IBM Storage FlashSystem grid

Barry Whyte
Principal Storage Technical Specialist

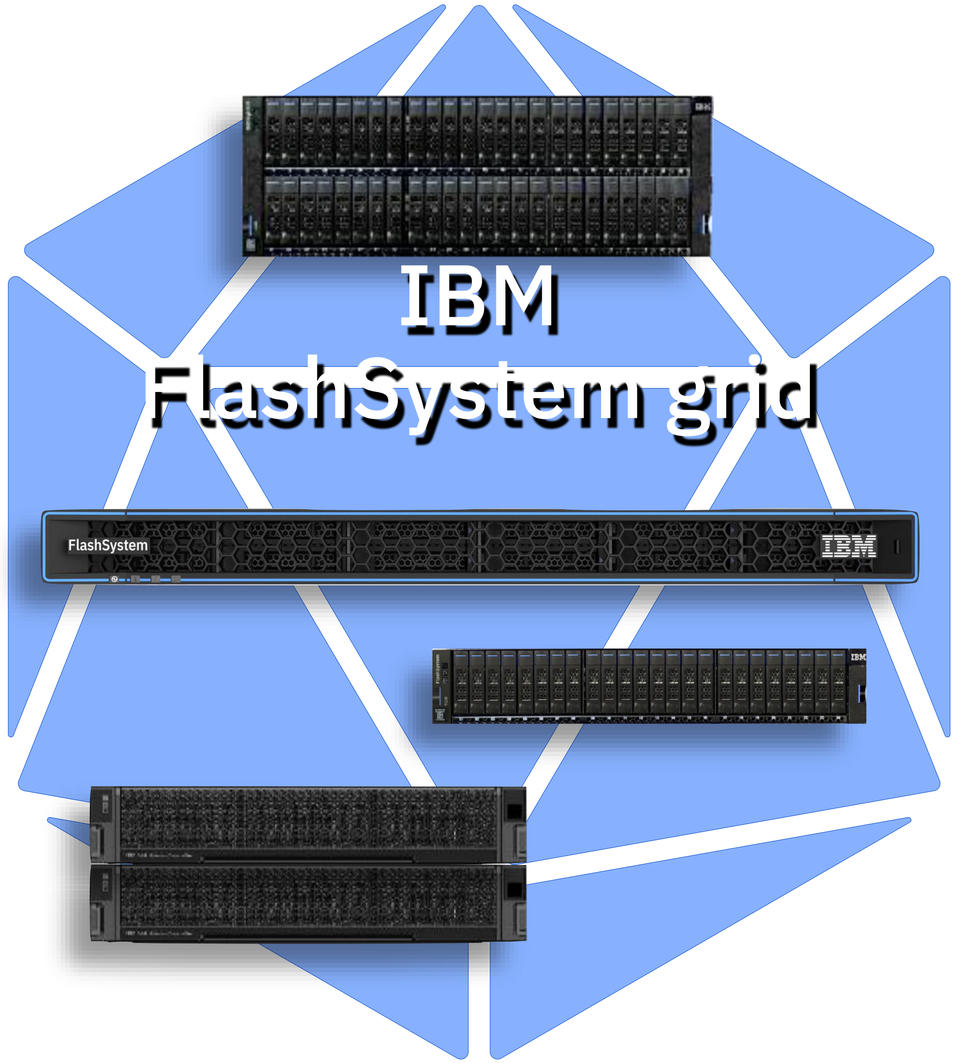
Spring Storage Day

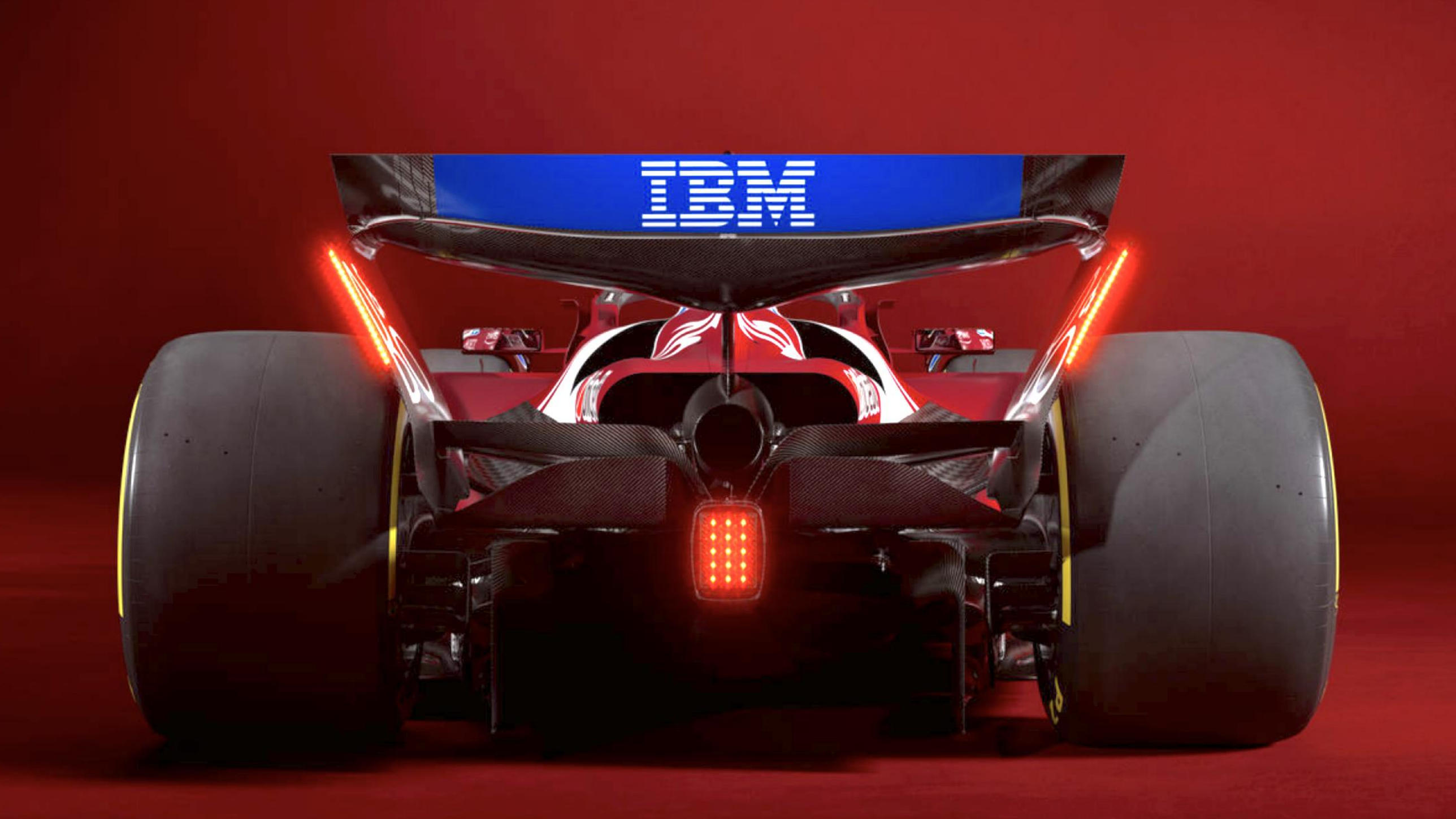
April 2025

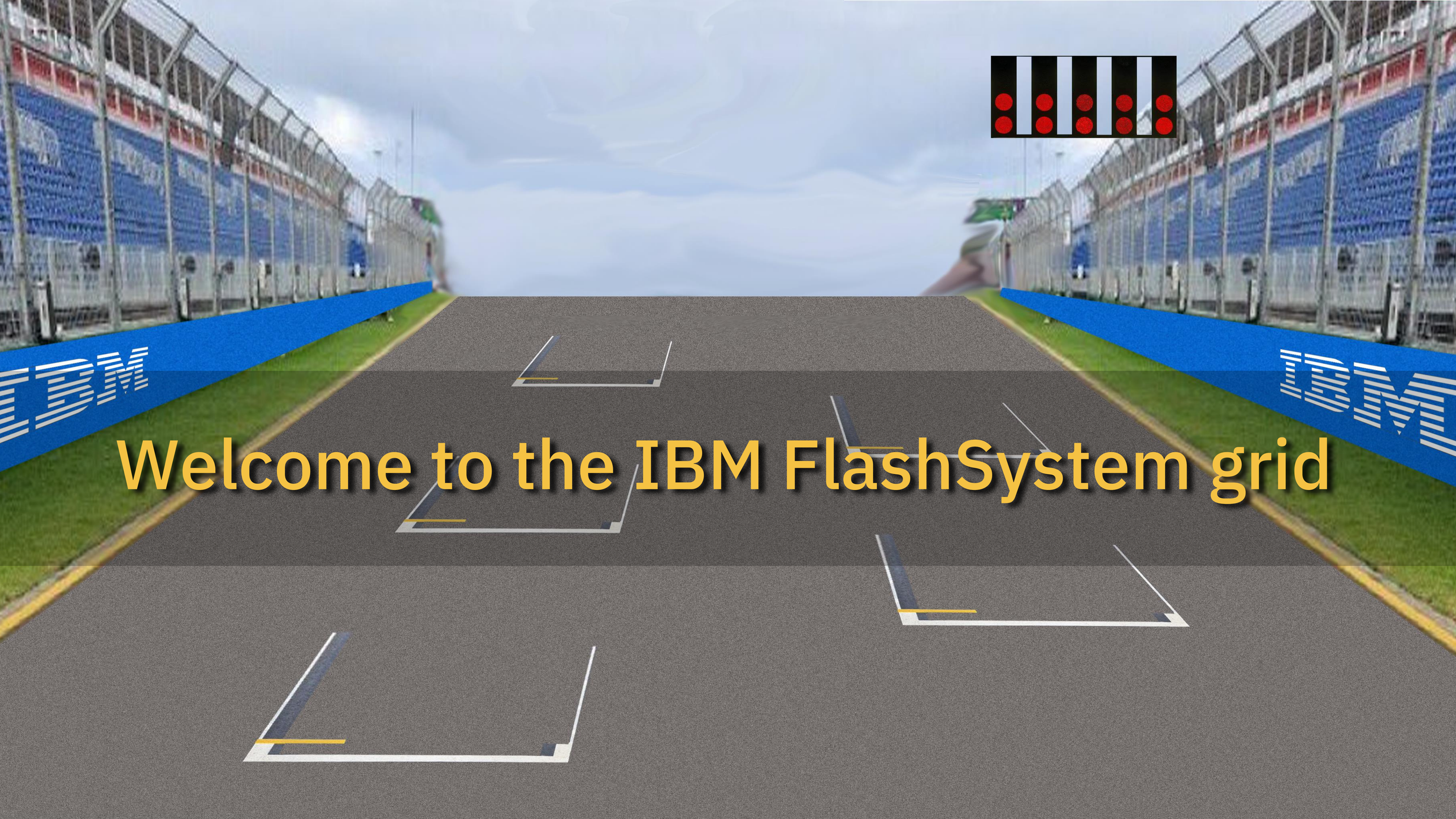
IBM FlashSystem grid

Introducing a new way to address scale-out

A **FlashSystem grid** is a scalable storage platform comprising of multiple **FlashSystem** or **SVC** systems that enables operational simplicity through federated management, AI-powered data placement recommendations and flexible deployment options.

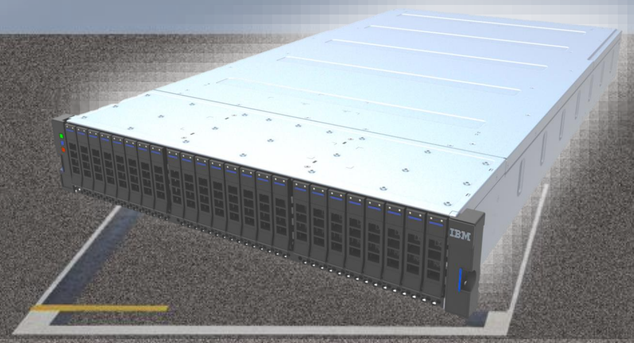



IBM



Welcome to the IBM FlashSystem grid

Welcome to the IBM FlashSystem grid



#5 FlashSystem C200

New for 2025 :

Capacity optimised flash for general purpose and secondary storage

Stats : ~1ms, 200K IOPs, 23GB/s

Single config : 2RU - 1.1PB raw, 2.3PBe

FCM4 features at better than Near-line latency.

Welcome to the IBM FlashSystem grid

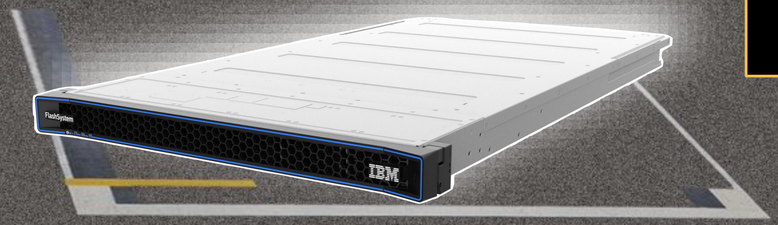
#4 FlashSystem 5300

Compact 1U design for data centre and hybrid cloud efficiency and constrained edge environments

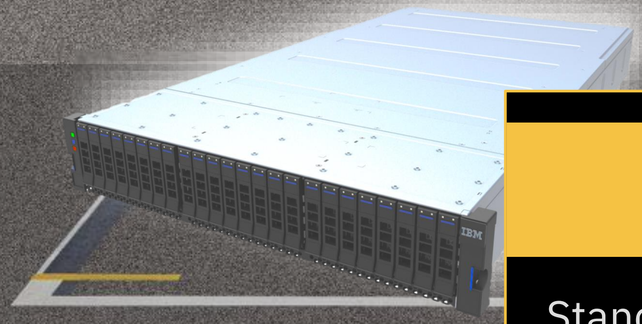
Stats : ~70us, 400K IOPs, 30GB/s

Single system: 1RU - 0.35PB , 1PBe

FCM4 performance at entry price point



Welcome to the IBM FlashSystem grid



#3 FlashSystem 7300

Standard 2U24 design. The workhorse midrange system evolved from the highly regarded Storwize V7000

Stats : ~70us, 580K IOPs, >50GB/s

Single system: 2RU - 0.77PB , 2.3PBe

FCM4 performance and capacity

Welcome to the IBM FlashSystem grid

#2 FlashSystem 9500

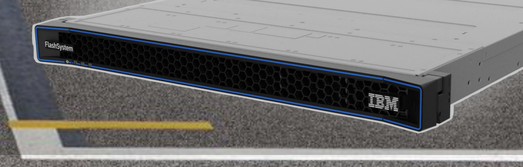
4U48 design. The “beast” boasting enterprise class reliability and redundancy with the expected feature set to provide data availability and resiliency - always.

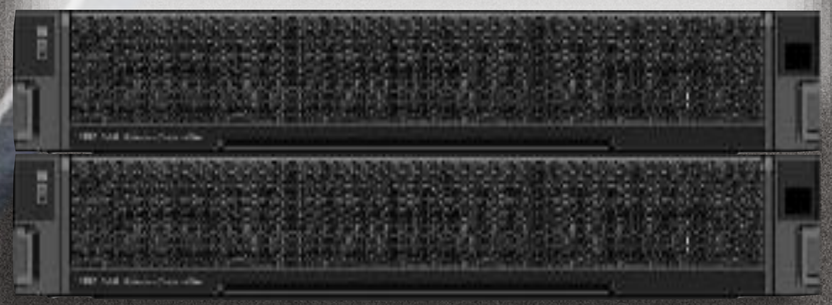
Stats : ~70us, >1.6M IOPs, >90GB/s

Single system: 4RU - 1.58PB , 4.74PBe

Ultimate FCM4 performance and capacity for most demanding workloads

Welcome to the IBM FlashSystem grid





#1 SVC - SV3

IBM SAN Volume Controller, where it all began over 20 years ago... Virtualising IBM and non-IBM environments today and tomorrow.

Stats : ~50us, >4M IOPs, >360GB/s

Single system: 32PB max

Bring your non-IBM storage to the grid!

Welcome to the IBM FlashSystem grid

Who can join a grid?

FlashSystem C200

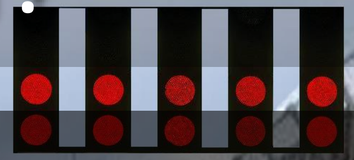
FlashSystem 5200,5300

FlashSystem 7200,7300

FlashSystem 9100,9200,9500

SAN Volume Controller SA2, SV2, SV3

Why IBM FlashSystem grid ?



Scale-out performance and resources as needed

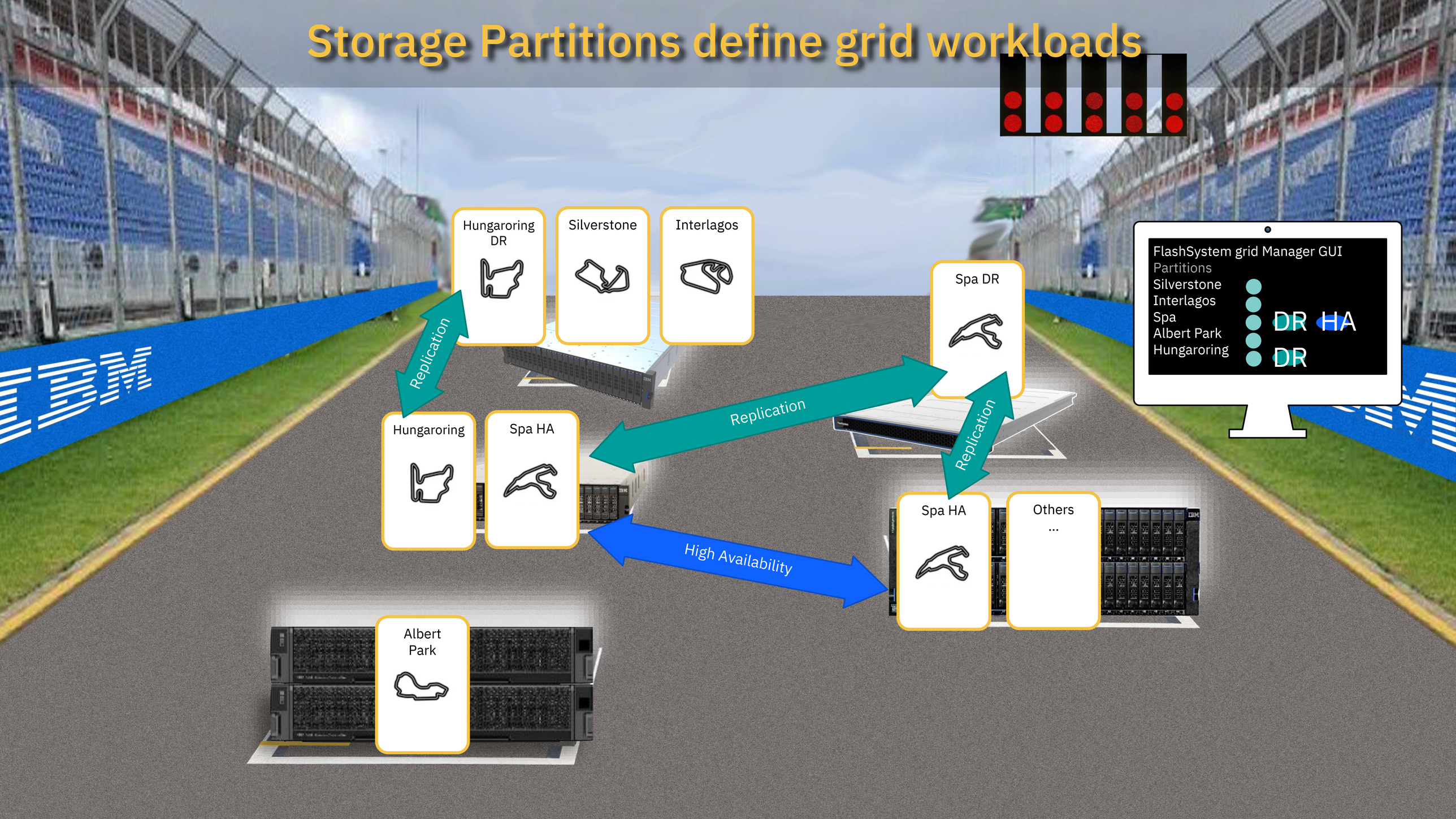
Federated ‘*at-a-glance*’ management of all grid members

Non-disruptive workload migrations within the grid

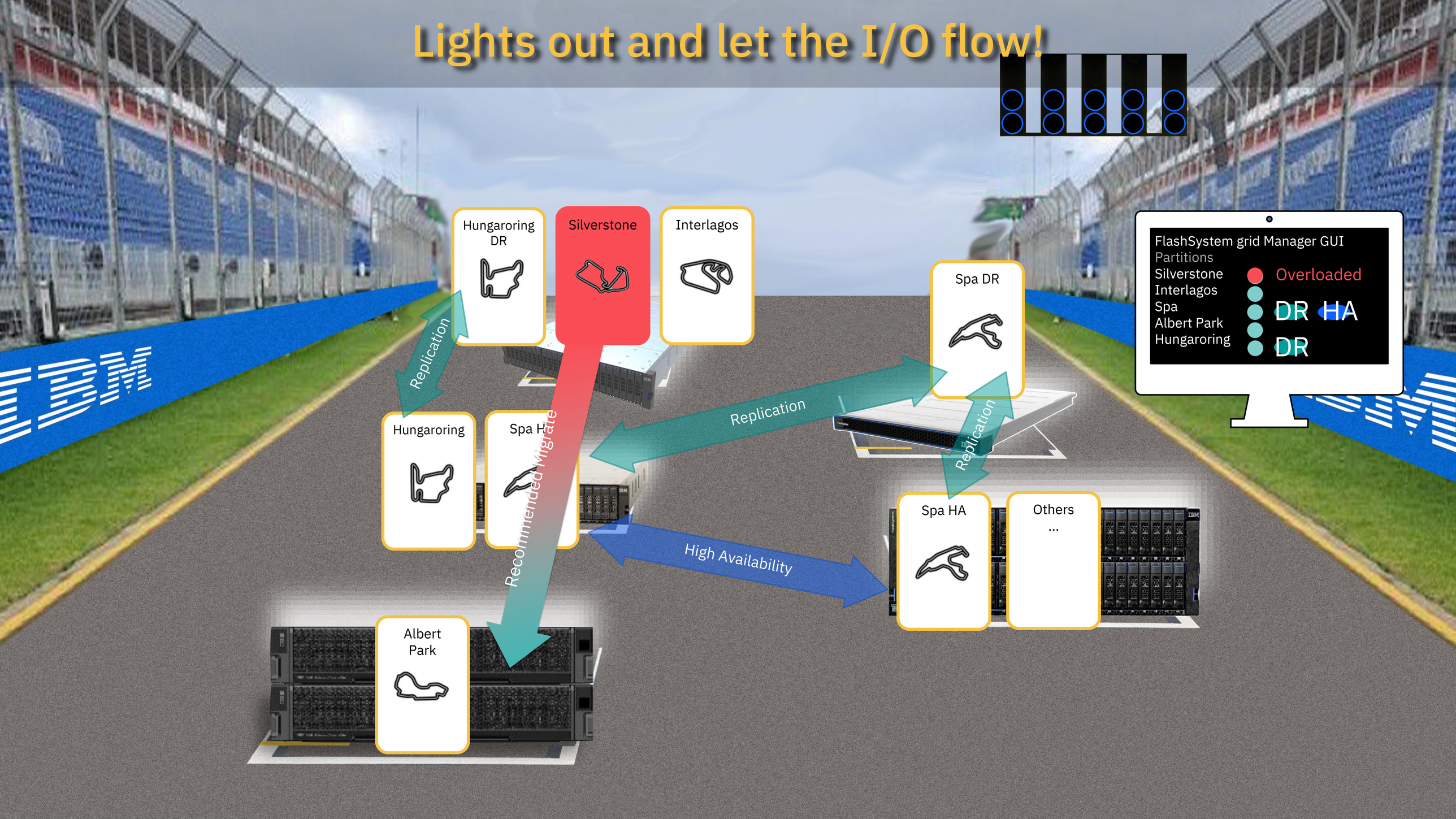
AIops data placement recommendations in the grid

Add capacity and performance at the right price point

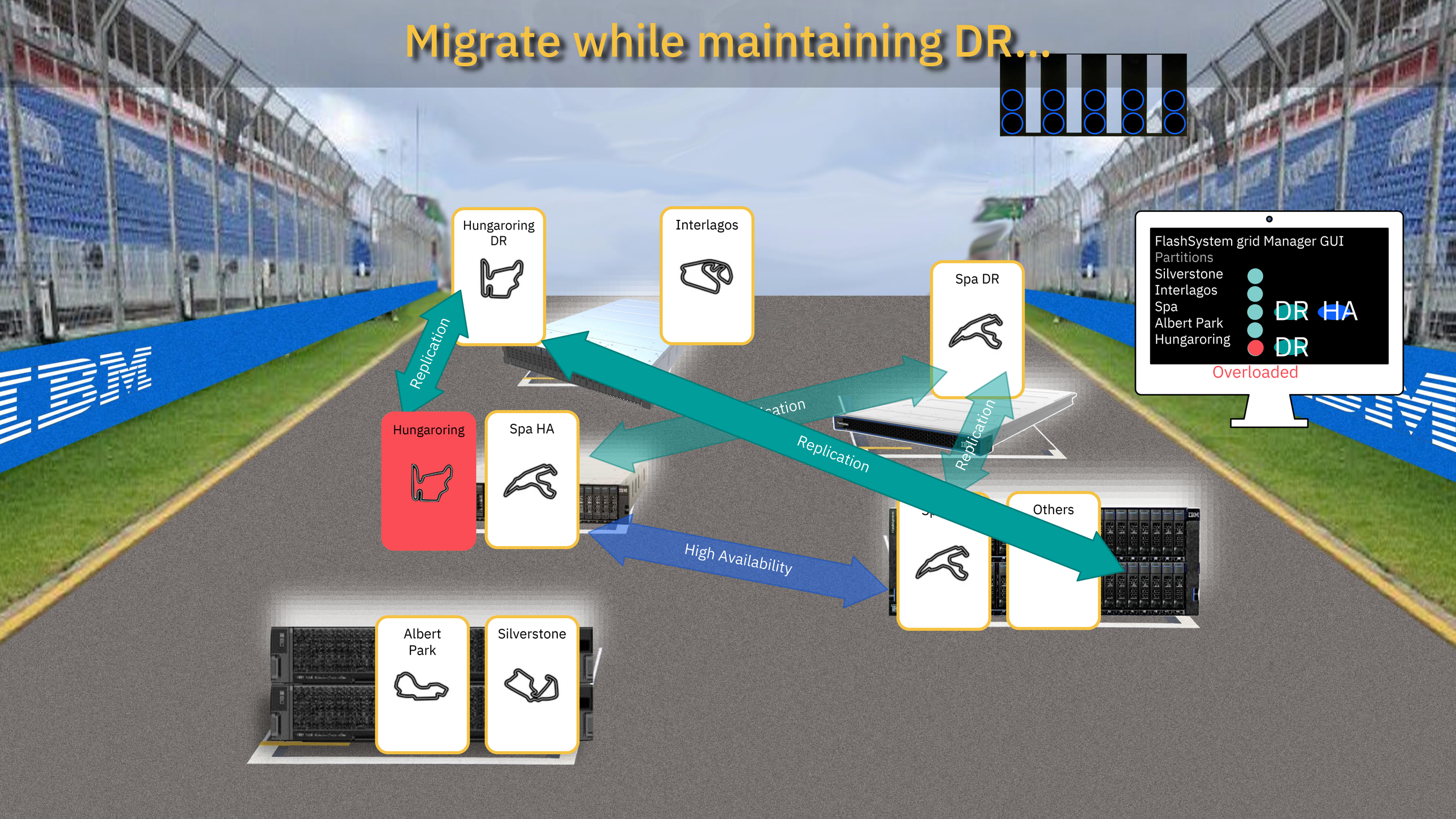
Storage Partitions define grid workloads



Lights out and let the I/O flow!



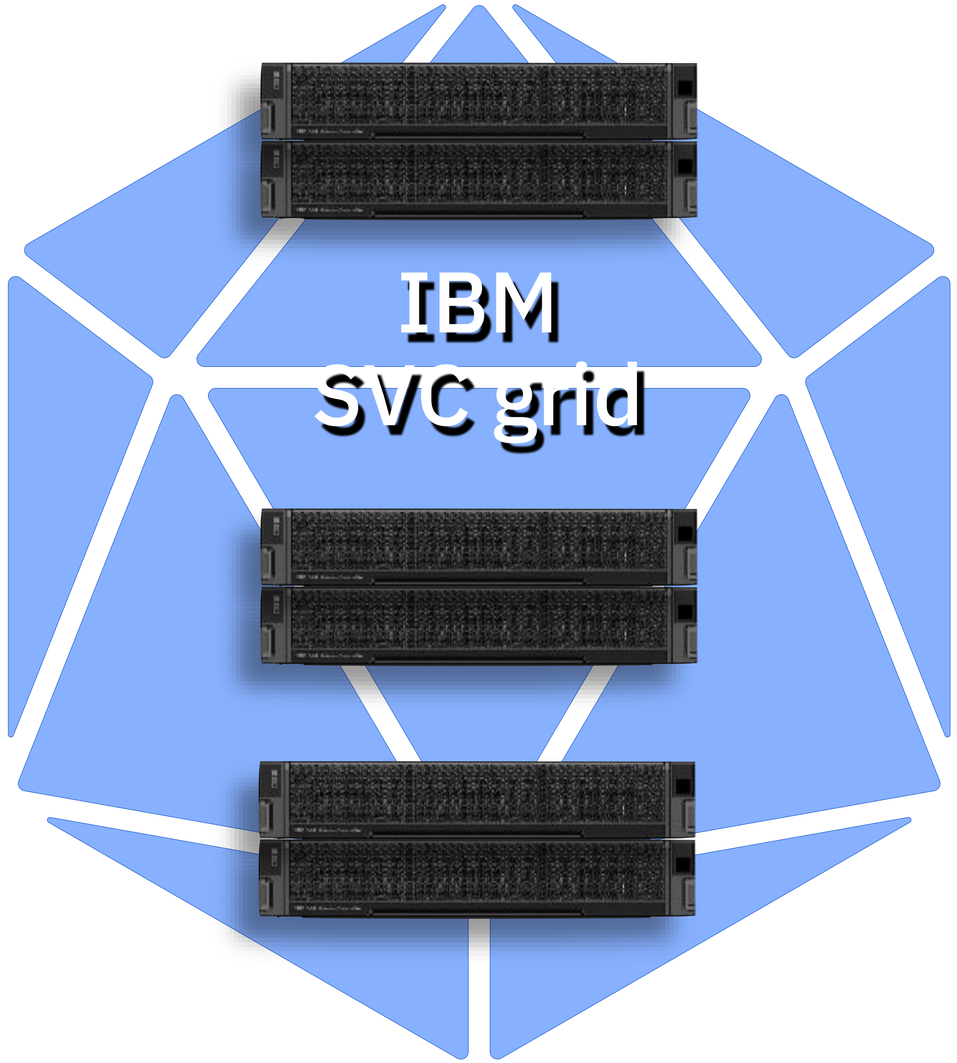
Migrate while maintaining DR...



IBM ~~FlashSystem~~ SVC grid

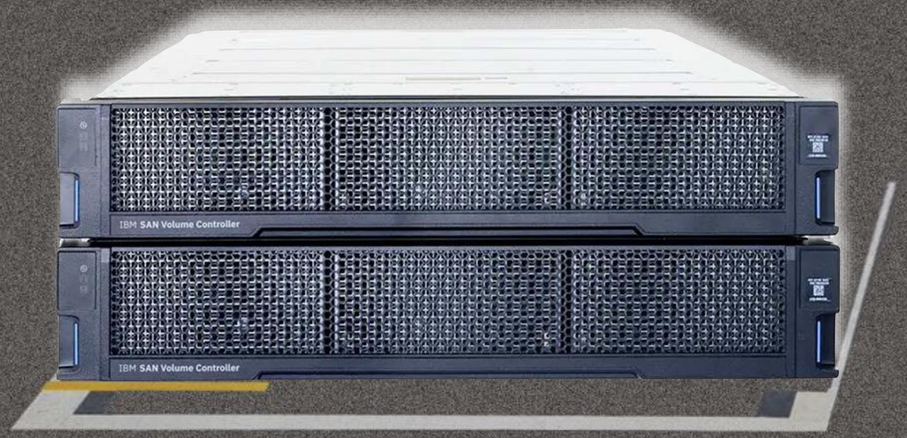
Introducing a new way to address scale-out

An **SVC grid** is a scalable storage platform comprising of multiple **SVC** systems that enables operational simplicity through federated management, AI-powered data placement recommendations and flexible deployment options.



Welcome to the IBM SVC grid

Welcome to the IBM SVC grid



#3 SVC SA2

Welcome to the IBM SVC grid

#2 SVC

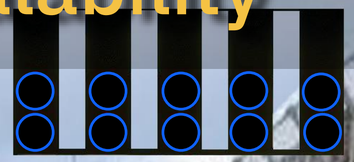
SV2

Welcome to the IBM SVC grid

#1 SVC

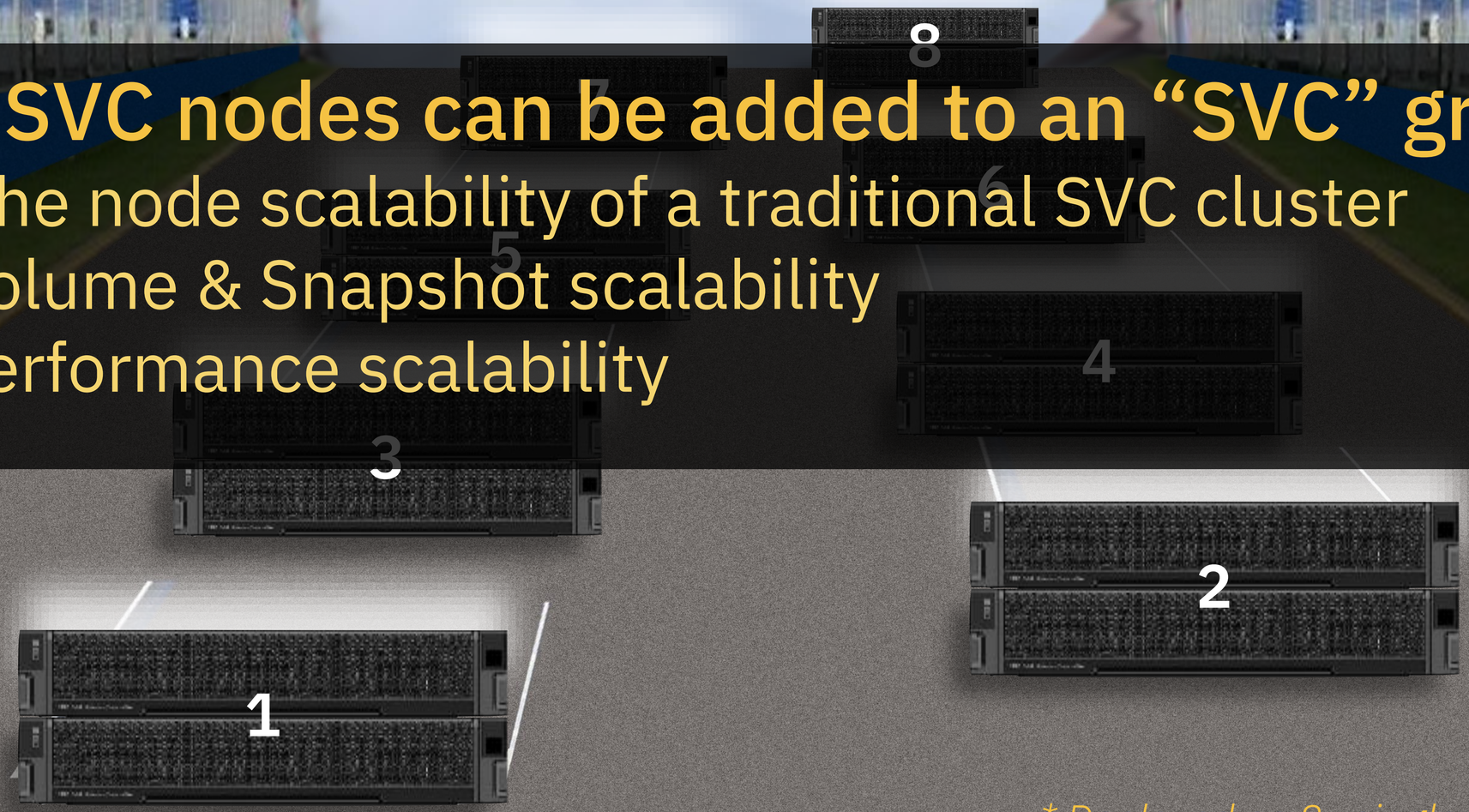
SV3

SVC grid - Extended SVC Scalability



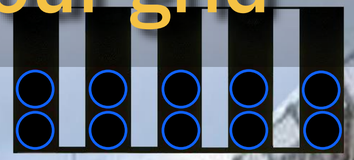
Up to 16 SVC nodes can be added to an “SVC” grid today*

- Double the node scalability of a traditional SVC cluster
- 8x the Volume & Snapshot scalability
- 2x the performance scalability



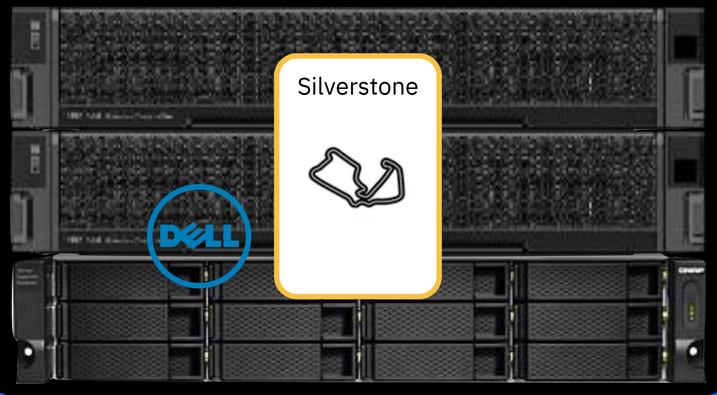
** Deployed as 8x single I/O Group systems*

SVC - Bring non IBM Storage to your grid



SVC can bring grid mobility and AIops to your non-IBM workloads. Migrate and lifecycle without disruption.

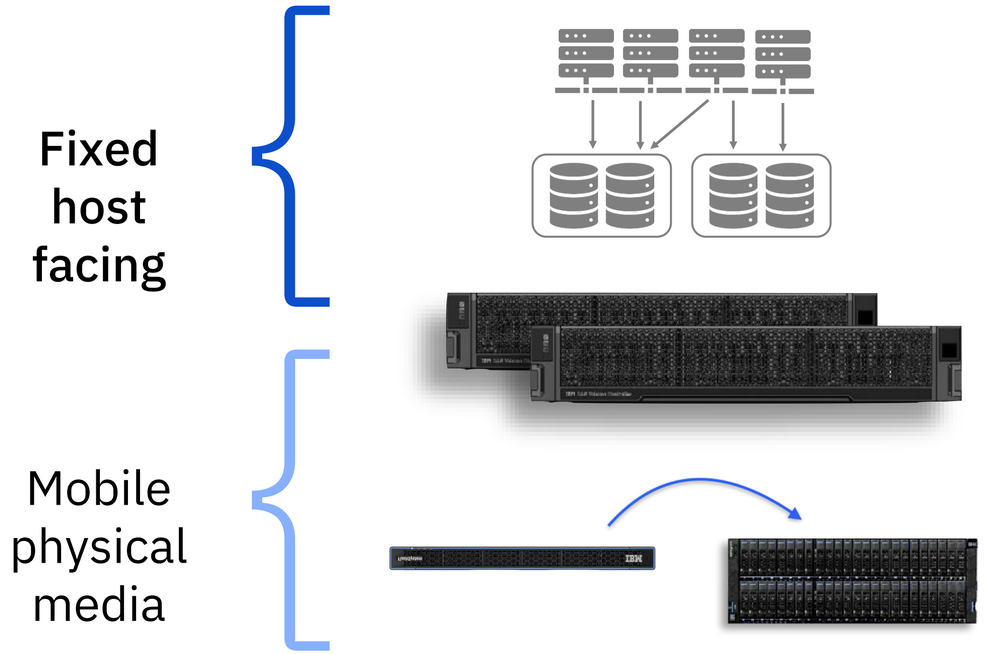
SVC Virtualizing non-IBM Storage



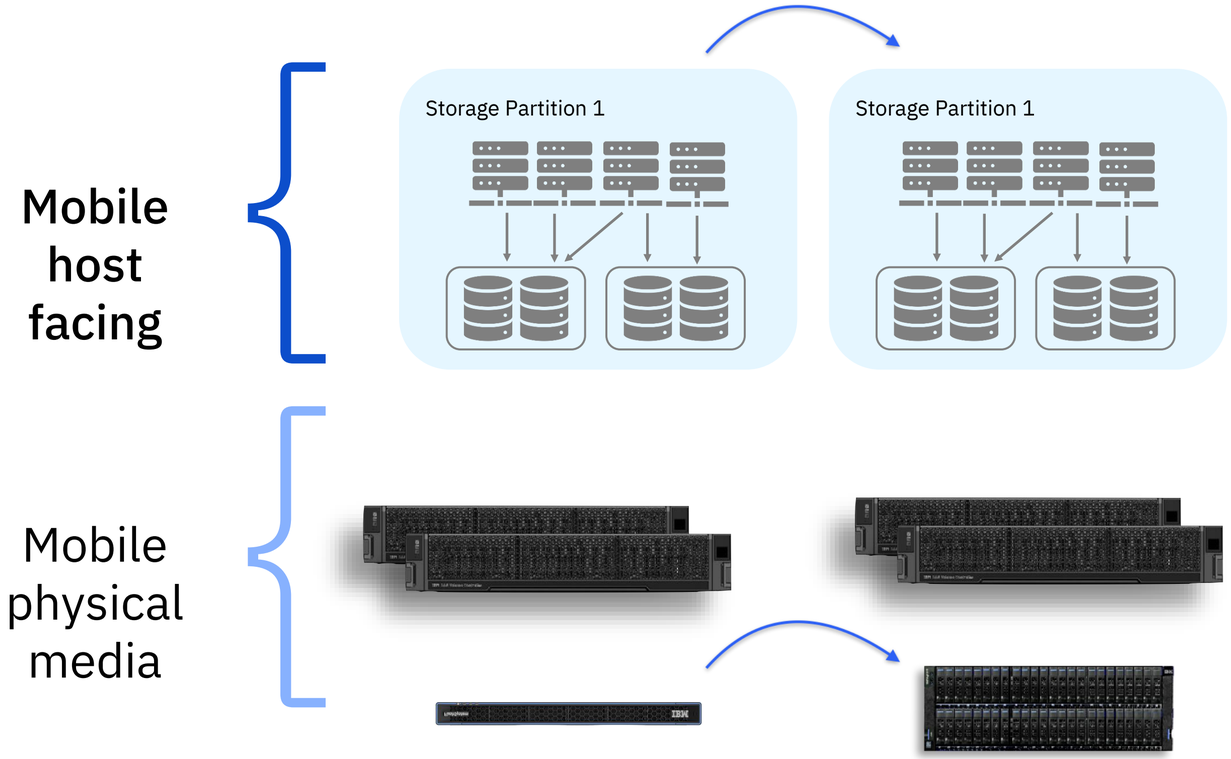
Why grid is better than clustering?

Partition Migration

Until now, externally virtualised storage **only** allowed for **physical media migration within a single system** - the host facing logical configuration remains unchanged.



Storage partitions can also migrate both the host-facing logical configuration and the backend data on physical media, **non-disruptively between systems!**



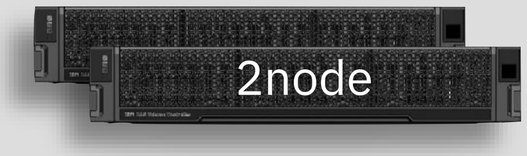
Why grid is better than clustering?

Scaling, Reliability & Simpler Code Development

As the cluster itself knows about all **logical objects**, some of these were **cluster limits**.
For example volumes and snapshots.

Fixed cluster limits

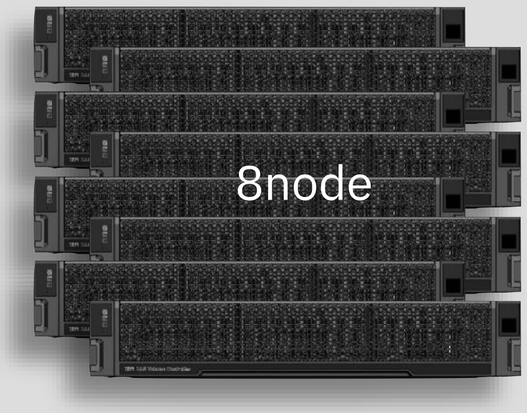
16K volumes
32K snapshots



Same limit:

16K volumes

32K snapshots



N-way code considerations

Where software needed to cater for more than 2node's, n-way considerations were complex and required additional bespoke software handling.

Shorter duration, and independent code upgrades

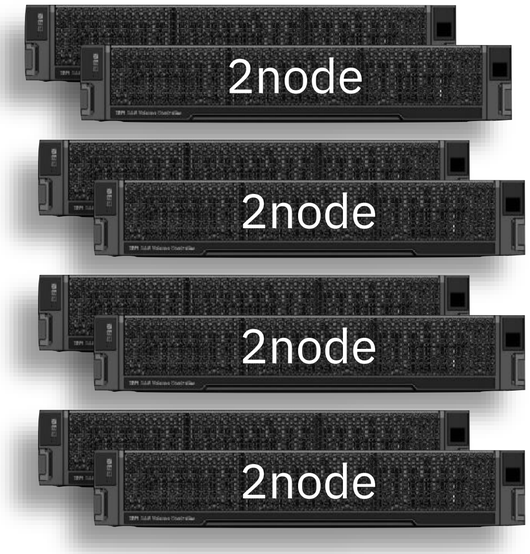
Individual 2node systems can be running different code levels and can be upgraded independently - less risk and less scope for problems

Scalable grid limits

Each 2node has up to

16K volumes

32K snapshots



**Total:*

64K volumes

128K snapshots

* This comparison shows like for like 8 nodes - however an SVC grid can scale to twice these numbers with 16 nodes.

SVC grid - The fine print

As with all grid members, single I/O group systems only (one SVC node pair)

Migrating an existing >2 node system

- 2026 roadmap item to enable partition migration from the I/O group with the lowest ID in the cluster
  - *Cycle through all volume groups in that I/O group to “empty” it*
  - *Remove that I/O group (re-deploy nodes as a new grid member)*
  - *Rinse repeat*

Deploying a new SVC system

- Consider deploying as multiple single I/O group systems from the start

Pool and Mdisk considerations

- All nodes in an SVC cluster today can see all mdisks - so pools are global
- With a grid, each node pair sees unique mdisks - not a major issue
- You can share an external controller across multiple SVC systems (different mdisks!)

Building blocks of a grid

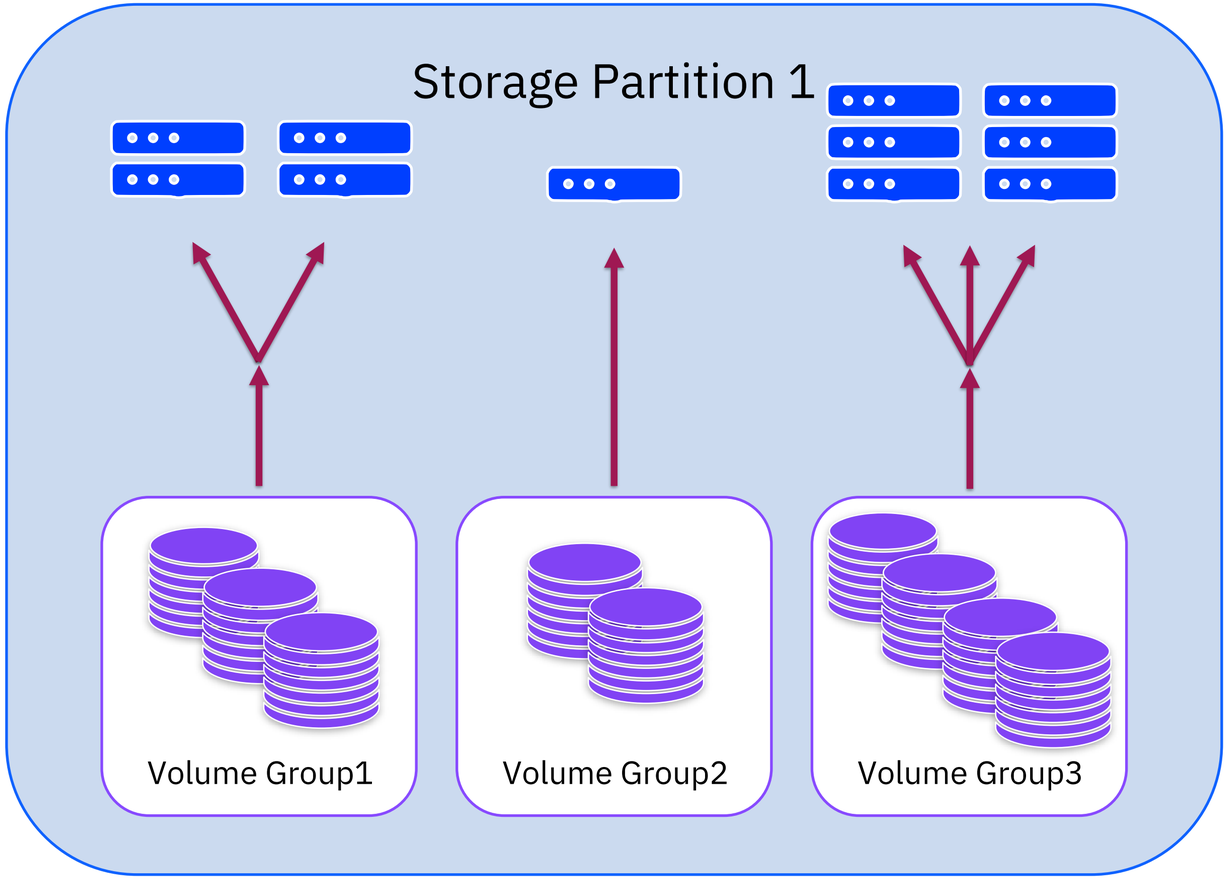
Introducing the Storage Partition

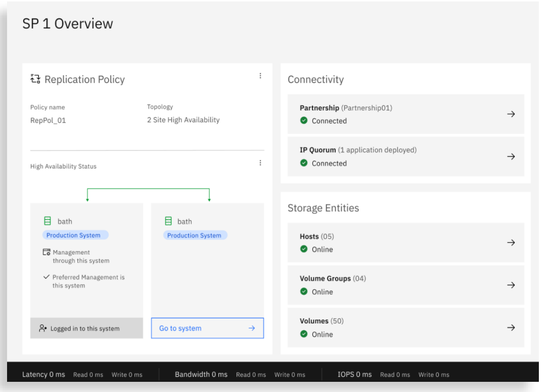
Purpose

We needed **an object** that could **encapsulate** not only an application's **data** but also all of the **logical configuration** required to make the data available to the sever.

i.e. the volumes (in volume groups), the logical host objects and the mappings that exist between them.

By creating a higher level object, ***the Storage Partition***, we can simplify the movement and replication of not only the data, but everything the server needs to access the data.

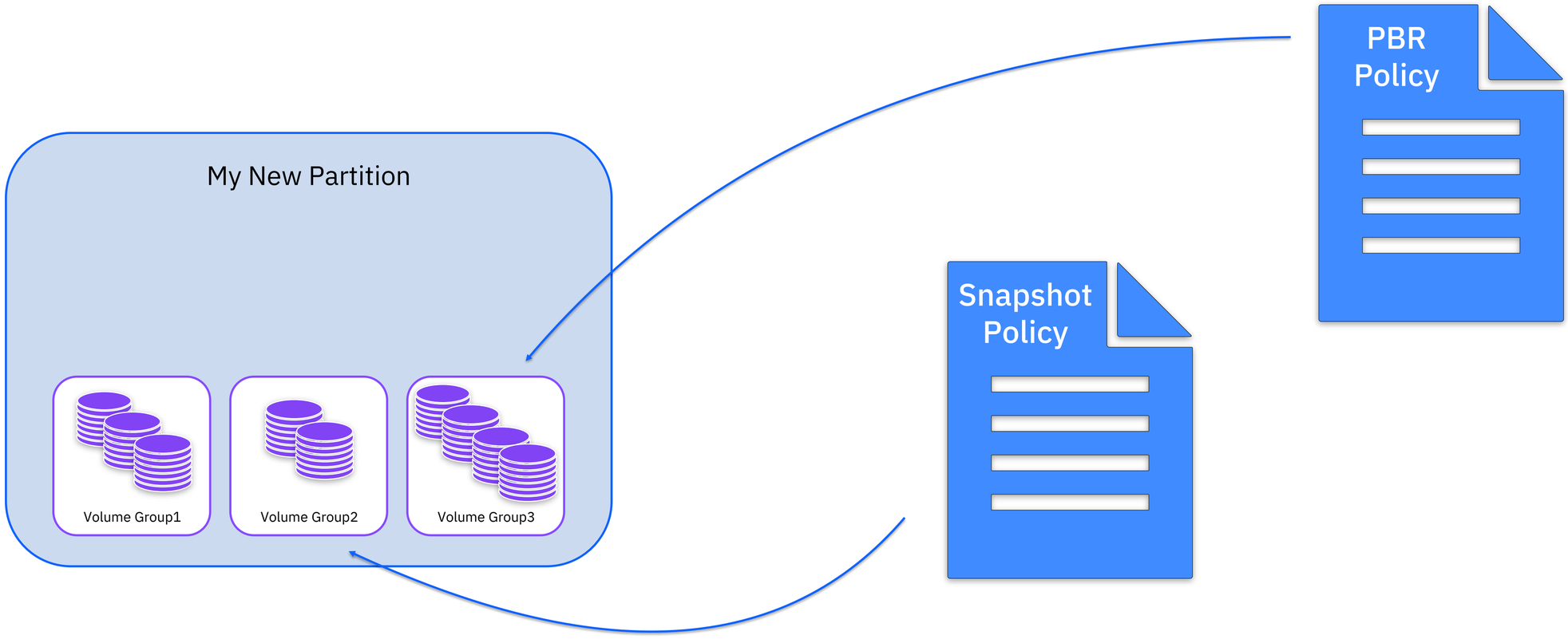




Storage Partition

Creation

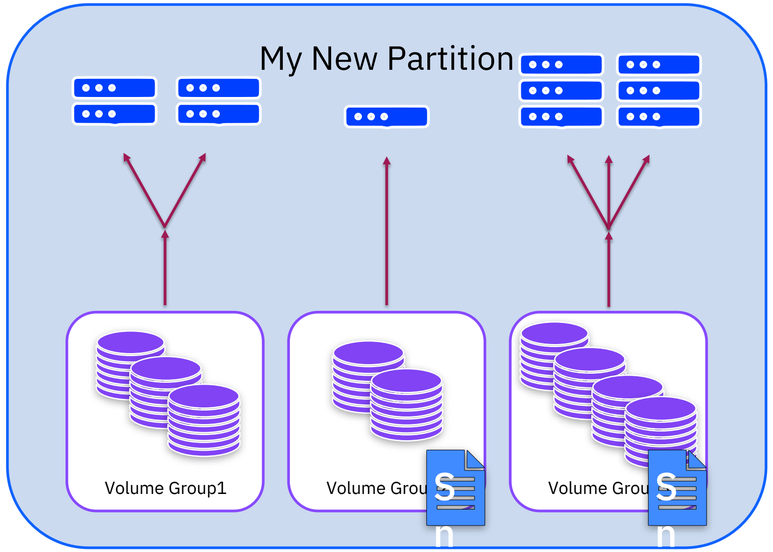
1. Create a new Storage Partition and add some Volume Groups



Storage Partition

Creation

2. Hosts that are mapped to those volumes are automatically added to the new “draft” Partition



Storage Partition

Ideology

*9.1.0 in mid 2025 will increase this to 16.

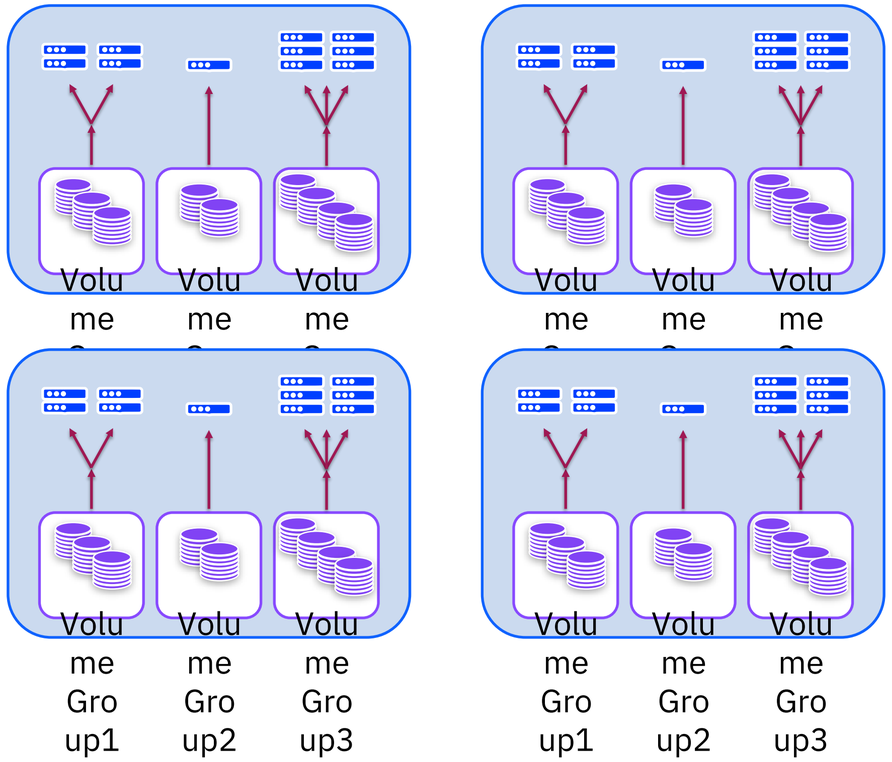
Today, a **single system** can contain at most **4 partitions***.

Storage Partitions have two primary use cases :

- To provide High Availability using PBHA
- To migrate data between systems

Why only 4 partitions today?

1. HA partition with preference to this system.
2. HA partition with preference to the remote system.
3. Temporary partition for migration purposes.
4. *Unused / spare*

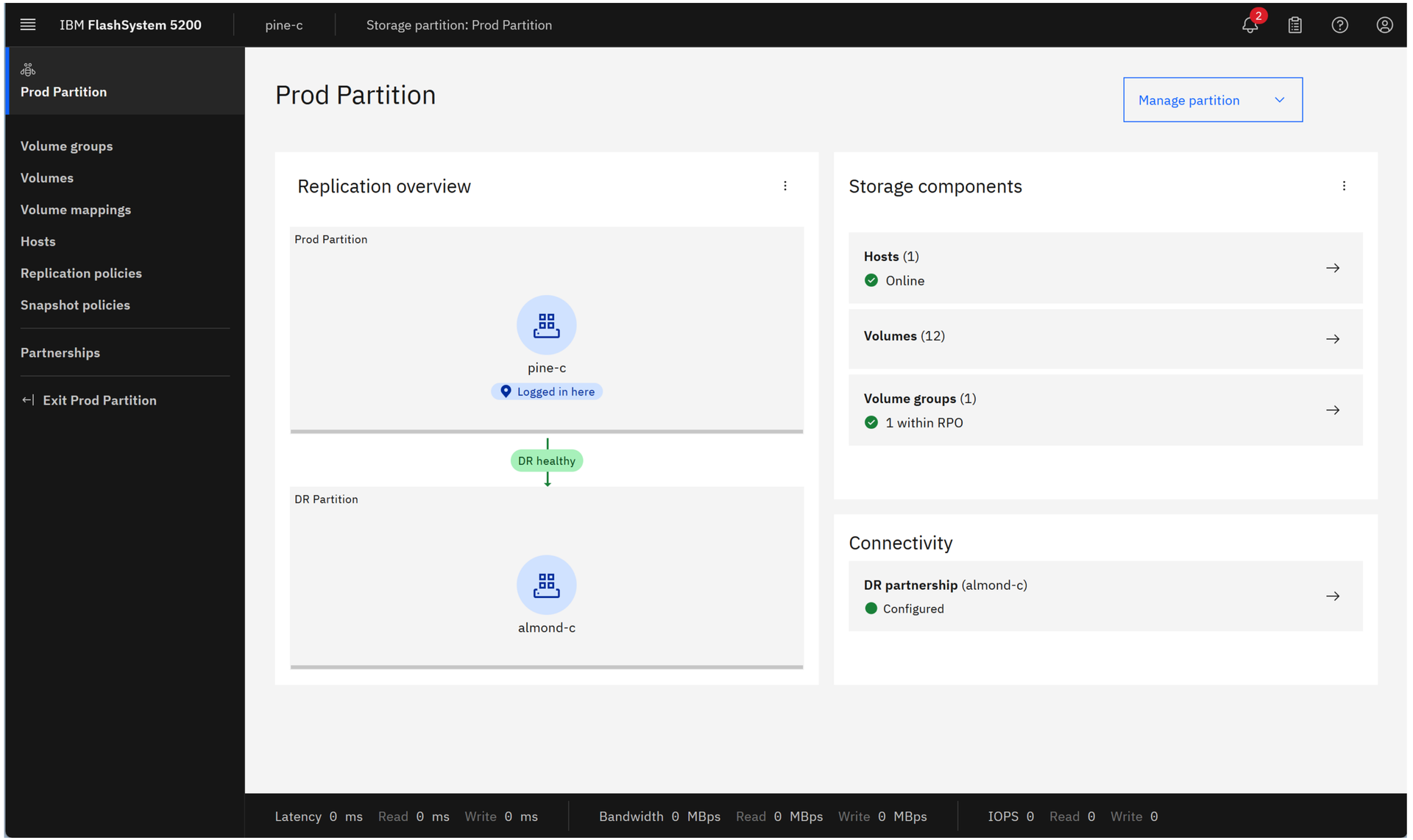


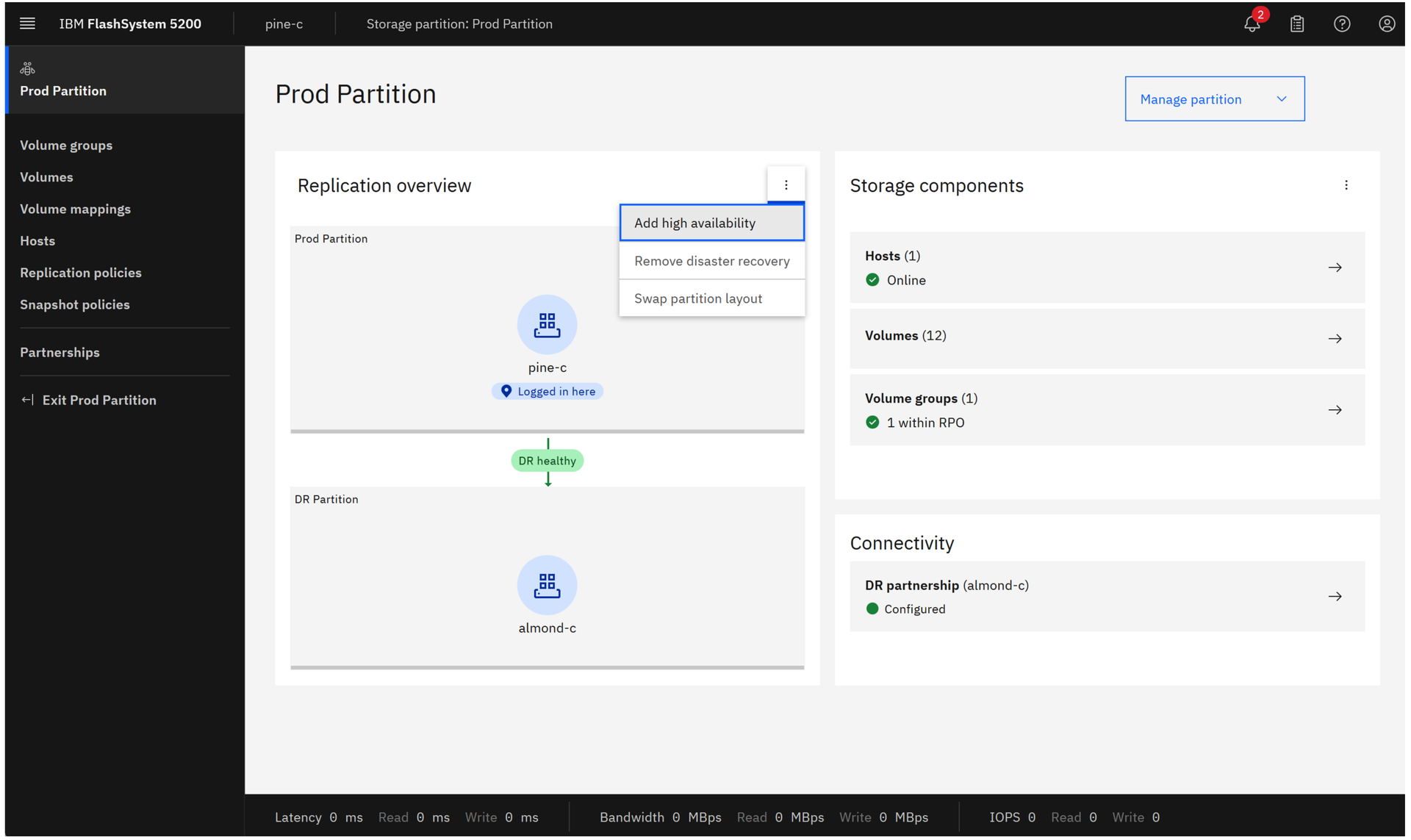
Volume Groups :

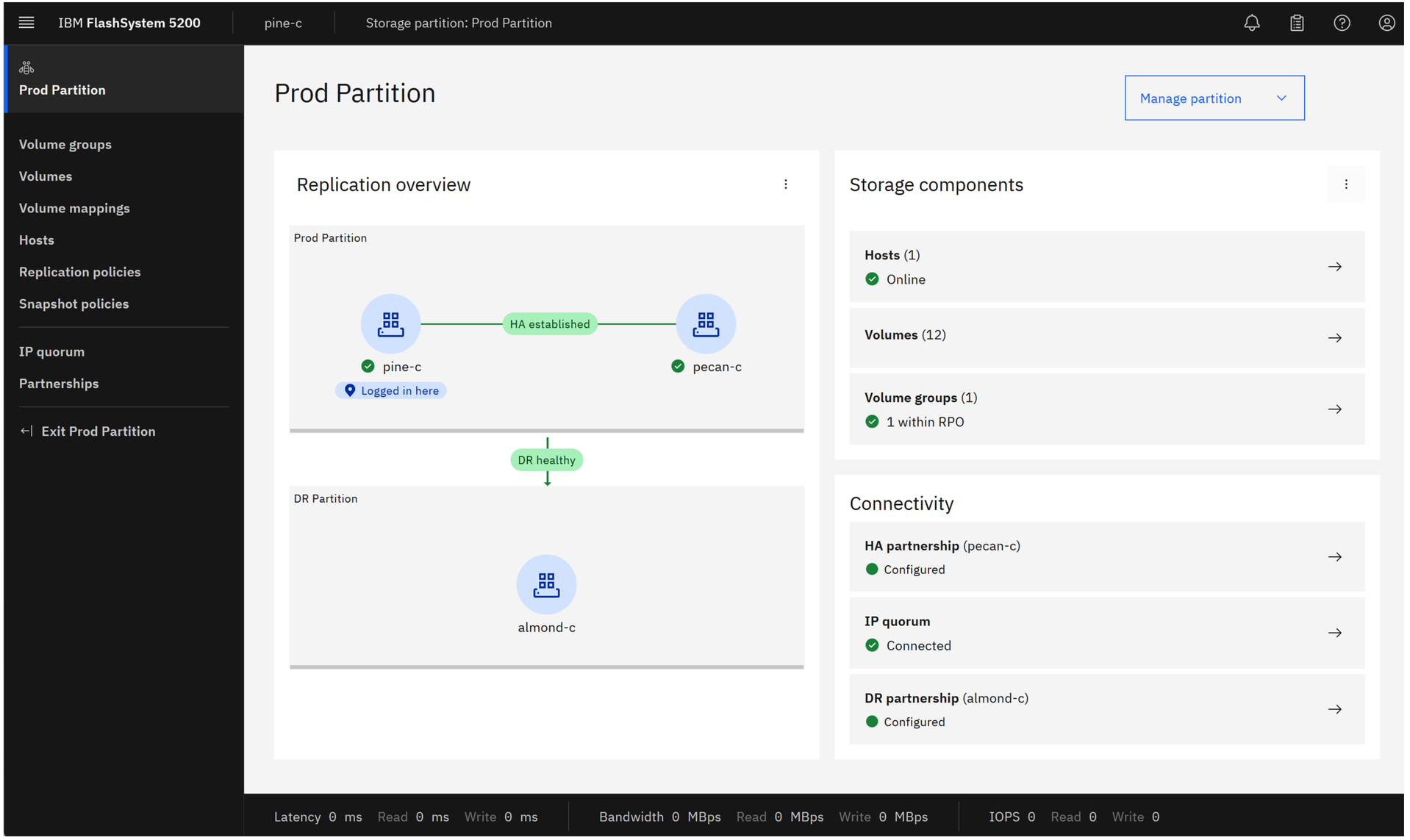
The unit of application cross volume consistency, a system will therefore have many. There could be a need for tens or hundreds of Volume Groups.

Storage Partitions :

Contains all the objects active on this system that require HA or need to be migrated, a system may need to only have only a few. There is no need for tens or hundreds of partitions.


IBM FlashSystem 5200
pine-c
Storage partition: Prod Partition
2
Prod Partition
Volume groups
Volumes
Volume mappings
Hosts
Replication policies
Snapshot policies
Partnerships
Exit Prod Partition
Prod Partition
Manage partition
Replication overview
Prod Partition
pine-c
Logged in here
DR healthy
DR Partition
almond-c
Storage components
Hosts (1)
Online
Volumes (12)
Volume groups (1)
1 within RPO
Connectivity
DR partnership (almond-c)
Configured
Latency 0 ms Read 0 ms Write 0 ms
Bandwidth 0 MBps Read 0 MBps Write 0 MBps
IOPS 0 Read 0 Write 0


IBM FlashSystem 5200
pine-c
Storage partition: Prod Partition
2
Prod Partition
Volume groups
Volumes
Volume mappings
Hosts
Replication policies
Snapshot policies
Partnerships
Exit Prod Partition
Prod Partition
Manage partition
Replication overview
Add high availability
Remove disaster recovery
Swap partition layout
Prod Partition
pine-c
Logged in here
DR healthy
DR Partition
almond-c
Storage components
Hosts (1)
Online
Volumes (12)
Volume groups (1)
1 within RPO
Connectivity
DR partnership (almond-c)
Configured
Latency 0 ms Read 0 ms Write 0 ms
Bandwidth 0 MBps Read 0 MBps Write 0 MBps
IOPS 0 Read 0 Write 0


IBM FlashSystem 5200
pine-c
Storage partition: Prod Partition
Prod Partition
Volume groups
Volumes
Volume mappings
Hosts
Replication policies
Snapshot policies
IP quorum
Partnerships
Exit Prod Partition
Prod Partition
Manage partition
Replication overview
Prod Partition
HA established
pine-c
Logged in here
pecan-c
DR healthy
DR Partition
almond-c
Storage components
Hosts (1)
Online
Volumes (12)
Volume groups (1)
1 within RPO
Connectivity
HA partnership (pecan-c)
Configured
IP quorum
Connected
DR partnership (almond-c)
Configured
Latency 0 ms Read 0 ms Write 0 ms
Bandwidth 0 MBps Read 0 MBps Write 0 MBps
IOPS 0 Read 0 Write 0

Building blocks of a grid

Analytics layer

Grid management layer

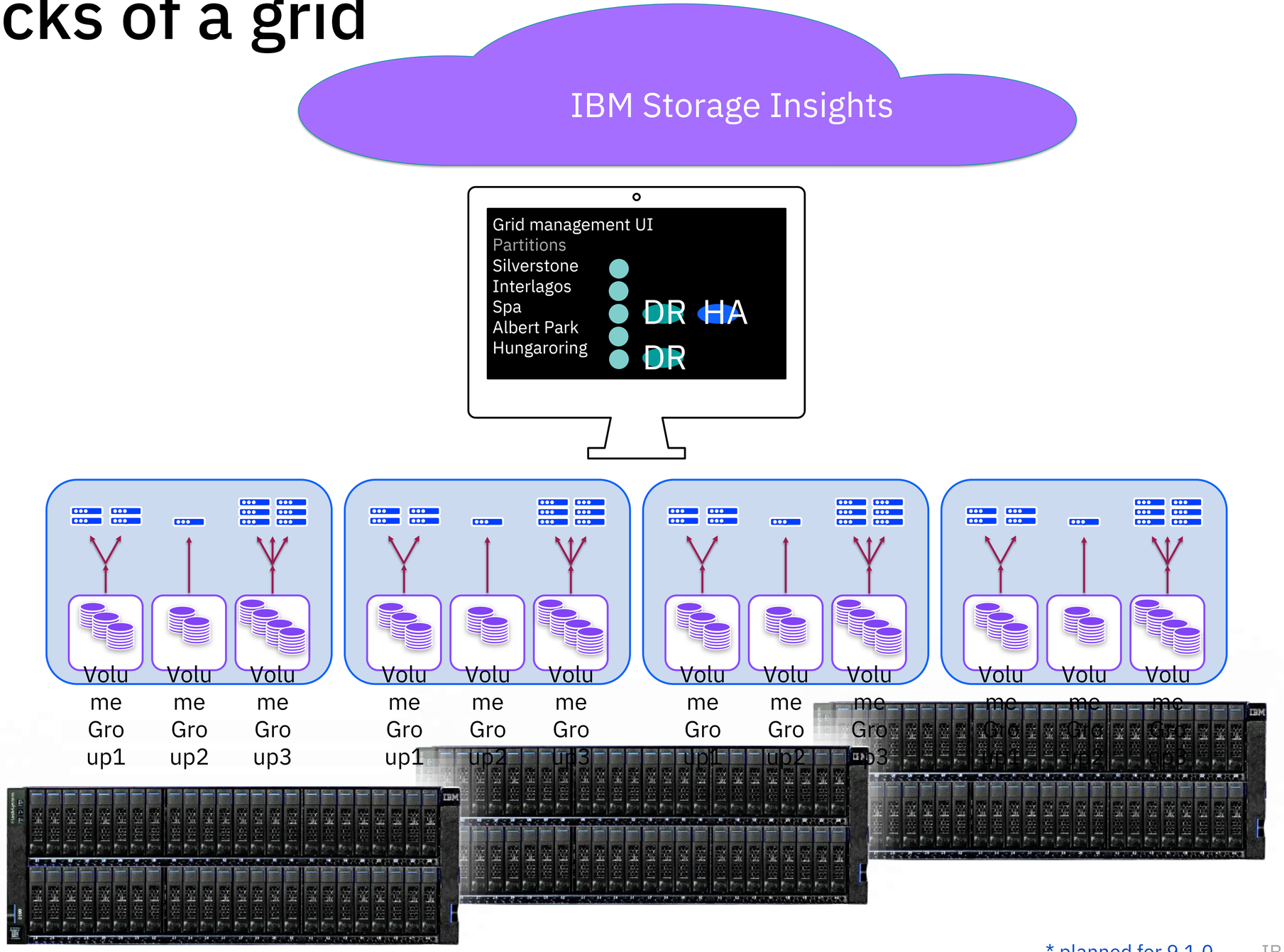
Onboard management from any system

Partitions layer

Up to 16* per system

Infrastructure layer

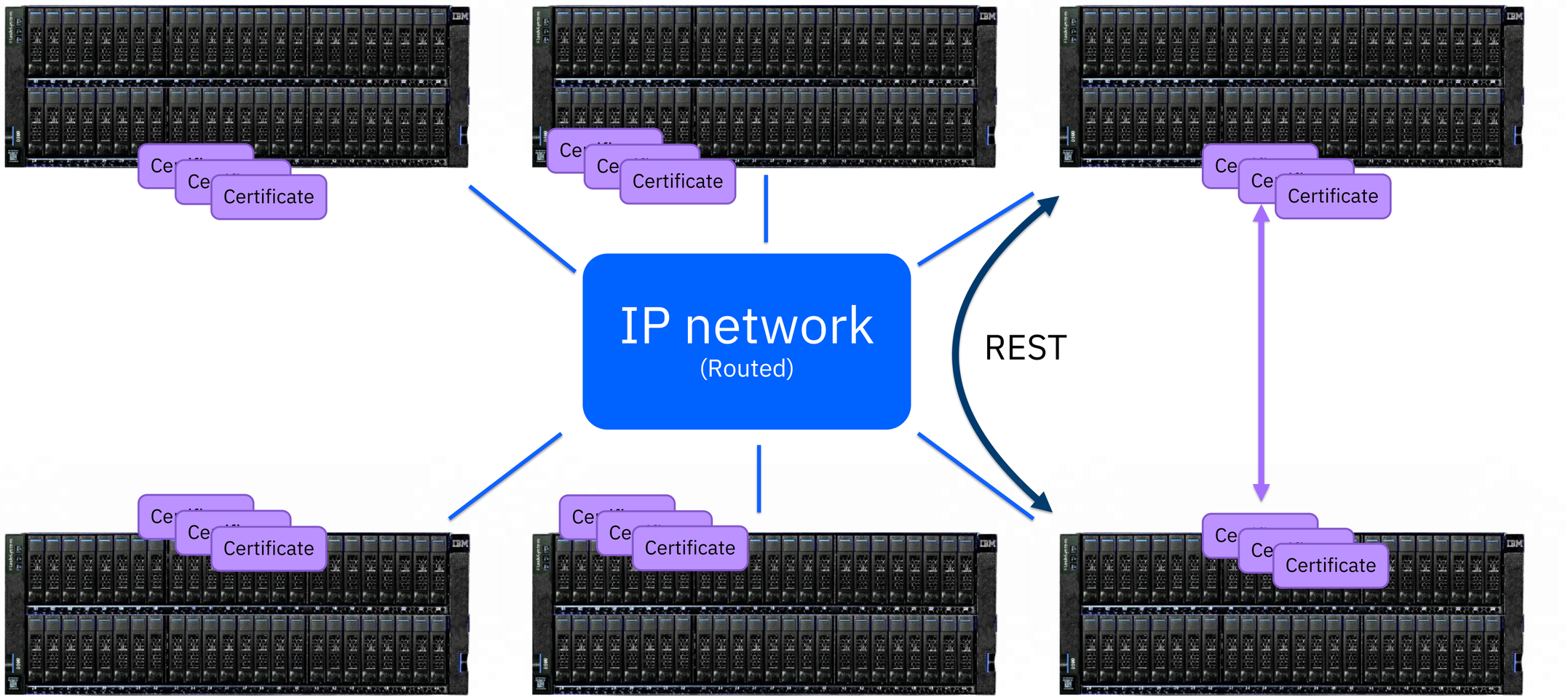
1-32* systems



* planned for 9.1.0

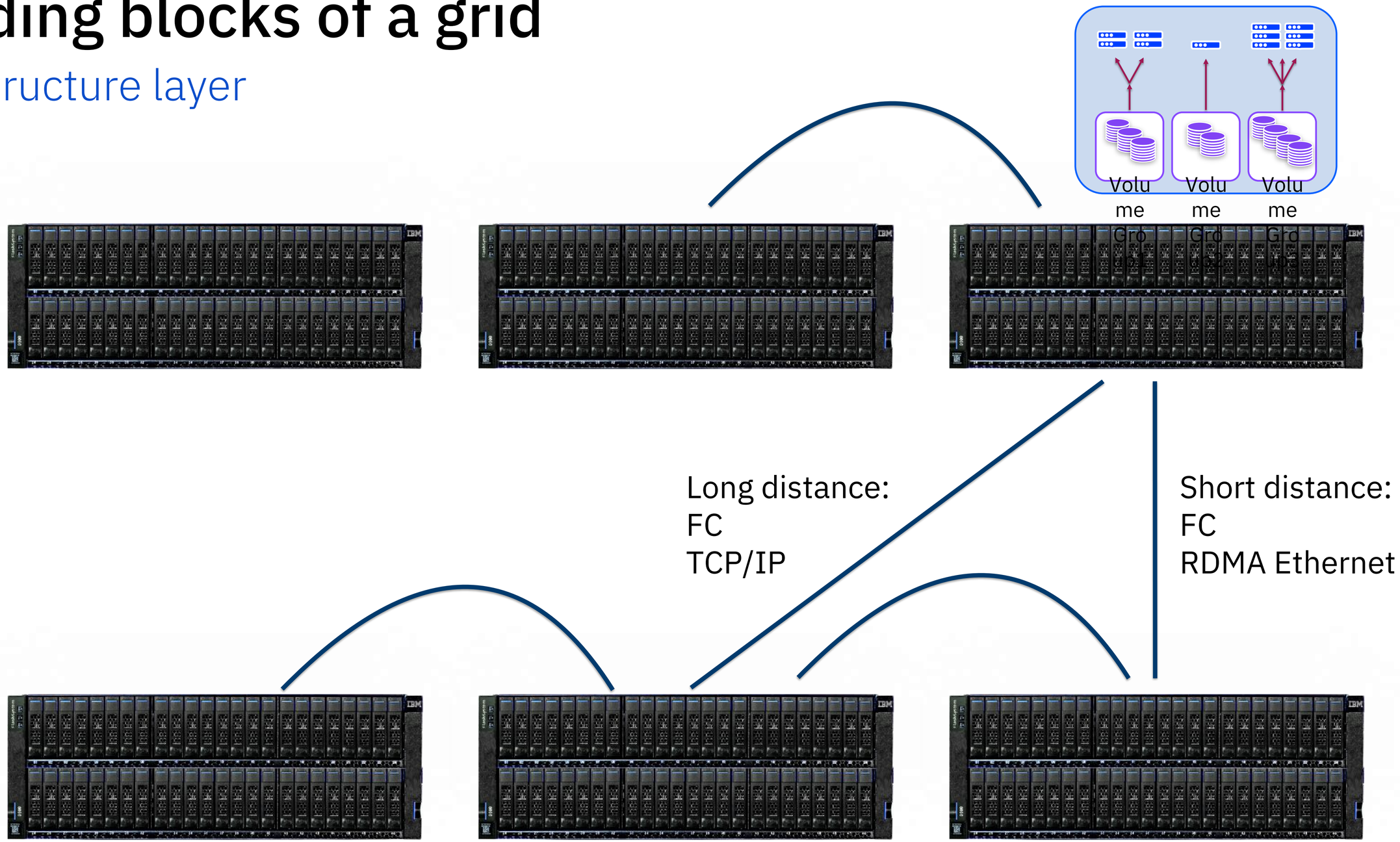
Building blocks of a grid

Infrastructure layer



Building blocks of a grid

Infrastructure layer



Management

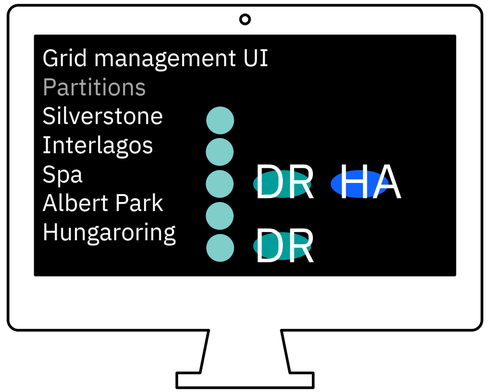
- Always-required
- Management traffic only
- Secured using exchanged certificates
- REST traffic between systems
- Outbound from the first system IP
- Inbound to the first system IP

Data

- Required to support data traffic between systems
- Conceptualised into a *partnership* concept
- Each system supports 3 partnerships, with at most 1 HA/migration partnership
- Created only when needed for replication, HA or migration
- Partnerships define how data should be sent between systems:
  - Networking type (Fibre Channel, TCP/IP or RDMA)
  - Bandwidth limit (asynchronous replication only)
  - Ports to use (Ethernet only - Fibre Channel uses zoning or port masking)
  - Compression/Encryption support (TCP/IP only)

Building blocks of a grid

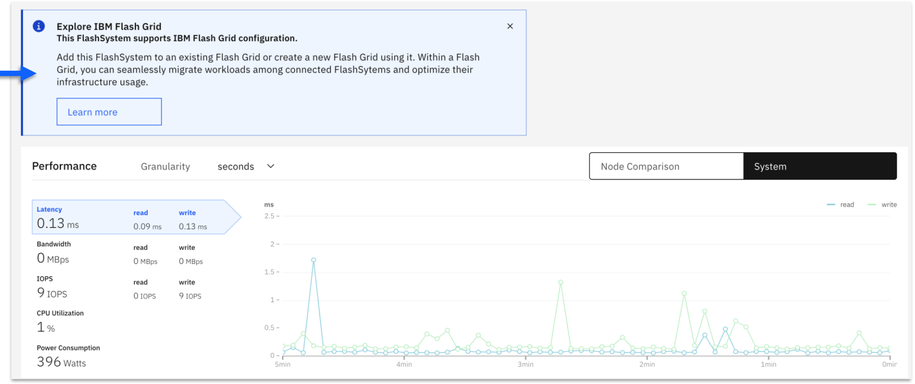
Grid management layer



Grid management

Systems that support the joining a grid will have a new notification

- *Notification can be disabled*



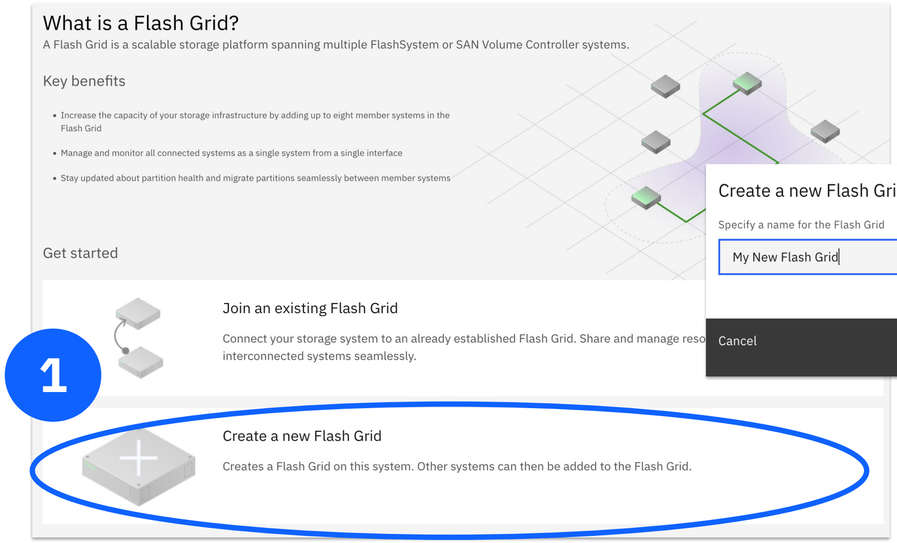
Create new FlashSystem grid

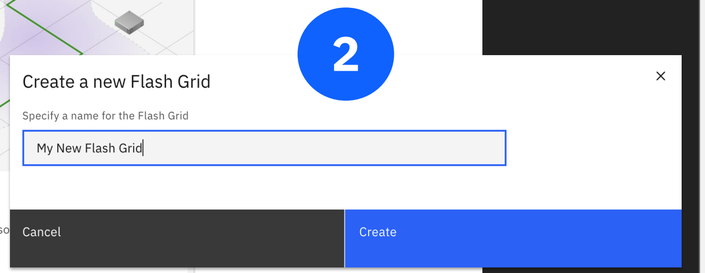
1. Click on Create a new grid
2. Name your grid
3. Successful creation will take you to the Grid dashboard
   - Single system added
   - Grid Coordinator

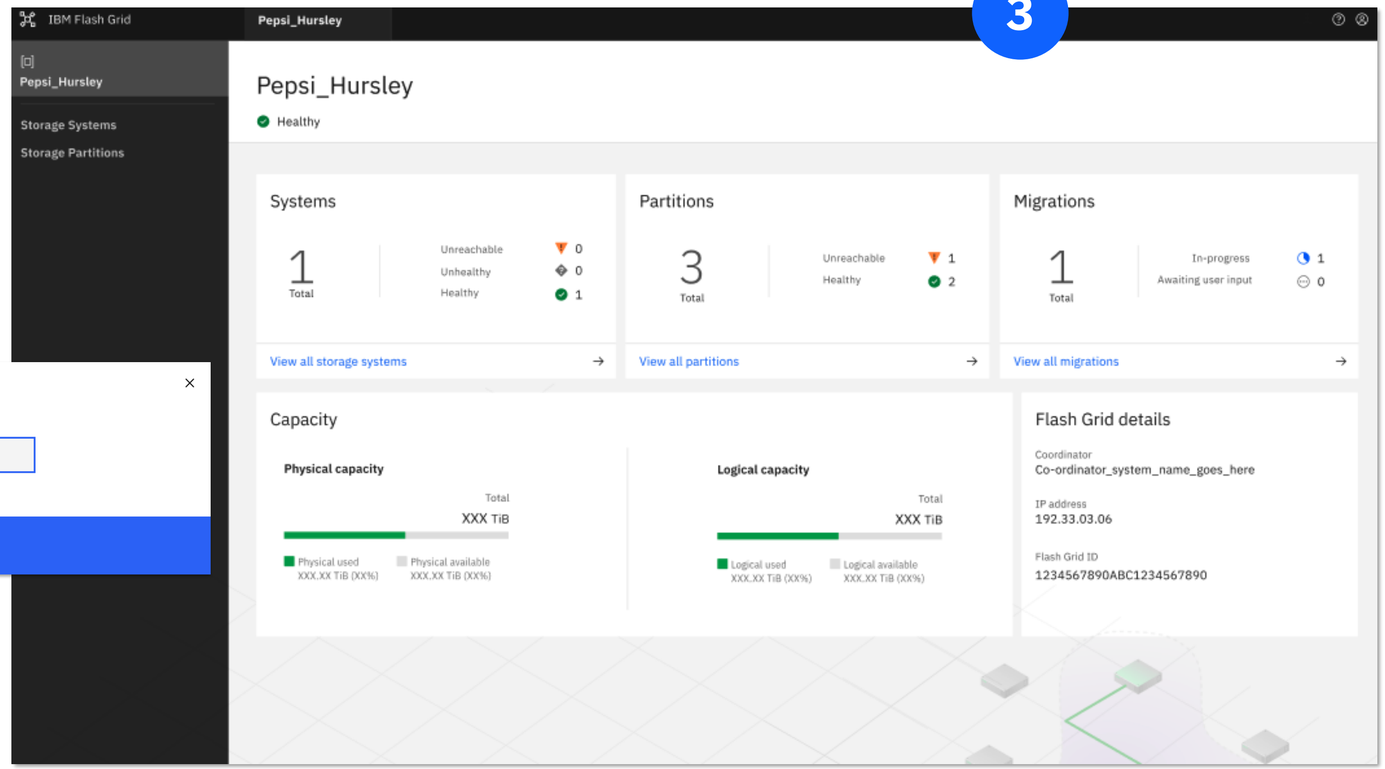
Grid Dashboard shows :

- Number of systems and partitions
- System / Partition Health
  - *Healthy, Unhealthy, Unreachable*
- Aggregated Capacity
  - *Physical capacity only in this release*
- Grid details

Side menus for Storage systems and Storage partitions







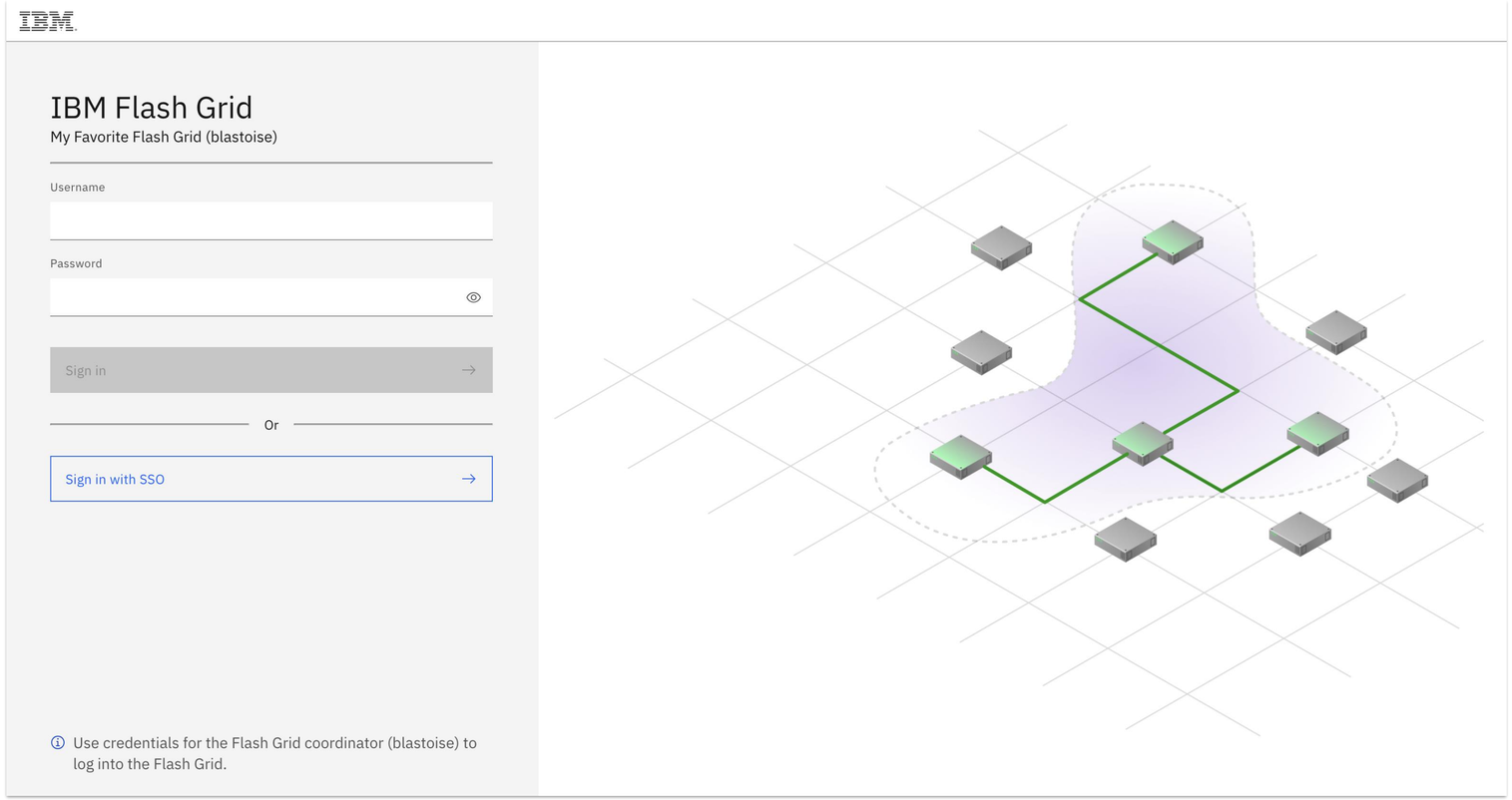
Grid Login Screen

Once configured as a member of a grid the native UI will have a new grid login screen.

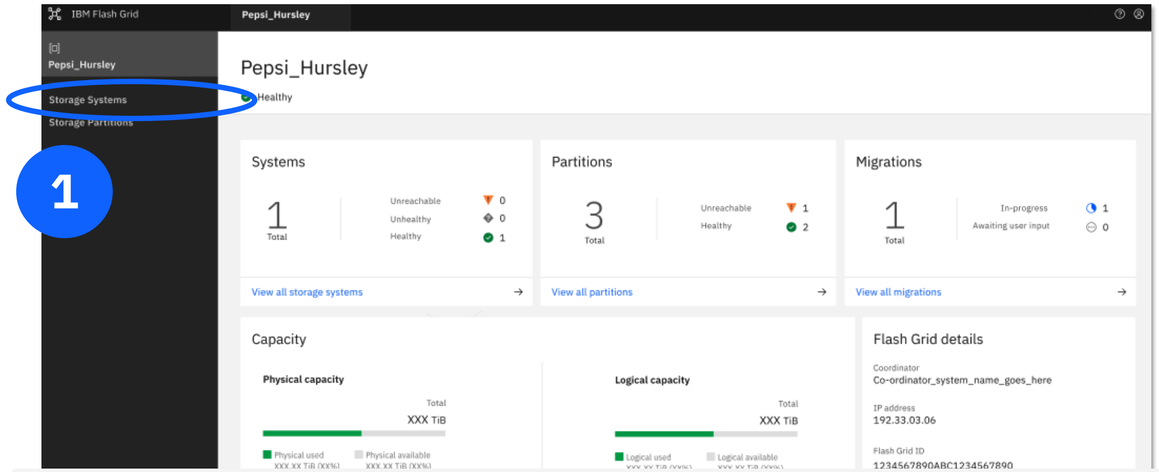
User credentials are those of the system you connect to

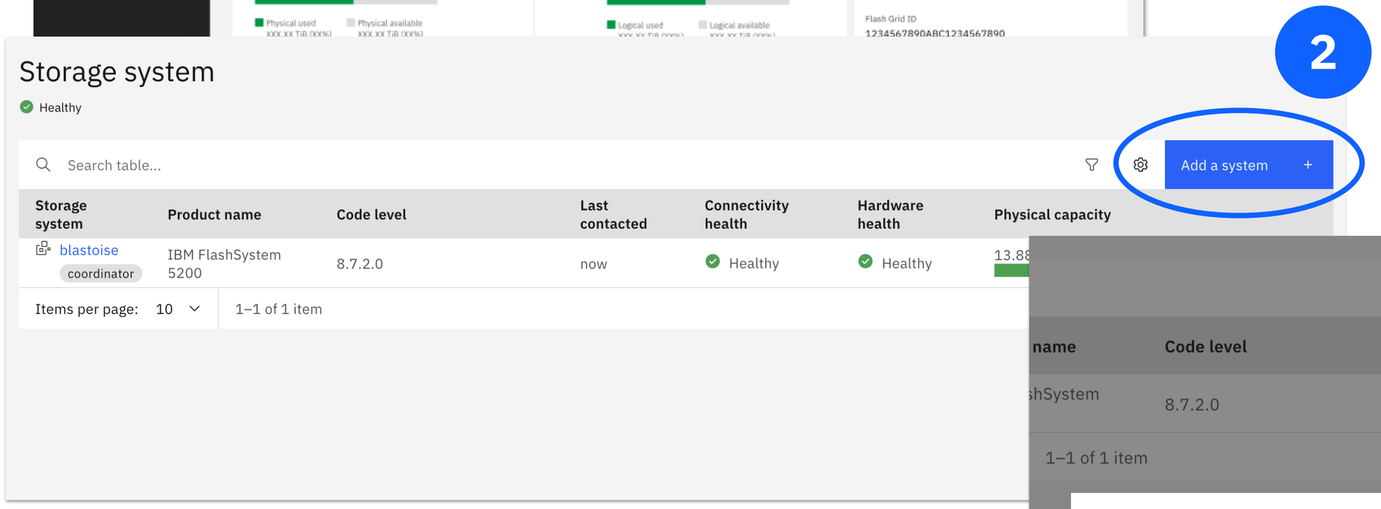
- Also supports LDAP, SSO

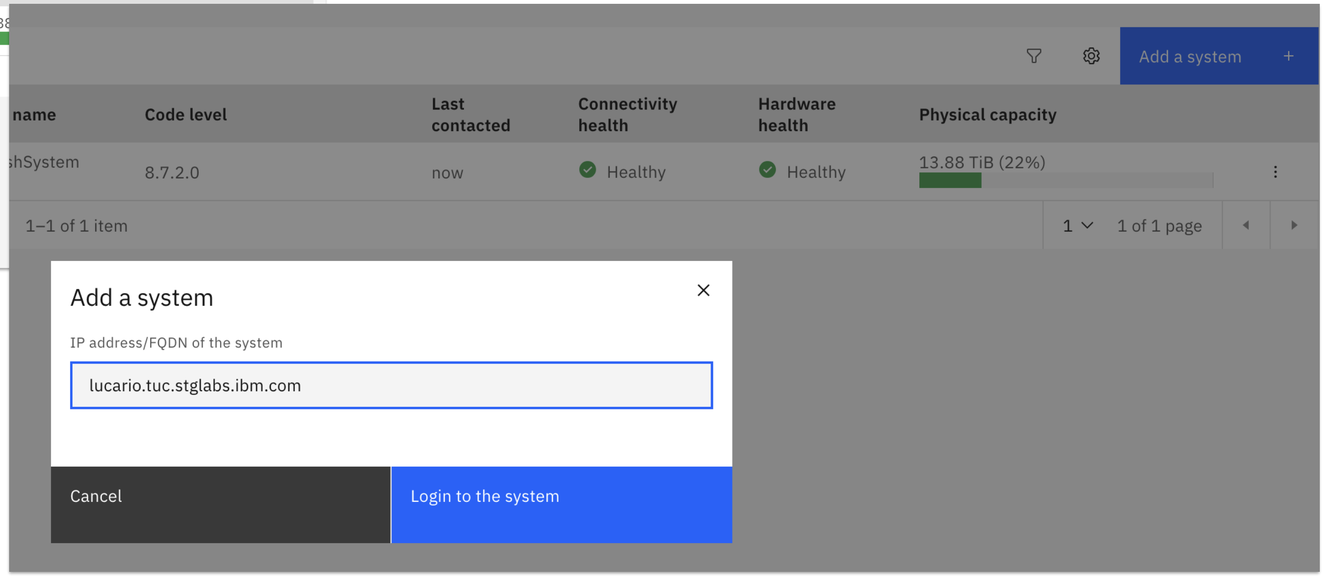
Logging in takes you initially to grid dashboard (not the system)



Add additional members to the grid



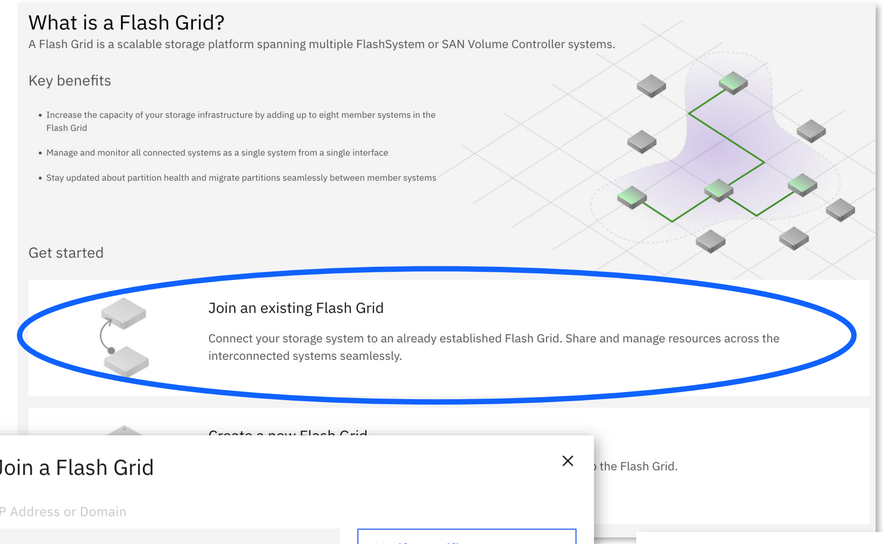




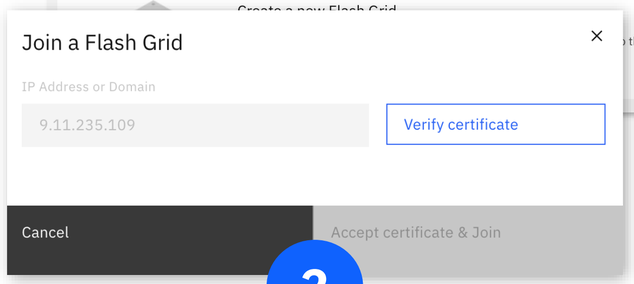
1. Select the storage systems list from the grid Dashboard
2. Click to Add a system
3. Enter the details

or...

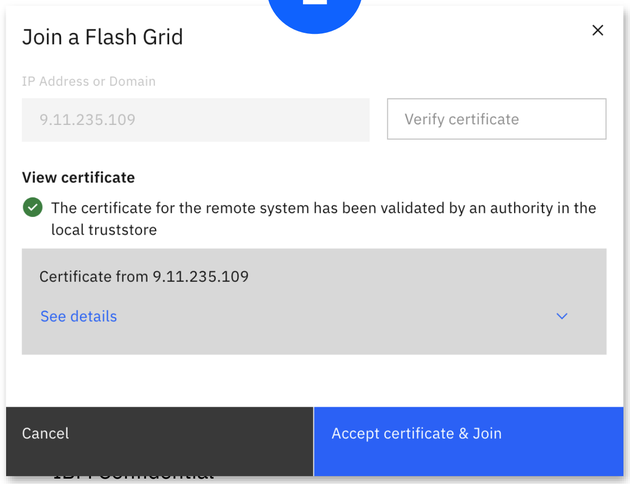
Add additional members to the grid



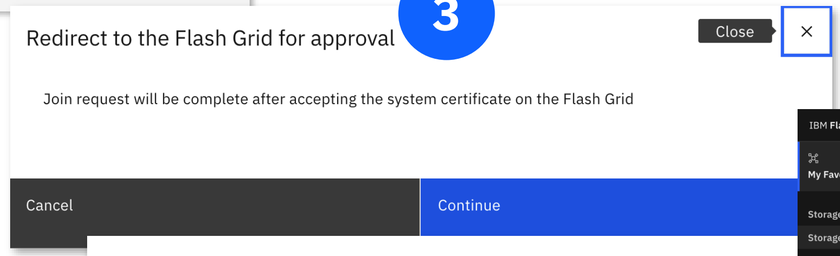
1



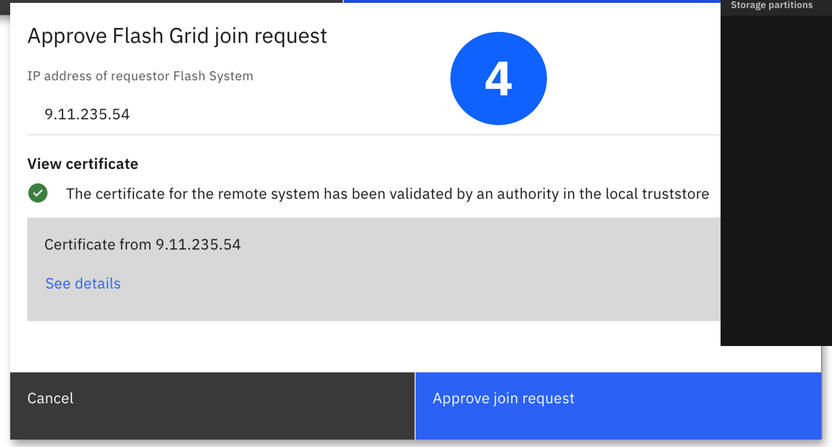
2



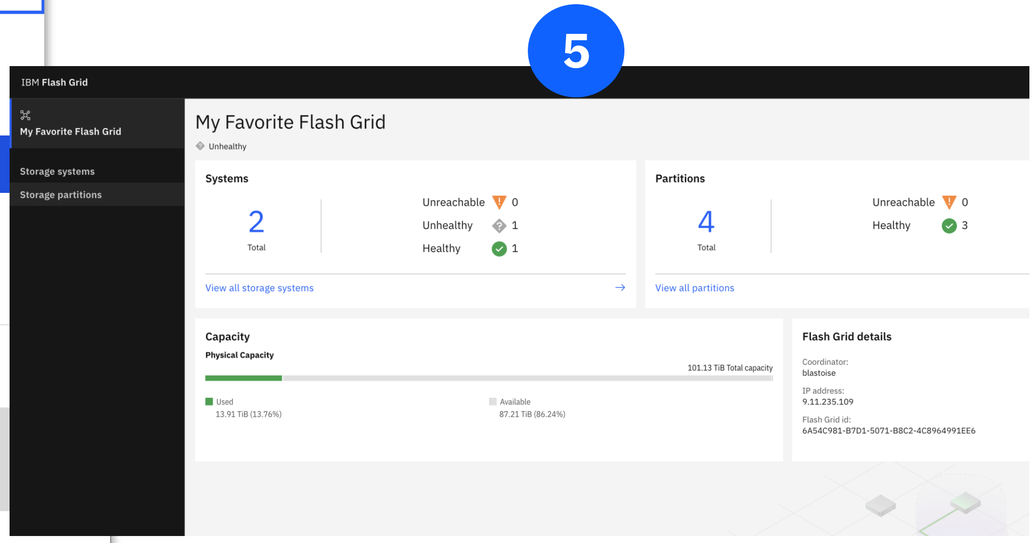
3



4

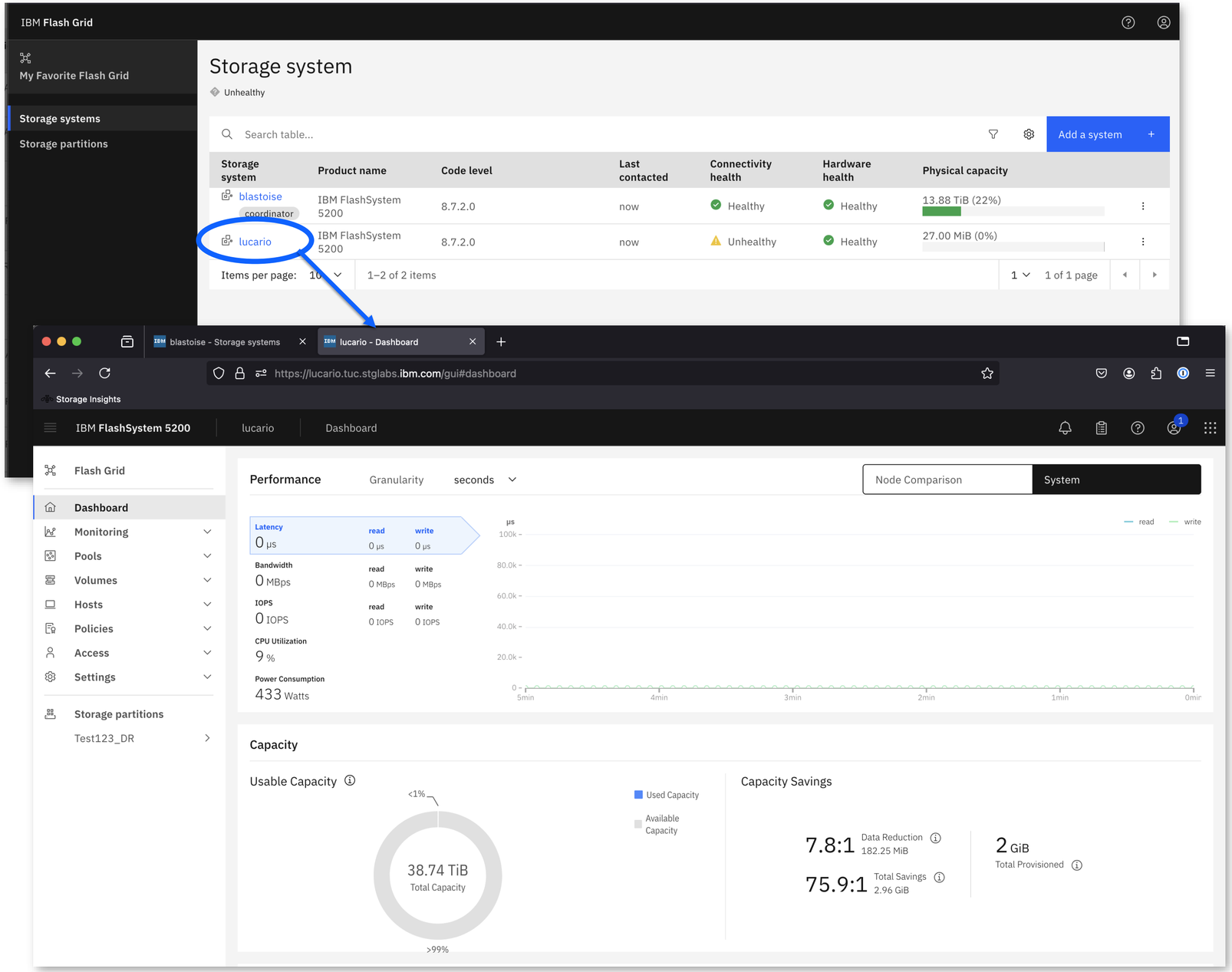


5



1. Select to join a grid from the new members grid configuration panel.
2. Verify and exchange certificates
3. UI automatically navigates to the grid coordinator to complete the join process
4. Approve grid join request and the new member is added
5. Dashboard now shows the additional systems and new aggregate capacity

Grid - Storage Systems View



All storage systems in the grid can be managed here

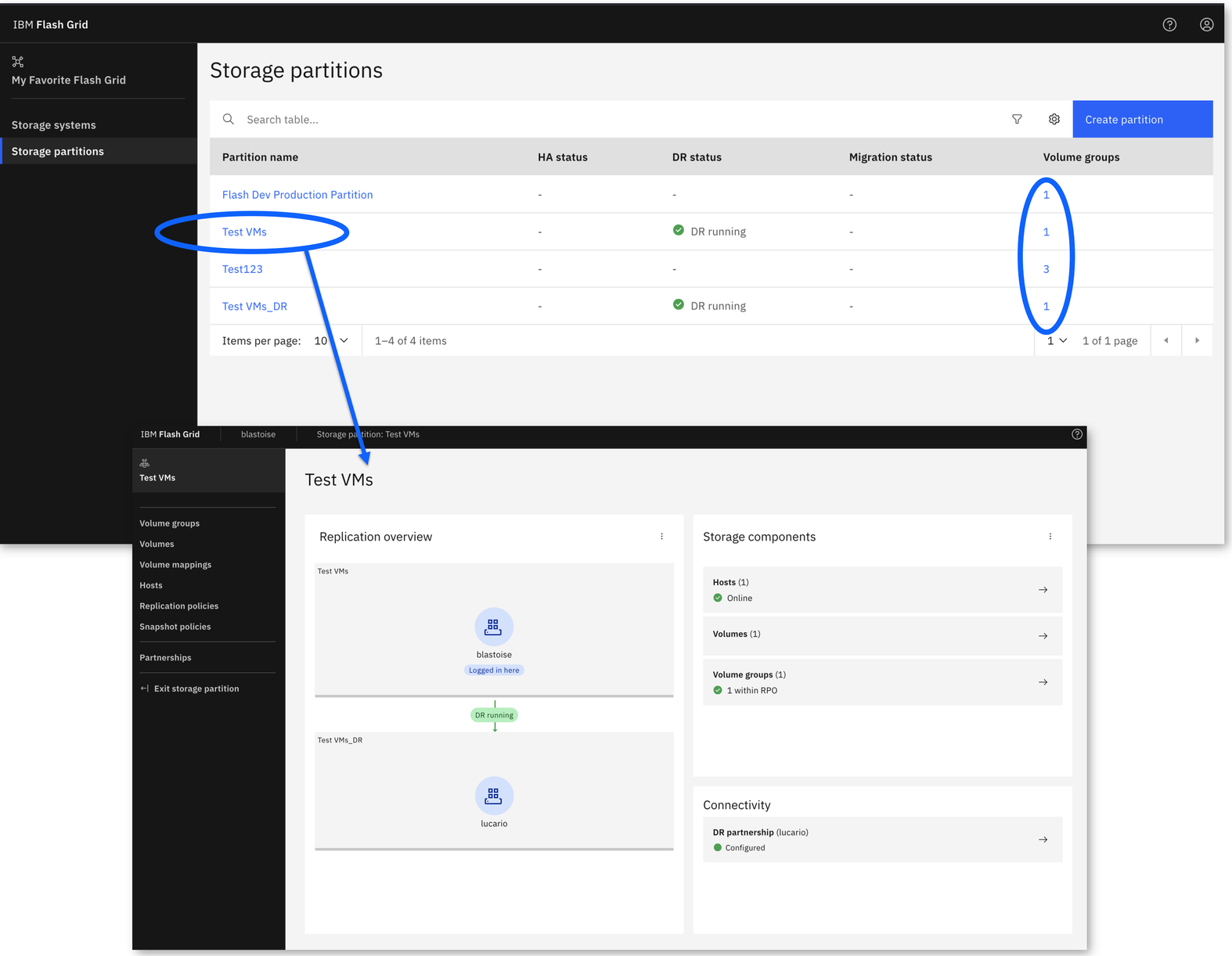
Product name, Code level details, Physical capacity

Last contacted shows last time the system was queried from the grid

Connectivity, Hardware health provides similar attributes to system health boxes

Clicking on storage system name will pop up system dashboard in a separate tab

Grid - Storage Partition View



Aggregated list of all storage partitions in the grid

Provides replication and migration status and number of volume groups

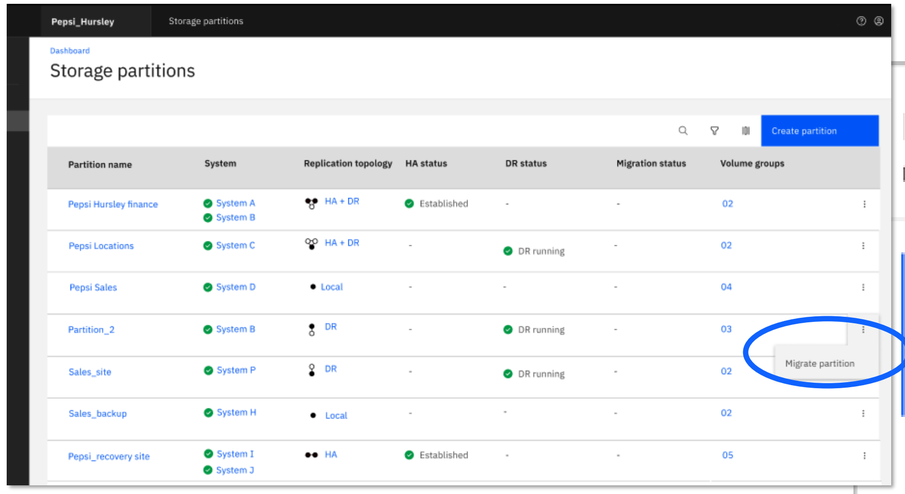
Partition and Volume Group links :

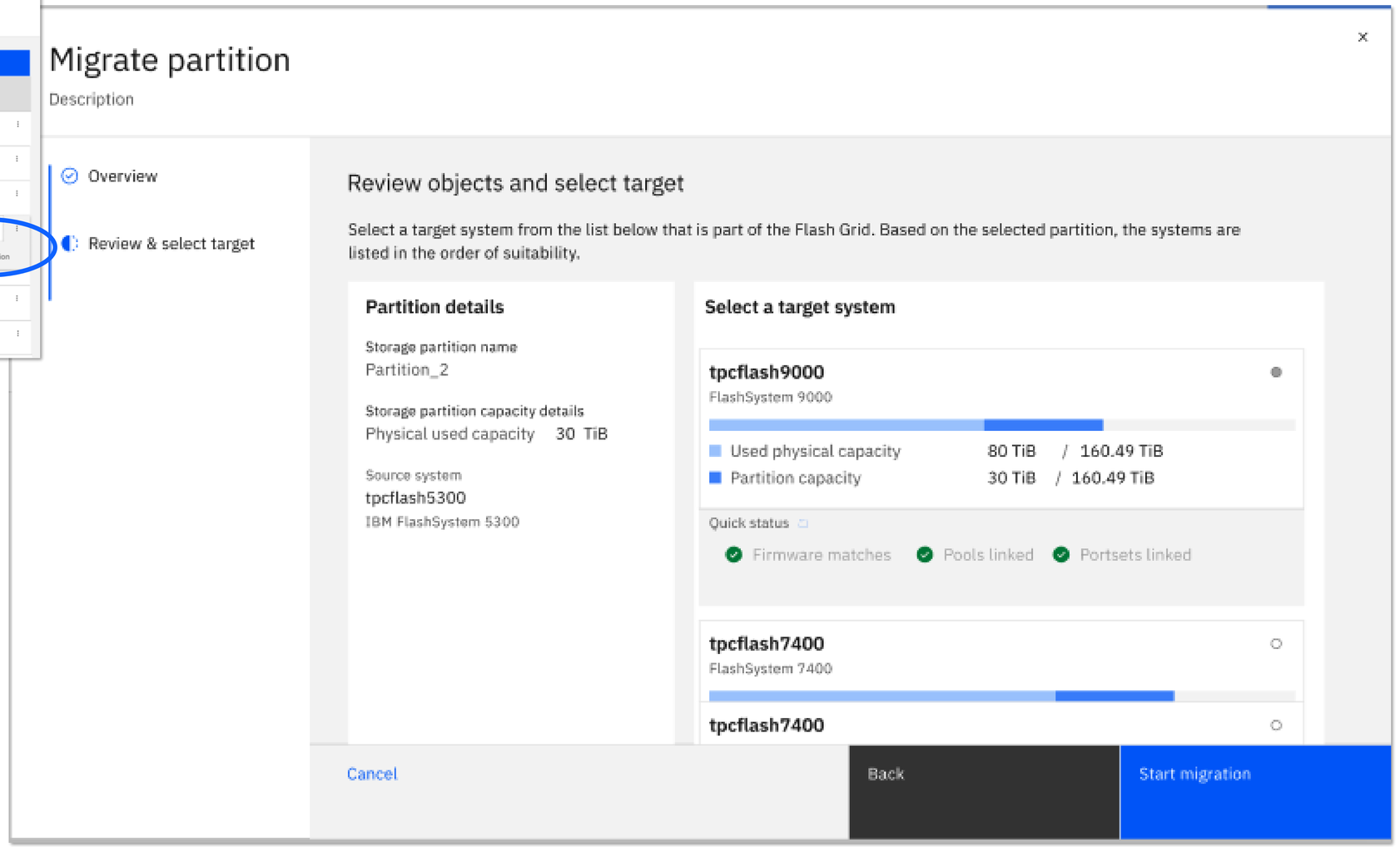
Clicking on links will launch to the partition or volume group

- Launches within the window, no additional tab
- May require login, SSO configured systems will auto log in

Exiting storage partition brings you back to the Storage partitions list

Partition Migration Initiation from GUI





Migration only available in the grid context.

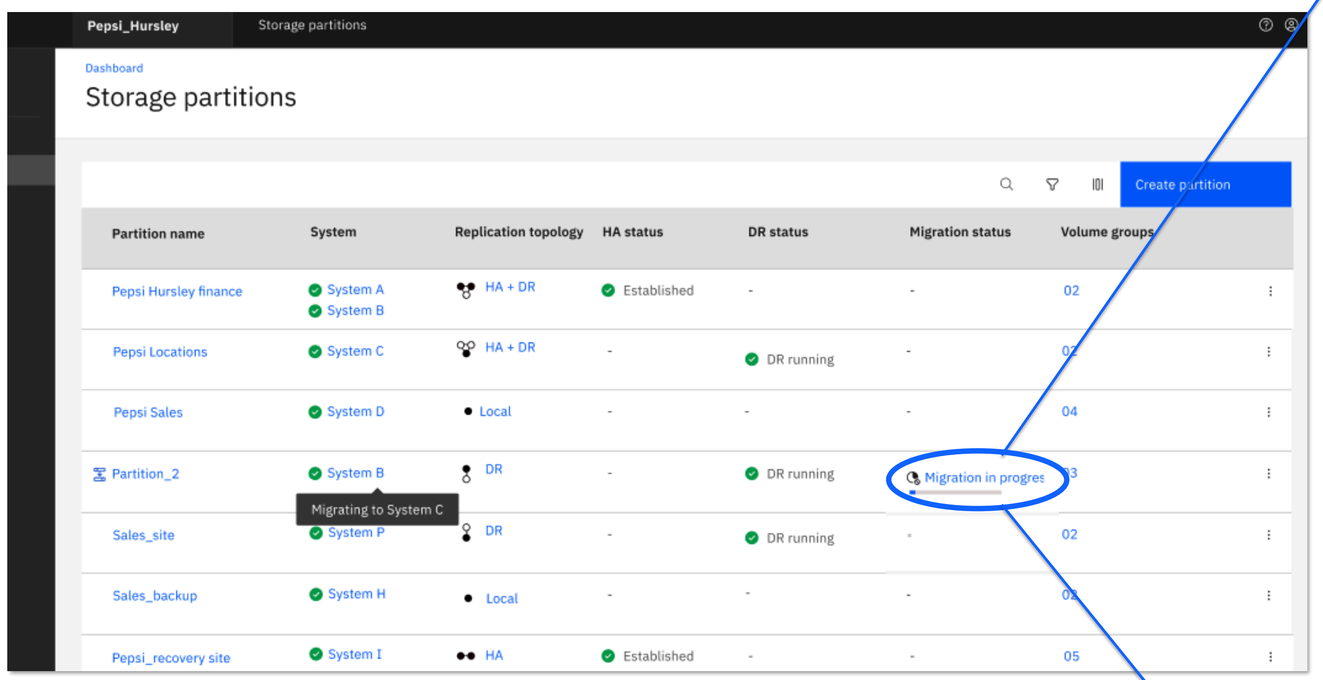
Both source and target must be members of the same grid.

Targets filtered automatically for available capacity.

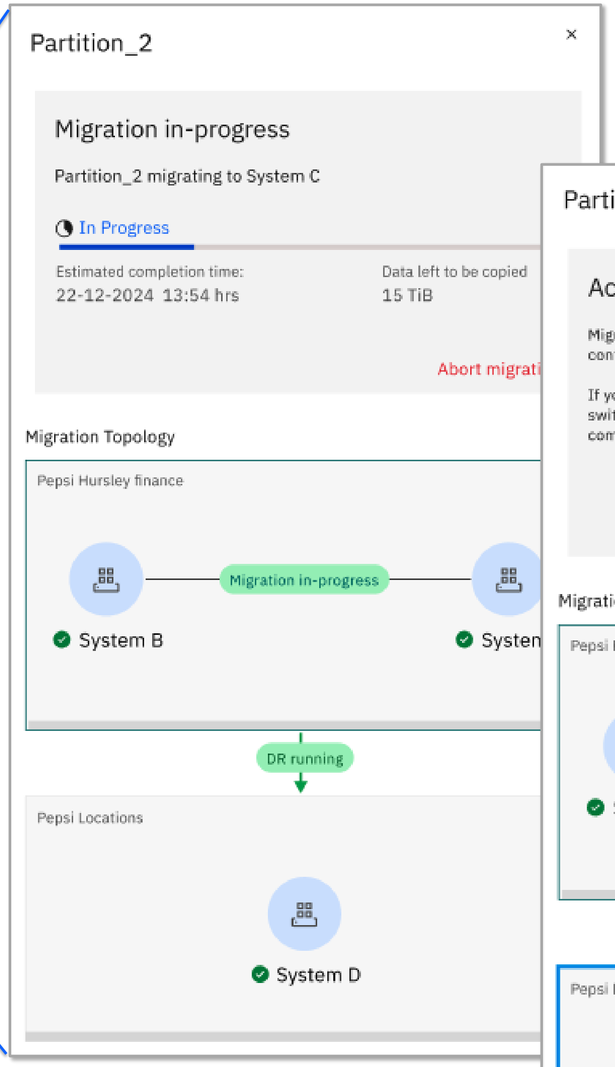
Pre-requisite checks for code level, pool and portset linkage.

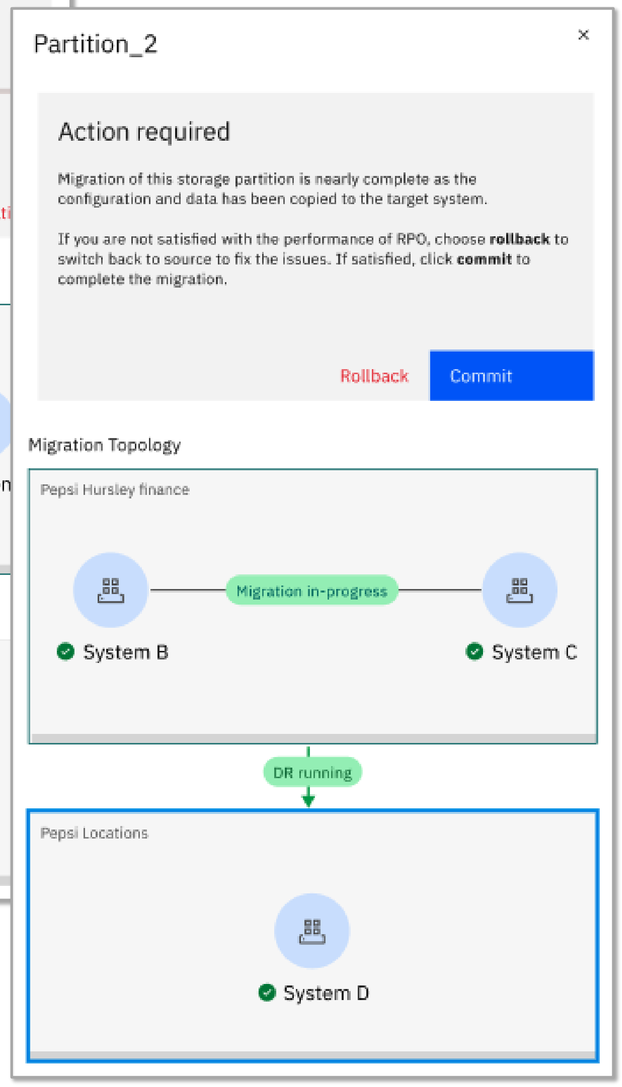
Partition Migration Monitoring from GUI

Grid GUI Monitoring



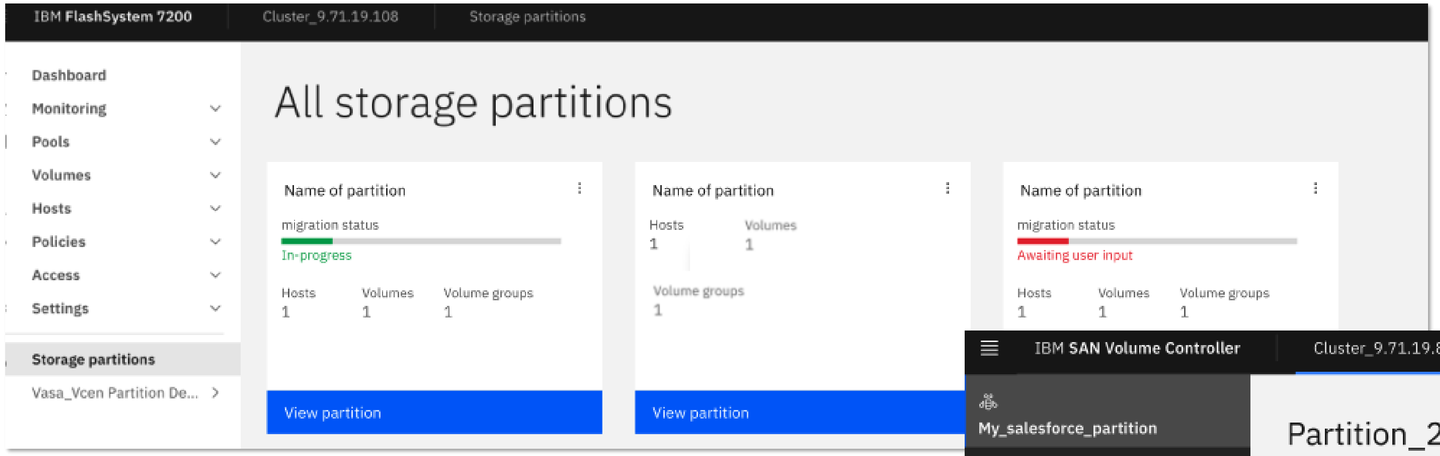
Side panel contains more details.



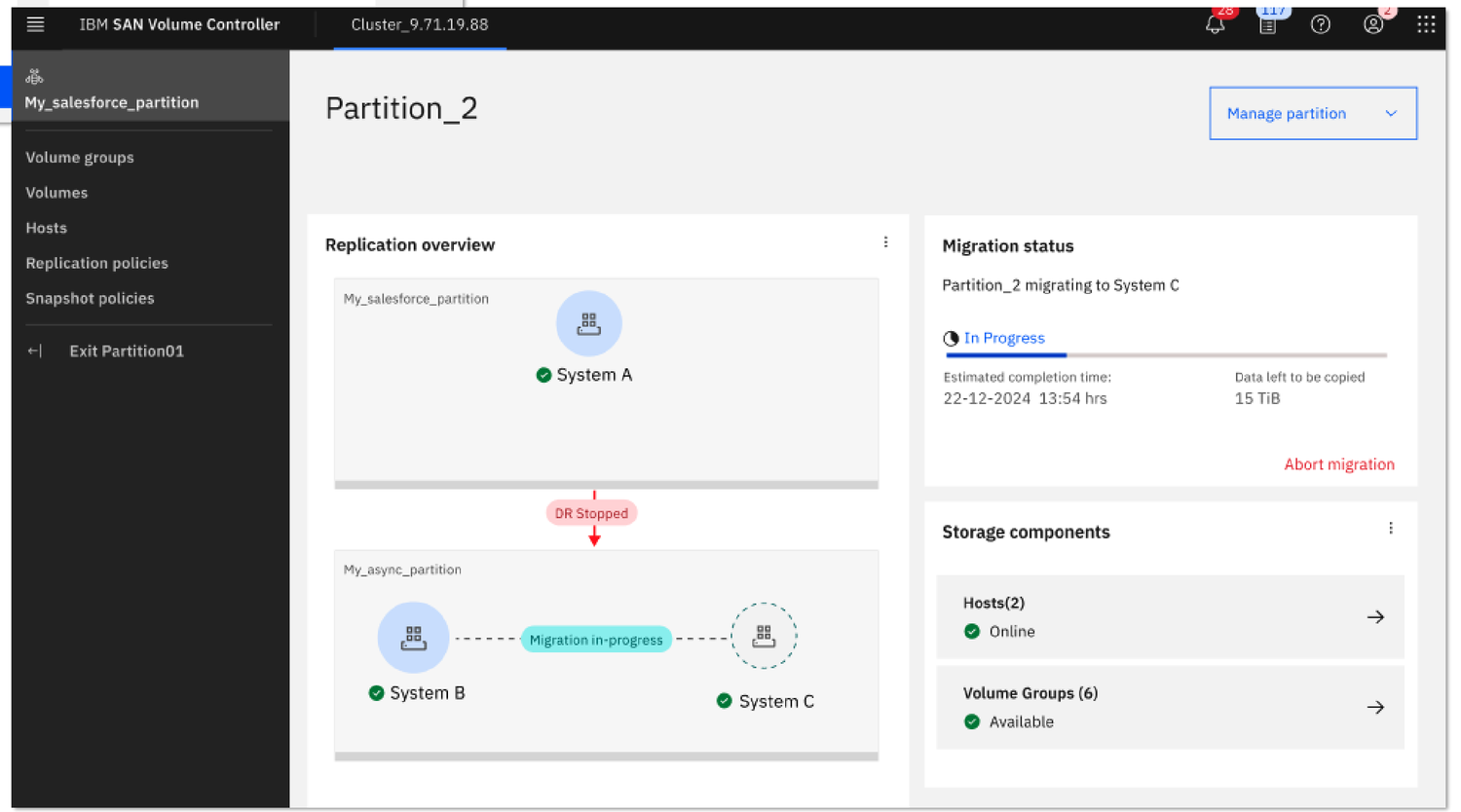


Partition Migration Monitoring from GUI

System Context GUI Monitoring



Partition details view in system context

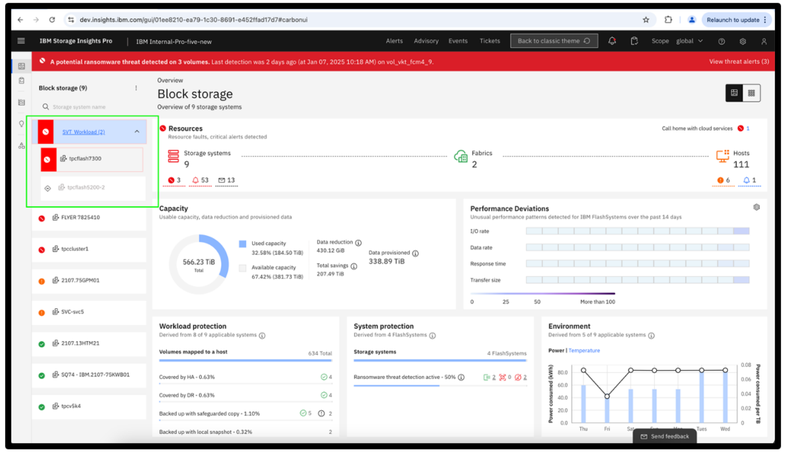


Building blocks of a grid

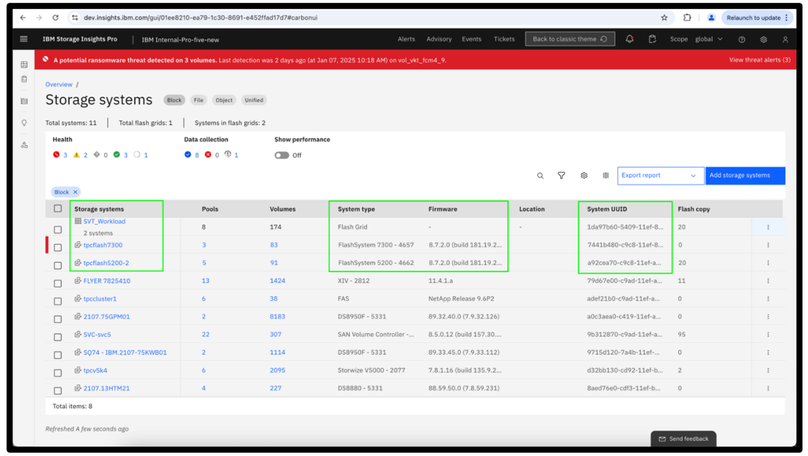
Analytics layer

Building blocks of a grid

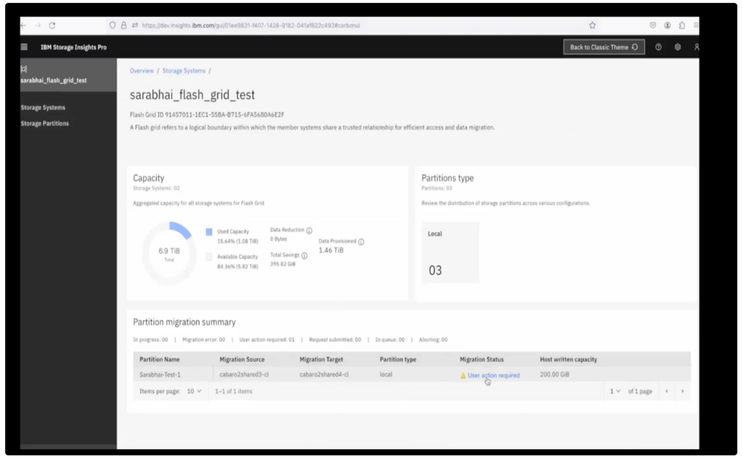
Analytics layer



Seamless Integration

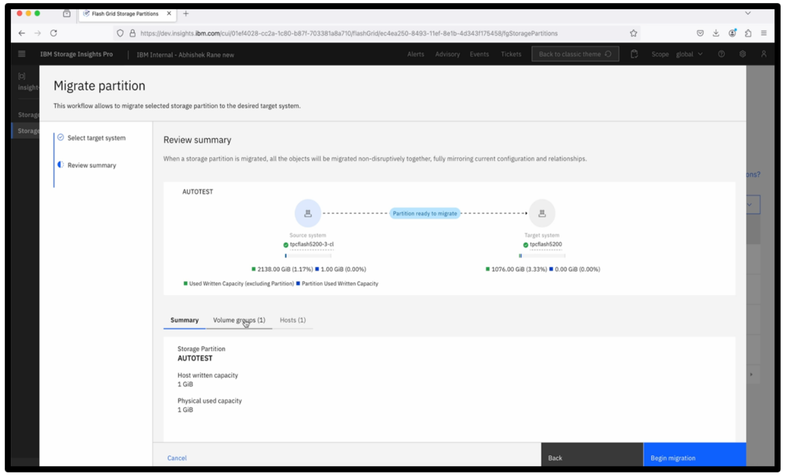


Single Pane Of Glass View

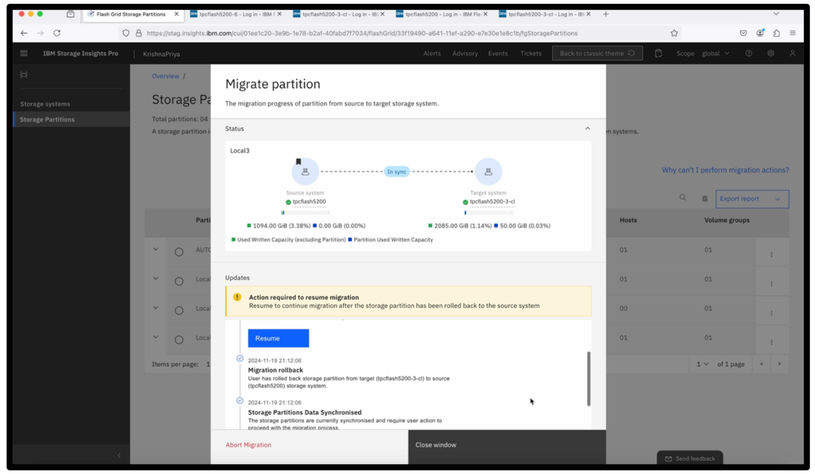


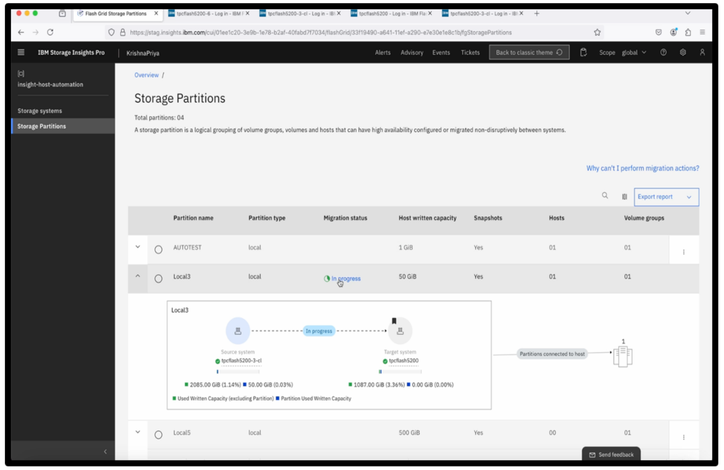
Detailed Operations View

Simplified Migration Workflows



Pause and Resume Migration



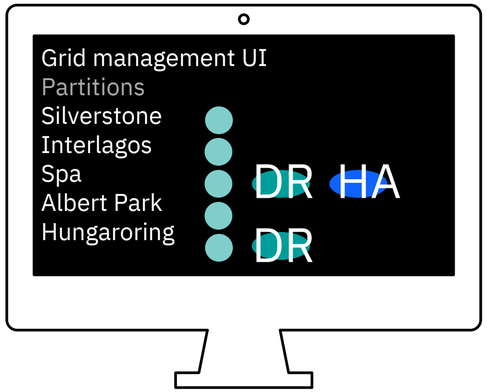


Building blocks of a grid

Analytics layer

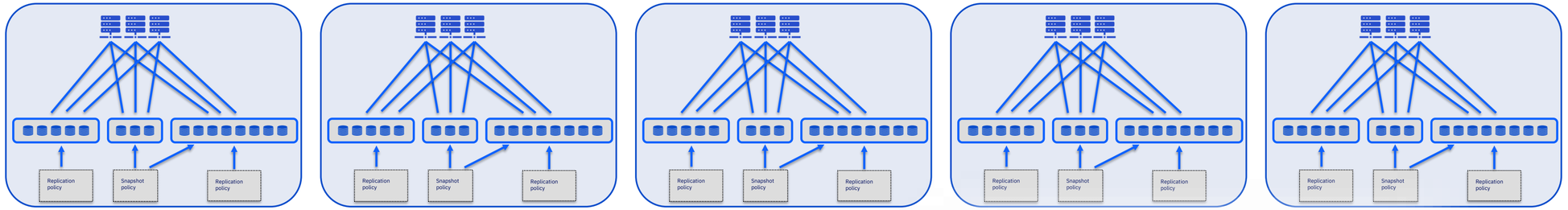
Grid management layer

Onboard management from any system



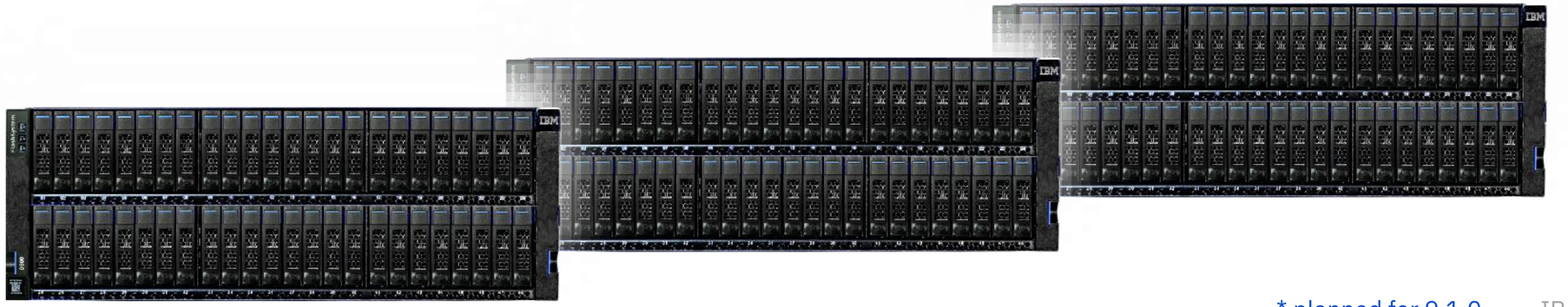
Partitions layer

Up to 32* per system



Infrastructure layer

1-32* systems



* planned for 9.1.0

IBM FlashSystem grid

Scale-out performance and resources as needed

Federated ‘*at-a-glance*’ management of all grid members

Non-disruptive workload migrations within the grid

AIops data placement recommendations in the grid

Add capacity and performance at the right price point