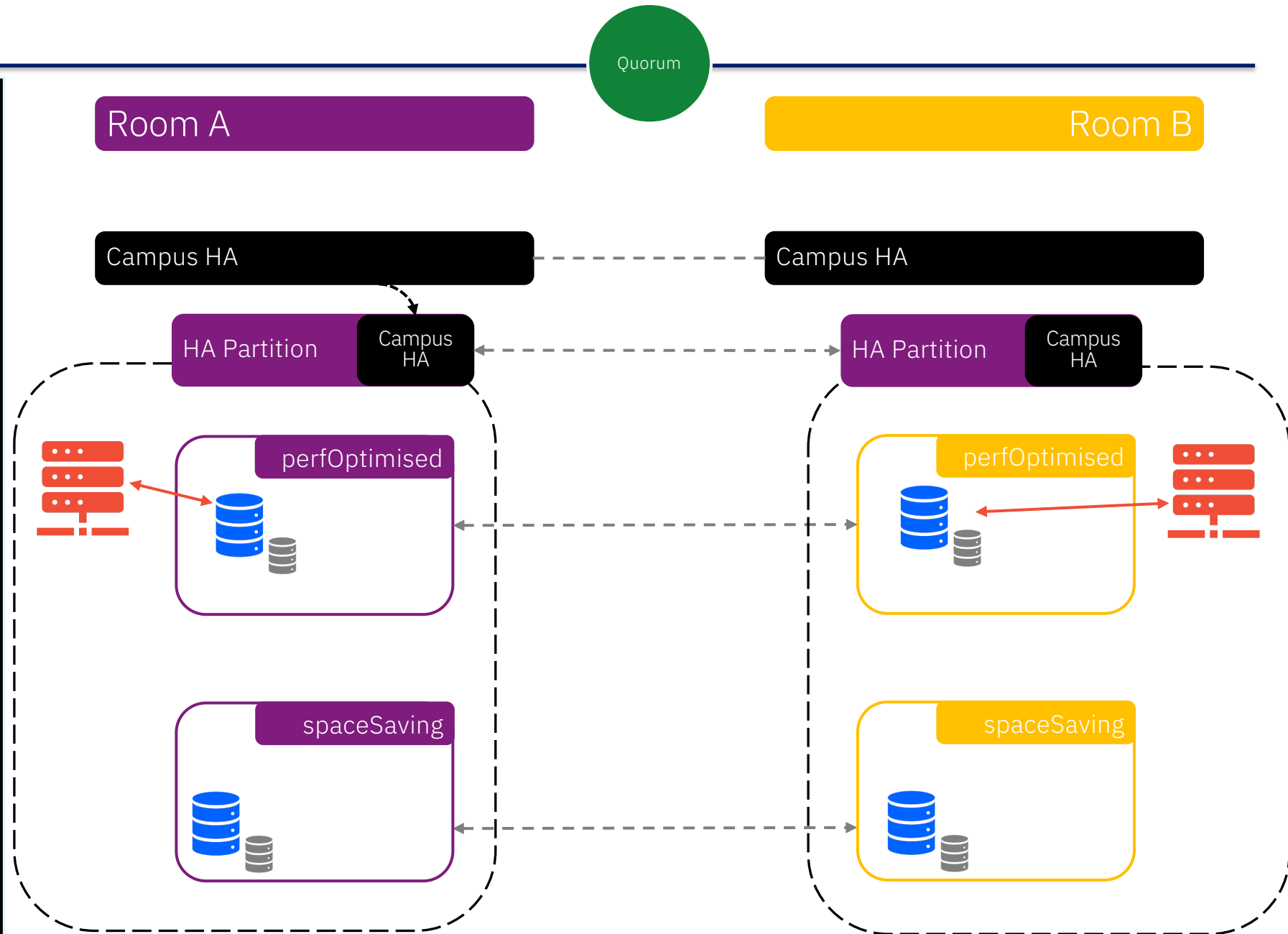
Storage admin maps the hosts to the volumes

- Volumes in a storage partition can be mapped to hosts in the same partition.
- The host map is automatically created on the partner system
- The host can discover paths to the volumes on both systems

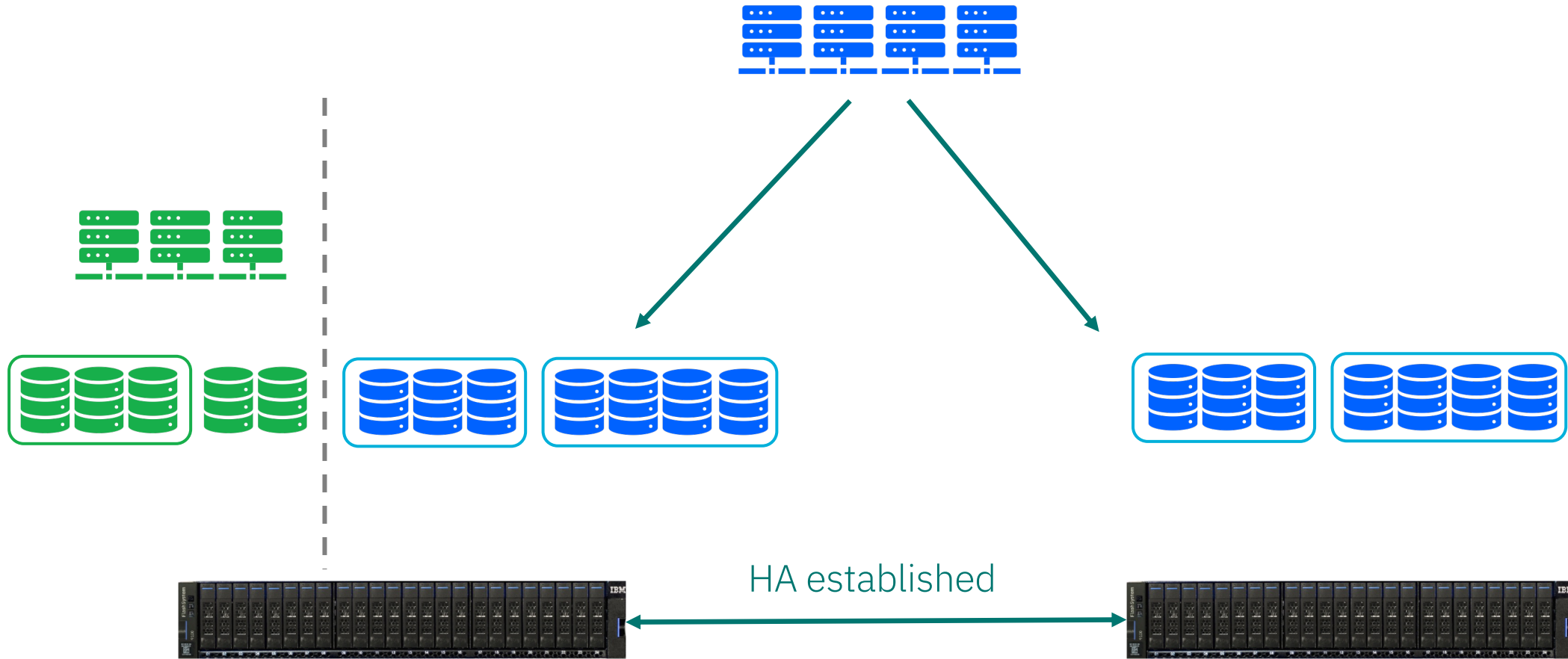


Storage admin configuration is now complete

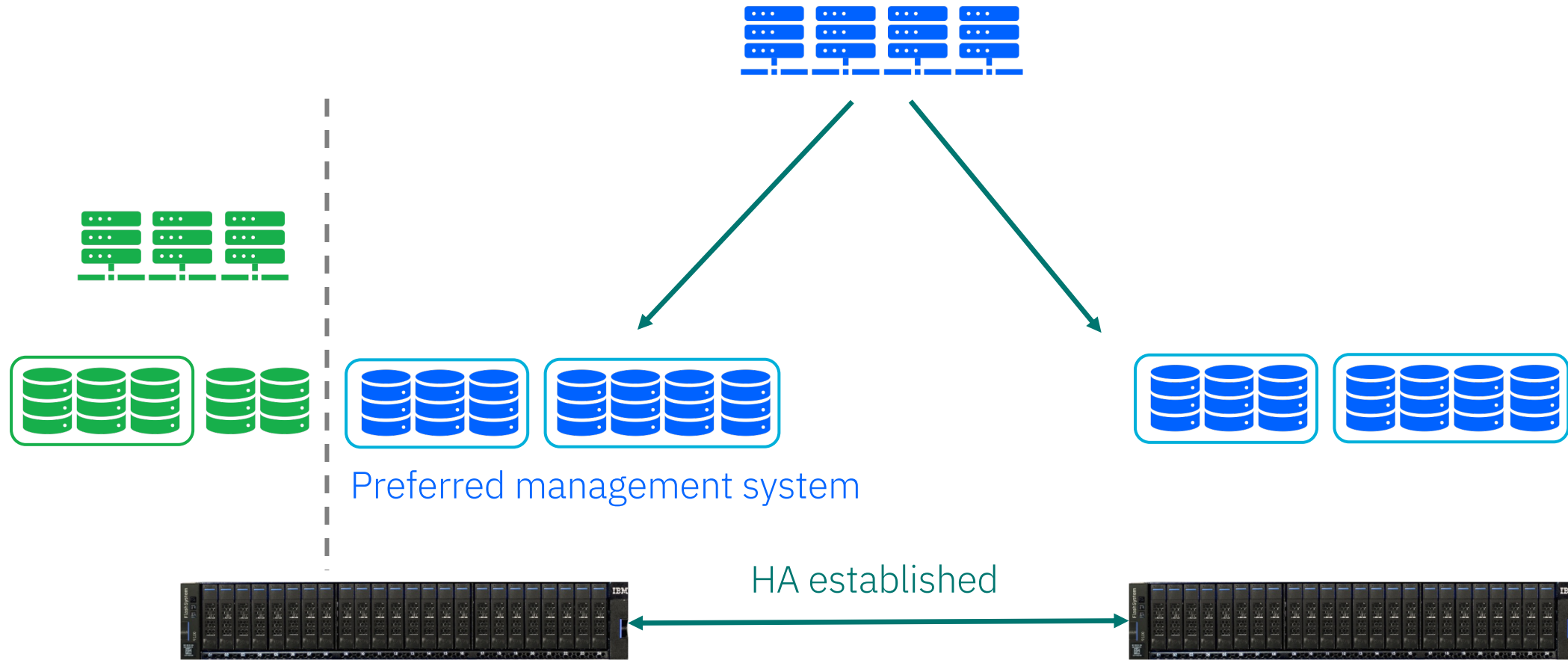
- The storage admin's initial provisioning is complete, and the configuration and volumes are highly available
- The storage admin can configure more hosts, volume groups and volumes as required



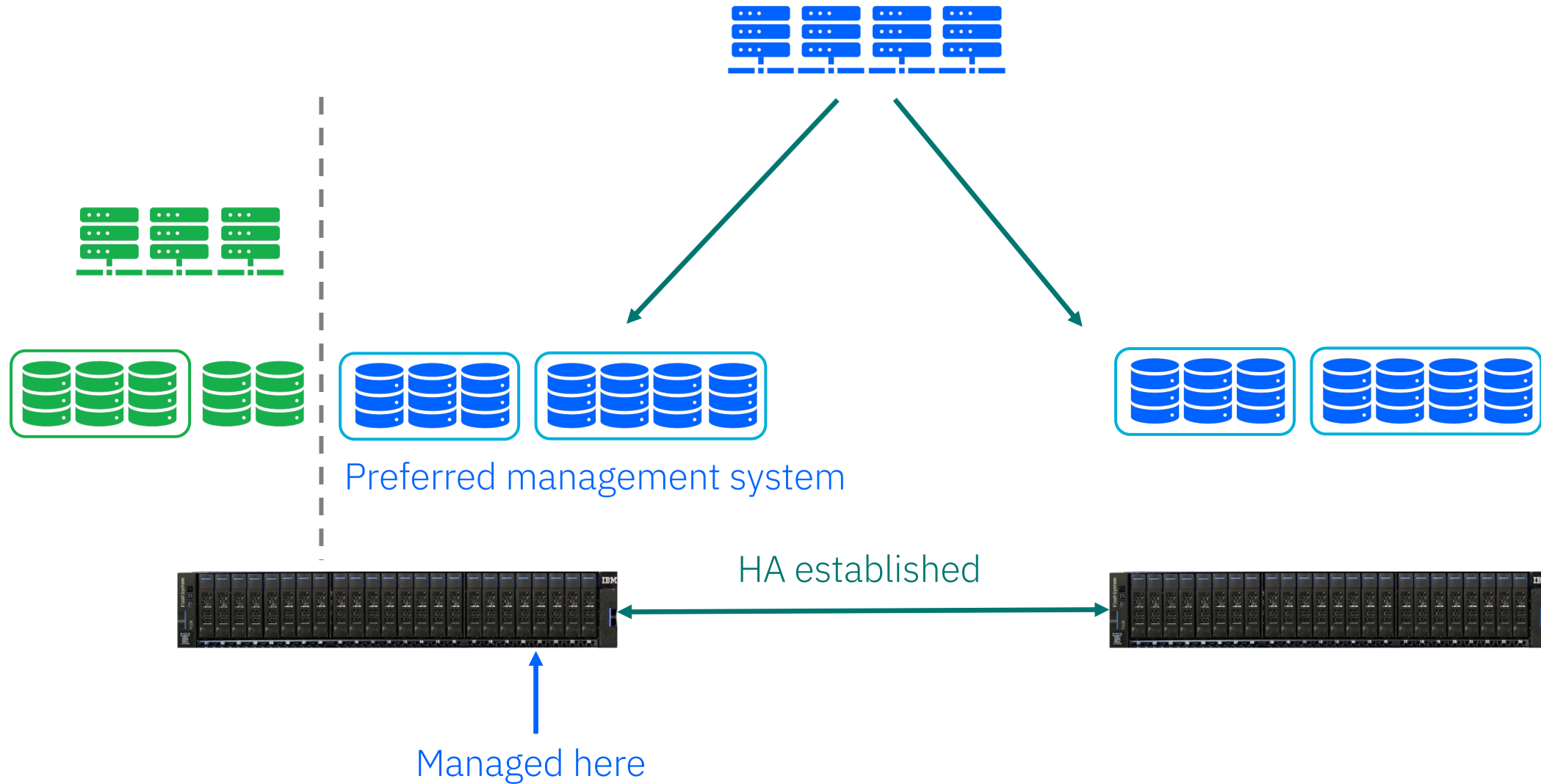
HA storage partitions



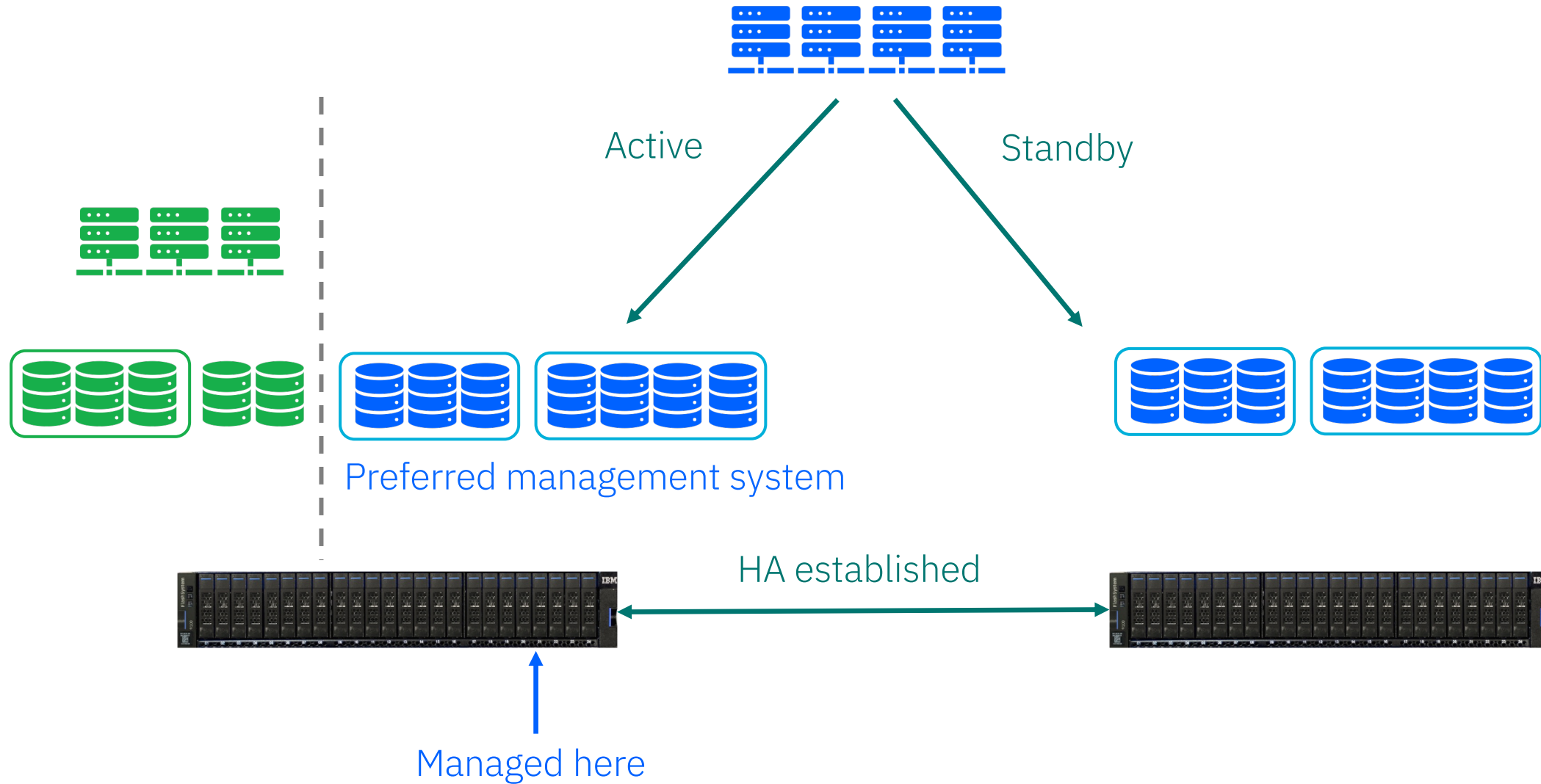
HA storage partitions



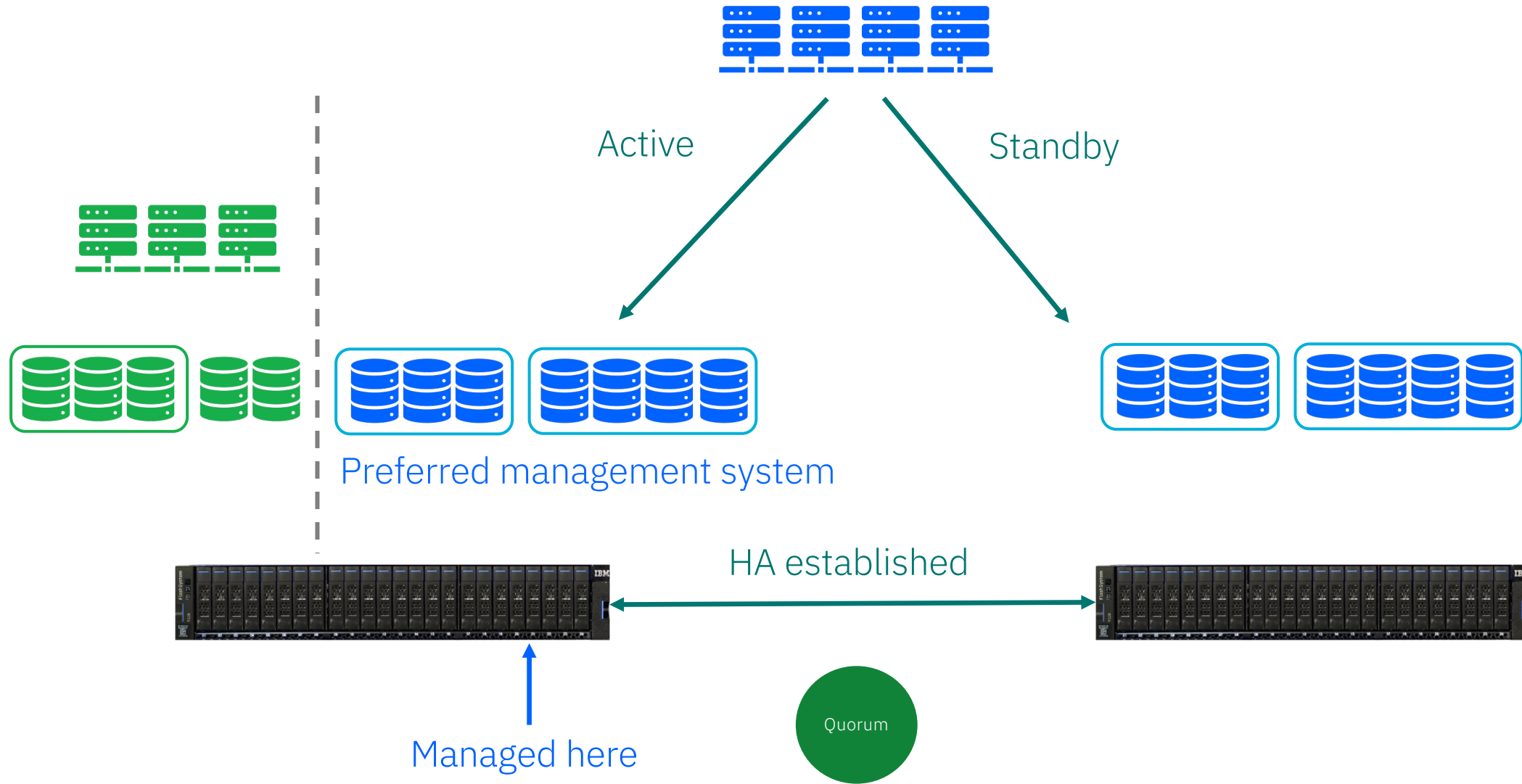
HA storage partitions



HA storage partitions



HA storage partitions



Agenda

Objectives

Feature overview

Quorum

HA behaviours

Host requirements

Limitations of first release

Useful information

Comparison vs HyperSwap and ESC

High Availability: IP Quorum

Loss of connectivity between systems

Policy-based High Availability requires an IP quorum application to arbitrate in the event of a partnership disconnect, site or system loss.

- The partnership must be configured before generating and deploying the IP quorum application
- Therefore, there must be a certificate in the truststore tagged with restapi on both systems
- The IP quorum application can be generated on either system in the partnership
 - Since these are separate systems, it is only used for a tie break in a quorum race for the highly available storage partition(s)
- The maximum number of IP quorum applications that can be deployed on a single system is five
- Only one instance of the IP quorum application per server, per system, is supported
 - Quorum server 1: FS7300-1, FS9500-1, etc.
 - Quorum server 2: FS7300-1, FS9500-1, etc.

IP Quorum QOL (Quality of Life) Improvements pt1

Lessen the need to redeploy the app

If you change the configuration by adding a node, deleting a node, changing a service or cluster IP address, the IP quorum application will automatically update itself with the new IP addresses.

If you change the SSL certificates the IP quorum application will automatically update itself with the new certificate.

IP Quorum QOL (Quality of Life) Improvements pt2

New options when deploying the app

Configurable logging:

- Specify the full, or relative path to store your log files
- Specify the maximum number of log files.
- Specify the log file size

Display the v.m.m.f. (version, major, minor, fix) of the system when the app was generated, the timestamp and the generation id.

IP Quorum QOL (Quality of Life) Improvements pt3

- Download the Quorum App through the RESTAPI
 - (and any other file that's in /dumps)
- Automatically Collect IP quorum logs
 - If there is an IP quorum app connected to a cluster when a snap is taken it will upload the IP quorum logs and include them in the snap.

Agenda

Objectives

Feature overview

Quorum

HA behaviours

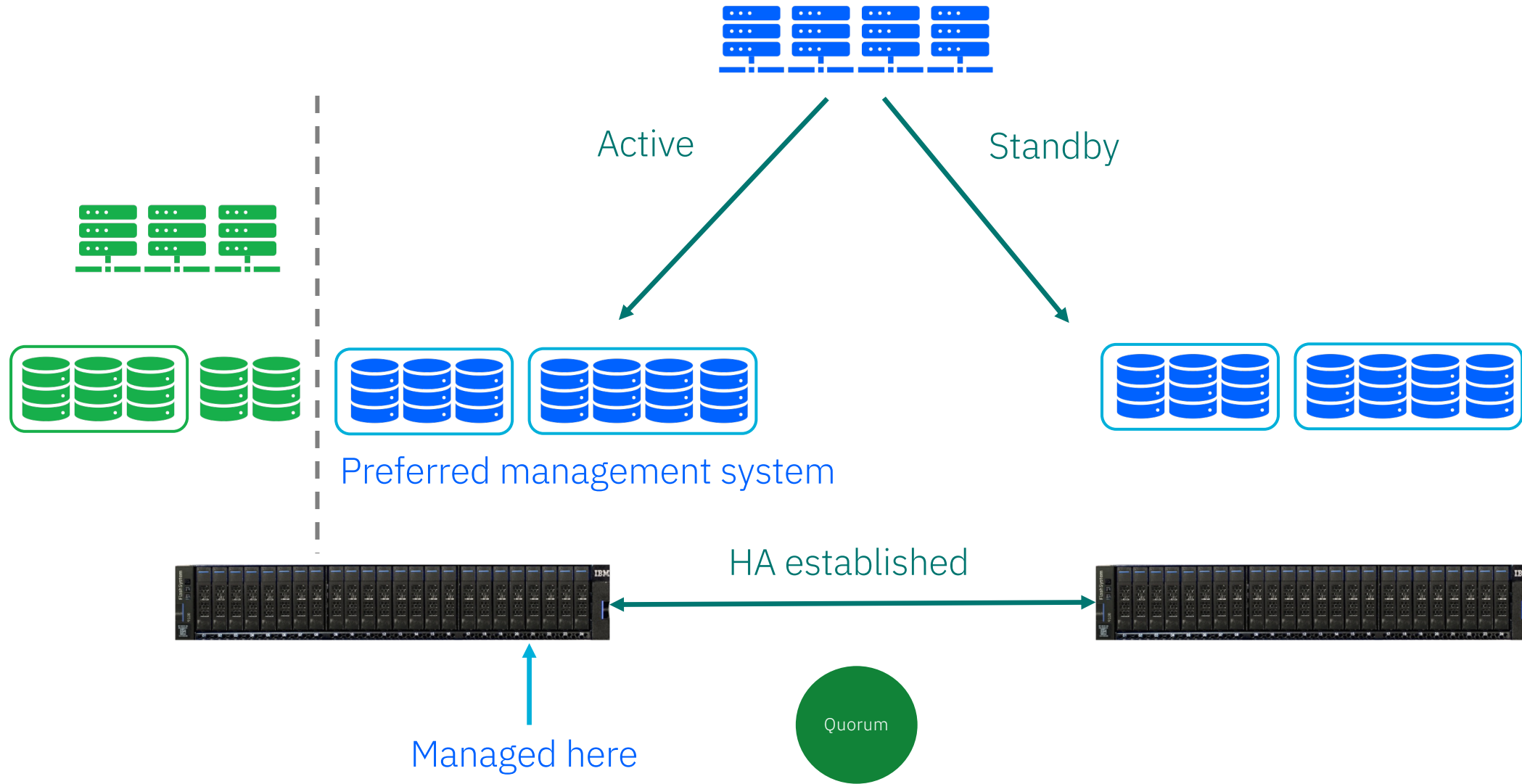
Host requirements

Limitations of first release

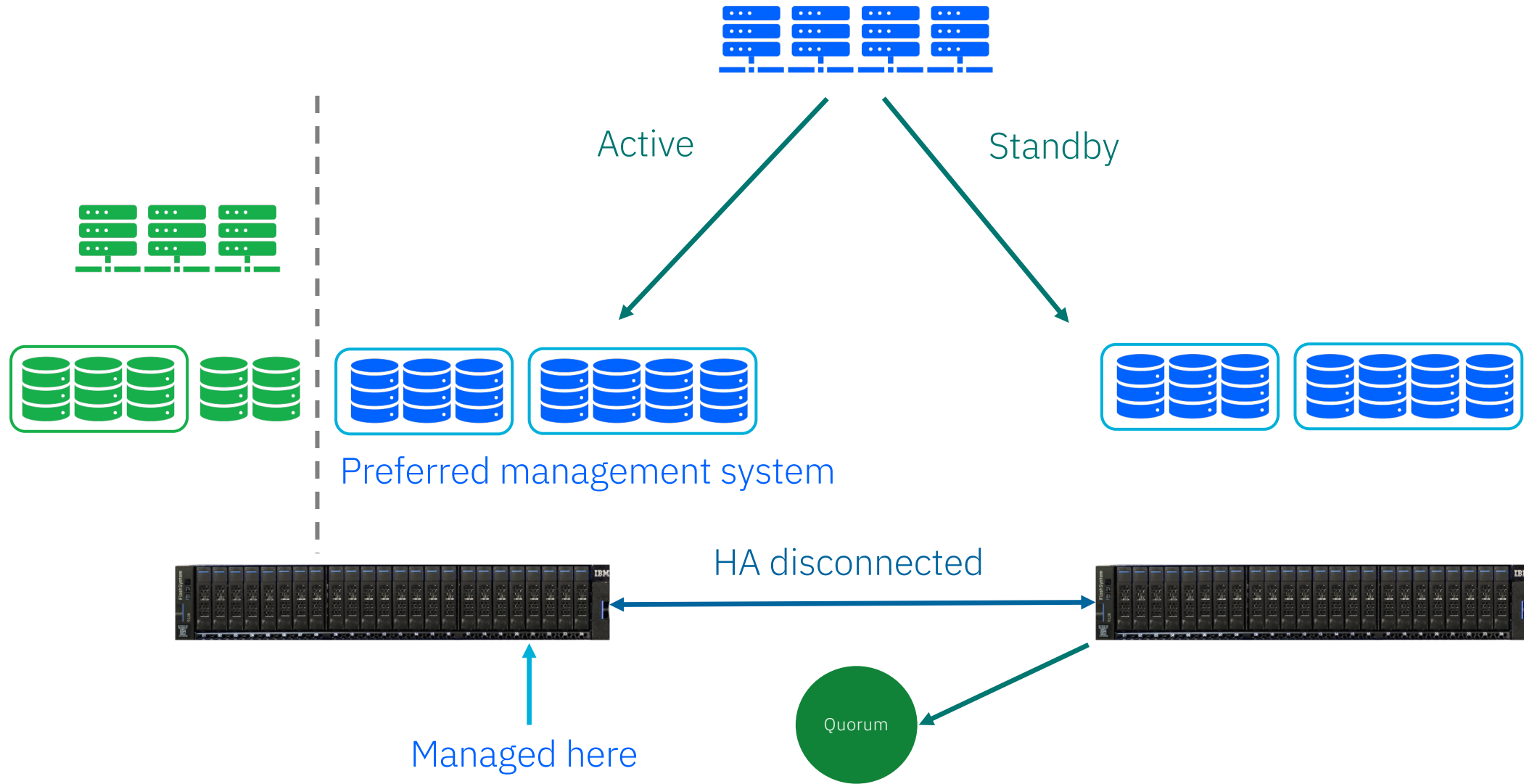
Useful information

Comparison vs HyperSwap and ESC

System Running Normally



HA Disconnected/System 1 cannot get to the quorum



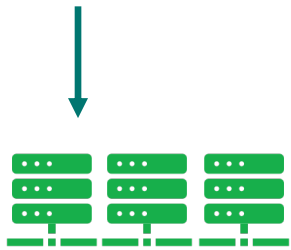
Operation While In Fault State

Non-HA volumes remain accessible



Offline

Active



Preferred management system

HA disconnected



Managed here

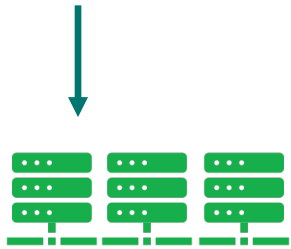
Problem Fixed - Resynchronizing

Non-HA volumes remain accessible



Standby

Active



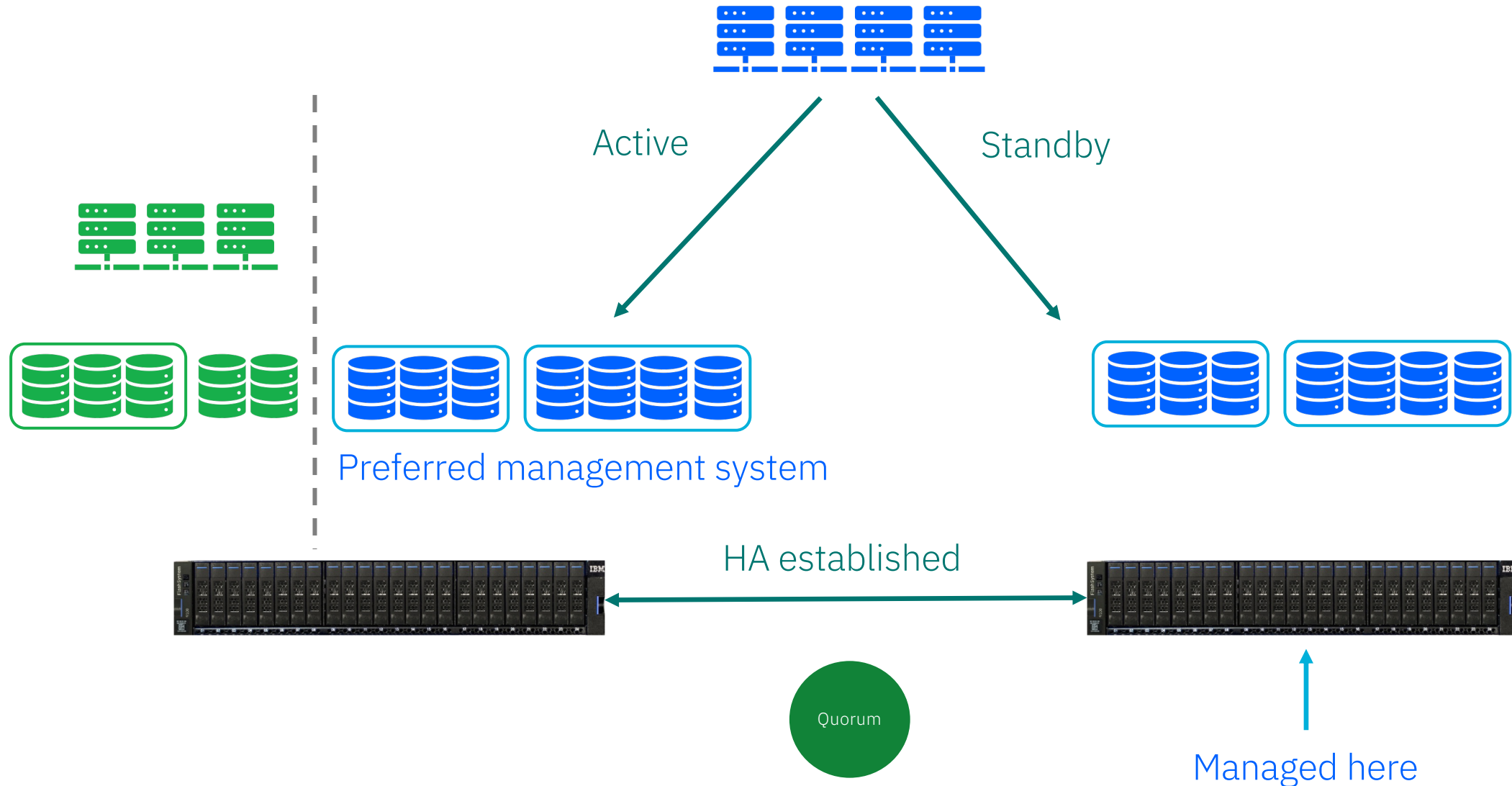
Preferred management system

HA establishing

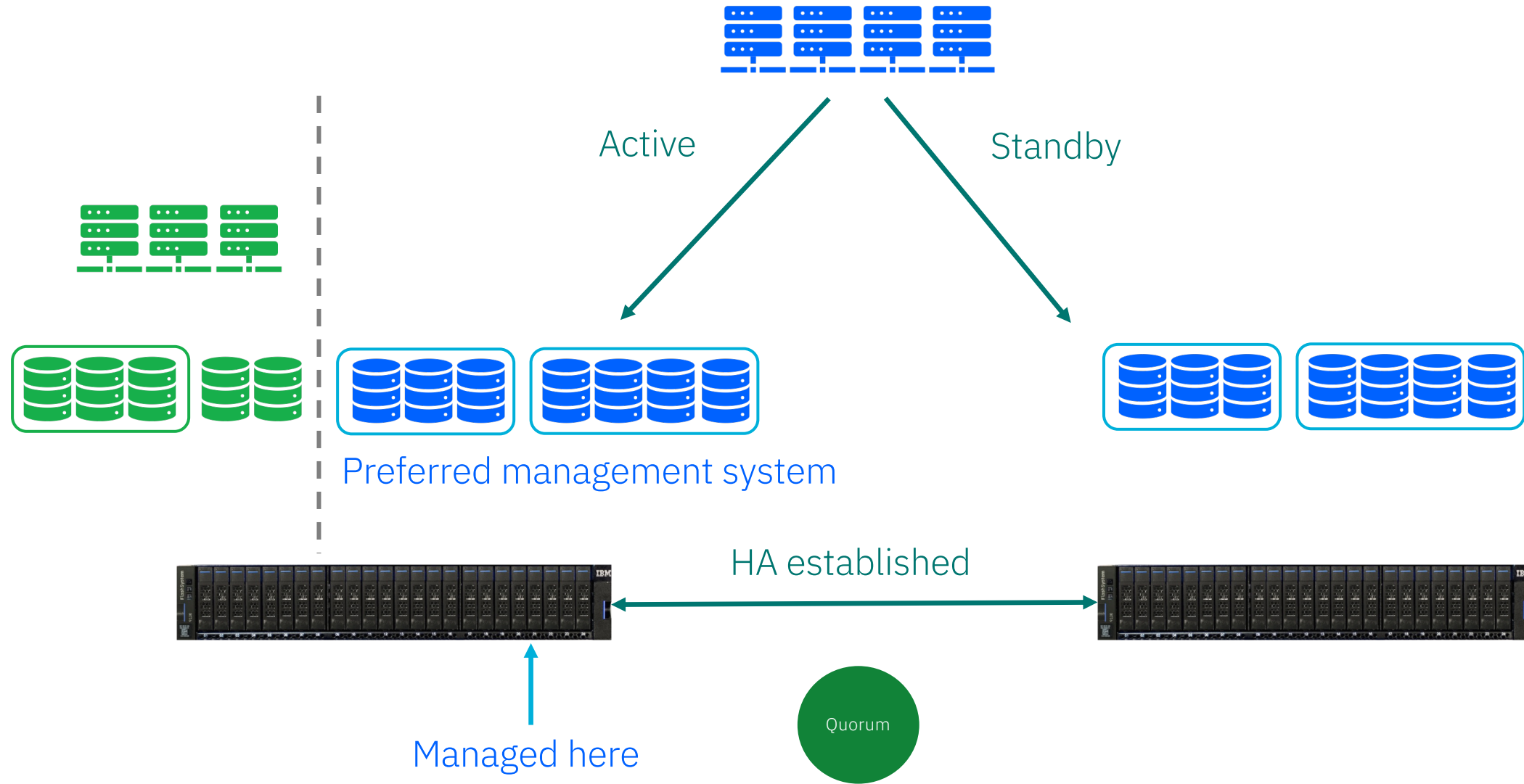


Managed here

Paths Return to Normal



Management Returns to Normal



How does HA behave?

Loss of connectivity between systems

Single node failure

Site failure (including dual-node failure)

Loss of all quorum apps

Offline arrays/pools

Slow/hung writes

System upgrades

Configuration problems

- If connectivity between the systems is lost for more than ~8 seconds, the systems will initiate a race for quorum
- Once a race is started, the partnership will be excluded and remain excluded for 15 minutes to protect against further link instability
- If both systems can access quorum, the I/O stall is around 10s (+/- 2s)
- If only one of the systems can access quorum, storage partitions may failover to the site that can access quorum. The system will be servicing I/O in about 15s (+/- 2s) from the disconnect
- HA is suspended and will attempt to re-establish automatically when the partnership reconnects
- Non-HA volumes have no impact or I/O stall

How does HA behave?

Loss of connectivity between systems

Single node failure

Site failure (including dual-node failure)

Loss of all quorum apps

Offline arrays/pools

Slow/hung writes

System upgrades

Configuration problems

- No change to existing behaviour
- Non-HA volumes have no impact
- HA will only suspend if there's a problem
- Slow write detection is suspended around node failover/failback to avoid false positives

How does HA behave?

Loss of connectivity between systems

Single node failure

Site failure (including dual-node failure)

Loss of all quorum apps

Offline arrays/pools

Slow/hung writes

System upgrades

Configuration problems

- Covers all cases where a storage array becomes unavailable
- A quorum race is triggered, and the timings are the same for the loss-of-connectivity case (10s/15s)
- **Quorum must be independent to the storage it's protecting – quorum must not be hosted on the storage it's protecting**
- HA will automatically attempt to re-establish when possible but manual intervention may be required to recover the system

How does HA behave?

Loss of connectivity between systems

Single node failure

Site failure (including dual-node failure)

Loss of all quorum apps

Offline arrays/pools

Slow/hung writes

System upgrades

Configuration problems

- Within a few seconds of the last quorum app disconnecting, the systems will switch to a mode without quorum that changes the HA behaviour
- In this mode, no quorum races happen and in the event of...
 - **disconnect** the storage partitions will continue to run on their current management system
 - **site failure** any storage partitions managed by the failed system will become unavailable
 - **offline storage** the storage partition will suspend HA and may failover, depending on which system has the fault
- An event is logged if there are no quorum apps

How does HA behave?

Loss of connectivity between systems

Single node failure

Site failure (including dual-node failure)

Loss of all quorum apps

Offline arrays/pools

Slow/hung writes

System upgrades

Configuration problems

- If a pool goes offline containing HA volumes, HA will suspend and may failover depending on which system has a problem
- All volumes in the storage partition failover together, even if only a subset of volumes are affected
- HA will attempt to recover automatically when the pool recovers

How does HA behave?

Loss of connectivity between systems

Single node failure

Site failure (including dual-node failure)

Loss of all quorum apps

Offline arrays/pools

Slow/hung writes

System upgrades

Configuration problems

- If a write takes more than 15s on either system, HA will suspend to protect against unresponsive backend storage
- Read latency is not tracked
- Slow write detection is disabled if the system is unstable

How does HA behave?

Loss of connectivity between systems

Single node failure

Site failure (including dual-node failure)

Loss of all quorum apps

Offline arrays/pools

Slow/hung writes

System upgrades

Configuration problems

- No affect on HA replication
- HA can be manually suspended by stopping the partnership during the upgrade
- Some configuration actions are blocked (from both systems) to HA storage partitions if it would require a configuration change on both systems

How does HA behave?

Loss of connectivity between systems

Single node failure

Site failure (including dual-node failure)

Loss of all quorum apps

Offline arrays/pools

Slow/hung writes

System upgrades

Configuration problems

- HA does not suspend if only management connectivity is lost between systems
- Loss of management connectivity or certificate errors prevent configuration changes if HA is established
- Some types of configuration changes will be failed, and rolled back, if the operation was not successful on both systems
- If HA is not established, configuration changes are permitted regardless of the remote system's state and the configuration changes will be rolled forward whenever possible

Agenda

Objectives

Feature overview

Quorum

HA behaviours

Host requirements

Limitations of first release

Useful information

Comparison vs HyperSwap and ESC

Host Requirements

Supported Operating Systems

Protocols

Interop

- GA code has limited host interop
- More interop is planned later in follow releases as testing allows.
- Supported OS types:
 - RHEL 7 and later
 - ESXi 7 and later

Host Requirements

Supported Operating Systems

Protocols

Interop

- GA code will only support the FC – SCSI protocol
- As with OS types we plan to introduce support for other protocols over time.
- Attachment
 - Fully zoned Fabric
 - I.e. hosts must be zoned to the virtual NPIV host ports (and able to login) to both storage systems
 - “Uniform” deployment model
 - Direct attach, useful across a very small area.

Host Requirements

Supported Operating Systems

Protocols

Interop

- Follow on releases will introduce support for
 - Host clusters
 - SCSI Persistent reserve commands
- RHEL Clustering is not supported at GA due to the Persistent reserve limitation.
 - Support tools can identify if PR has been attempted on a HA volume.
- Best failover times achieved with default settings in the multipathing:
 - DM multipath in linux should run with the default 5s polling time (do not increase to 30s).
 - Default ESXi settings

Agenda

Objectives

Feature overview

Quorum

HA behaviours

Host requirements

Limitations of first release

Useful information

Comparison vs HyperSwap and ESC

First Release Restrictions

- Cannot add existing volume groups to storage partitions, therefore no migration to policy-based HA
 - No host cluster support
 - No support for expanding volumes
 - No support for adding/removing host ports
 - Interoperability limited to a small set of FC SCSI host operating systems
 - Async replication cannot be used on another partnership while HA is configured
 - FC SCSI hosts only
 - FC partnerships only
 - No support for renaming HA objects
 - No support for multi-tenancy
- Development is investigating lifting restrictions as they can in coming releases

Agenda

Objectives

Feature overview

Quorum

HA behaviours

Host requirements

Limitations of first release

Useful information

Comparison vs HyperSwap and ESC

Reminder: Change Volumes

- Required for all replicated and HA volumes
- Automatically provisioned on both systems but hidden as the customer should never need to care about them
- Always thin-provisioned; compressed in DRP pools and dedupe'd if user-volume is dedupe'd
- Created in the same pool as the user's volume
- Has the same preferred node as the user's volume
- Capacity information visible as “protection capacity” for a volume/volume group/pool/system

Do not forget about change volumes when designing systems!

- They do not require or consume any licenses
- They do consume capacity and will grow and shrink if HA needs to resynchronise
 - Change volumes being offline due to any condition may prevent HA re-establishing – ensure you plan for their capacity needs
- They require twice the FlashCopy bitmap capacity as the volume's virtual capacity
 - E.g. A 1TiB replicated volume requires 2TiB of FlashCopy capacity on both systems (1MiB of bitmap capacity if using 256KB grains)

Partnerships

Data

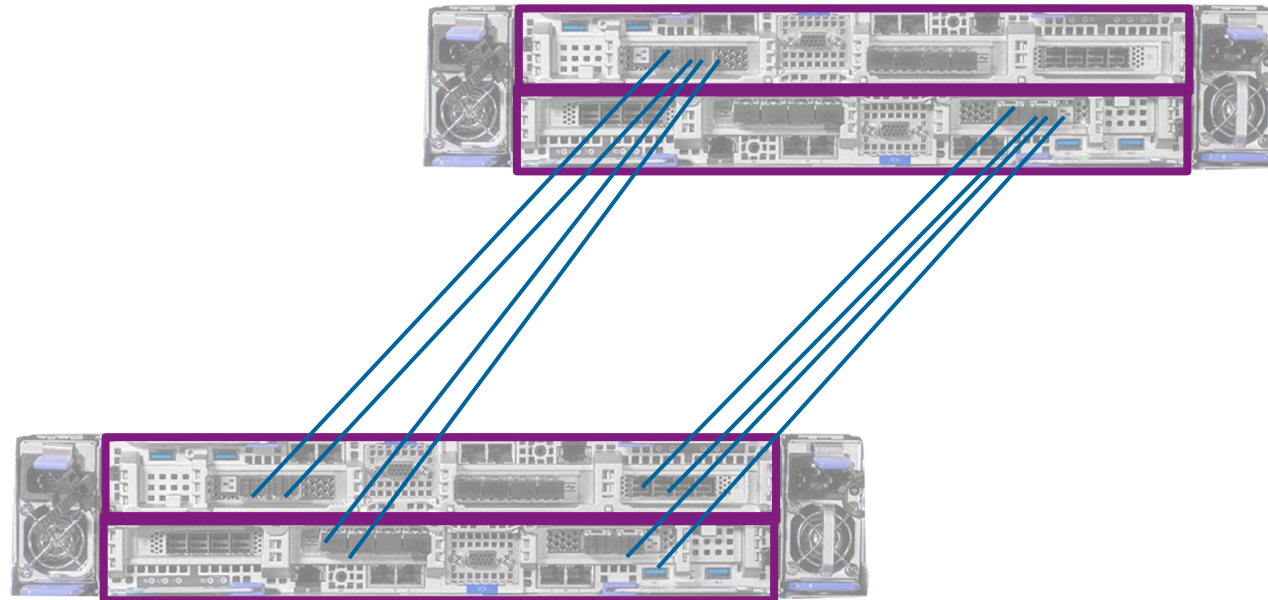
- FC partnerships supported up to 1ms RTT (~100km)
- Dedicate ports for HA traffic between systems
- Partnership bandwidth settings are not used to limit HA traffic
- You need high bandwidth, high quality links between the systems to avoid disruption

Management (same as async)

- IP connectivity between management IPs is required. Both IPv4 and IPv6 supported
- Firewall rules must allow for outbound and inbound traffic between the management IP addresses; port numbers are in the documentation
- All management traffic authenticated using certificates, configured during partnership setup
- Self-signed and CA-signed certificates supported

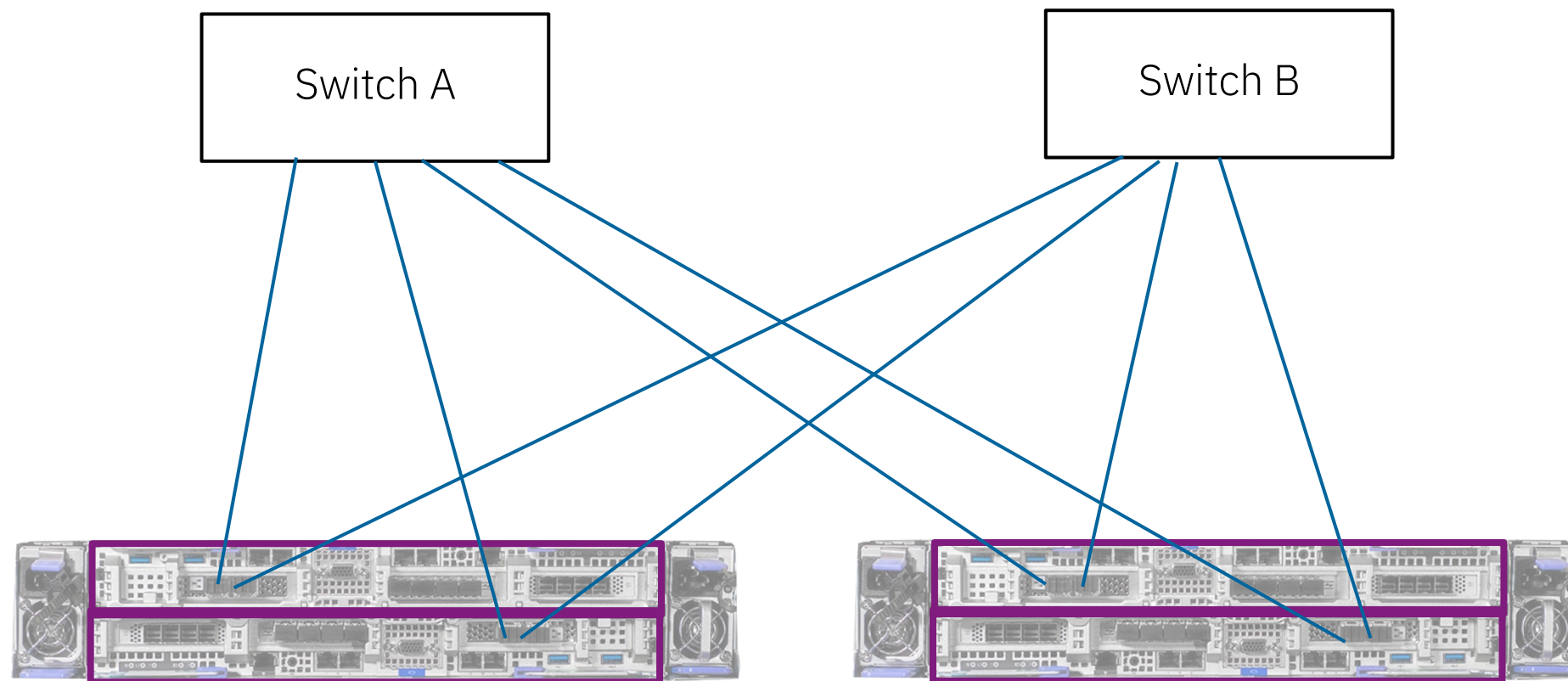
Supported connectivity: Direct attached FC

- Two cables between each pair of nodes recommended for redundancy
- Requires the same port speed on all nodes



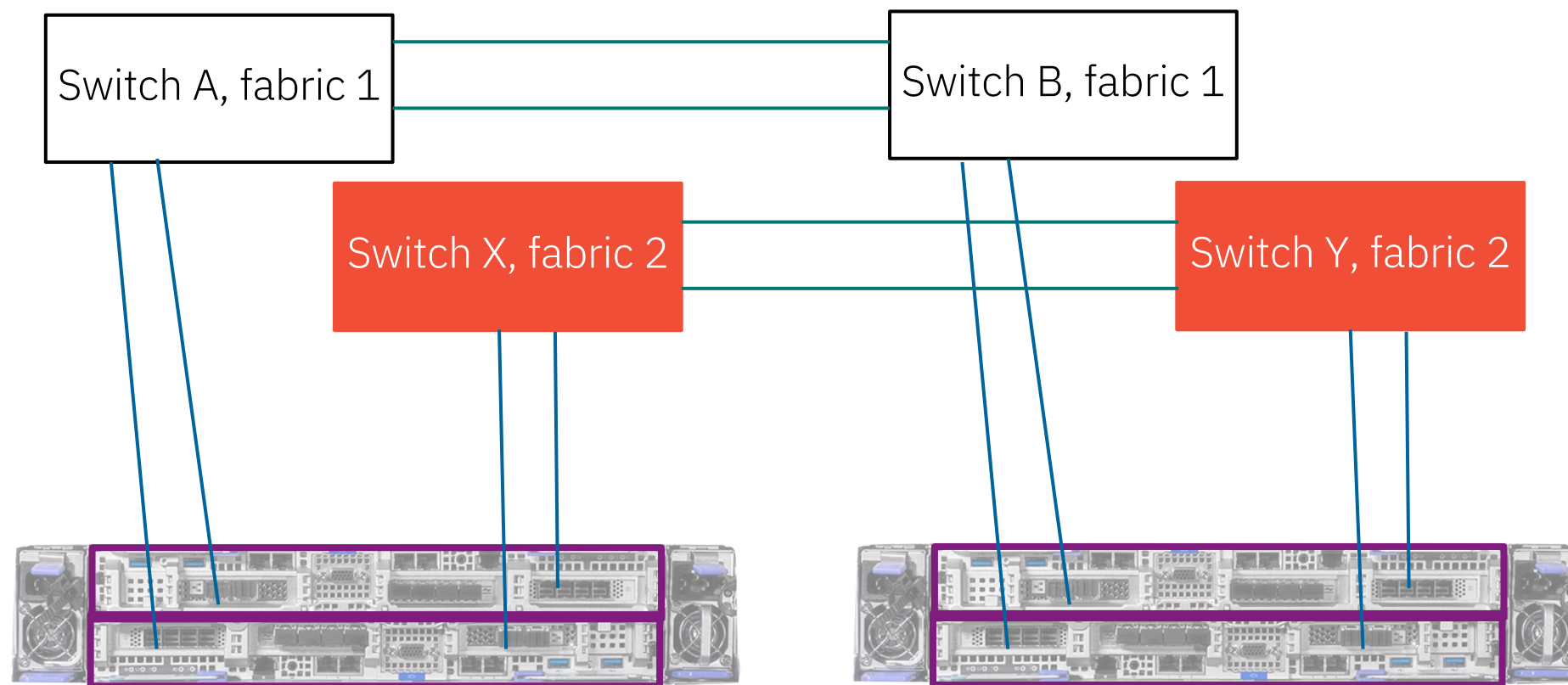
Supported Connectivity: Switch Attached

- Dedicate ports for connectivity between the systems using zoning
- Ensure redundant connectivity
- Minimum of 2 paths between each pair of nodes
- 4 paths between each pair of nodes recommended



Supported connectivity: ISL attached

- Dedicate ports
- Minimum of 2 paths between each pair of nodes
- 4 paths between each pair of nodes recommended
- Avoid sharing redundant DWDMs between fabrics
- Using virtual fabrics or QoS to isolate/prioritise HA traffic
- Private SAN used for partnership traffic/Public SAN used for host traffic



Planning for Policy-Based HA

- The feature uses the existing **remote copy** license and the existing rules apply. Does not require or consume a FlashCopy license
 - This only applies to SVC or virtualized storage, as all FlashSystems that are capable of Policy-based HA have all-inclusive licensing
- Remember change volumes!
- Bitmap capacity for policy-based replication/HA is consumed when the first replication policy is created for an I/O group
 - Shares bitmap with other features (Remote Copy, Mirroring, RAID)

Object limits:

- 32 replication policies per system
- 4 storage partitions per system (future increase planned)
 - Policy-based HA partitions can be replicating in different directions
- No imposed limit on the number of volume groups, hosts or host mappings per partition

Support for HA

Model	HA volume count	HA capacity
FlashSystem 50x5	Not supported	Not supported
FlashSystem 5200*	7932	1024 TiB
FlashSystem 7200	7932	2048 TiB
FlashSystem 7300	15864**	2048 TiB
FlashSystem 91x0/9200	7932	2048 TiB
FlashSystem 9500	15864**	4096 TiB
SAN Volume Controller	7932**	2048 TiB**
SV Public Cloud	Not supported	Not supported

* Requires minimum of 64GB of memory per node. ** Increase being investigated

Miscellaneous Info

Volume groups

Naming

Synchronization

Interaction with other policies

System rules

Interaction with other features

- Unlike asynchronous replication, HA is managed on the storage partition and affects all volume groups in that partition
- The volumes for a host or application should be in the same volume group
- Volumes cannot be moved between replicated volume groups
- Existing volumes cannot be added to HA volume groups (create them in the volume group)
- Volumes must be in a volume group to be HA (no standalone HA volumes)
- Volumes in the remote system are always provisioned by the system; existing volumes will not be used
- **Plan to use volume groups and storage partitions, even if you don't need HA today to simplify future feature adoption...**

Miscellaneous Info

Volume groups

Naming

Synchronization

Interaction with other policies

System rules

Interaction with other features

- Volume, volume group, host and partition names must be unique only on the system they are created on and the remote system if HA
- Naming conflicts will cause configuration errors
- HA volumes, volume groups, hosts and partitions cannot be renamed
- The name is required to be unique on both systems so it can be uniquely identified on both systems (as the ID is not guaranteed/unlikely to be the same)

Miscellaneous Info

Volume groups

Naming

Synchronization

Interaction with other policies

System rules

Interaction with other features

- HA is only established if every volume in the partition is synchronised
- Configuration actions that break HA unintentionally are generally prevented
- Creating volumes in HA volume groups skips the need for a full synchronisation
- Change volumes are used automatically during synchronization to protect the consistent image at the recovery location
- In the case of rolling failures where HA was not fully re-established between failures, DR-like access can be enabled to the most recent consistent copy of volumes in the storage partition

Miscellaneous Info

Volume groups

Naming

Synchronization

Interaction with other policies

System rules

Interaction with other features

- Volume groups can be associated with snapshot or SafeGuarded policies, but they apply only on the system on which they are assigned
- Clones and thin-clones cannot be added to storage partitions
- Asynchronous replication policies cannot be added to HA volume groups (...yet)
- Asynchronous replication policies cannot be assigned to storage partitions and HA policies cannot be assigned to volume groups

Miscellaneous Info

Volume groups

Naming

Synchronization

Interaction with other policies

System rules

Interaction with other features

- Policy-based HA can only be used with systems in ‘standard’ topology – only a single type of HA is permitted on a system at a time (otherwise HA behaviours are indeterminate)
- Only a single I/O group permitted in a system using policy-based HA
- HA is permitted only on a single Fibre Channel partnership
- Replicated volumes must have cache enabled
- No support for HA vVols
- Image-mode volumes are not supported
- HA and async replication are not permitted on the same partnership

Miscellaneous Info

Volume groups

Naming

Synchronization

Interaction with other policies

System rules

Interaction with other features

- HA volumes cannot be replicated using Remote Copy
- No migration path from Metro Mirror or HyperSwap to policy-based HA
- HA volume groups cannot use Transparent Cloud Tiering
- Restore-in-place using volume group snapshot is not supported
- Volumes that are the target of a legacy FlashCopy map cannot be in a storage partition
- Replicated volumes cannot be resized

Agenda

- Objectives
- Feature overview
- Quorum
- HA behaviours
- Host requirements
- Limitations of first release
- Useful information
- Comparison vs HyperSwap and ESC

HA comparison

	Policy-based HA	HyperSwap	Enhanced Stretched Cluster
Supported on	SVC and NVMe products. 1 I/O group systems	2+ I/O group systems	2+ I/O group SVC only
Maximum HA capacity	Up to 4PiB	Up to 2 PiB per pair of I/O groups, 4PiB per system	40 PiB per I/O group (SV3), 1PiB (SV2, SA2)
Maximum HA volume count	15864 (more planned)	2000	7932
Non-HA volumes	Unaffected by HA problems	Link problems cause offline volumes	Link problems cause offline volumes
Protects against outage of entire system	Yes	No	No
Protects (mutual) consistency during resynchronisation	Yes	Yes (mutual consistency with consistency groups)	No
Host interoperability support	Limited (with planned increase in support)	Full	Full
HA snapshots (including Safeguarded)	No*	No*	Yes

* Volume Group Snapshots would be taken on both sides rather than making the snapshots HA

HA Comparison

	Policy-based HA	HyperSwap	Enhanced Stretched Cluster
3-site support	No	Yes – externally orchestrated async	Synchronous with Metro Mirror, async with policy-based replication
Volume resize support	No	Expand thin volumes	Expand/shrink
Quorum	App only, preferred-site behaviour per storage partition	App or controller, default/preferred/winner per system	App or controller, default/preferred/winner per system
Maximum nodes per site	2 (1 I/O group)	4 (2 I/O groups)	4 (half of each I/O group)
Mixed hardware models	Yes, unrestricted	Limited (must cluster)	Limited (must cluster)
Mixed software levels	Yes (future, within range)	No	No
Licensing	Remote, included in LMC	Remote, included in LMC	Base

Policy Based HA Performance Summary

- Policy-based HA offers significant higher performance with lower latency over HyperSwap. The performance improvement is more distinguishable on the higher end systems with more CPU cores than the lower end systems. Consider the following:
 - Better throughput/Bandwidth compared to HyperSwap (up to 8x throughput improvement)
 - Better latency compared to HyperSwap (up to 50% latency reduction)
 - Exceptional lower overhead compared to no replication performance.
- Policy-based HA is a better mirror option than HyperSwap.

Support Statements



VMware vVols

- 8.6.0 Long Term Support Release Support
 - IBM Storage (Spectrum) Connect
 - Embedded VASA Provider
 - Default for new configurations
- 8.6.1+ Continuous Development Release Support
 - Embedded VASA Provider Only

Background

- LTS 8.6.0.0 supports v1 and v2 metadata volumes
- A new metadata volume created on 8.6.0 will be version 2 by default if the system supports the new VASA provider
 - No reconfiguration is required before upgrading to 8.6.0
- A metadata volume created on earlier versions uses version 1
 - **Reconfiguration is required before upgrading to 8.6.1+**

Upgrade Test Utility

- The Upgrade Test Utility will check the configuration and prevent upgrade beyond 8.6.0.x if a version 1 metadata volume exists
- The system identified that a version 1 metadata volume exists for VMware Virtual Volumes (vVols). Please see the following web page for the actions that must be completed before this upgrade can be performed:
 - <https://www.ibm.com/support/pages/node/6983196>
- No restrictions on upgrades to 8.6.0.x PTFs

Documentation

- For systems already using the embedded VASA provider, the 8.6 Docs have a page with the instructions required to migrate:
- [Migrating from VASA 1 to VASA 2 Metadata](#)
- Migration is a simple single-step process and does not require any VMware reconfiguration

Documentation

- For systems still using Spectrum Connect, the following Redbook has instructions to migrate to the embedded VASA provider:
- [IBM FlashSystem and VMWare Implementation and Best Practices Guide](#)
- Migration is a multi-step process and requires reconfiguration in VMware vCenter

Limitations

- The following products do not support the embedded VASA provider:
 - IBM FlashSystem 5015
 - IBM FlashSystem 5035
 - IBM FlashSystem 5045
- These systems cannot upgrade beyond 8.6.0 with vVols configured.
- These systems can continue to use vVols with Spectrum Connect on 8.6.0.x

VMware Integration Modernization

Feature	Spectrum Connect	New World	Notes
Virtual Volumes (vVol)	Spectrum Connect VASA Provider	Embedded VASA Provider	SAN Volume Controller, FlashSystem 5200,7xxx,9xxx 8.5.1.0 and later
vSphere Plugin	IBM Storage Enhancements for vSphere Web Client (vWC)	IBM Storage Virtualize Plugin for vSphere	https://ibm.biz/ibmvsphereplugin https://ibm.biz/vspherevideo GA - V1.1 June 2023
vRealize Operations	IBM Storage vROps Adapter	Management Pack for IBM SVC and Storwize	https://ibm.biz/ibmvrops 3 rd party plugin
vRealize Orchestrator	Spectrum Connect vRealize Orchestrator/ Automation	-	

NPIV

- Starting with 8.6.1+ Storage Array NPIV will be required to be ON
 - Transitional Mode will NOT work
 - 95% of our clients are already there
 - Long term SVC clients may have some work to do to get ready

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

