

Université **IBM i**

19 et 20 novembre 2024

IBM Innovation Studio Paris

S26 – Disponibilité IBM i : Quelle solution de résilience choisir ?

19 novembre 16:00 - 17:00

Nicolas FRAYSSE

CAPIDP

fraysse@capidp.com



uui2024

#ibmi

#uui2024

The classic IBM logo, consisting of the letters 'IBM' in a bold, sans-serif font with horizontal stripes.The logo for 'common FRANCE', with 'common' in a stylized, lowercase font and 'FRANCE' in a smaller, uppercase font below it.

Sommaire

- La résilience : combien de 9 ?
- LUN-level switching
- Full System HyperSwap
- FlashCopy : vue d'ensemble et exemples de configuration
- Aperçu : IBM i Mirroring
- Aperçu : Metro Mirror
- Aperçu : Global Mirror
- Aperçu : Haute disponibilité avec VM Recovery Manager
- Aperçu : DB2 Mirror for i



La résilience : combien de 9 ?

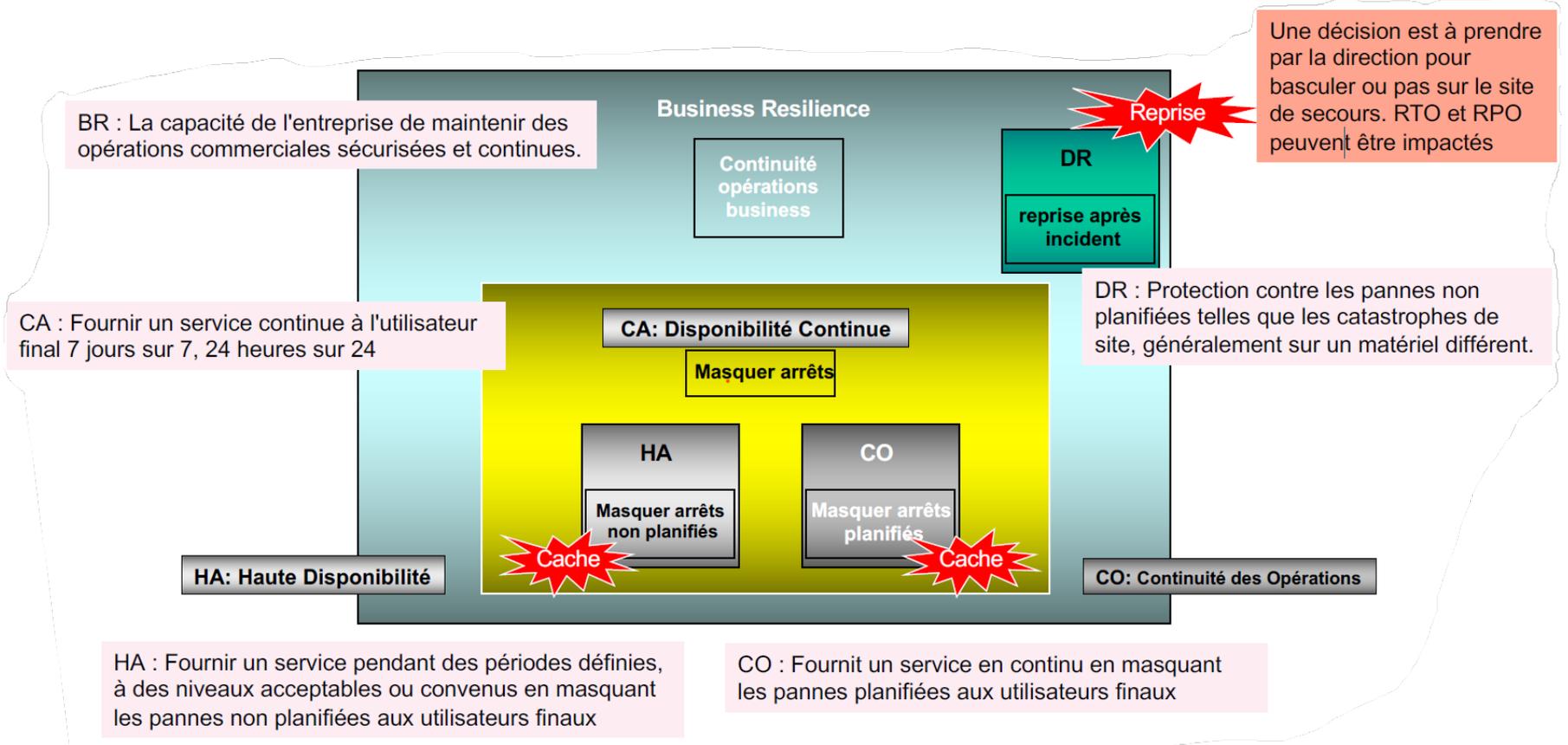
Enquête ITIC 2023

- L'enquête 2023 Global Server Hardware, Server OS Reliability de l'ITIC a également révélé qu'une majorité de 88 % des utilisateurs des derniers serveurs IBM Power10 (livrés depuis septembre 2021) déclarent que leurs organisations ont atteint un temps de disponibilité de huit neuf - 99,999999 %.
- Cela représente 3,15 millisecondes de temps d'arrêt non planifié, par serveur et par an, dû à des défauts du système sous-jacent ou à des défaillances de composants.
- Ainsi, les entreprises ayant un serveur IBM Power10 ne dépensent que 7,18 dollars par serveur et par an pour remédier aux pannes de serveur non planifiées.

Reliability %	Downtime per year	Downtime per month	Downtime per week
90% (one nine)	36.5 days	72 hours	16.8 hours
95%	18.25 days	36 hours	8.4 hours
97%	10.96 days	21.6 hours	5.04 hours
98%	7.30 days	14.4 hours	3.36 hours
99% (two nines)	3.65 days	7.20 hours	1.68 hours
99.5%	1.83 days	3.60 hours	50.4 minutes
99.8%	17.52 hours	86.23 minutes	20.16 minutes
99.9% (three nines)	8.76 hours	43.8 minutes	10.1 minutes
99.95%	4.38 hours	21.56 minutes	5.04 minutes
99.99% (four nines)	52.56 minutes	4.32 minutes	1.01 minutes
99.999% (five nines)	5.26 minutes	25.9 seconds	6.05 seconds
99.9999% (six nines)	31.5 seconds	2.59 seconds	0.605 seconds
99.99999% (seven nines)	3.15 seconds	0.259 seconds	0.0605 seconds
99.999999% (eight nines)	3.15 milliseconds	26.25 milliseconds	6.05 milliseconds
99.9999999% (nine nines)	31.56 milliseconds	2.63 milliseconds	606.9 microseconds

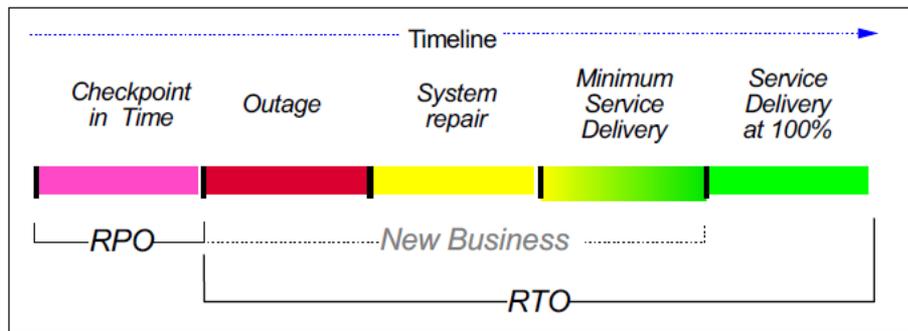
<https://www.ibm.com/downloads/cas/EVMWVOE9>

Définitions standard de l'industrie pour la résilience

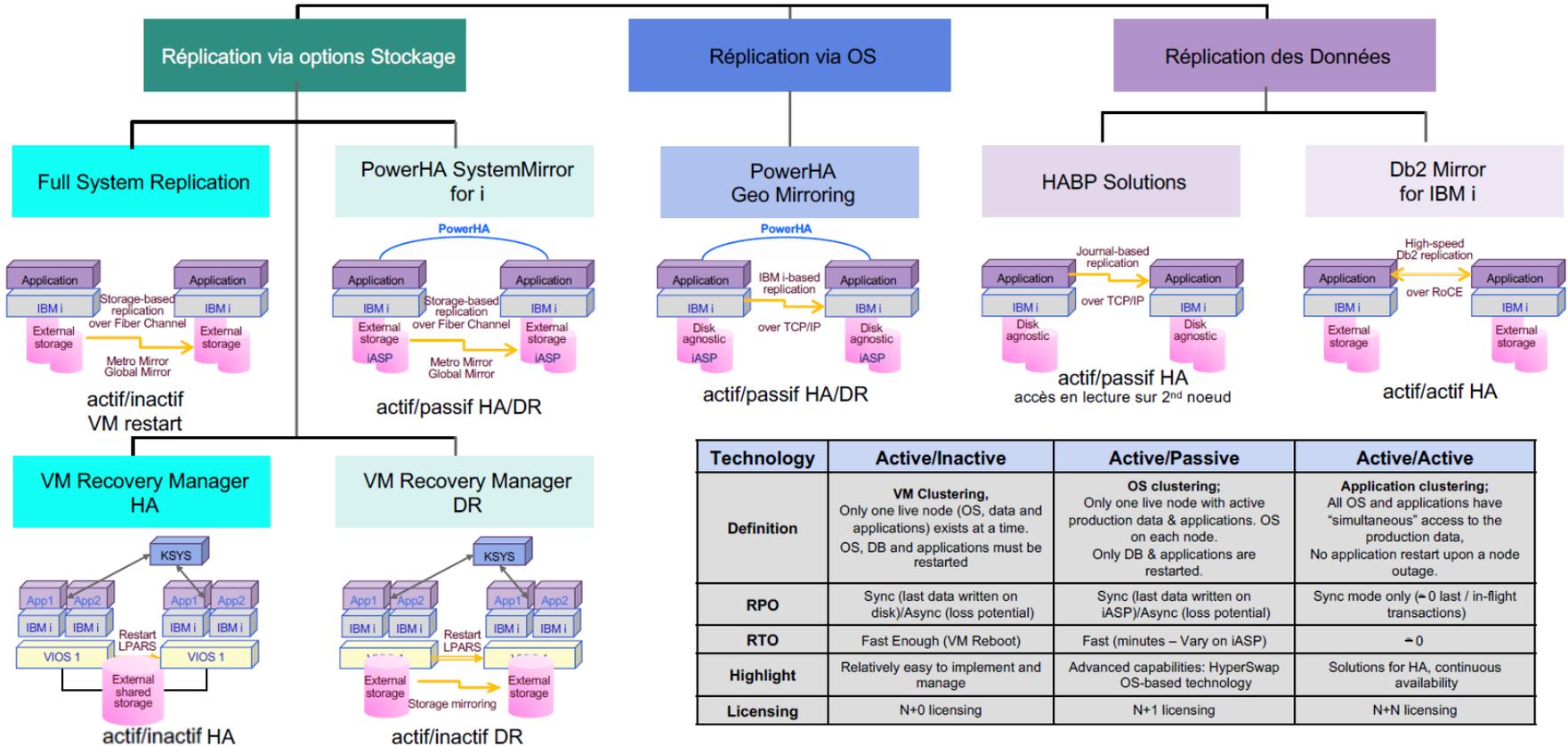


Quels objectifs ?

- **Assurer une plus grande disponibilité des serveurs Power Systems IBM i face à différents types de situation :**
 - Arrêts planifiés : par exemple une mise à jour système d'exploitation ou logiciels...
 - Arrêts non planifiés : par exemple, une défaillance matérielle.
 - Réduire les fenêtres de sauvegarde : par exemple, sauvegarde sur base de données répliquées ou via snapshot.
 - Sinistres : par exemple, une perte d'un Data Center, coupure générale d'alimentation électrique,...
- **Déterminer les objectifs de disponibilité de votre système d'informations basé sur Power i :**
 - Recovery Time Objective – RTO : Temps maximum supportable pour remettre à disposition des utilisateurs un environnement applicatif.
 - Recovery Point Objective – RPO : Niveau acceptable de perte de données.



IBM i et principales solutions HA/DR

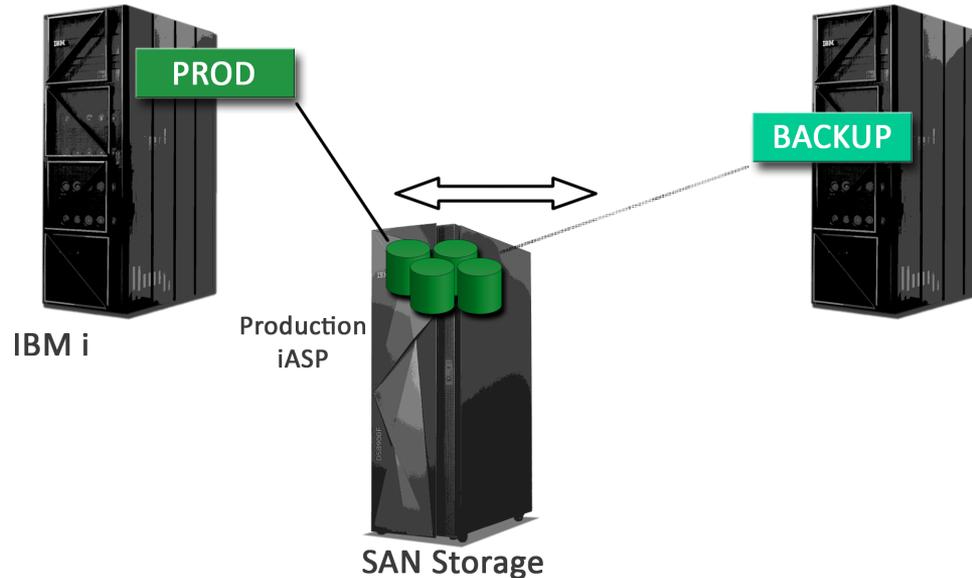




LUN-level switching

Fonction LUN-level switching

- Deux partitions IBM i ou plus (même CEC physique ou différent).
- Un ensemble de disques/LUN définissant le pool de stockage.
- La fonction est une caractéristique des sous-systèmes de stockage externes.



Résumé du LUN-level switching

Avantages

- Solution de haute disponibilité (HA)
- Impossible de déplacer les ressources en cas de défaillance de la production
- Basculement automatisé grâce à Power HA (depuis la V7.1)
- Géré par PowerHA (DS8K, SVC, Storwize) ou ICSM (DS8000 uniquement)
- Quantité limitée de disques nécessaires

Limitations

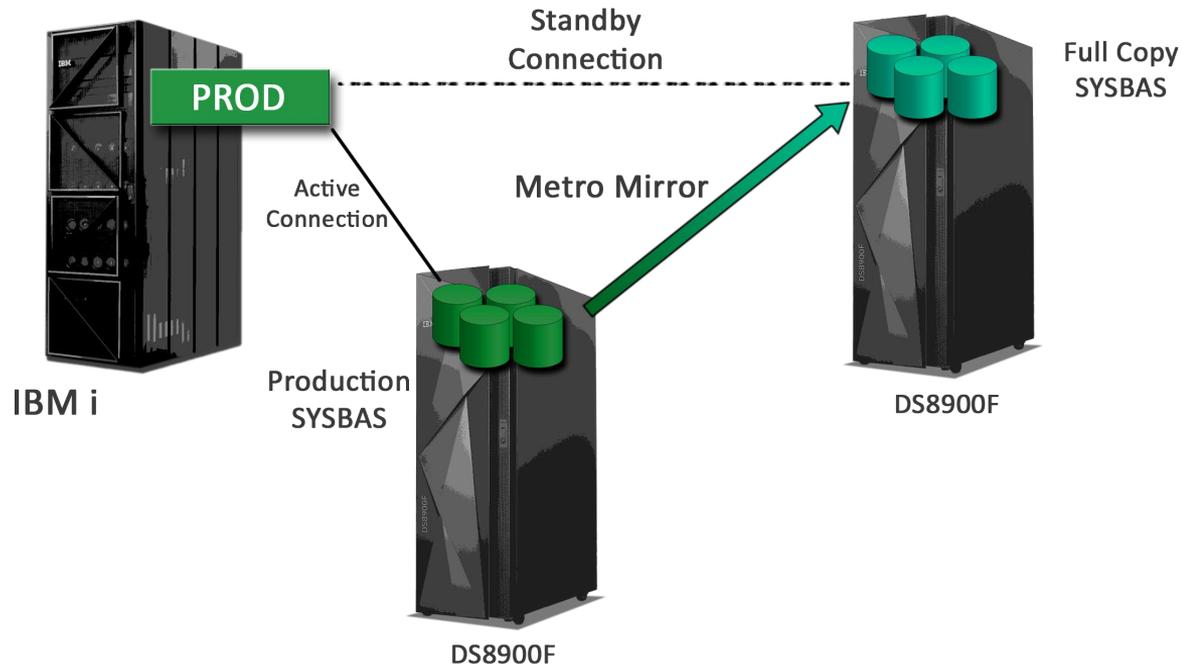
- Copie unique des données
- En fonction du RAID/de la parité utilisé(e), un nombre suffisant d'erreurs de disque peut entraîner la perte de l'ensemble de l'iASP.
- Ne constitue pas une solution de reprise après sinistre (DR)
- Limité aux DS8000, SVC, Storwize



Full System HyperSwap

Fonction HyperSwap

- Une partition IBM i.
- Deux ensembles de disques/LUN, copies miroir l'un de l'autre à partir de Metro Mirror.



Résumé de l'HyperSwap

Avantages

- Solution système complète, qui n'est pas limitée aux environnements iASP
- IBM i 7.2 PowerHA Express edition intègre la prise en charge d'HyperSwap
- Solution de haute disponibilité (HA) pour la redondance du stockage
- Deux copies synchronisées des données stockées sur disque
- Peut être géré par des commandes IBM i dans SST
- La réplication est inversée automatiquement en cas de permutation planifiée.
- La permutation planifiée ou le basculement peuvent se produire très rapidement (<1 seconde).

Limitations

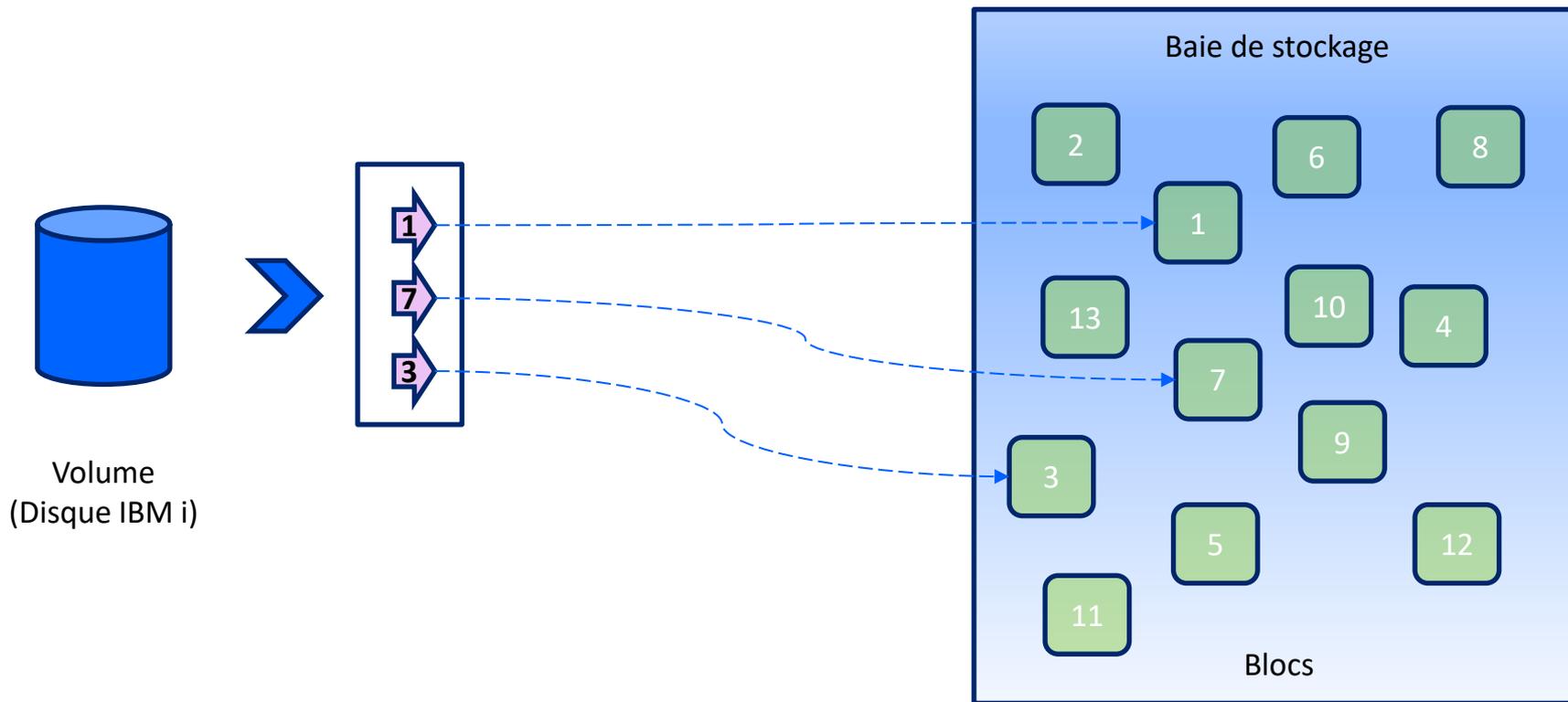
- Pas de solution de reprise après sinistre (DR)
- En cas de sinistre généralisé sur un site, il faudrait au moins qu'un serveur hors site (et le DS de secours) se trouve à un autre endroit.



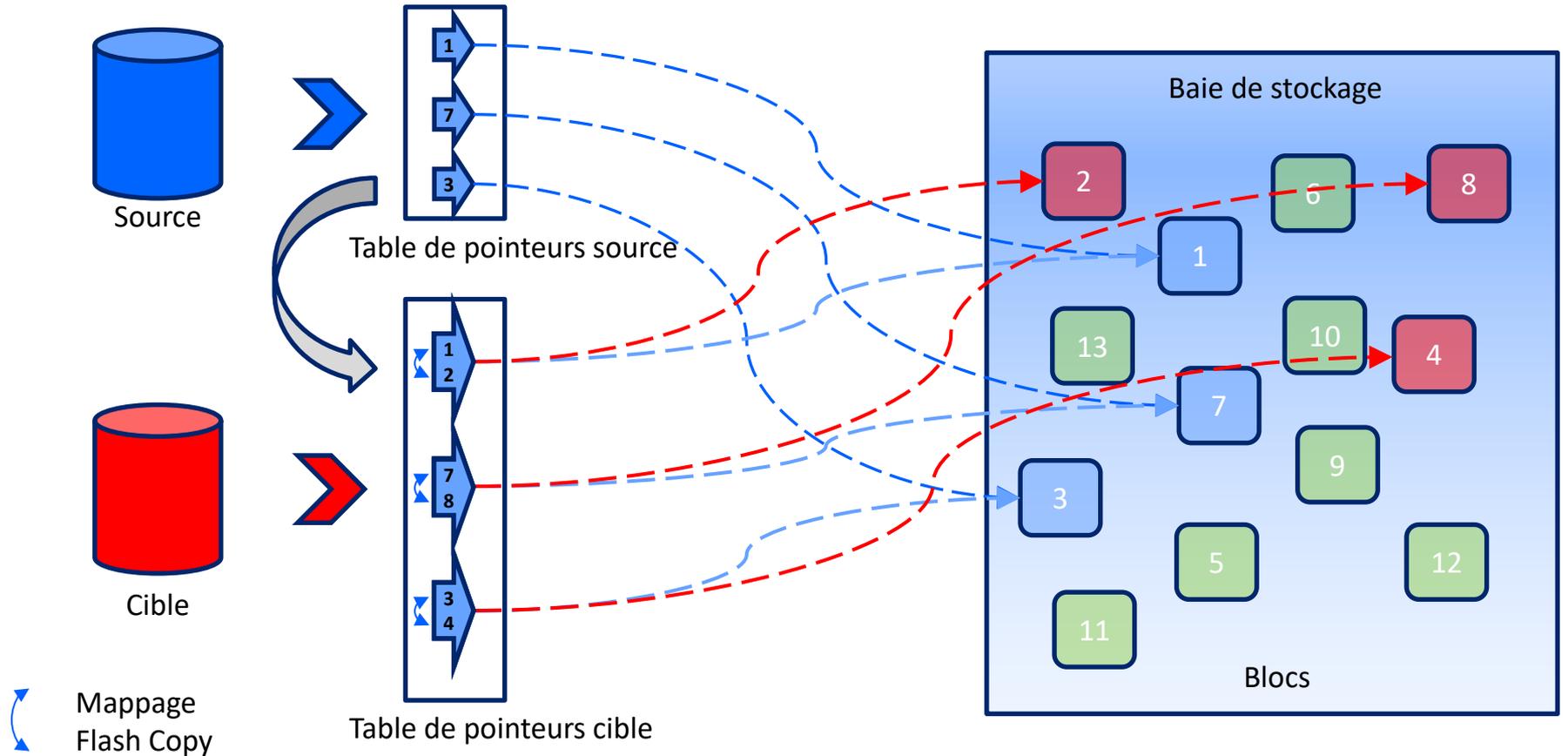
FlashCopy : vue d'ensemble et exemples de configuration

Fonctionnement FlashCopy

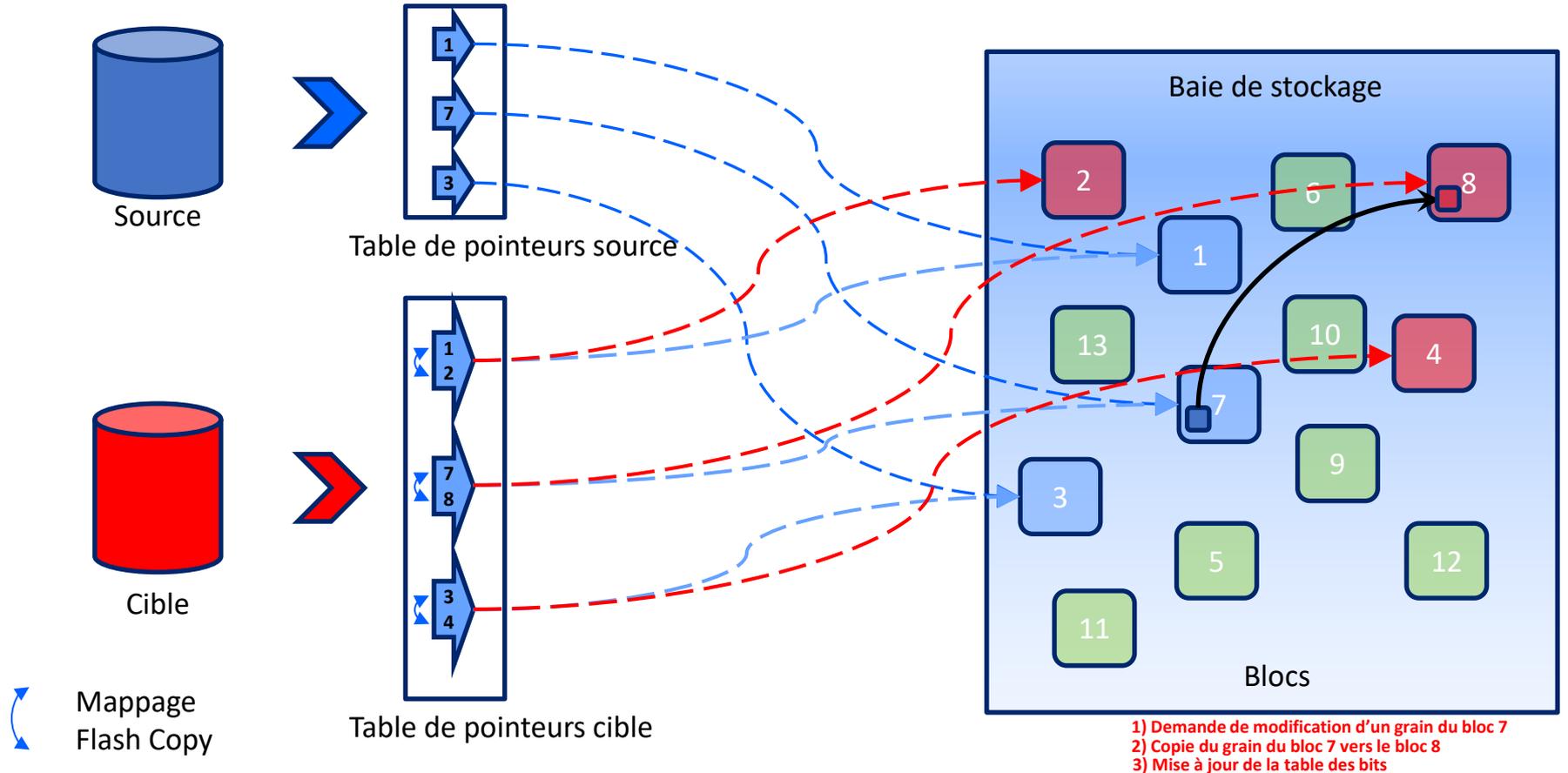
Le volume est associé à une table de pointeurs



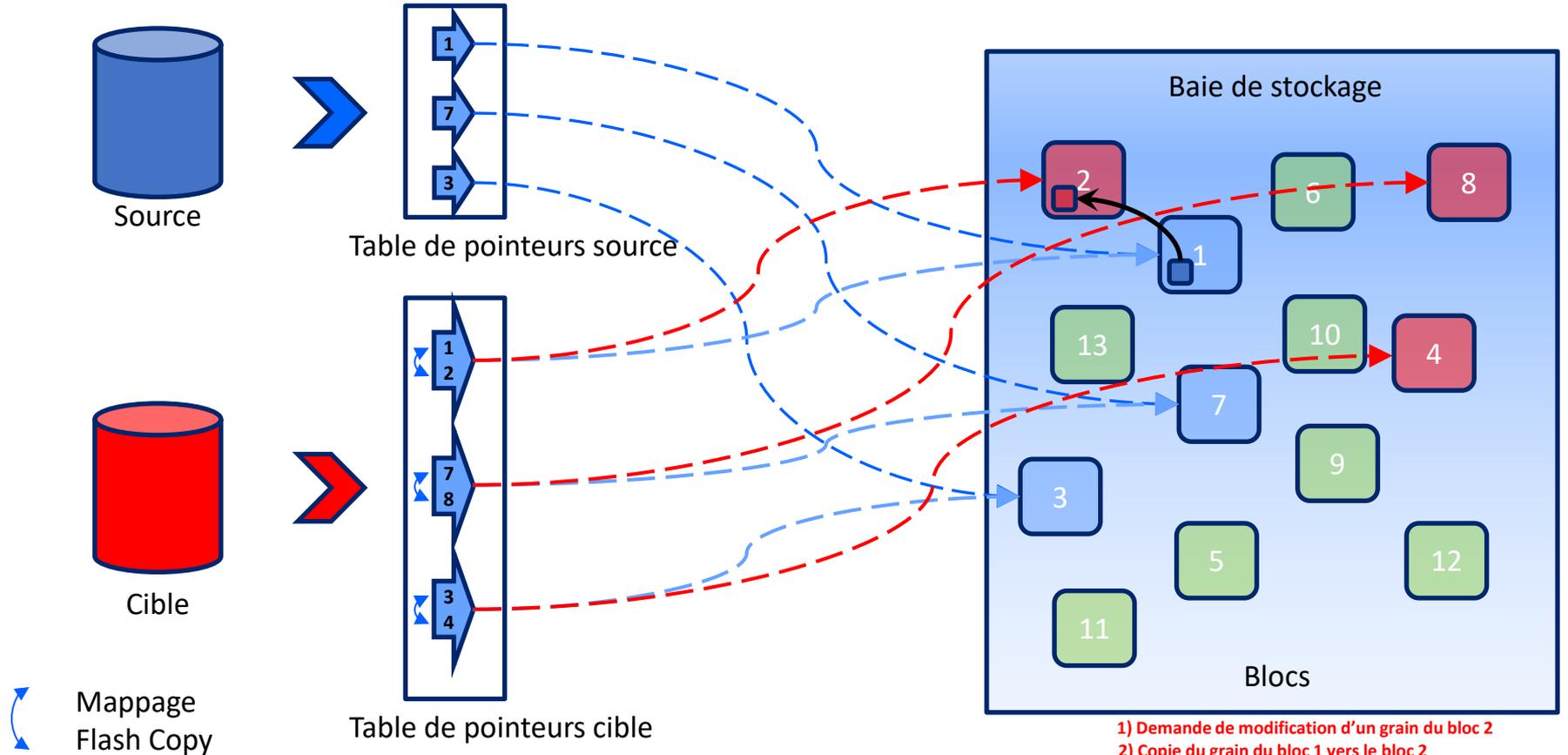
FlashCopy d'un volume : copie des pointeurs



Modification des données du volume source



Modification des données du volume cible



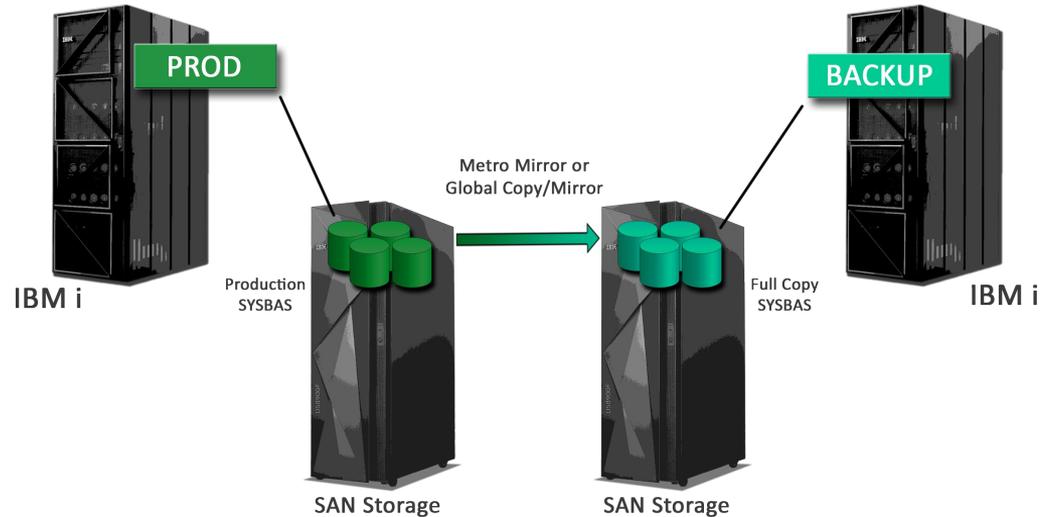
- 1) Demande de modification d'un grain du bloc 2
- 2) Copie du grain du bloc 1 vers le bloc 2
- 3) Mise à jour de la table des bits

Rappel : PowerHA SystemMirror

- Solutions complètes IBM Power Systems intégrées de bout en bout pour HA DR
 - Focus : disponibilité des applications 24 x 7 en cas de panne planifiée ou non planifiée
 - Extension profondément intégrée d'IBM i (implémentée dans le LIC et le système d'exploitation)
- Technologie de clustering
 - Fournit aux applications une infrastructure de résilience complète
 - Contrôle et gestion des ressources primaires et secondaires pour les opérations de HA et de DR
- Résilience des données basée sur le stockage
 - La résilience des données est une extension de l'architecture de gestion du stockage du système hôte.
 - Les volumes de stockage sont commutables ou mis en miroir entre les nœuds du cluster.
 - Services de réplication basés sur le matériel pour les opérations multisites
 - Réplication basée sur l'hôte (mise en miroir géographique pour IBM i)
 - Réplication basée sur le stockage (Metro Mirror ou Global Mirror)
- Caractéristiques de la solution globale
 - Automatisation, implication minimale des opérations informatiques
 - Les données entre les nœuds primaires et secondaires sont toujours synchronisées et toujours prêtes pour un événement de basculement.

Rappel : Full System Replication

- Deux partitions IBM i (différents « sites »)
- Un « site de production » et un ou deux ensembles de disques/LUN pour la copie miroir et (si nécessaire) la copie du groupe de cohérence.
- Le système d'exploitation cible est hors ligne, un IPL est nécessaire pour le basculement.
- Nécessite un stockage externe pour la réplication
- Nécessite un cluster PowerHA

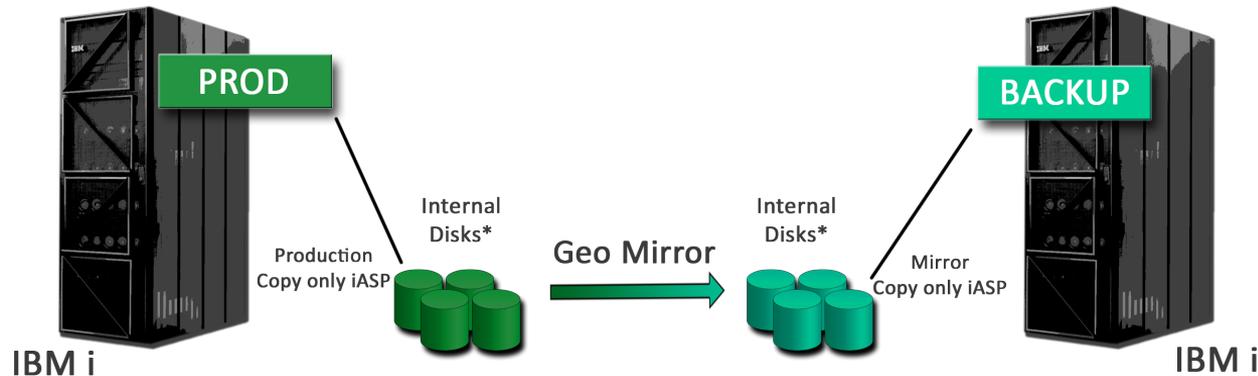




Aperçu : Geographic Mirroring

Fonction Geographic Mirroring

- Deux partitions IBM i (différents « sites »)
- Un iASP, deux copies (deux jeux de disques/LUN)
- Typiquement pour le stockage interne
- La réplication est gérée au niveau de la gestion du stockage IBM i SLIC.



Résumé du Geographic Mirroring

Avantages

- Solution de haute disponibilité (HA) et de reprise après sinistre (DR)
- La réplication est gérée au niveau de la gestion du stockage SLIC, par opposition au système d'exploitation et à la journalisation à distance (par exemple, les solutions de réplication logicielle).
- Gérée à l'aide de PowerHA
- Véritable option de transmission asynchrone sur IBM i

Limitations

- Nécessite une bande passante suffisante entre les sites pour maintenir des copies cohérentes et éviter les problèmes de suspension automatique.
- Il est recommandé d'avoir la même quantité, la même capacité et le même type de disques sur les deux sites pour maintenir la cohérence.
- Les ports de réplication utilisent des ports TCP éphémères aléatoires (1024+) et non des ports spécifiques.



Aperçu : Metro/Global Mirror

Fonction Metro/Global Mirror (MM/GM)

Avantages

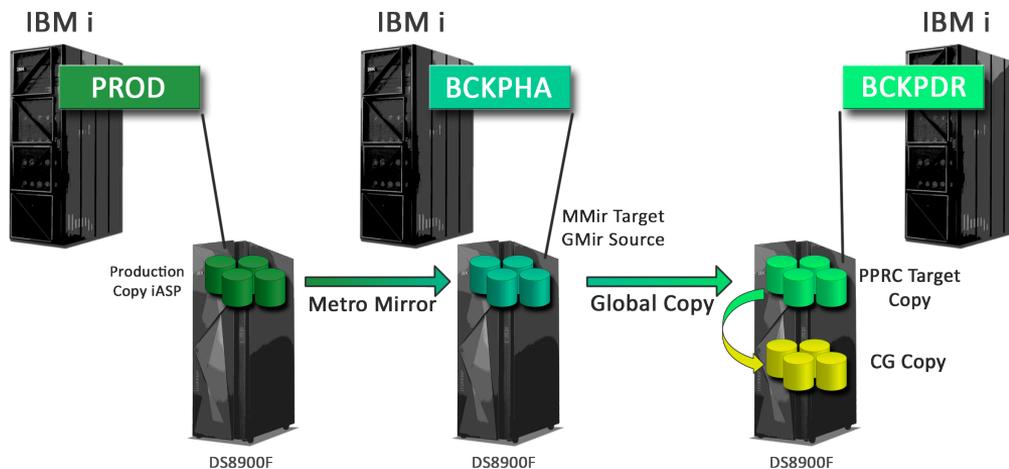
- Solution de haute disponibilité (HA) et de reprise après sinistre (DR)
- Vérification automatique de l'état de préparation à la FlashCopy Émission d'une FlashCopy
- Variable sur l'iASP jusqu'au nœud de FlashCopy de sauvegarde
- Facilité d'utilisation des utilitaires de gestion sur plusieurs nœuds du cluster PowerHA
- Complète et améliore les fonctionnalités d'IBM i PowerHA
- Intégration et automatisation des fonctions de sauvegarde du BRMS avec l'iASP
- Permet des améliorations personnalisées des environnements FlashCopy (flash inversé, flash à partir de HA).

Limitations

- Nécessite PowerHA, cluster
- Requiert un stockage externe
- Réplication iASP uniquement
 - données SYSBAS partiellement répliquées par « Administrative Domain » du cluster

Fonction Metro/Global Mirror (MM/GM)

- Trois partitions IBM i (trois « sites »)
- Un iASP, trois copies complètes, plus un ensemble supplémentaire de disques/LUN pour le Consistency Group (CG) FlashCopy.
- Nécessite un stockage externe DS8000 et TPC-R



Résumé des répliquions Metro/Global Mirror (MM/GM)

Avantages

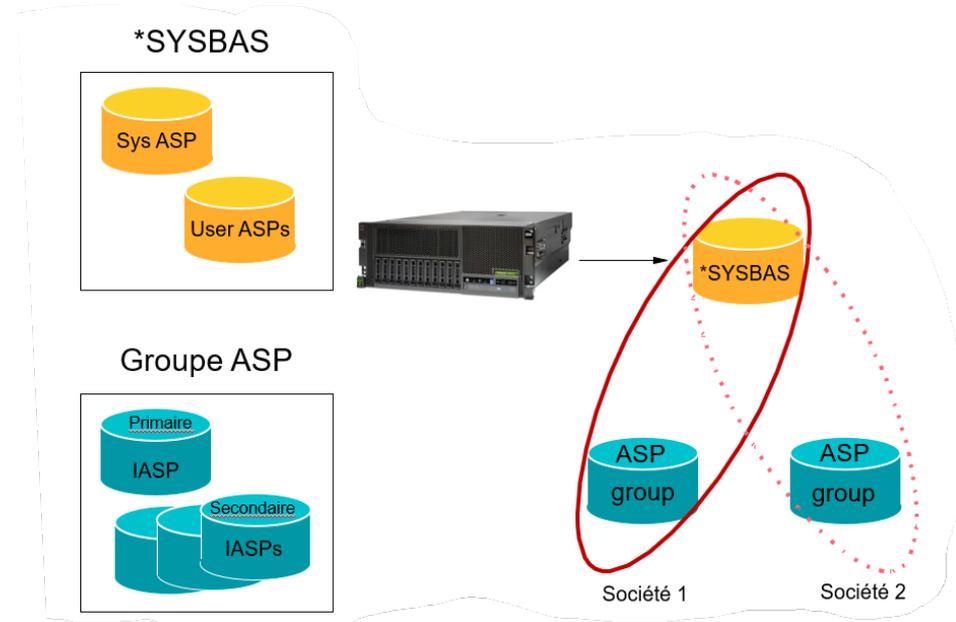
- Solution de haute disponibilité (HA) et de reprise après sinistre (DR)
- Le site local dispose d'une copie synchronisée et cohérente des données, tandis que le site distant fournit une solution de reprise après sinistre presque cohérente.
- basculement et reprise rapides
- Rétablissement rapide de la mise en miroir sur 3 sites sans interruption de la production
- Possibilité de tester le DR sur un site distant, tout en conservant une solution locale de DR
- L'ICSM peut gérer l'environnement pour faciliter l'utilisation quotidienne.

Limites

- Nécessite un stockage supplémentaire et des jeux de disques.
- Nécessite TPC-R pour la gestion
- Nécessite un stockage DS8000

Rappel : *SYSBAS & iASP

- Plusieurs iASP peuvent exister sur le même système. Chaque iASP primaire et toutes les iASP secondaires associées constituent une base de données distincte.
 - Étant donné que les tâches ou les threads ne peuvent avoir accès qu'à un seul groupe ASP à la fois, la duplication des noms de bibliothèques est autorisée entre les différents groupes ASP.
 - Étant donné que TOUS les travaux du système ont une visibilité sur les ASP 1 à 32, les noms de bibliothèques en double ne sont pas autorisés entre les iASP du système et le SYSBAS (ASP 1 à 32).
- Les iASP doivent être activées pour pouvoir être utilisées. Le système n'a aucune information sur ce qui se trouve sur l'iASP lorsqu'il est désactivé.
- Il doit être explicitement varié, ce qui est généralement fait dans le programme de démarrage.
- Le système monte également l'IFS pour l'iASP à un point de montage sur le système de fichiers racine qui porte le même nom que la description du périphérique : /<i>iASPNAME</i>





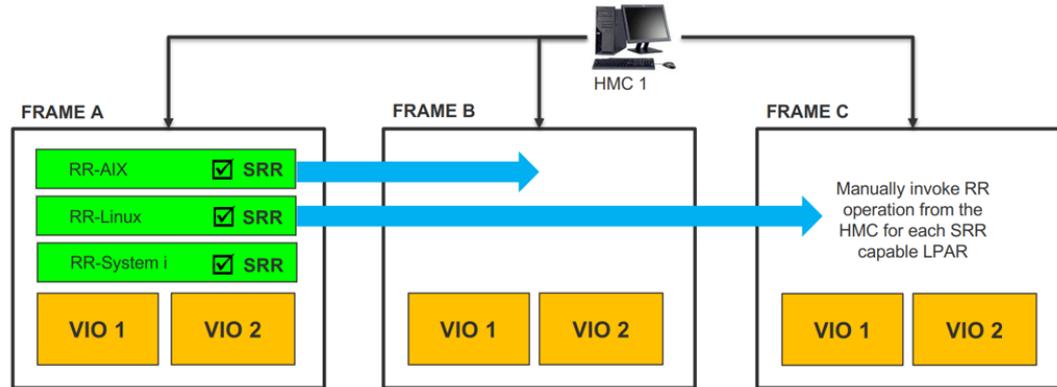
Aperçu : Haute disponibilité avec VM Recovery

Fonction VM Recovery

- VM Recovery est une variante améliorée de la réplication matérielle basée sur un SAN. Elle utilise la même technologie de réplication, mais intègre également un basculement automatisé via VM Recovery Manager.
- Cette solution est disponible en deux versions :
 - VM Recovery HA (site unique, stockage unique)
 - VM Recovery DR (multi-site, stockage multiple).
- Le schéma de fonctionnement de la réplication (DR) est similaire à celui de la réplication matérielle basée sur un SAN.
- L'automatisation des tâches est gérée par une VM appelée KSYS, qui est un AIX.

SRR : Simplified Remote Restart

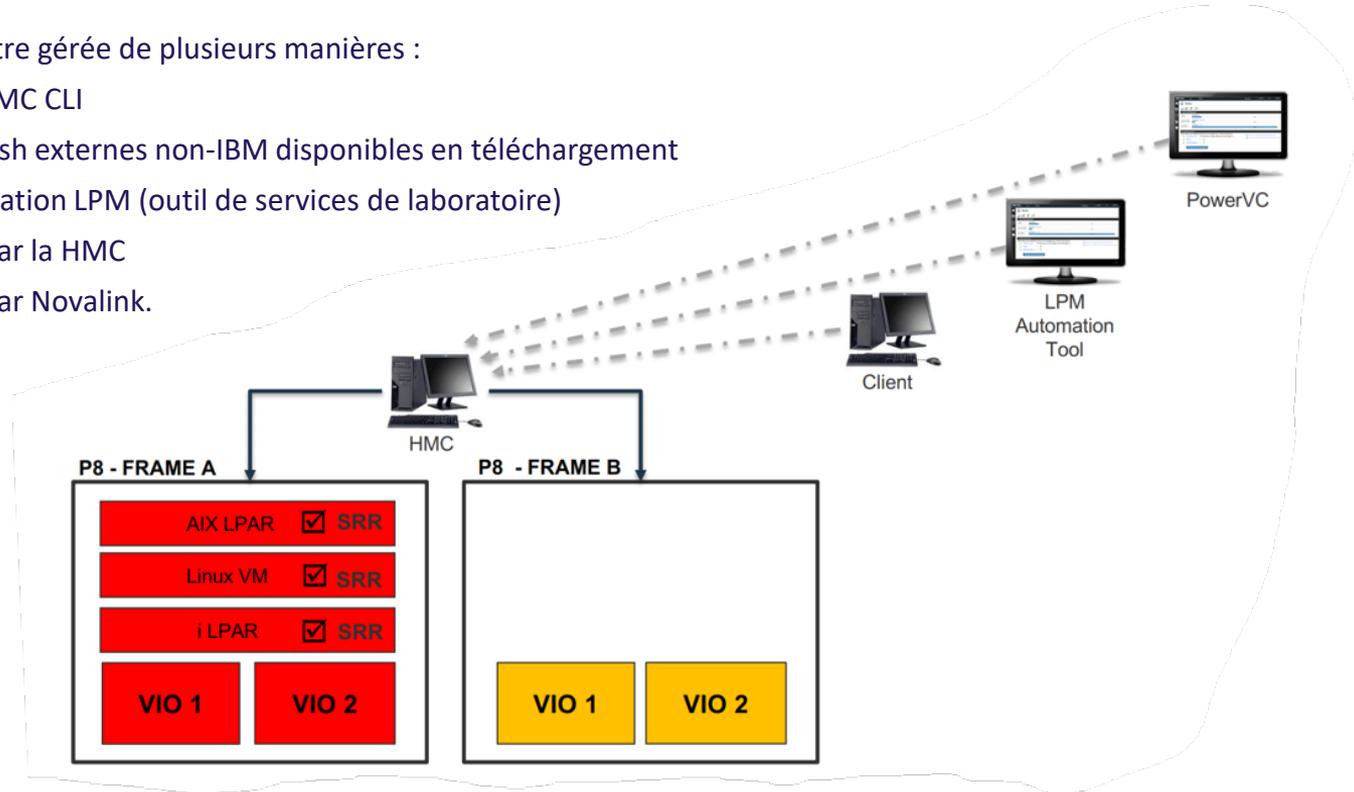
- **Qu'est-ce que c'est ?**
 - Méthode pour redémarrer des LPARs ailleurs si un serveur entier tombe en panne.
 - Disponible depuis les serveurs P8 avec PowerVM Enterprise Edition.
 - Permet le redémarrage d'une VM/Partition exécutant n'importe quel OS sur Power Systems.
- **Différence avec LPM :**
 - Les serveurs VIO ne sont pas disponibles
 - Le niveau de code HMC dictera le niveau de fonctionnalité.
 - L'utilisateur doit invoquer « manuellement » les commandes de redémarrage à distance.



SRR : Gestion et facilité d'utilisation

La fonction SRR peut être gérée de plusieurs manières :

- manuel ssh via HMC CLI
- scripts rrmonitor.sh externes non-IBM disponibles en téléchargement
- Outil d'automatisation LPM (outil de services de laboratoire)
- PowerVC - géré par la HMC
- PowerVC - géré par Novalink.



PowerVC : Redémarrage à distance automatisé (ARR)



PowerVC

- Le redémarrage à distance automatisé doit être activé au niveau du groupe d'hôtes.
- L'administrateur doit spécifier l'intervalle d'exécution et la période de stabilisation pour le moniteur.
 - Ajuster la sensibilité du redémarrage à distance automatisé, c'est-à-dire la durée pendant laquelle un hôte doit être « hors service » avant que l'automatisation ne soit déclenchée.
- Dans le cadre du redémarrage à distance automatisé, une VM d'un hôte en panne peut être redémarrée à distance sur un hôte (dans le même groupe d'hôtes) qui n'a pas activé le redémarrage à distance automatisé.

IBM PowerVC Configuration Messages DRO Events

Hosts > Create Host Group

Create Host Group

Host groups allow you to logically partition your hosts for placement targeting.

[Learn more about host groups](#)

6

Host group name:
AutoRR-Group

7

Placement policy:
Striping

Automated Remote Restart

[Learn more about automated remote restart](#)

4

Enable automated remote restart

Run interval (minutes): 1

Stabilization (times): 5

Dynamic Resource Optimizer

[Learn more about DRO settings](#)

Configure Dynamic Resource Optimizer (DRO)

Members

+ Add - Remove

60

Create Host Group Cancel

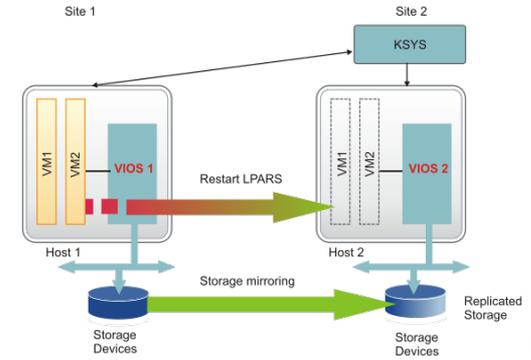
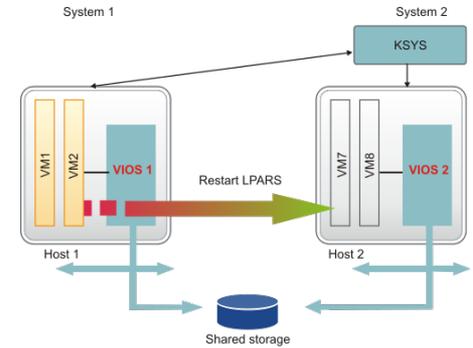
VM Recovery Manager 1.8

VM Recovery Manager for HA (VMR HA) Les machines virtuelles défaillantes sont redémarrées sur un serveur adjacent. Il s'agit d'une topologie de stockage partagé qui est utilisée en conjonction avec LPM.

- Le serveur/la partition secondaire est inactif(ve) jusqu'à ce que les machines virtuelles soient redémarrées (démarrées) ou si LPM est utilisé pour un événement de panne planifiée.
- Il y a une copie physique et une copie logique (sauf si elle est déployée dans une configuration de type HyperSwap, auquel cas il y a deux copies physiques des données).

VM Recovery Manager for HA (VMR DR) est basé sur la réplication des machines virtuelles vers un site distant.

- Le serveur de reprise après sinistre est inactif jusqu'à ce que les machines virtuelles répliquées soient redémarrées.
- Si le système de production tombe en panne (ou doit être testé pour la conformité DR), les VM sont redémarrées sur un système secondaire dans le cluster.
- Dans cette configuration particulière, il y a deux copies physiques des VM et une copie logique



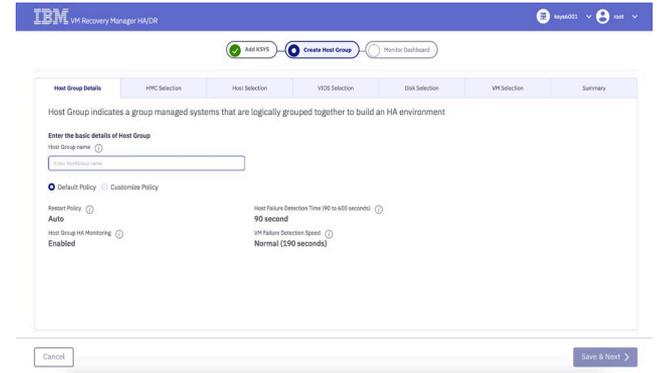
gpr02

gpr02

VM Recovery Manager : Gestion graphique

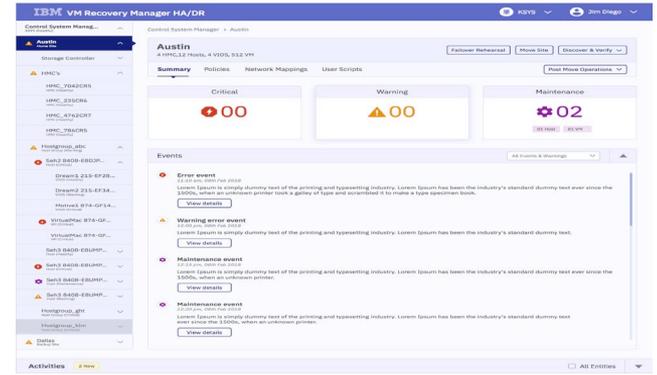
Gestion HA

- Déploiement du groupe d'hôtes en toute simplicité (Assistant)
- Surveillance de la santé de tous les groupes d'hôtes de manière centralisée
- Fonctions administratives à partir de l'interface utilisateur
- LPM une ou plusieurs VMs, libérer ou restaurer l'hôte

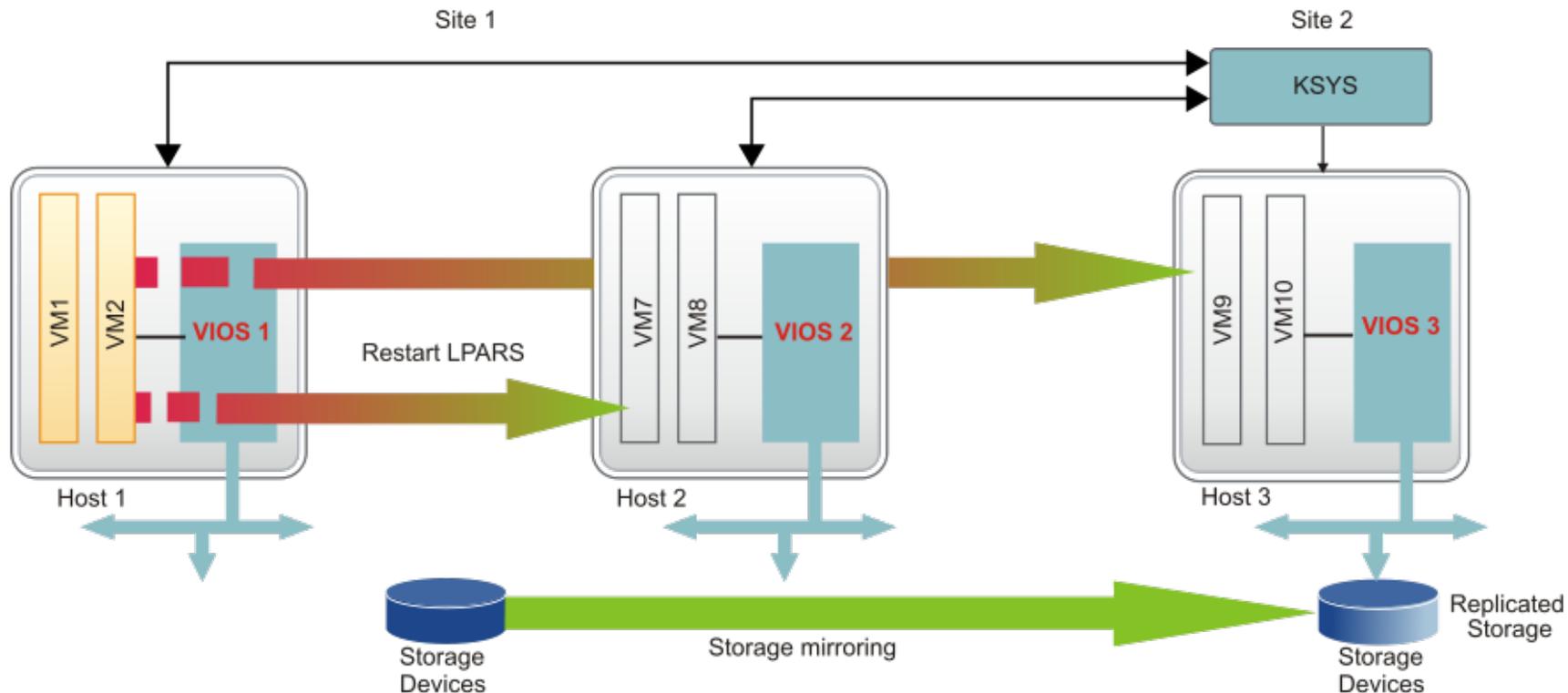


Gestion Disaster Recovery

- Exécution du basculement d'un groupe d'hôtes ou d'un site
- Assistant de répétition du basculement
- Validation de l'environnement



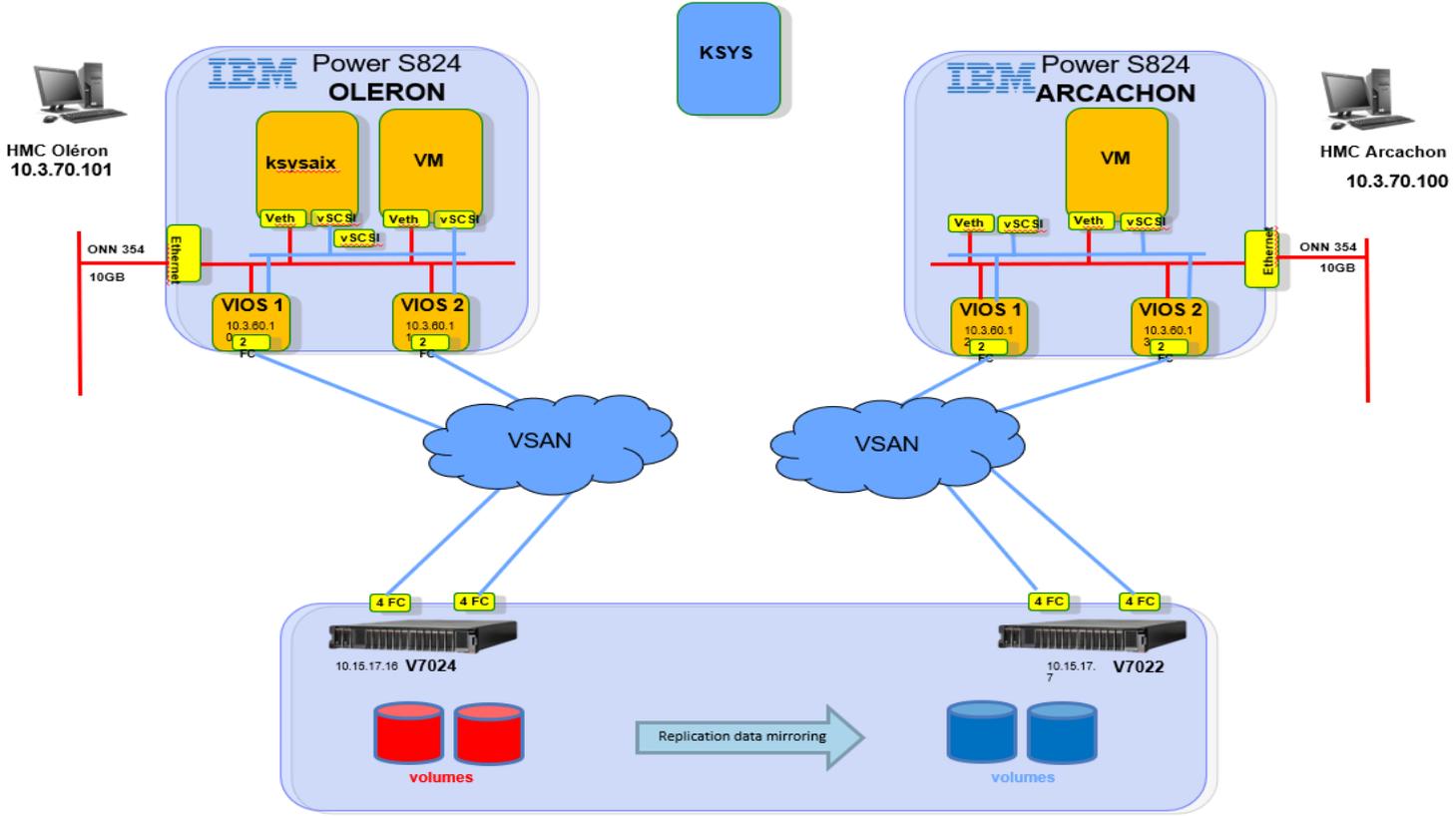
VM Recovery Manager for DR (supporte à la fois le redémarrage HA et DR)



VM Recovery Manager for DR permet de gérer à la fois le HA et le DR à partir d'un seul KSYS.

<https://www.ibm.com/docs/fr/vmrmdr/1.7?topic=overview>

Infrastructure de démonstration VM RECOVERY MANAGER DR



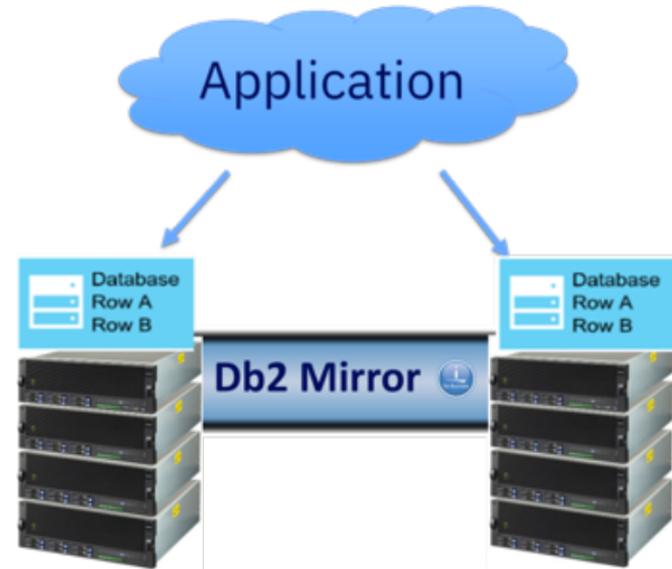


Aperçu : DB2 Mirror for i

Db2 Mirror - Actif/Actif DB2 for i, pour une Disponibilité Continue

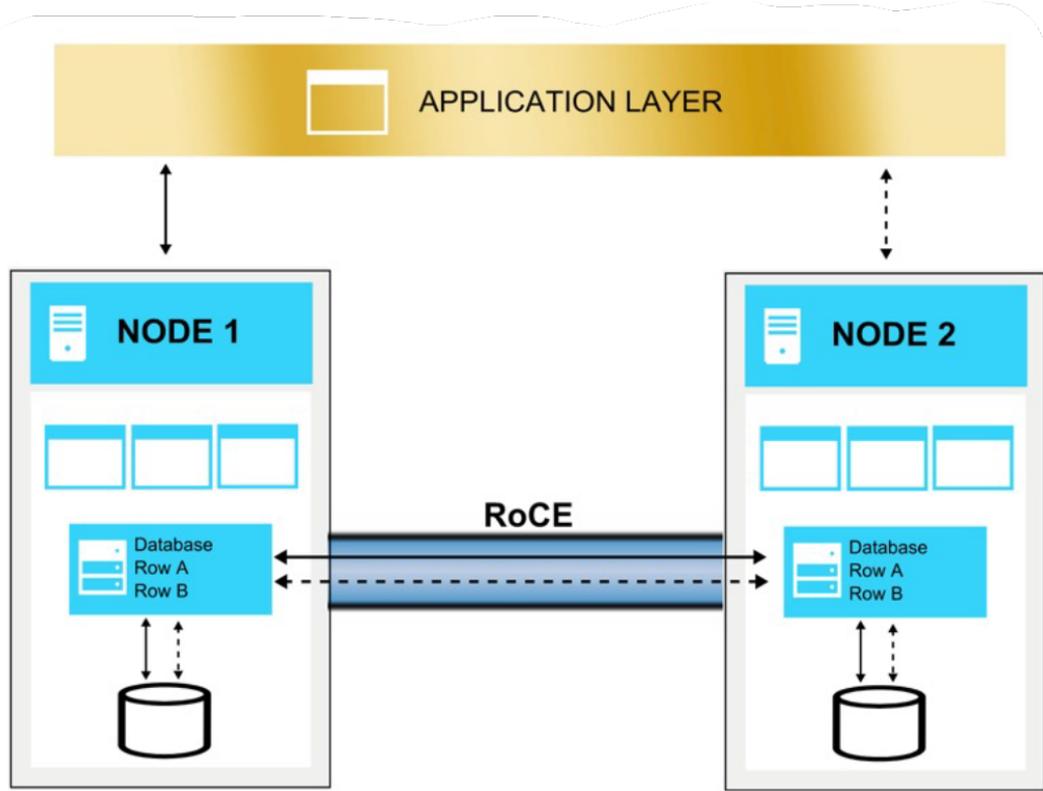
IBM Db2 Mirror for i

- Solution Data Center pour la **Disponibilité Continue**
- Réplication synchrone grande vitesse pour DB2 for i
- **Activation de la disponibilité de l'application**
 - Deux Lpars lisent et écrivent sur les mêmes fichiers DB
 - Déplacement rapide de tous les travaux vers une Lpar, pour maintenance planifiée ou défaillance d'un nœud
- **Mises à niveau du système non bloquante**
 - OS et Serveurs peuvent être de générations différentes
 - Mises à niveau par roulement sans temps d'arrêt
- **Nouvelle technologie intégrée de synchronisation IBM i**
 - N'utilise pas la journalisation DB2 comme certaines technologies existantes

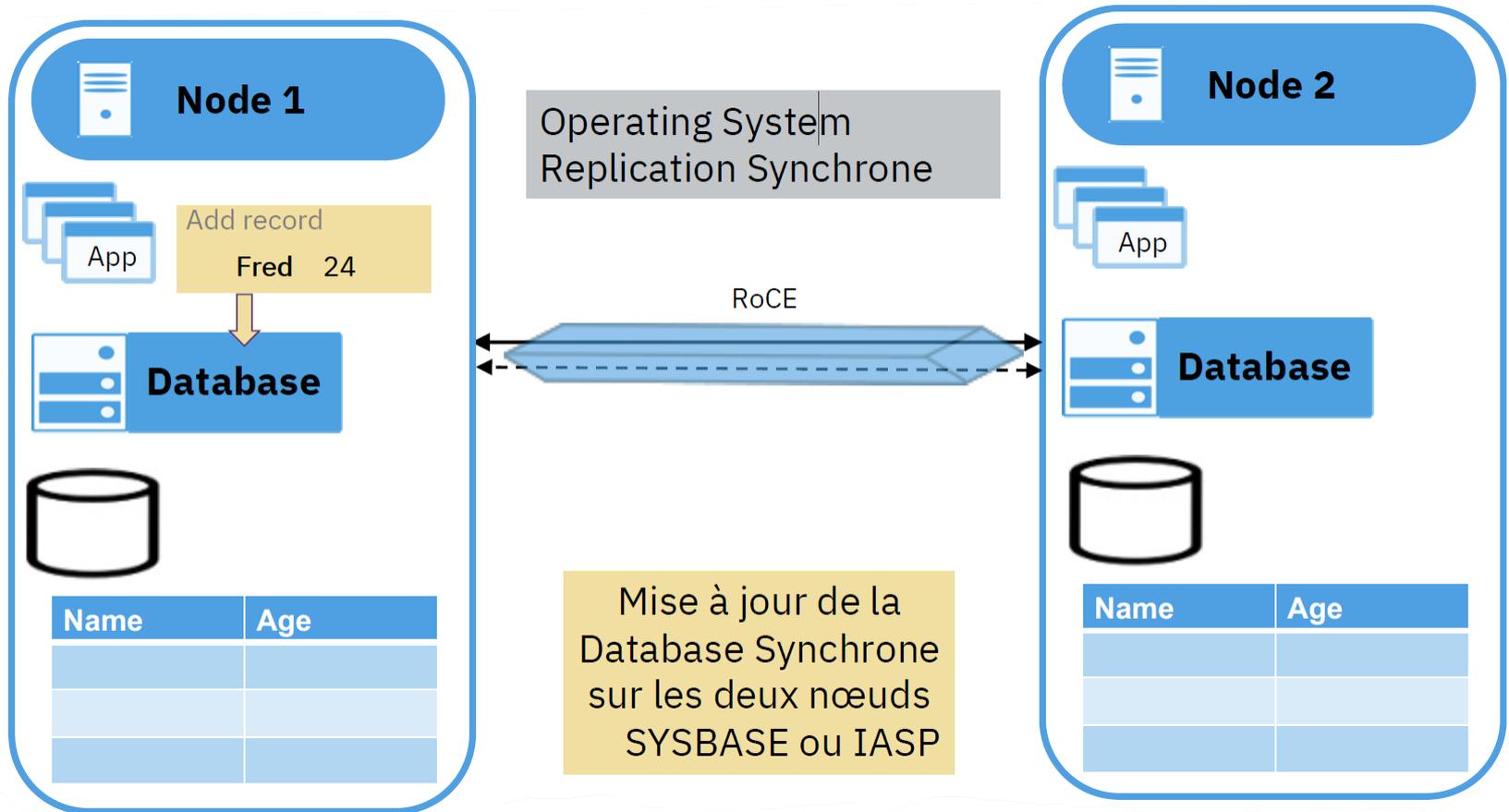


Architecture

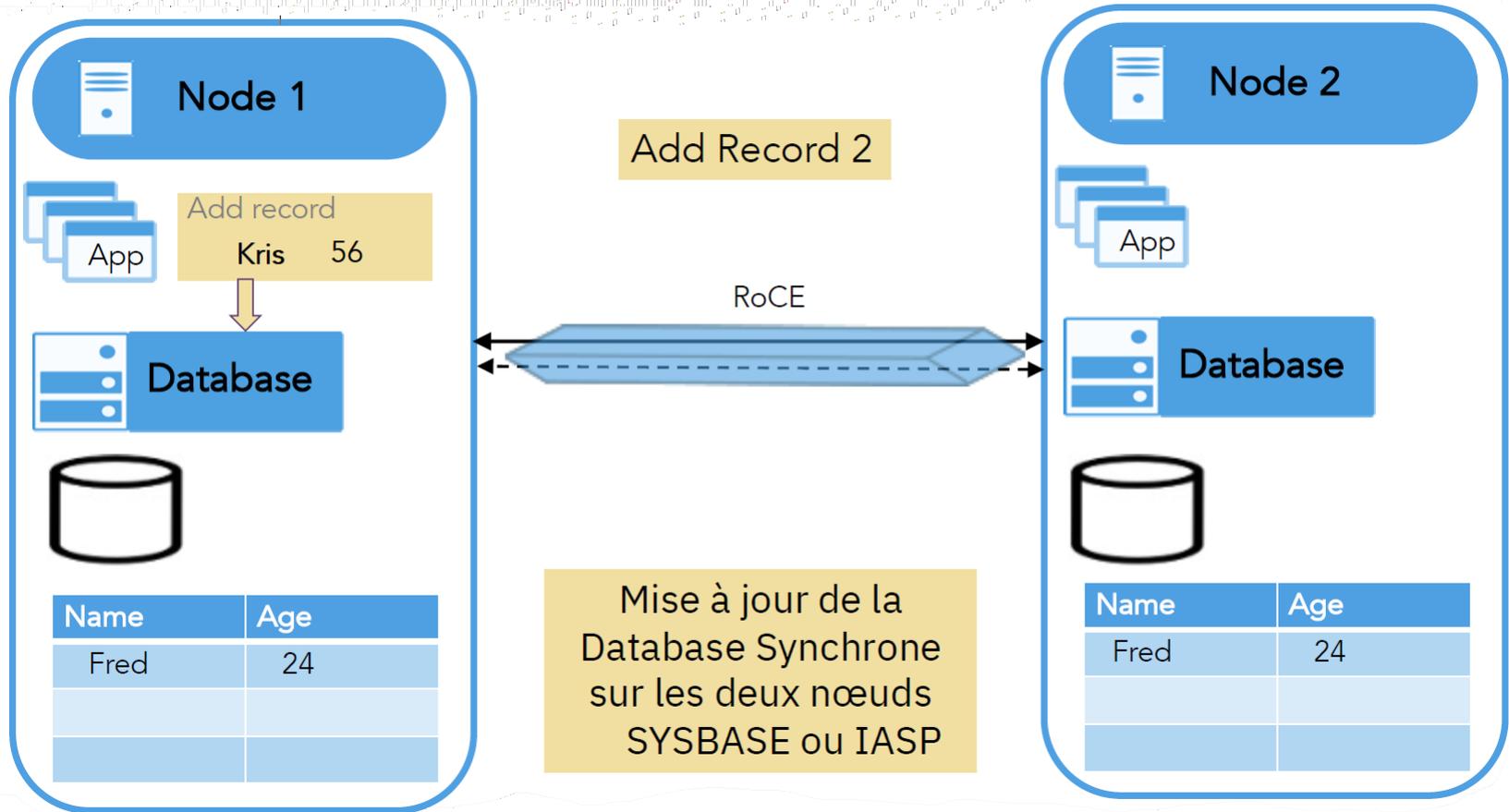
- Remote Direct Memory Access (RDMA)
- RDMA over Converged Ethernet (RoCE)
- Distance Maximum 100m ou 200m



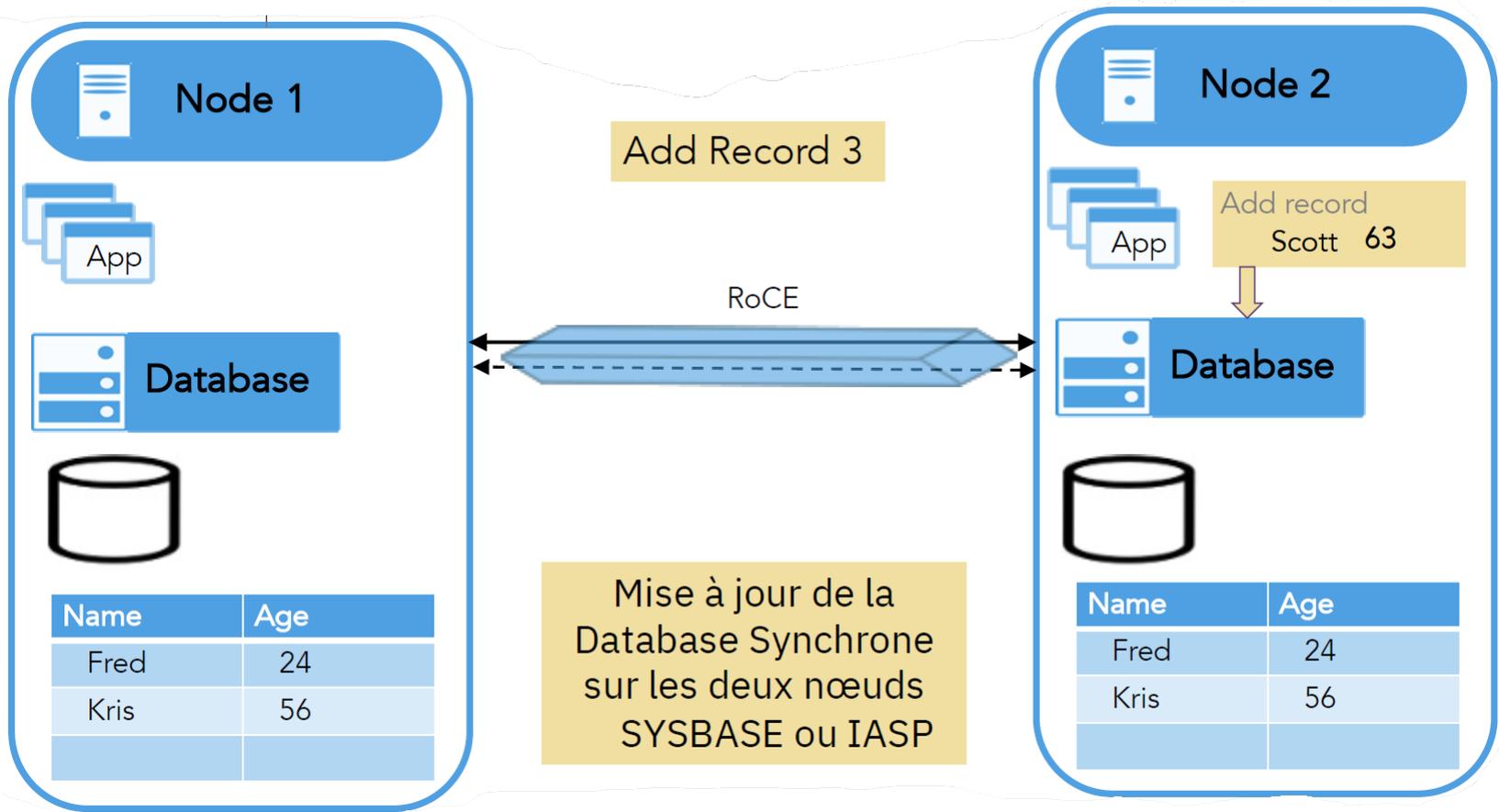
Db2 Mirror – Active Active



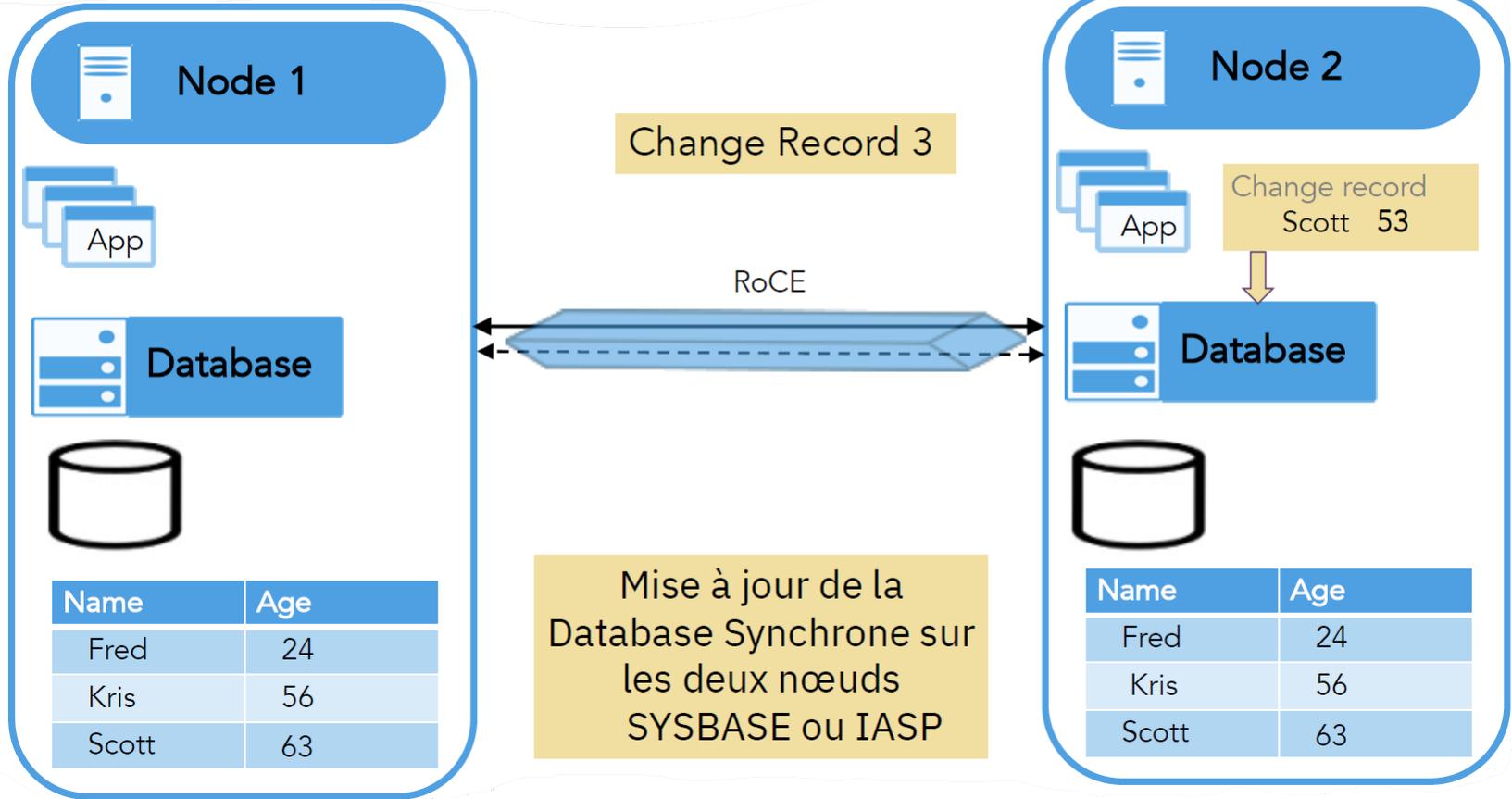
Db2 Mirror – Active Active



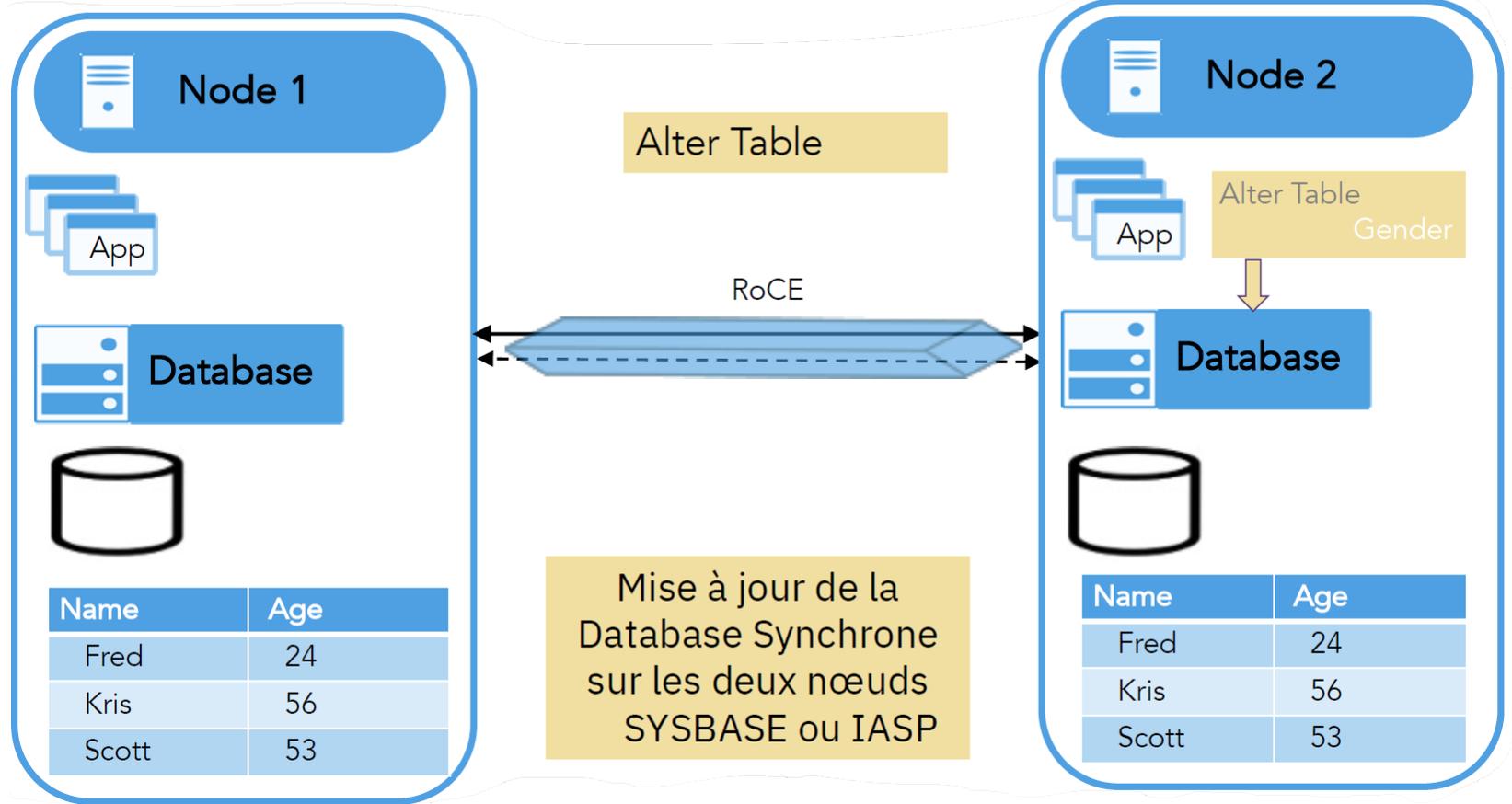
Db2 Mirror – Active Active



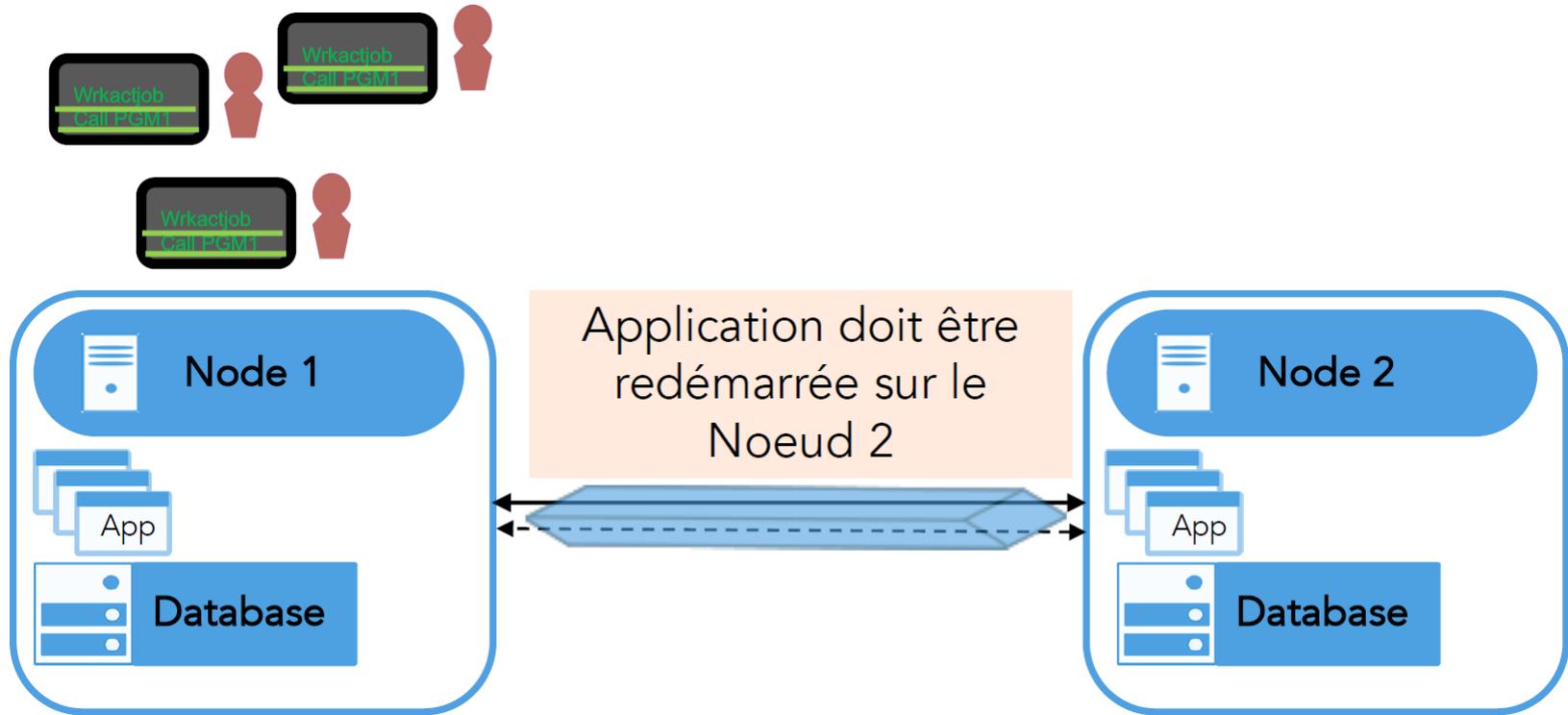
Db2 Mirror – Active Active



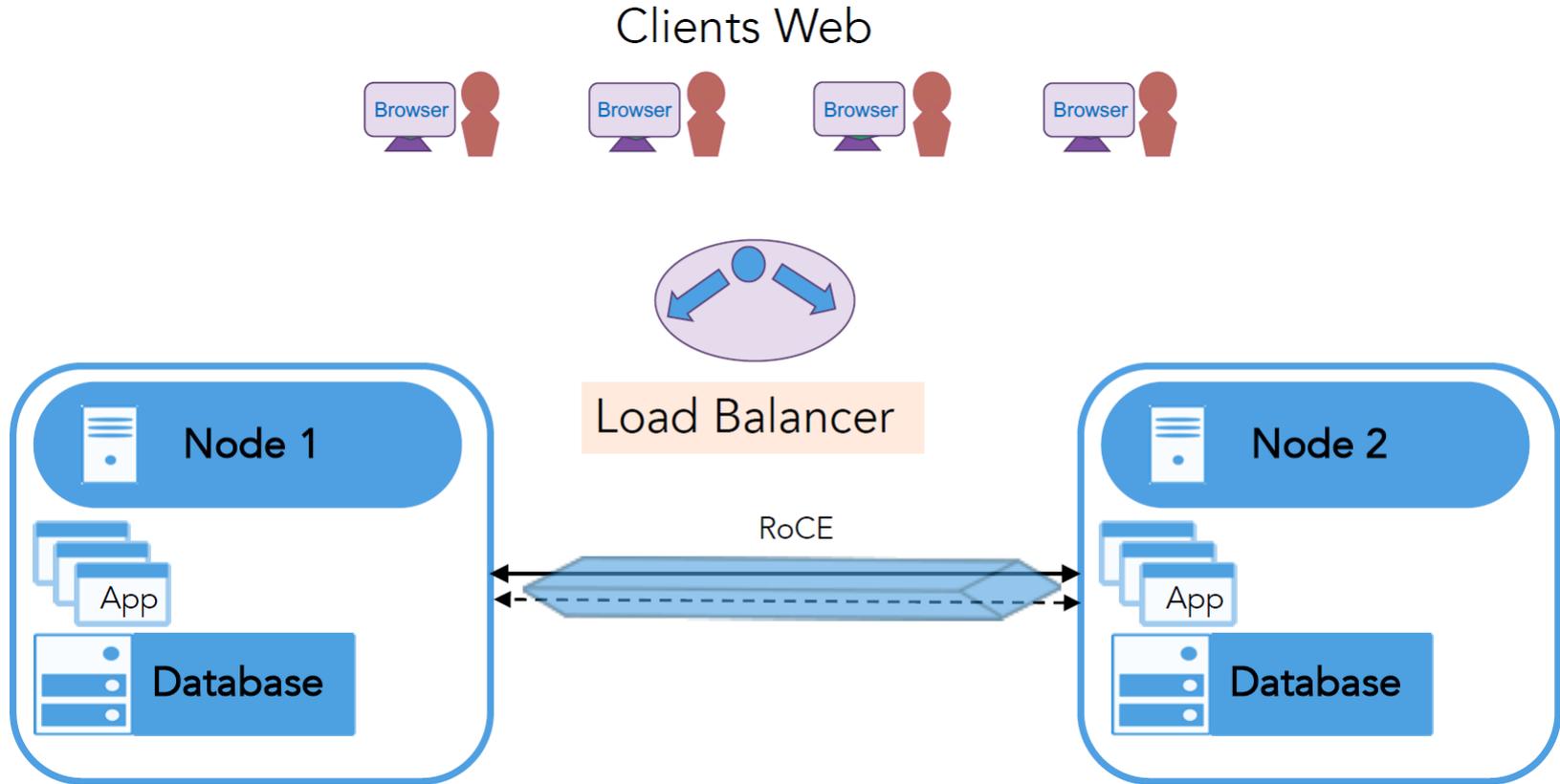
Db2 Mirror – Active Active



Db2 Mirror – Active Active, Green Screen Application



Db2 Mirror – Active Active, Clients Web avec Load Balancer



Règles de Réplication

IBM Db2 Mirror for i Primary: ZZ2P28 Secondary: ZZ2P29 User: qsecofr IBM GUI Build Time: 2019-04-14 10:34:34

Manage Replication List - Rules

Primary - ZZ2P28 Secondary - ZZ2P29 Rules Inspect Active Pending Active/Pending Select

Add a Rule test Object Type Object Name Exclude Include Definition Only

User Defined Rules Only Default Inclusion State: Exclude

Status	Library Name	Object Name	Replication State	Rule Group	Action	Rule Source	IASP Name
	Filter			All		All	All
	QCMD325283	*JRN	Exclude	Active		User	*SYSBAS
	QCMD333586	*JOBQ	Exclude	Active		User	*SYSBAS
	QCMD542101	*PGM	Exclude	Active		User	*SYSBAS
	QCMD681367	*SQLPKG	Exclude	Active		User	*SYSBAS
	QCMD759797	*SQLUDT	Exclude	Active		User	*SYSBAS
	QCMD821435	*SQLXSR	Exclude	Active		User	*SYSBAS
	QCMD980732	*SRVPGM	Exclude	Active		User	*SYSBAS
	QDEXDATA	*ALL	Include	Active		User	*SYSBAS
	QDEXDATA01	*ALL	Include	Active		User	*SYSBAS
	SPLMR000KW	*ALL	Include	Active		User	*SYSBAS
	TRANS1000	*ALL	Include	Active		User	*SYSBAS

1 300 Total Rows: 14

État d'inclusion par défaut (choix à l'installation)
INCLUDE - Inclure par défaut
EXCLUDE - Exclure par défaut

Détermination des règles de réplication

- Liste des critères de réplication
 - Moteur basé sur des règles
 - Hiérarchique
- État d'inclusion par défaut
 - Inclure - réplique tous les objets éligibles existants ainsi que tous les nouveaux objets au fur et à mesure de leur création
 - Exclure - ne rien répliquer
- Ajouter ensuite des règles hiérarchiques aux niveaux inférieurs
 - Une bibliothèque
 - Type d'objet dans la bibliothèque
 - Objet spécifique

— Exemple 1 :

- État d'inclusion par défaut : exclure
- Inclure tout dans la bibliothèque Sales sauf les files d'attente de sortie, mais inclure les reçus de file d'attente de sortie

Library	Object Type	Object	Incl/Excl
Sales	*all	*all	Include
Sales	*OUTQ	*all	Exclude
Sales	*OUTQ	Receipts	Include

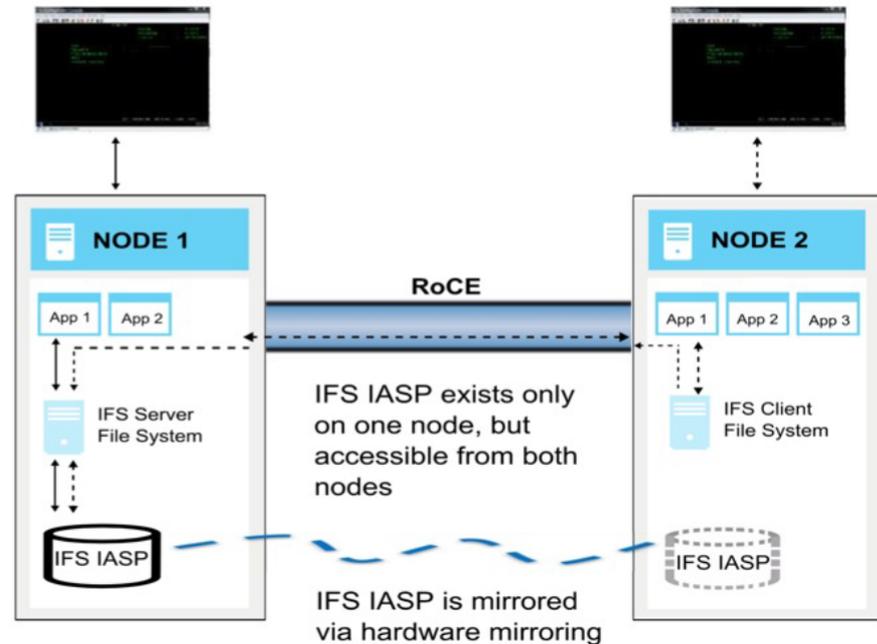
— Exemple 2 :

- État d'inclusion par défaut : inclure
- Inclure tout dans la bibliothèque Sales sauf les files d'attente de sortie, mais inclure les reçus de file d'attente de sortie

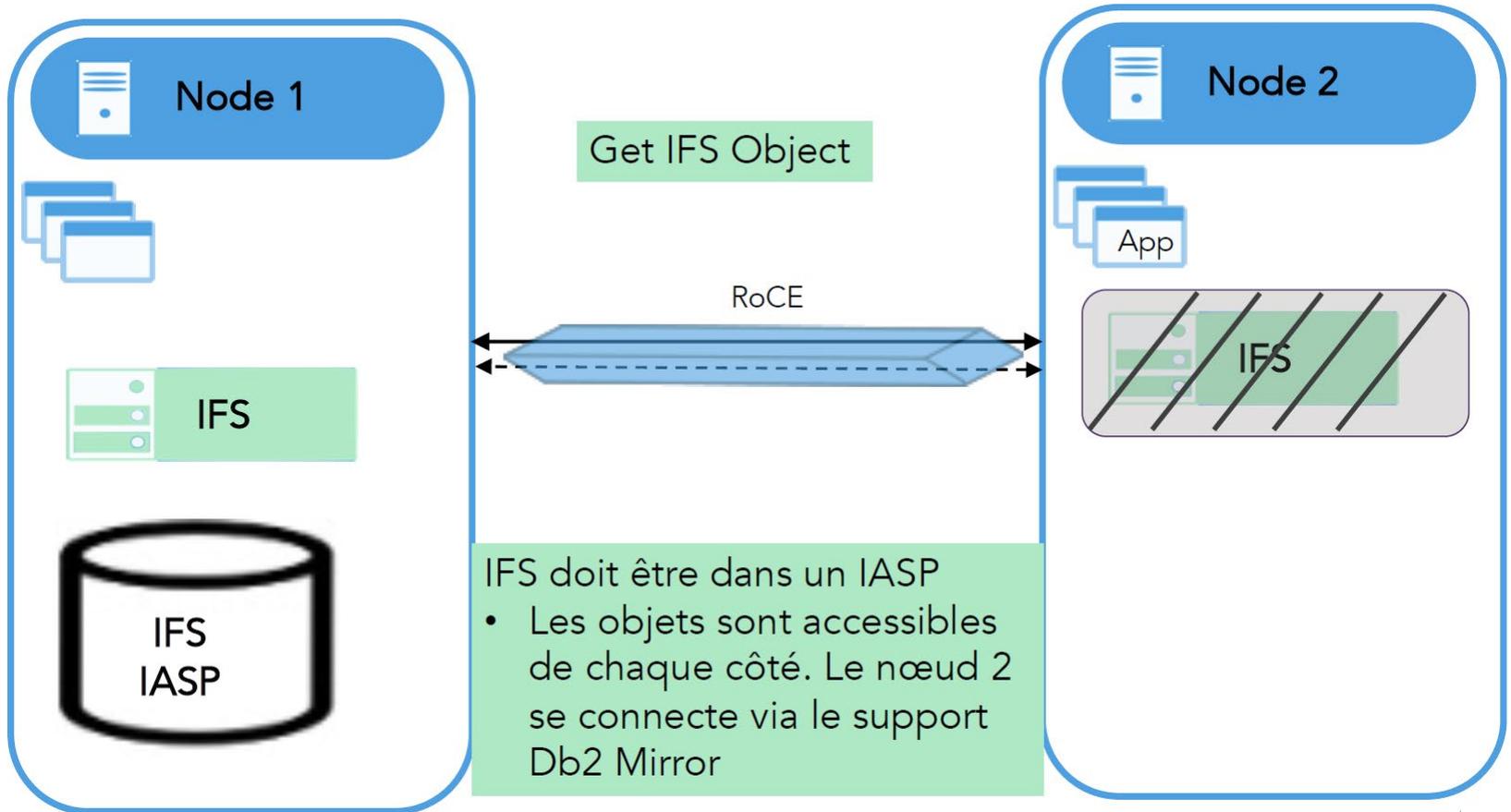
Library	Object Type	Object	Incl/Excl
Sales	*OUTQ	*all	Exclude
Sales	*OUTQ	Receipts	Include

Support IFS

- Nécessite IASP
- IFS accessible sur les deux nœuds (R/W)
- Nécessite PowerHA
 - Db2 Mirror fournit l'accès simultané.
 - PowerHA commute l'IASP
- Le système de fichiers «mute» automatiquement lorsque le stockage est basculé



Db2 Mirror – IFS



Db2 Mirror GUI

L'interface graphique fonctionne sur IBM i

L'interface graphique peut s'exécuter sur les nœuds Db2 Mirror

L'interface graphique peut s'exécuter en dehors des nœuds Db2 Mirror et gérer plusieurs paires



<http://systemname:2006/Db2Mirror>

WHERC

