

Power Week 2025

#pw2025

18 - 19 - 20 novembre 2025
IBM Innovation Studio Paris

Faire confiance au stockage IBM Flashsystem pour répliquer ses données.

PBR / PBHA

Antoine Maille
IBM France
amaille@fr.ibm.com

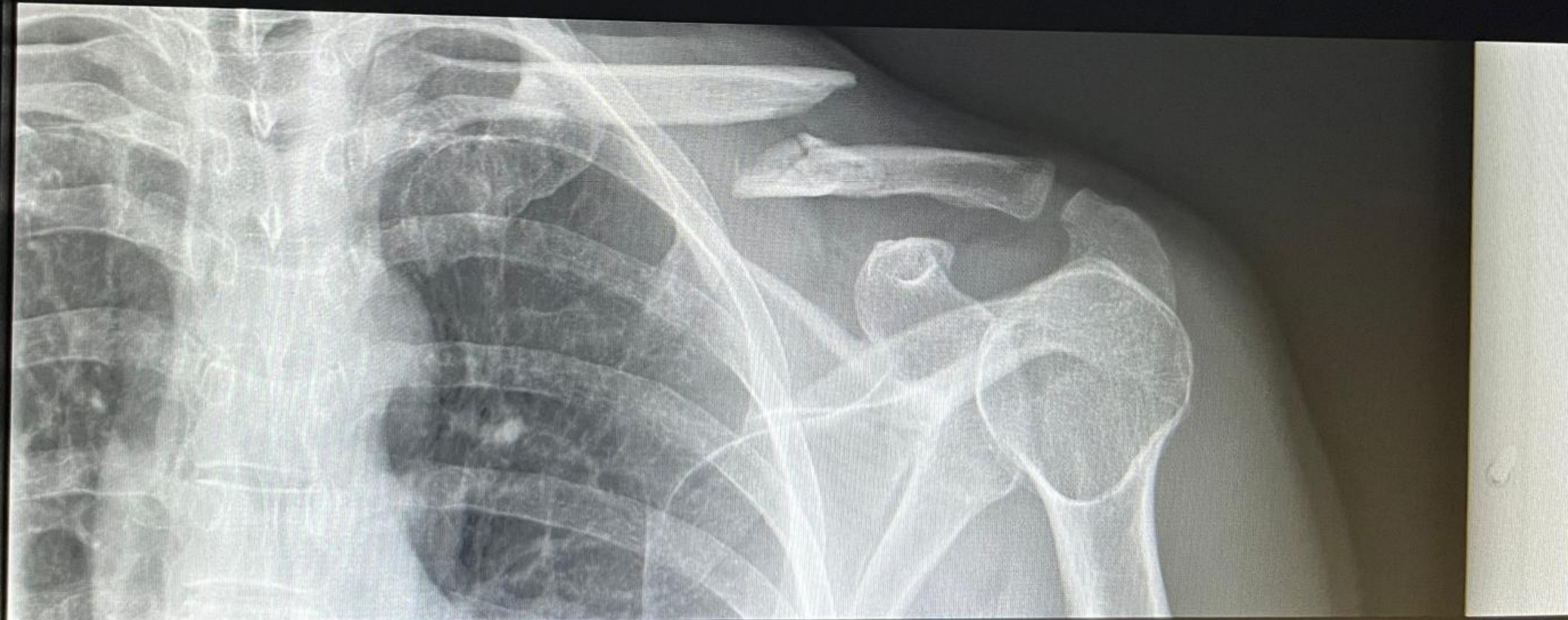
Guillaume Legmar
IBM Techzone
guillaume.legmar@fr.ibm.com

IBM

common
FRANCE

MAILLE, ANTOINE
09-05-1976

Clinique du Parc
FLUOROSPOT COMPACT
STEFANOVIC
15-11-2025
13:54:40



kV: 56.9

Résilience des données avec IBM FlashSystem

« S'assurer que vos **données restent disponibles et accessibles** pour vos applications est la **fonction principale du stockage** »

Chiffrement des données

(FIPS certified drives)



Aucun impact de performance

Réplication Distante

Metro Mirror
(RPO=0)
Policy-based Replication (RPO>0)



Synchrone ou Asynchrone

Haute-Disponibilité

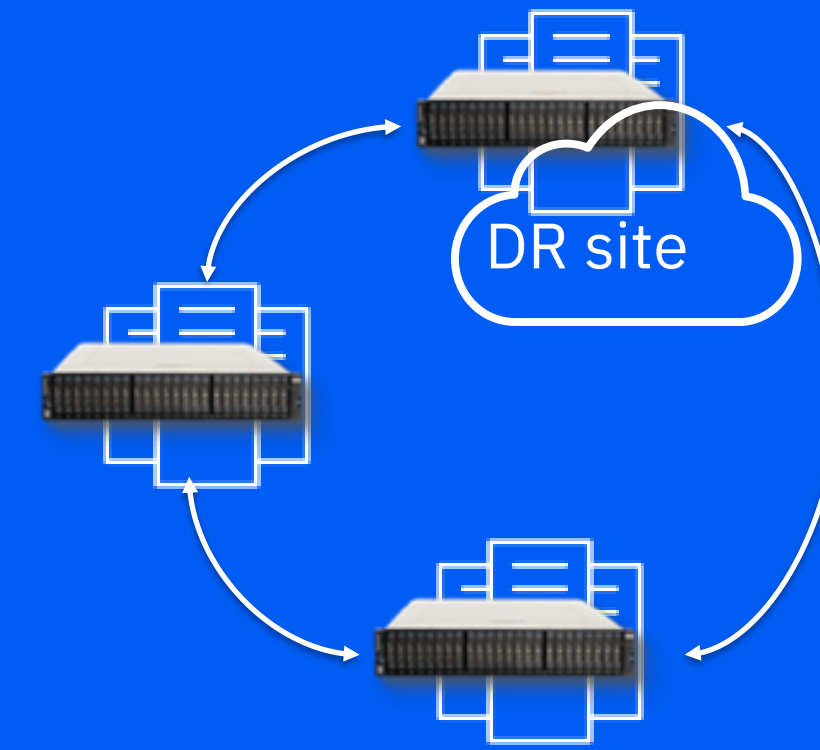
Policy-based HA
RPO=0 / RTO=0



Volumes logiques répliqués

100% de disponibilité

Architecture 3 copies (HA+DR)



Prises d'images instantanées

FlashCopy
+
Safeguarded Copy



Protection contre les corruptions logiques, Test/Dev

Storage Virtualize

Evolution des mécanismes de réplication

Sync

2003 - MetroMirror

Presque inchangé à ce jour

Async

2005 - GlobalMirror

RPO très court, bons liens nécessaires

2010 - GlobalMirror w Change Vols

Configuration complexes et sujets aux erreurs
RPO de quelques minutes, tolérance de liaison plus élevée

Limites de
capacité

Limites de
performance

HA

2007 - SVC Stretched Cluster

SVC only - site enhanced 2009

2014 - HyperSwap

Basé sur du bi-directional Sync mirror

3-site

201x - Hyperswap + scripts vers 3ème site

Storage Virtualize

Evolution des mécanismes de réplication

Sync

2026 (Version PBR à venir)
Presque inchangé à ce jour

Async

2005 - GlobalMirror
Introduit en 2022
PBR - Async pour remplacer les deux
liaison plus élevée

HA

2007 - SVC Stretched Cluster
Introduit en 2023
PBHA – conçu pour remplacer les deux
Basé sur bi-directional Sync mirror

3-site

Introduit en 2024
PBHA 3 sites

Services de copie : anciens ou nouveaux

Anciens services de copie

- FlashCopy 
- Global Mirror 
- Global Mirror with Change Volume 
- Metro Mirror
- 3-Site Mirror with Orchestrator 
- HyperSwap 

Nouveaux services de copie

- Policy Based Snapshots (8.5.1)
- Policy Based Replication (8.5.2)
- **PB HA with async DR -> new in 8.7.1**
- Policy Based High Availability (8.6.1 / 8.7.0)
- Storage Partition Migration (8.6.3)

8.7.0 is the final release

Note: Storage Partitions ne sont pris en charge que sur les système mono IO Group..
(PB HA, Partition migration, PB HA + DR three site replication)

Replication basée sur des Politiques

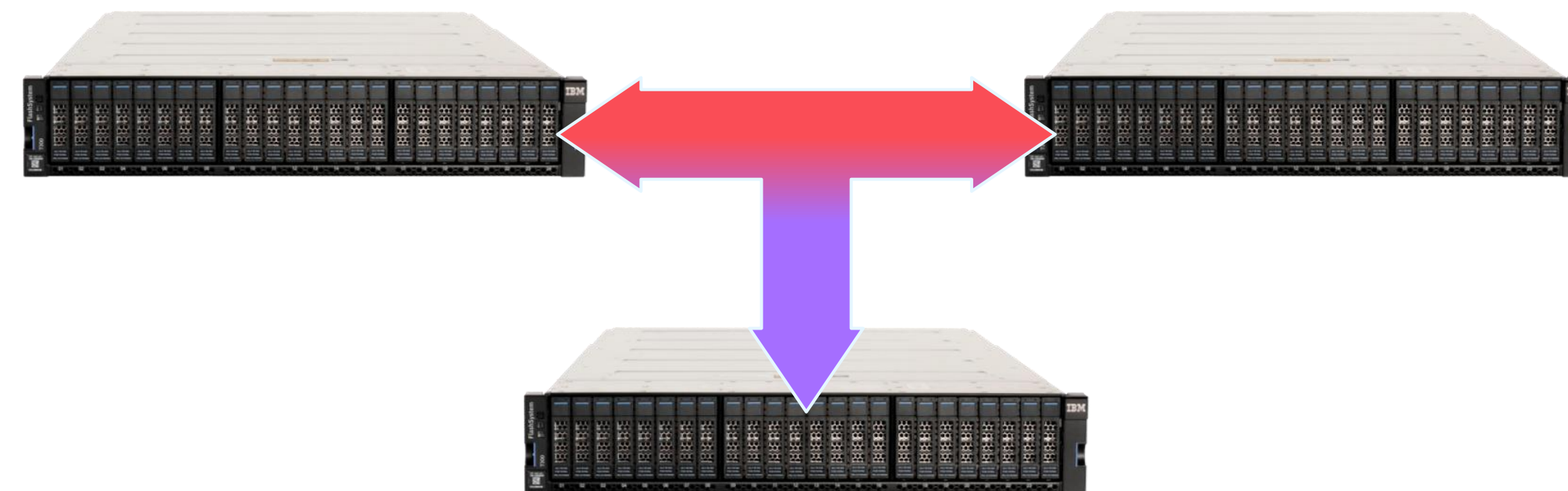
Asynchrone
PBR



Synchrone/HA
PBHA

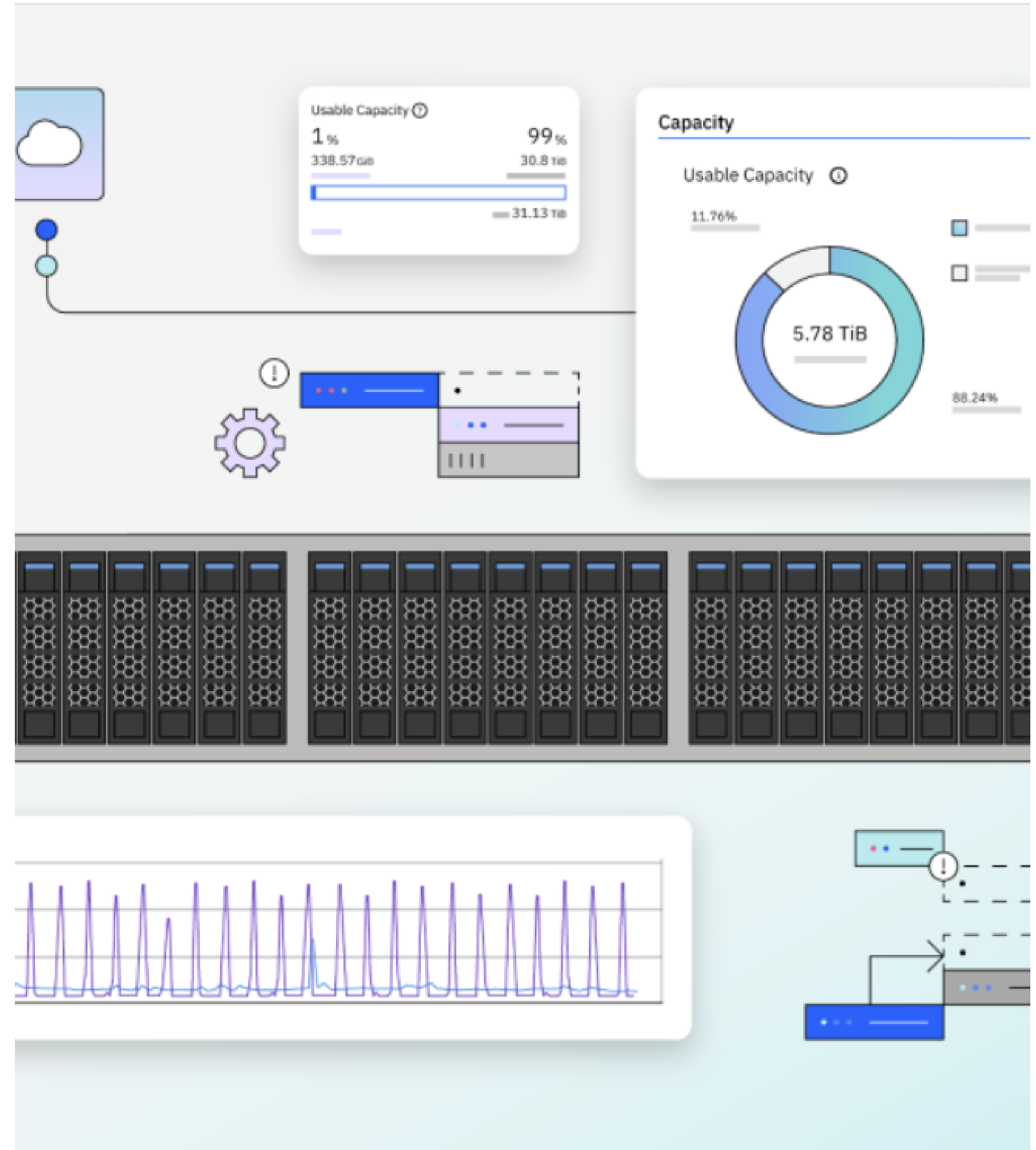


HA + Asynchrone
PBHA 3-site

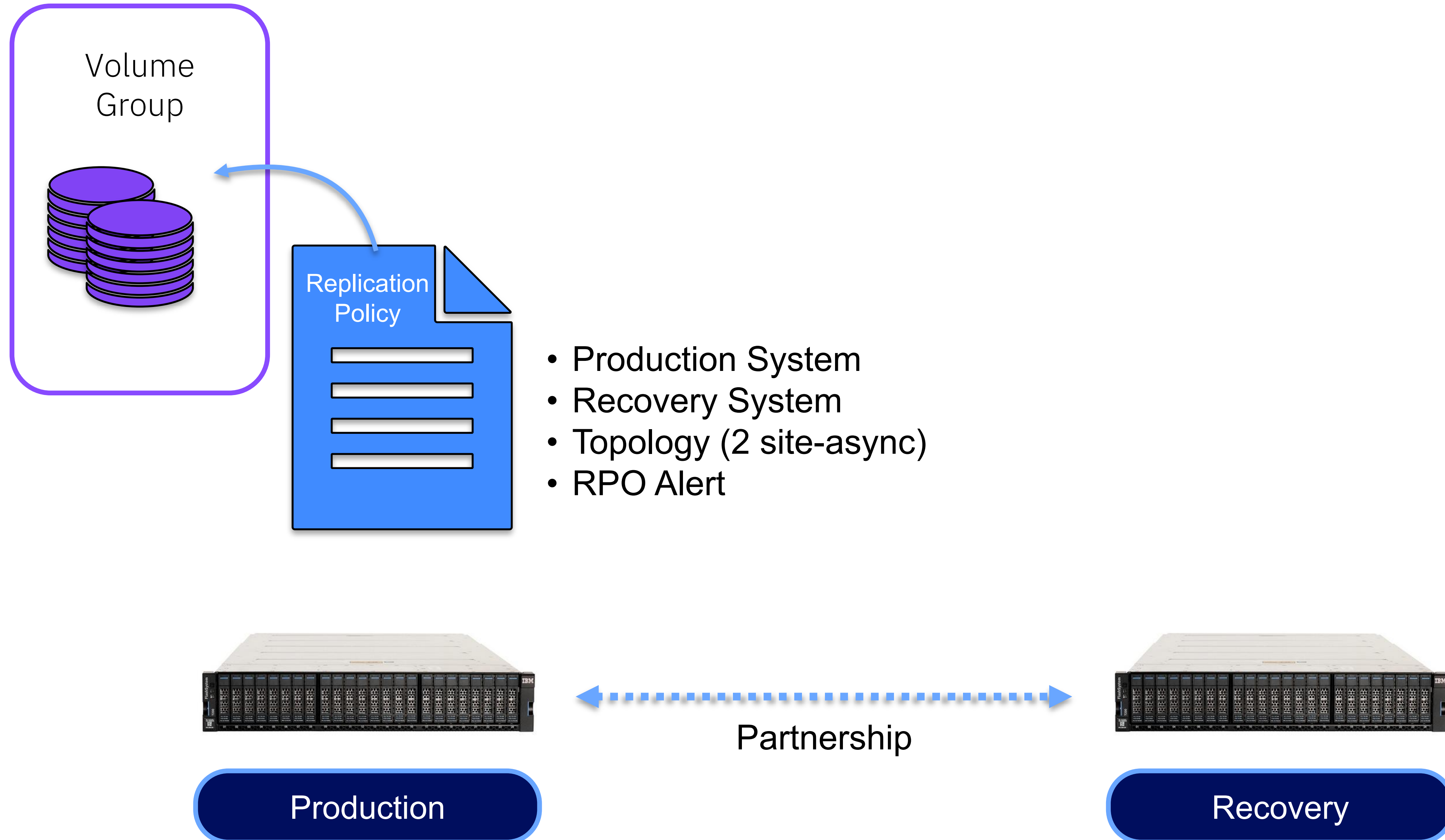


IBM FlashSystems

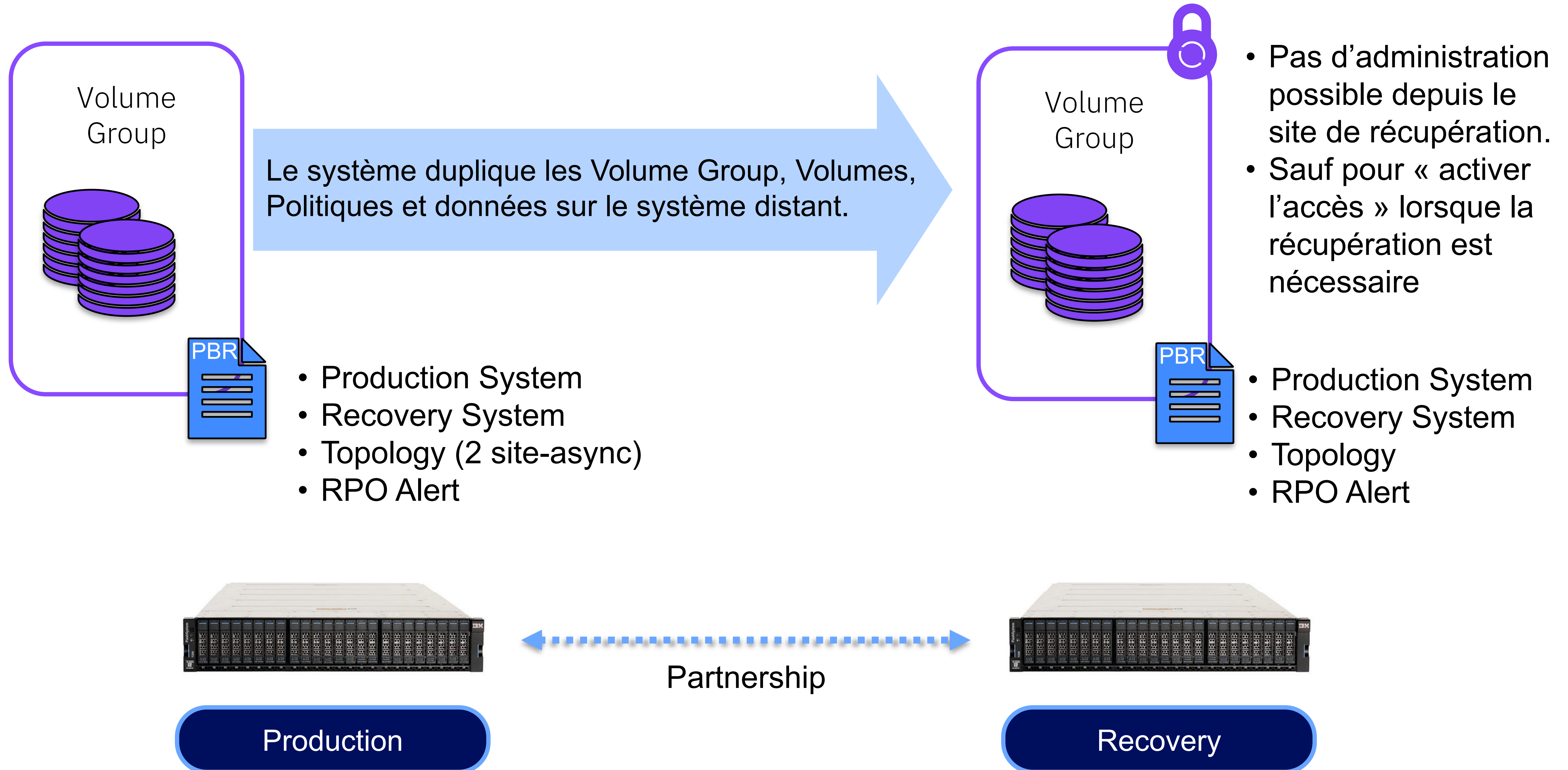
PB Replication Review



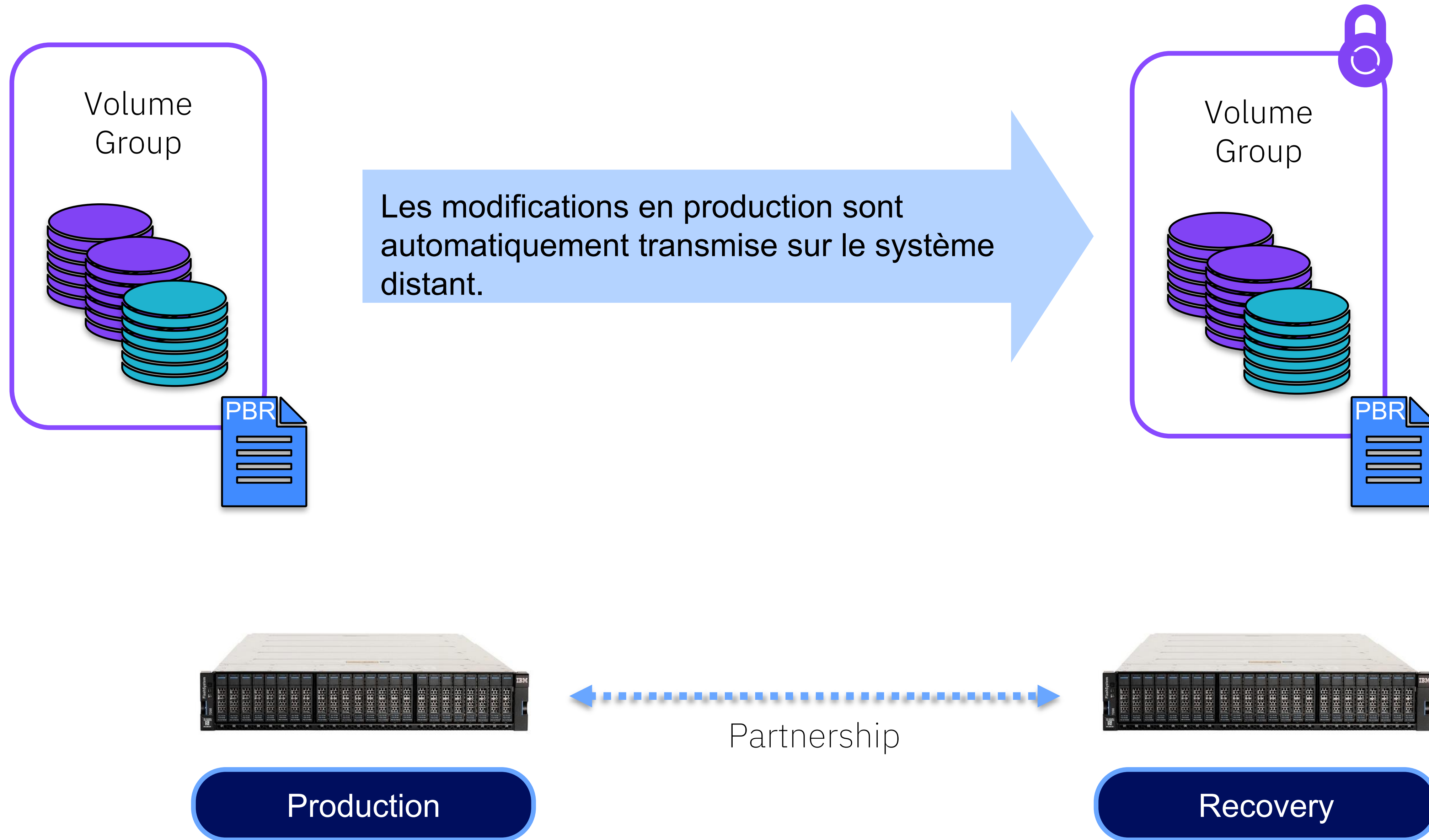
Introducing Policy-based Replication (PBR)



Introducing Policy-based Replication (PBR)

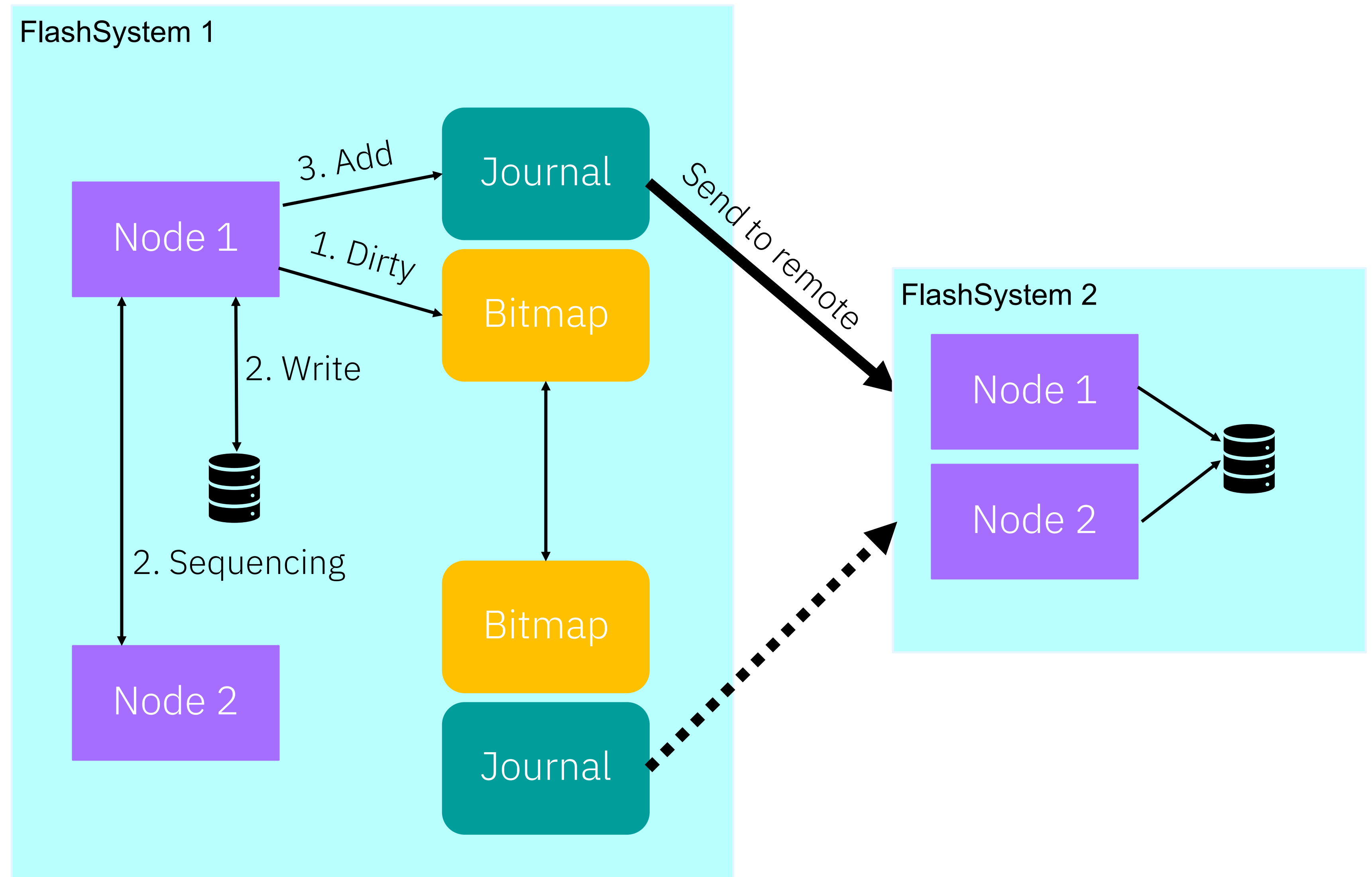


Introducing Policy-based Replication (PBR)



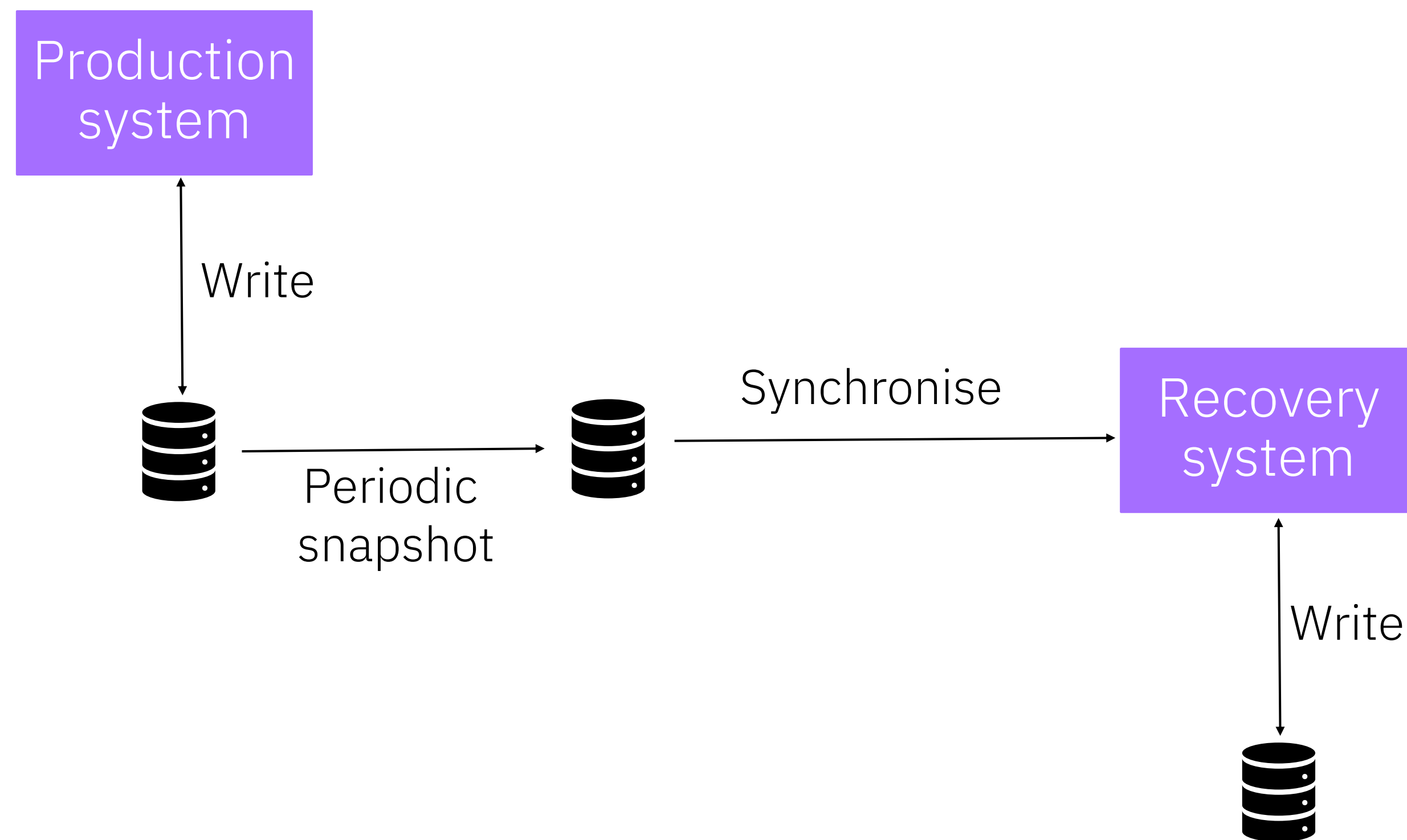
Journaling mode

- Les écritures sont mises en mémoire tampon dans un journal en mémoire qui permet de stocker de nombreuses secondes d'écritures pour absorber les problèmes de performances de réplication sans impact sur l'hôte.
- La taille des journaux varie de 8 Gio à 32 Gio par nœud.
- Les volumes groupes essaieront toujours de se répliquer dans ce mode, car il offre un meilleur point de récupération.



Cycling mode

- Utilise des snapshots locaux pour stocker les modifications qui seront copiées du système de production vers le système de DR.
- La fréquence des cycles est dépendante du temps nécessaire à la transmission des informations.
- Réduit la bande passante requise pour la réplication.
- Les groupes de volumes peuvent être basculés vers ce mode si le mode de journalisation ne peut pas être maintenu.
- Dimensionnement : chaque volume PBR nécessite un volume Snapshot supplémentaire masqué et la capacité associée



Policy-based replication propose une **réplication adaptative**

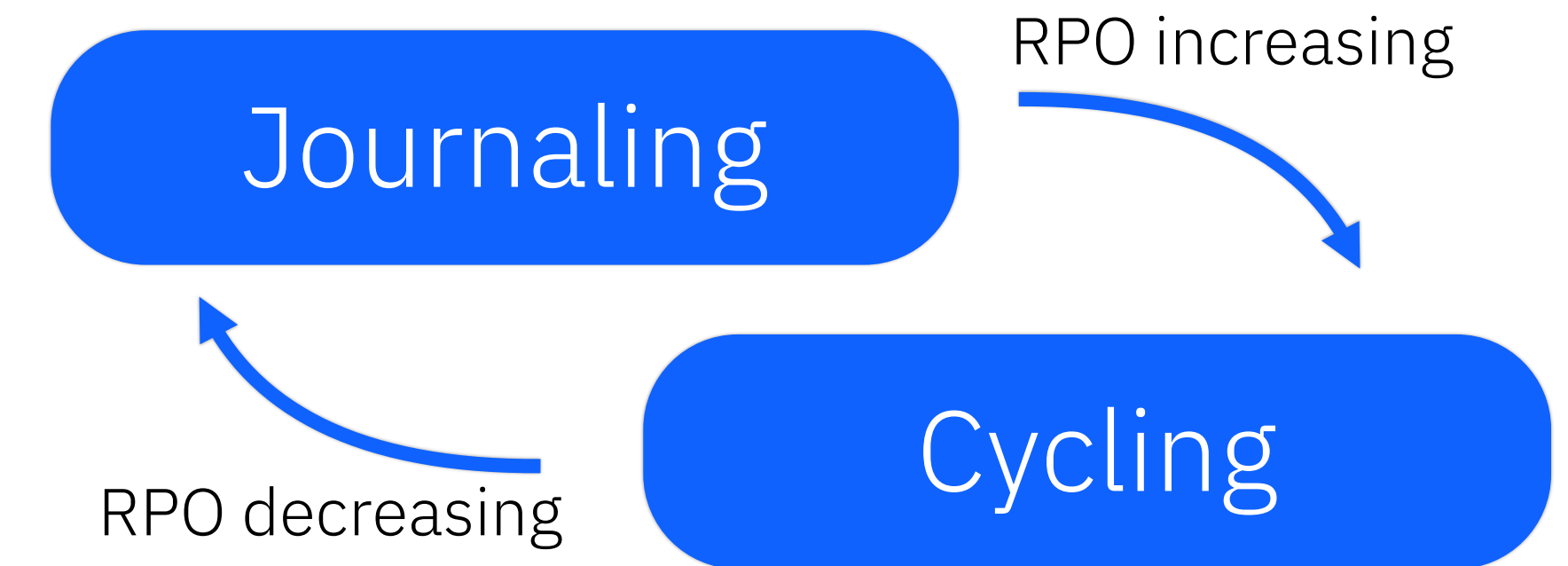
Groupes de volumes répliqués :

- Préfèrent toujours le mode journalisation
- Peut passer en mode cycling
(si le débit de l'hôte est supérieur au débit de réplication)
- Ceux qui ont des objectifs RPO plus long passeront d'abord en mode cycling.
(si tous les RPO sont identiques, ceux qui ont la bande passante d'écriture la plus élevée basculeront en premier)

Le système vise toujours à respecter le RPO pour tous les groupes de volumes, quel que soit le mode.

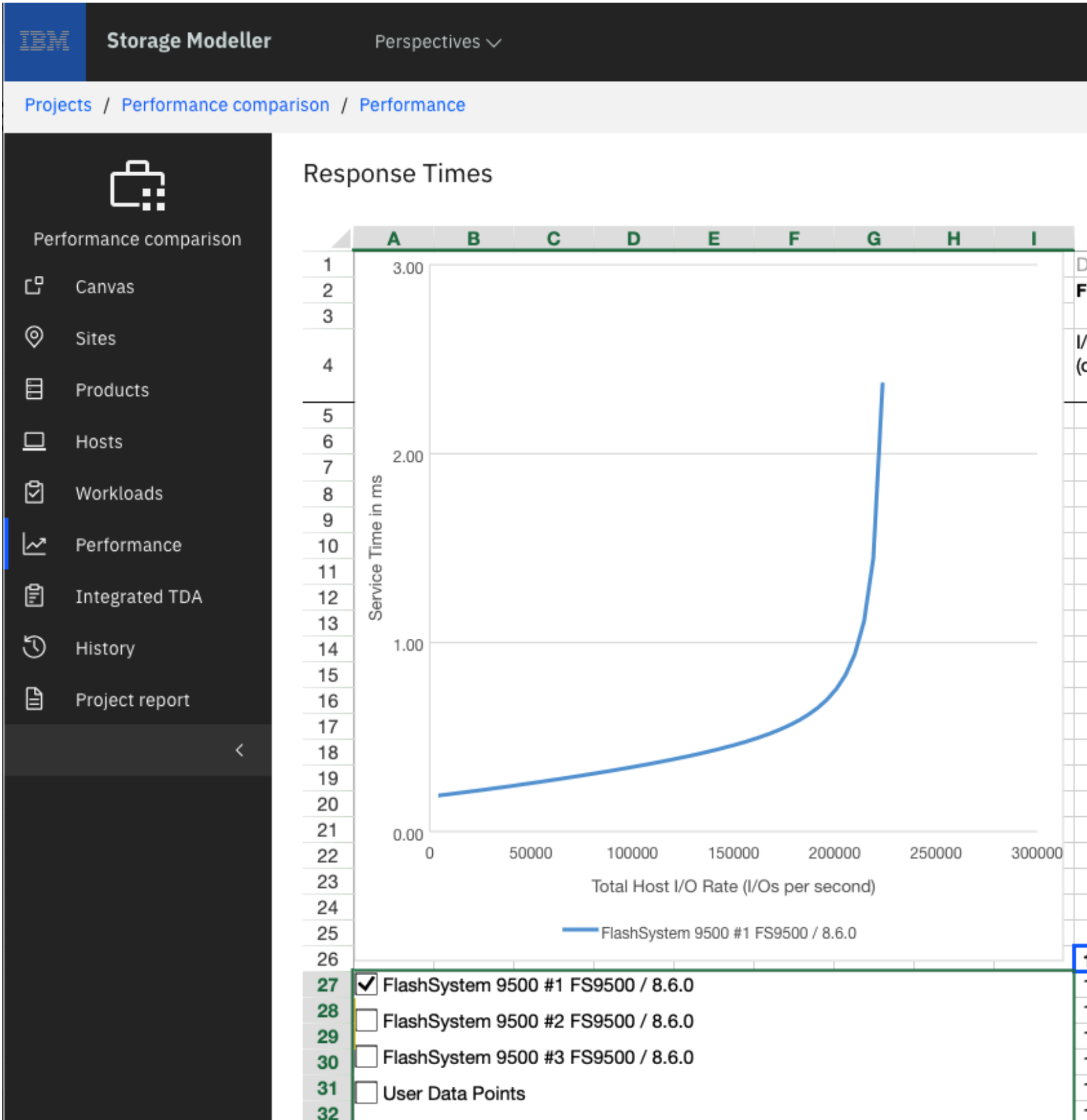
Le « recovery point » d'un groupe de volumes est signalé.

Il n'y a aucun contrôle ou visibilité de l'utilisateur sur le mode.

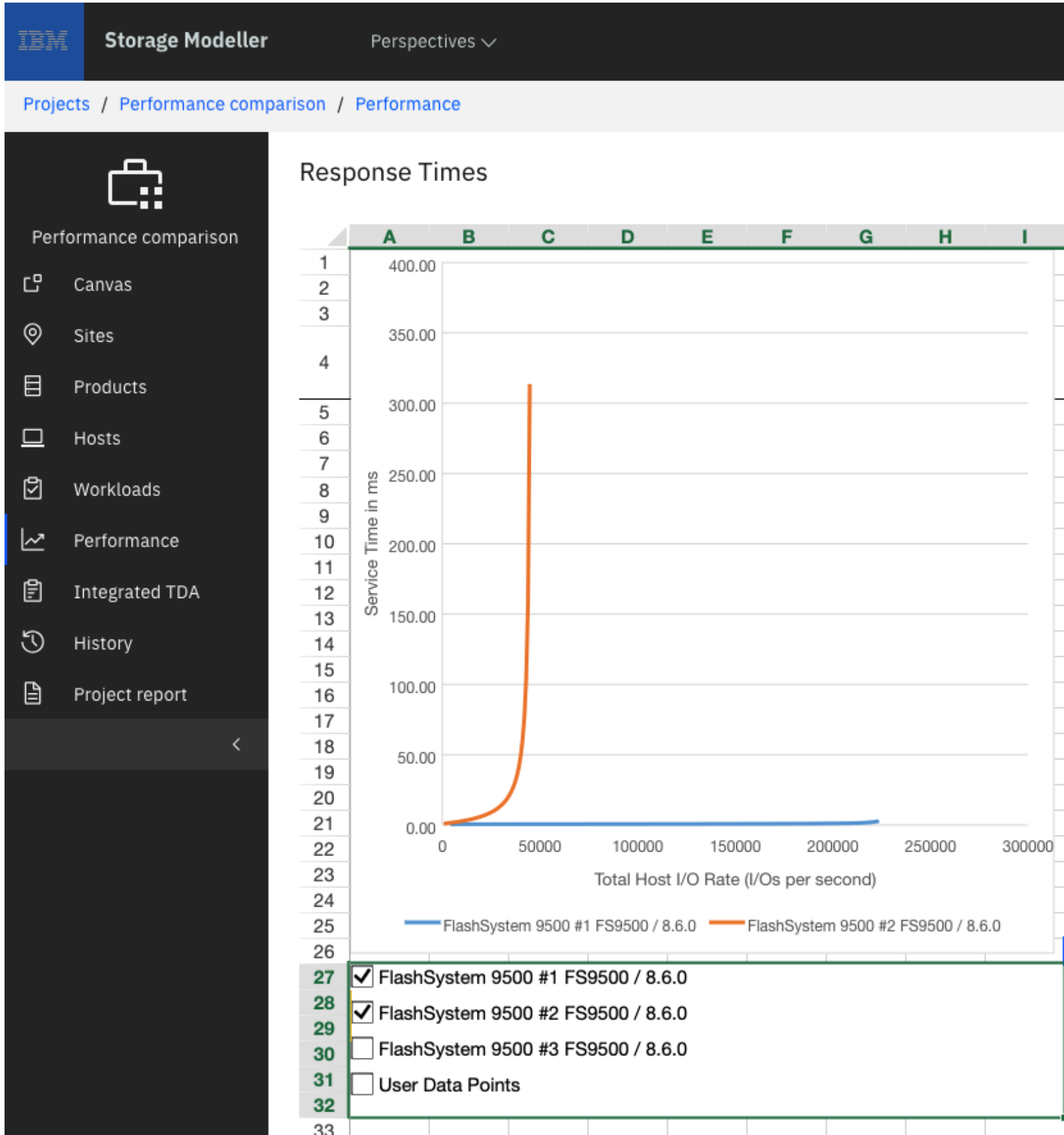


Performance: IBM Storage Modeller, FS9500, 8.6.0

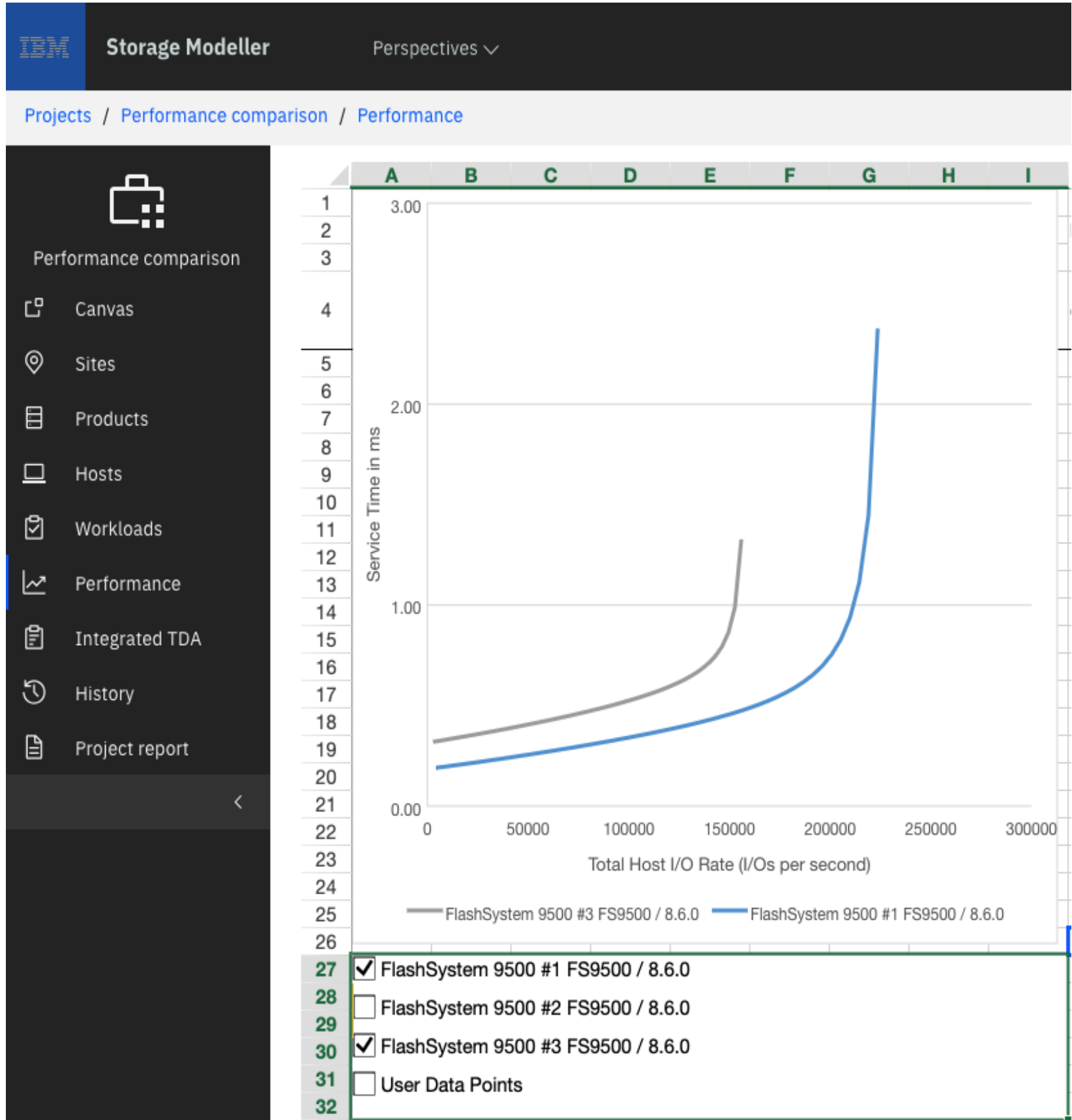
No Mirror



Global Mirror

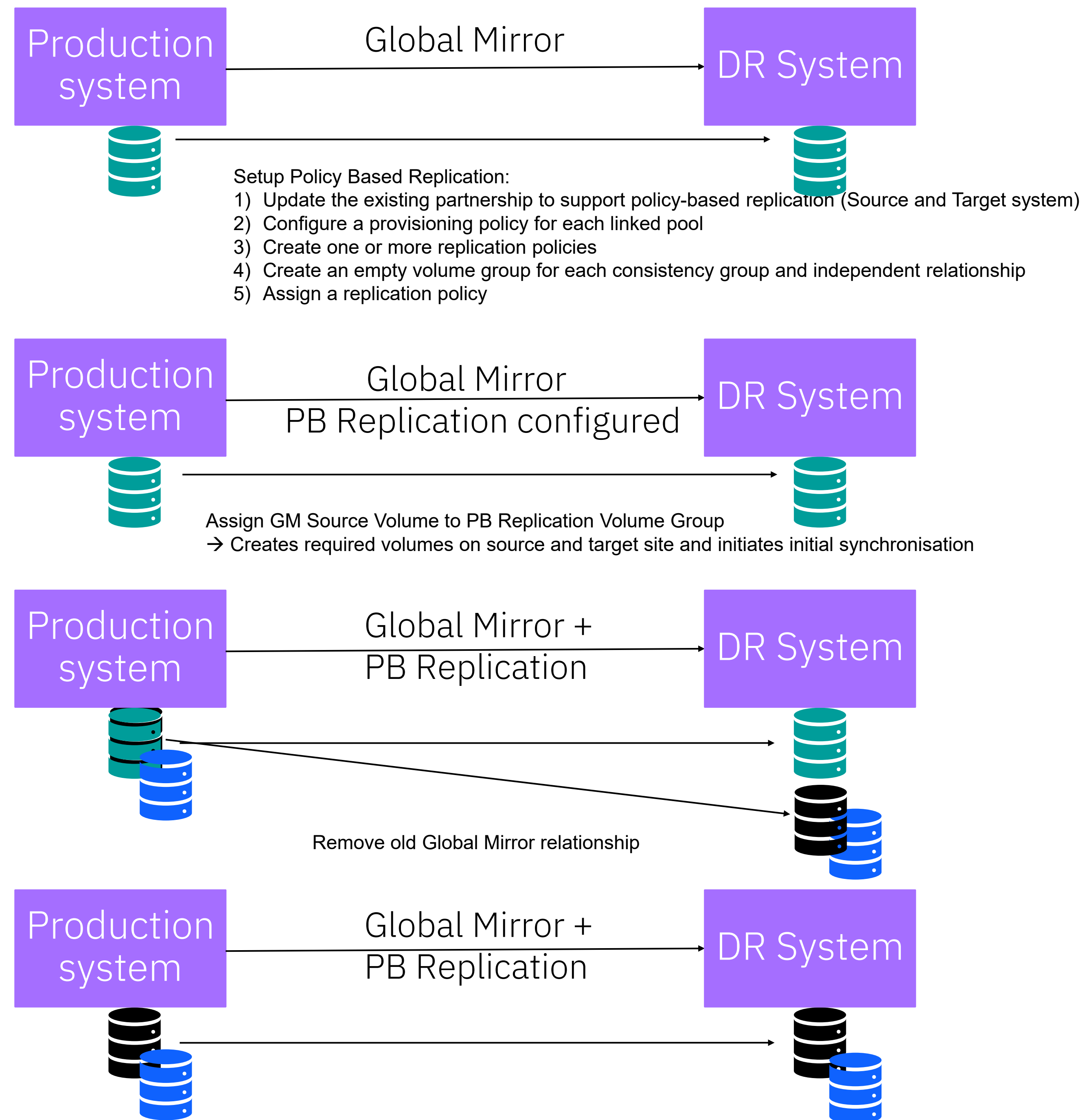


PB Mirror



Conversion MM / GM / GMwCV to PB Replication

- La conversion est toujours lancée au niveau du volume source
- Les volumes avec des FlashCopies associées ne peuvent pas être convertis
- Migration de volume de GMwCV vers GM requise
- Synchronisation complète avec le nouveau volume
- Notez les exigences de capacité
- Consultez la documentation pour connaître les exigences et les restrictions actuelles



Policy-based Replication (PBR) Timeline of feature Introduction

	8.5.2	8.6.0	8.6.2	8.7.0	8.7.1	8.7.2	8.7.3	9.1.0
PBR Async - with Adaptive Replication								
PBR Sync								
Supported on NVMe Systems (250ms RTT max)								
Supported on 5045 (80ms RTT max)								
Same src vol with GM and PBR								
FC & IP Partnerships	3	3	3	3	3	3	3	3
Indirect Partnerships not counted								
RDMA IP Partnerships			1	2	2	2	2	2
Only partner with 8.7.0 and above								
DR Test Mode								
Checkpointing								
3site support with PBHA and PBR								
Partition level PBR policies								
Snapshot restore to PBR production volume ¹				8.7.0.5+				
Automatic pool linking								
Partition Migration of PBR production or recovery VGs								
Rename PBR objects ⁴								
Expand PBR volumes								
Recovery site volumes read only rather than offline ³								
VMware VASA2 (Replication) Support								
Number of I/O groups per system using PBR ²	2		4					
Number of replication policies	32							132 ⁵

1. Feature back port to 8.7.0.5 and subsequent 8.7.0 PTFs - not available in 8.7.1

2. Single I/O group can be replicating with at most 6 other I/O groups in partnered systems

3. Only when PBR used with VGs inside a partition.

4. volumes, volume groups, hosts, host clusters & partitions

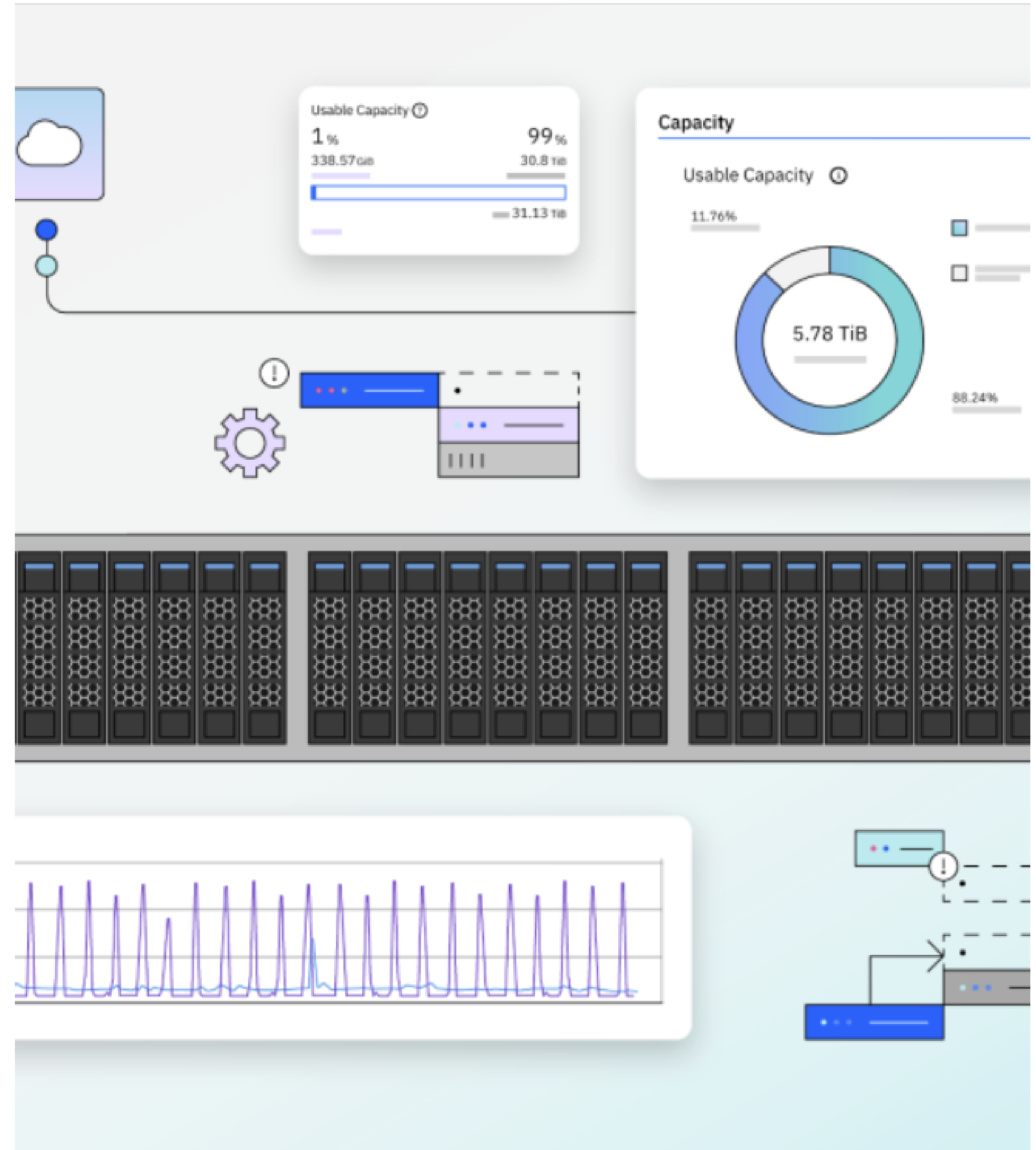
5. 5045 remains at 32

Show do not Tell !



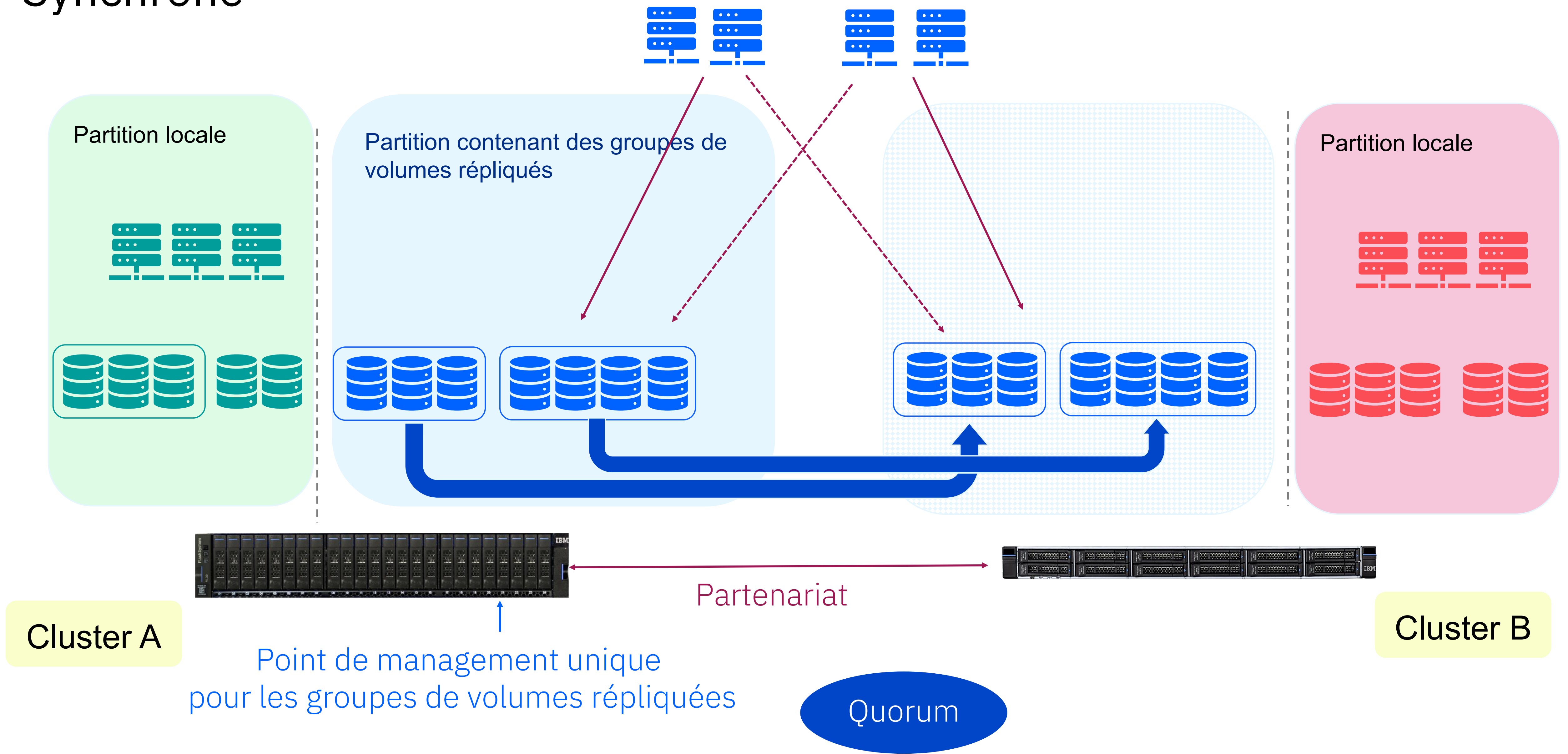
IBM FlashSystems

PB-HA Review

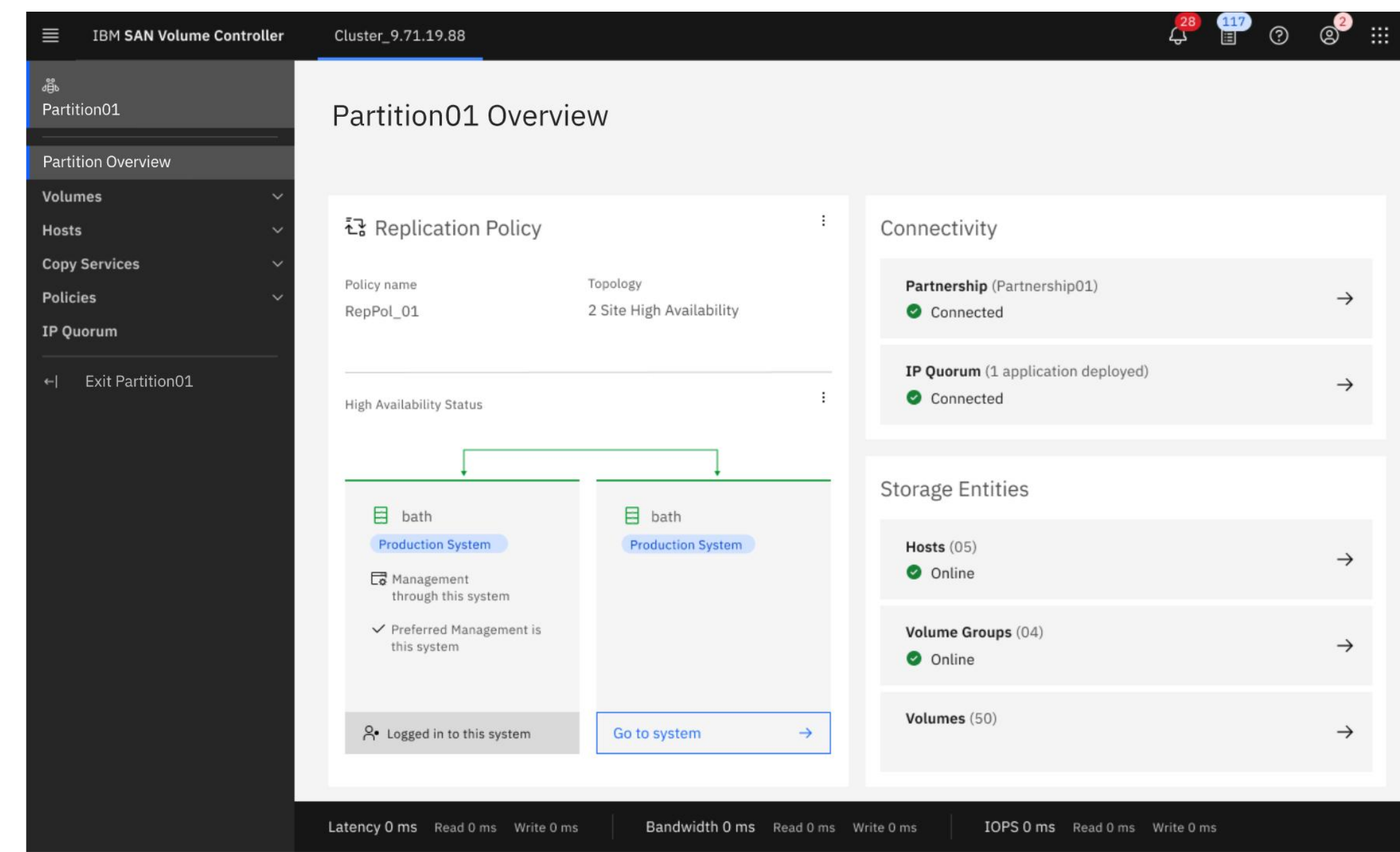


Réplication basée sur des politiques “Policy Based Replication HA”

Synchrone



Les partitions de stockage simplifient la configuration, la gestion et la surveillance, avec un point de contrôle unique pour la haute disponibilité



As seen in 8.7.0

Create and populate

Créez une partition de stockage en quelques clics et créez ou ajoutez un nombre quelconque d’hôtes, de groupes de volumes et de volumes.

Policy-based

Ajoutez une stratégie de réplication HA à la partition et tout ce qu’elle contient sera automatiquement configuré pour être hautement disponible, avec un provisionnement à distance géré automatiquement.

Hautement disponible

En cas de sinistre, la partition de stockage gère automatiquement sa disponibilité en veillant à ce que les applications restent accessibles, avec un contrôle par partition sur le système préféré en cas de perte de connectivité.

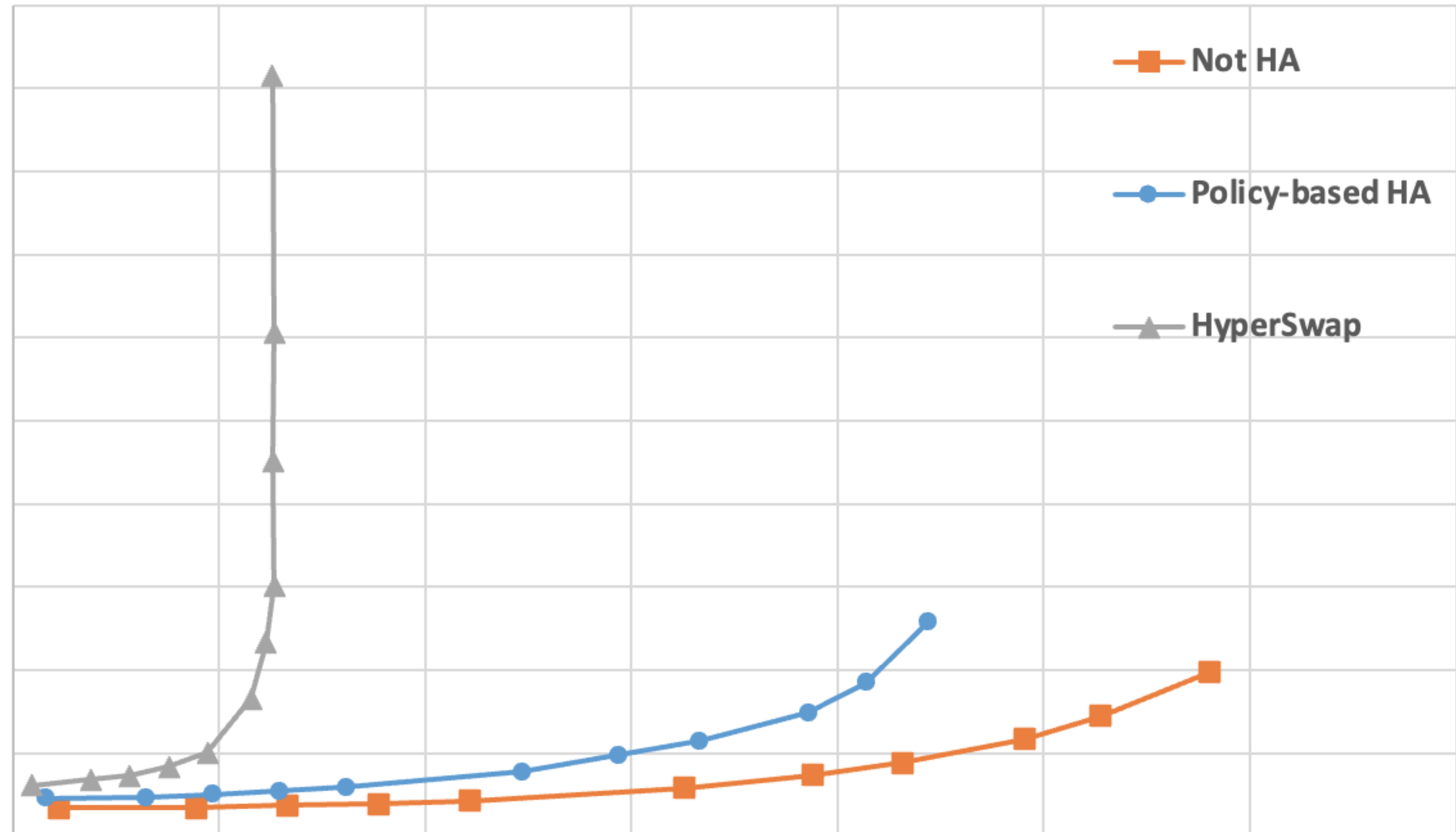
Delimité

Les problèmes de haute disponibilité n’ont aucun impact sur les volumes non HA : les stockages HA et non-HA peuvent coexister sur le même système.

Performances

À l'instar de la réplication basée sur des règles, une réécriture de fond en comble offre une augmentation significative des performances par rapport à HyperSwap. Les performances de l'AH sont proches des performances non HA, ce qui maximise l'investissement.

Deux systèmes FlashSystem 9500 (mono I/O group) connecté via FC, 16KiB 50%r/50%w/50% hit à une distance négligeable



HA Comparison - HA Feature Set

Improved

Current Restriction

	Policy-based HA	HyperSwap	Enh Stretched Cluster
Non-HA volumes (local)	Unaffected by HA problems	Link problems cause offline vols	Link problems cause offline vols
Non-HA volumes mapped to HA host	Future	Yes	Yes
Volume resize support	Yes (9.1.0)	Expand thin volumes	Expand/shrink
Host clusters	Yes (9.1.0)	Yes	Yes
Modify hosts in HA	Future	Yes	Yes
Host locations	Active/active with optional host-site awareness (8.7.0)	Active/active with mandatory host-site awareness	Active/active with optional host-site awareness
Protects (mutual) consistency during resynchronisation	Yes	Yes (mutual consistency with consistency groups)	No
3-site support	Yes (8.7.1)	Yes	Synchronous with Metro Mirror, async with policy-based replication
3-site orchestration	Not required, (8.7.1)	External SW	Not required
> 3-site support	Options being considered	No	Yes
HA snapshots (including Safeguarded)	Yes (8.7.3)	No	SA2,SV2,SV3 8.6.0
HA snapshot restore	Yes (8.7.2)	No	Yes
HA snapshot thin/clone workflows	Future	Yes	Yes
PBR support on same system	To different partner (8.7.0)	No	Yes

	Policy-based HA	HyperSwap	Enh Stretched Cluster
Protects against outage of entire system	Yes	No	No
Protects against outage of IO Group	Yes	Yes	No
Implications of ISL loss	Only non-preferred management location partitions go offline	Entire non-preferred site (IOGs) go offline	Half of each IO group goes offline at non-preferred site. Cache loss
Performance	>5x HyperSwap	HyperSwap code limited	Node, Cache and ISL RTT limited
Internal Messaging Round Trips (per write)	Worst case 2	Worst case 2	At least 6 (more with replication etc)
Performance impact under failure	None	No cache loss	Loss of write caching, latency implications
Site Link Networking	[Partnership] Fibre Channel RDMA HSE Ethernet	[Clustering] Fibre Channel RDMA iSER Ethernet	[Clustering] Fibre Channel RDMA iSER Ethernet
Site Link Distances	Recommend < 1ms RTT (~100km)*	< 3ms RTT (~300km)	< 3ms RTT (~300km)

* SCORE requested for > 1ms RTT, to ensure end user aware of write latency implications.

HA comparison - Host OS & Protocol Interop

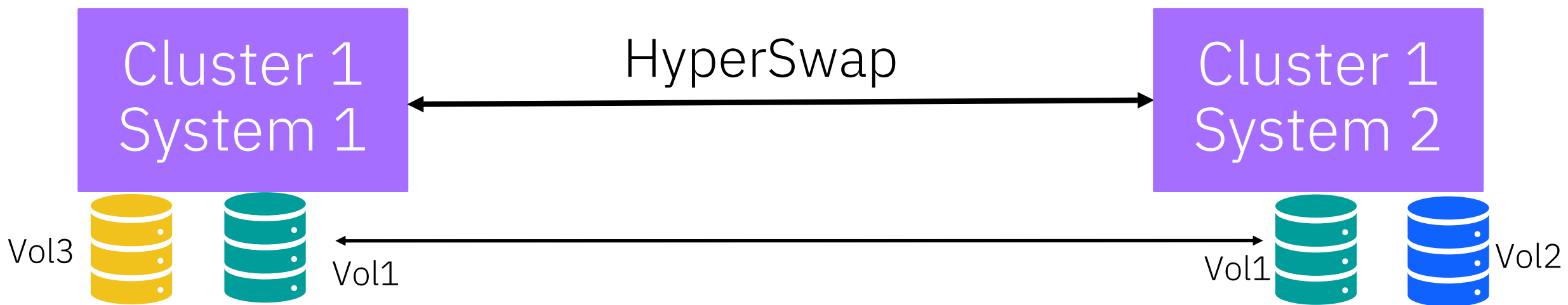
Current Restriction

Host Protocol support	Policy-based HA	HyperSwap	Enh. Stretched Cluster
Fibre Channel - SCSI	Yes	Yes	Yes
Fibre Channel - NVMe	9.1.1	No	Yes
SCSI-3 Persistent Reservations	8.7v.2	Yes	Yes
Ethernet - iSCSI	8.7.2	Yes	Yes
Ethernet - NVMe	9.1.1	No	No
Ethernet - iSER	No	Yes	Yes - up to 8.7.0

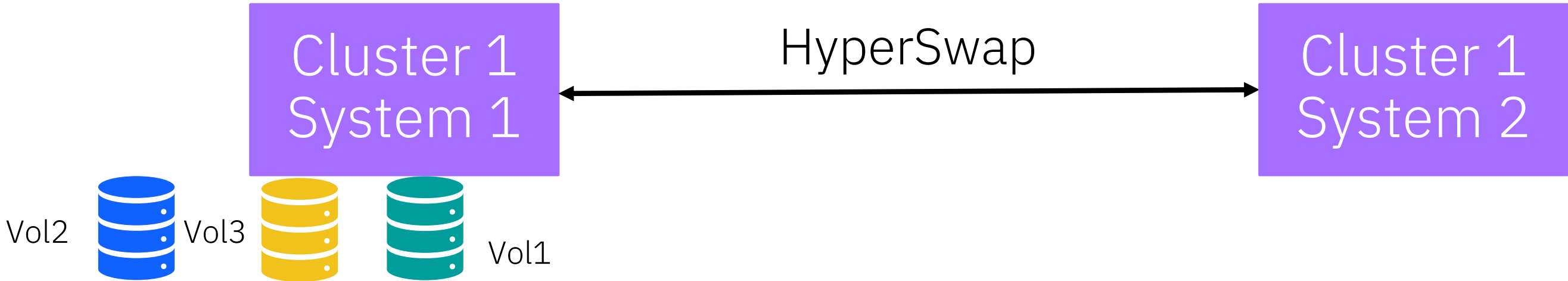
Host OS Support	Policy-based HA		HyperSwap	Enhanced Stretched Cluster
	Fibre Channel	iSCSI		
8.7.0	RHEL (+KVM), SLES VMware ESXi AIX 7.2+ zLinux	No	Full	Full
8.7.2	<i>Additionally :</i> Solaris 11 Windows Server (+HyperV) IBM I (via SCORE)	RHEL (+KVM), SLES VMware ESXi (excluding vVols)	8.7.1+ does not support HyperSwap	SVC Clustering only supported on LTS releases
8.7.3	<i>No additional</i>	<i>Additionally :</i> Windows Server (+HyperV)		

Conversion à partir de HyperSwap vers PB High Availability

HyperSwap topology:



Transformer les volumes Hyperswap en volume standard sur un seul iogrp



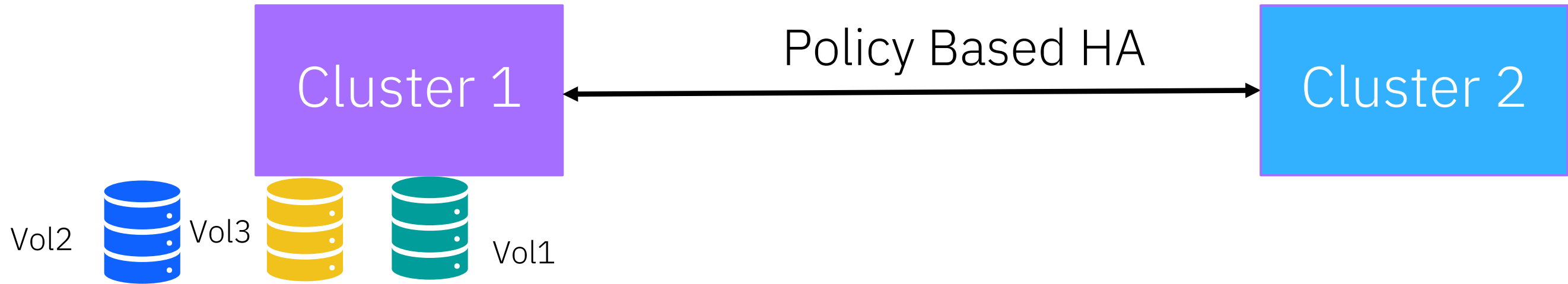
Supprimer les nodes “vides” du cluster



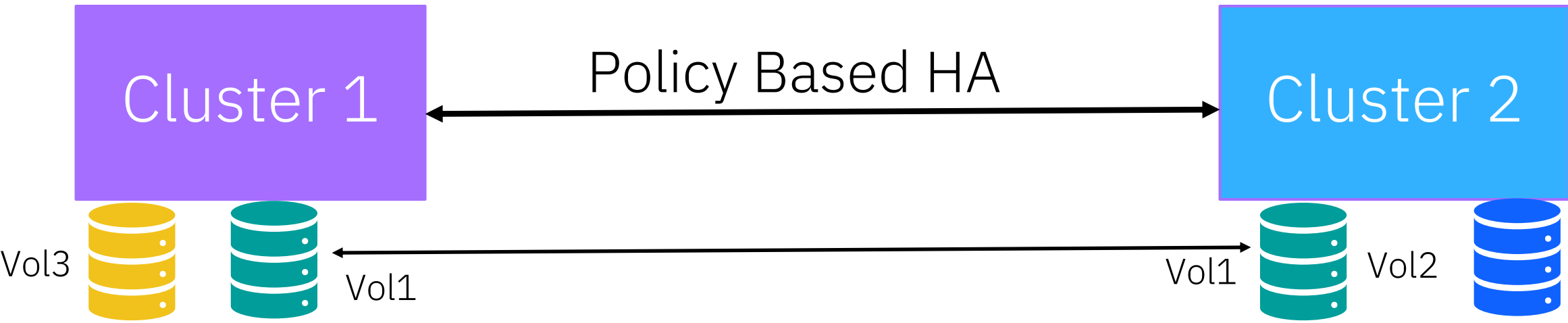
Conversion from HyperSwap to PB High Availability



Créer un second cluster avec les nodes candidats
Configurez le nouveau système, effectuez la configuration de base (sécurité, callhome) et configurez le pool et le stockage



Utiliser le GUI sur le système d'origine pour créer une nouvelle partition avec les volumes existants,
Ajoutez HA et transmettez toutes les données au nouveau système partenaire



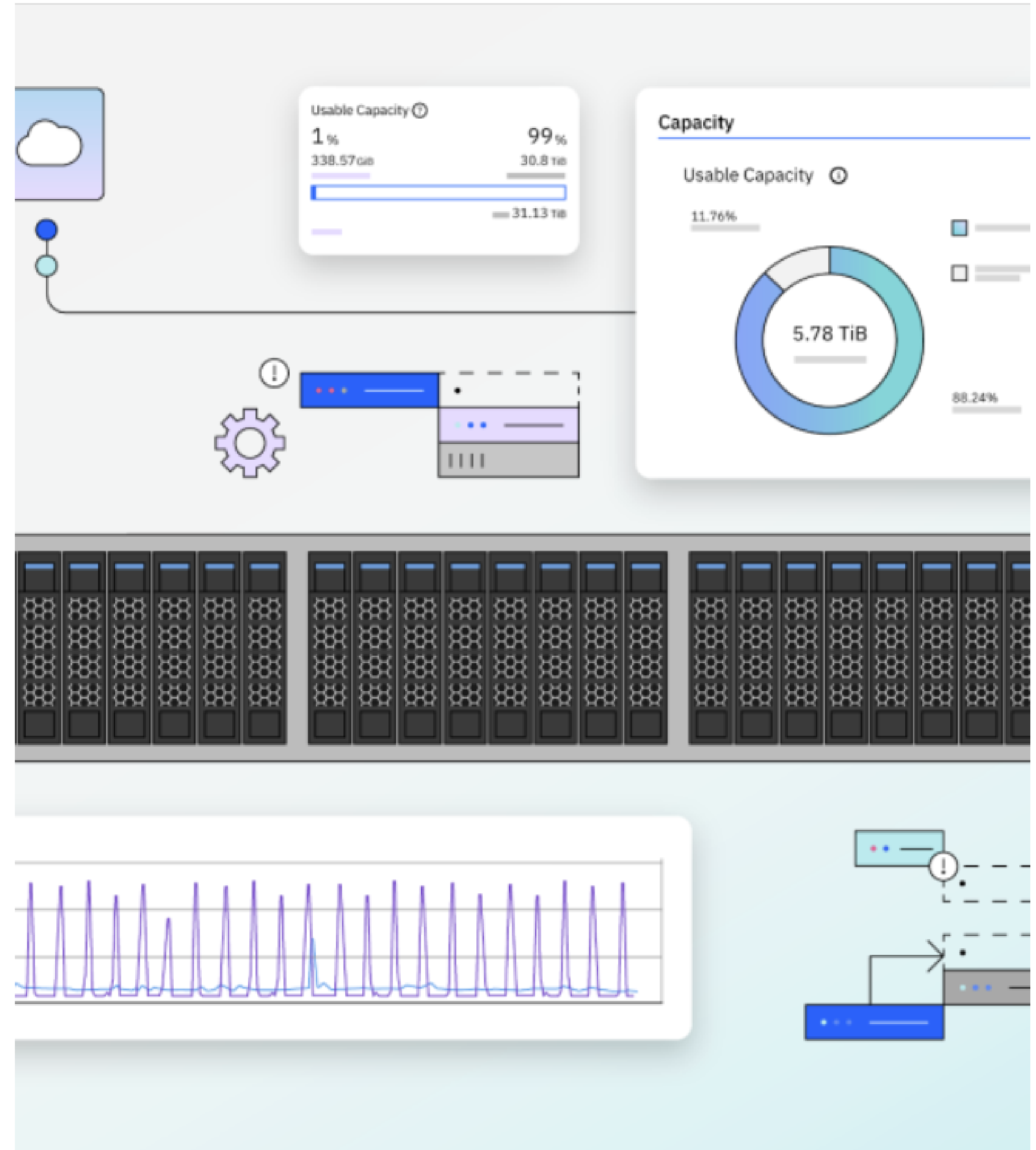
Show do not Tell !



IBM FlashSystems

Policy-based High Availability with Async DR

8.7.1



Les répliquions n'ont n'a jamais été aussi facile

Configurez HA, DR ou HA+DR
dès l'installation

Ou commencez par HA ou DR
sur 2 sites aujourd'hui et
ajoutez un 3ème site à l'avenir

Configuration guidée étape par
étape

IBM FlashSystem 5200

pine-c

Storage partition: myPartition

myPart

Volume

Volume

Volume

Hosts

Replica

Snapsh

Exit

Configure policy-based replication for myPartition

Complete the steps to configure replication for the storage partition

Select replication type

Configure replication

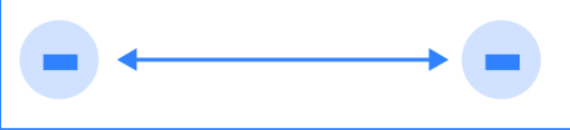
Summary Review

Select replication type

Select the replication types to be used by this storage partition.

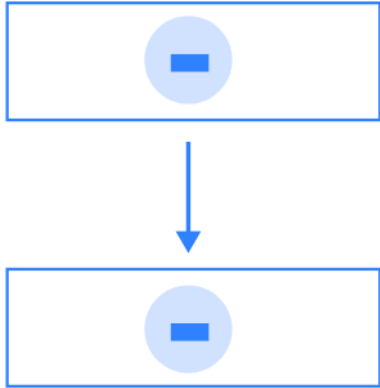
High availability

Active/active high availability between two independent storage systems.



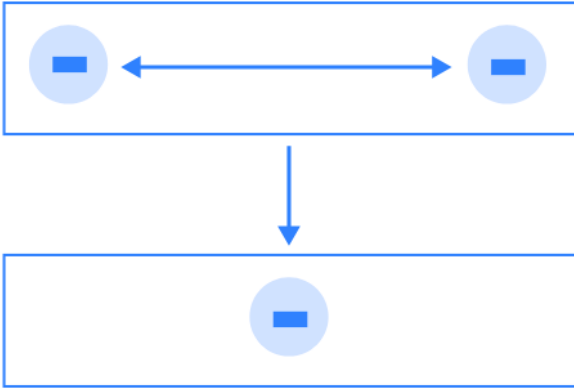
Disaster recovery

Asynchronous replication to another storage system for disaster recovery.



High availability and disaster recovery

Active/active high availability between two independent storage systems, with asynchronous replication to a third system for disaster recovery.



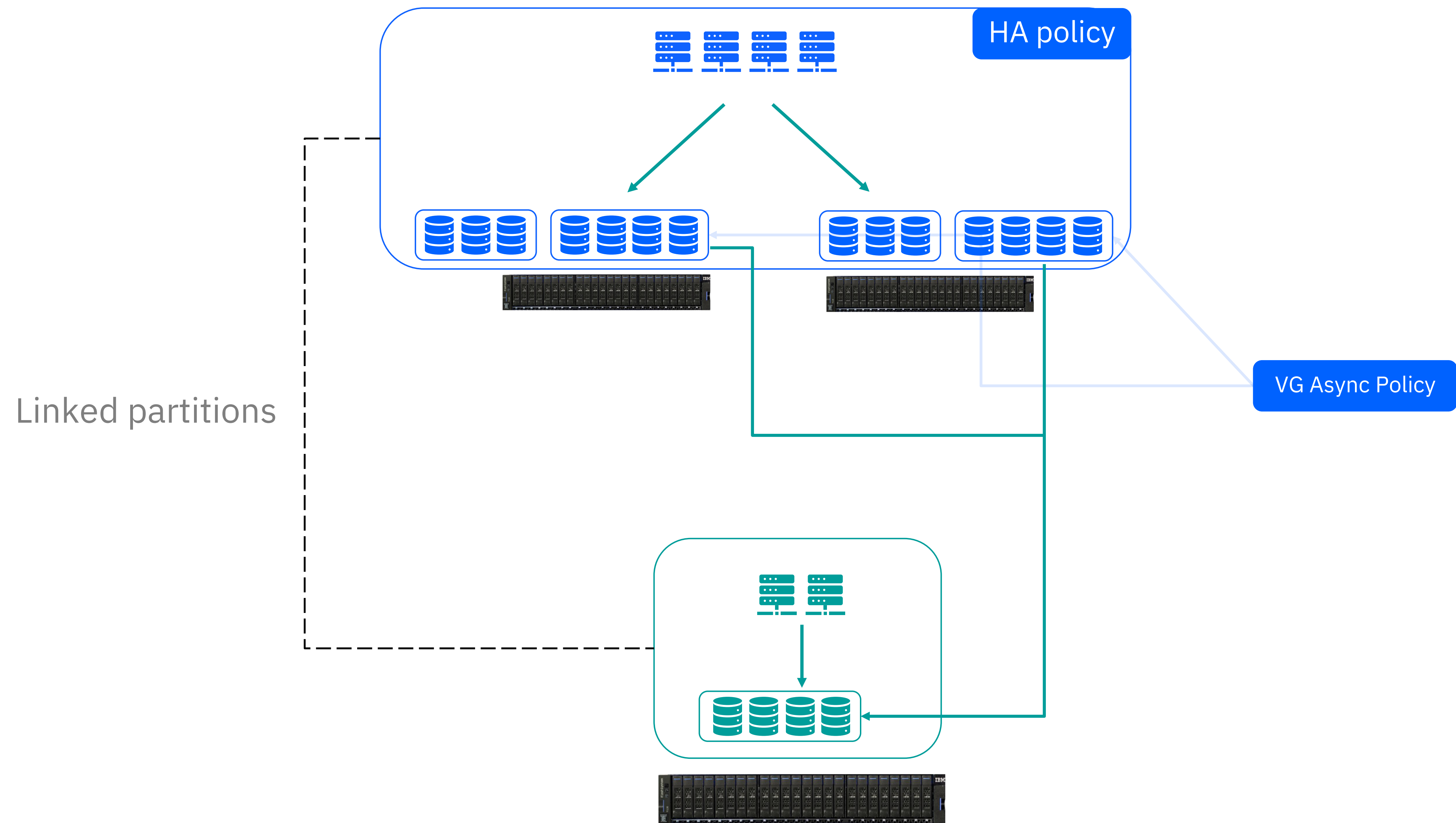
Cancel

Skip

Continue

30

HA Storage Partition With Async DR



Flexibilité
d'utiliser
différentes
politiques de
reprise après
sinistre ou aucune
politique pour
répondre à
différentes
exigences

IBM FlashSystem 5200

pine-c

Storage partition: myPartition - Volume groups

myPartition

Volume groups

Volumes

Volume mappings

Hosts

Replication policies

Snapshot policies

IP quorum

Partnerships

Exit myPartition

Volume Groups

Search table...

Create Volume Group

Name	Replication State	Volume Count	Replication Policy
appVG1	Running	10	Gold Tier DR
appVG2	Running	5	Gold Tier DR
appVG3	Running	12	Silver Tier DR
appVG4	Running	5	Silver Tier DR
demoVG	Running	0	Silver Tier DR
dev and test	-	14	-

Items per page: 10

1-6 of 6 items

1 of 1 page

Latency 0 ms

Read 0 ms

Write 0 ms

Bandwidth 0 MBps

Read 0 MBps

Write 0 MBps

IOPS 0

Read 0

Write 0