

# Smarter Storage for Smarter Computing: What's Next. Ready Now

1	<i>La stratégie IBM sur la technologie Flash</i>
2	<i>La proposition de valeur de FlashSystem</i>
3	<i>Exemples de mise en œuvre</i>
4	<i>Gains FlashSystems et retour d'expériences</i>
5	<i>FlashSystems dans un environnement IBM i</i>
6	<i>Points techniques</i>

Herve.Giop@fr.ibm.com



*IBM FlashSystem...  
Enterprise Storage in microseconds, not milliseconds*

# TMS : Les raisons de l'acquisition par IBM

- TMS a été créée en 1978, à Houston, Texas.
- TMS a, de fait, une très grande expérience dans la conception de systèmes SSD/Flash systems.
- TMS a d'ores et déjà de nombreuses références en production de par le monde.
- TMS a des spécificités techniques uniques sur le marché

**RamSan-720:** 5/10 TB SLC Flash, 4 FC (8 Gb)/IB (QDR)

**RamSan-710:** 5 TB SLC Flash, 4 FC (8 Gb)/IB (QDR)

**RamSan-640:** 8 TB SLC Flash, 10 FC (8 Gb)/IB (QDR)

**RamSan-620:** 5 TB SLC Flash, 8 FC (4 Gb)

**RamSan-440:** 512 GB RAM, 8 FC (4 Gb)

**RamSan-400:** 128 GB RAM, 8 FC (4 Gb), 4 IB (4x)

**SAM 500:** DSP/SSD, 64 GB RAM, 15 FC (1 Gb)

**SAM-2000:** DSP system

**Company founded by Holly Frost**

2012

2011

2010

2009

2008

2007

2006

2005

2004

2003

2002

• • •

1997

• • •

1990

• • •

1978

**RamSan-820:** 12-24 TB eMLC Flash, 4 FC (8 Gb)/IB (QDR)

**RamSan-810:** 10 TB eMLC Flash, 4 FC (8 Gb)/IB (QDR)

**RamSan-70:** 900 GB SLC Flash, PCIe x8 2.0

**RamSan-630:** 10 TB SLC Flash, 10 FC (8 Gb)/IB (QDR)

**RamSan-20:** 450 GB SLC Flash, PCIe x4

**RamSan-500:** 2 TB SLC Flash, 64 GB RAM, 8 FC (4 Gb)

**RamSan-320:** 64 GB RAM, 8 FC (2 Gb)

**RamSan-210/220:** 32 GB RAM, 4 FC (2 Gb)

**SAM-350/SAM-450:** DSP system

*Custom systems for seismic industry*

**CMPS:** custom SSD for Gulf Oil



## Quelques références TMS et leurs domaines d'activité (1/2)

- Envoyer un texto
- Parier en ligne
- Réserver un voyage
- Utiliser un GAB
- Trader en ligne
- Acheter en ligne
- Utiliser du wifi pre-payé
- Jouer en ligne
- Traiter du Big Data 'sensible'
- Acquisition de data en temps réel



TMS, c'est déjà :

Les plus **grosses installations** de Flash en production dans le monde

- En opération dans **10 places financières** majeures dans le monde
- Utilisé par **7 des 11 plus gros opérateurs telcos** dans le monde
- **300 clients** en production dans *35 pays*

### TMS / FlashSystem est DÉJÀ PARTOUT

### ...ou la PERFORMANCE EST CLE pour le Business!

# Quelques références TMS et leurs domaines d'activité (2/2)



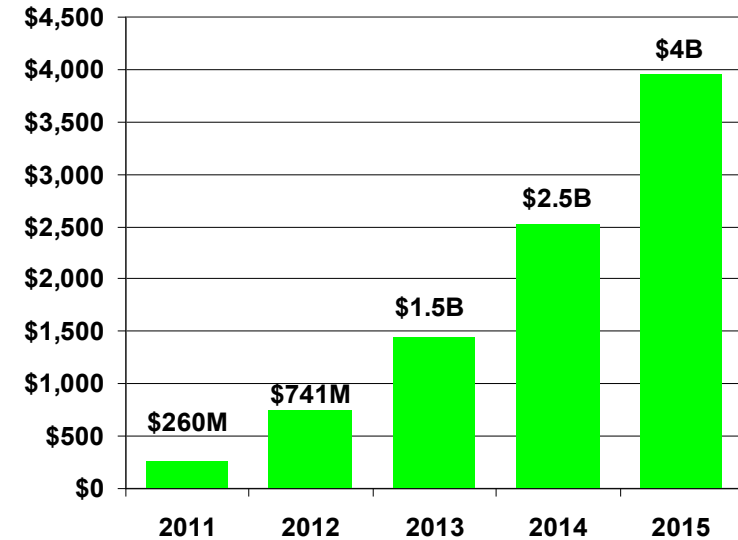
SUNGARD®



# La stratégie d'IBM suite au rachat de TMS

- Capitaliser sur la très grande expérience de TMS dans le domaine des Flash systems et sur sa base installée.
- Déployer des solutions Flash systems intégrées dans les différentes parties du portefeuille matériel/appliance d'IBM : serveurs, stockage, appliances.
- Se positionner en tant que leader sur un marché en très forte croissance selon Gartner.
- Amplifier l'avance technologique sur nos concurrents qui ont tous fait l'acquisition de solutions beaucoup moins éprouvées.

Global Flash Market Growth



Source: Gartner ETA Enterprise Solid-State Appliance Nov 2011



Ambuj GOYAL, GM, System Storage & Networking

'FLASH will change the structure of our future systems...  
We will be able to offer our clients flash optimized systems  
by integrating this technology in our overall portfolio'



Texas Memory Systems, Inc.  
An IBM Company

« IBM mise 1 milliard de dollars sur la technologie Flash »

# Les cibles pour la technologie FlashSystem ?

## “Partout où le temps a une importance”



### OLTP Bases de données

Finance, gaming, facturation temps réel, trading, real-time monitoring, accélération de requêtes...

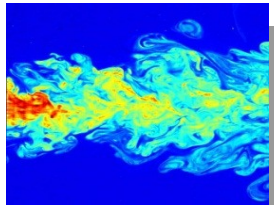
### Infrastructures virtualisées

VDI, consolidation d'infrastructures virtualisées



### Analytique

Business intelligence, traitement de batch, système ERP, reporting, ...



### Applications de calcul, HPC...

Simulation, modeling, rendering, FS metadata, scratch space, video on demand...

### Infrastructures Cloud

On-demand computing, Distribution de contenu, web, caching, active file management...



Financial

Government

E-Commerce

HPC

Telecom

FINANCIAL

GOVERNMENT

E-COMMERCE

HPC

TELECOM



## Exemple de cible : Accélération Oracle

Problème client : La plupart des applications Oracle sont très intensives en lecture (R/W > 70/30).

L'ajout de processeurs sur le serveur n'a alors que peu d'effet.

En orientant les lectures sur un système Flash à faible latence, les lectures seront exécutées beaucoup plus rapidement en boostant les performances **jusqu'à 12 fois, sans tuning ni changement d'architecture.**

### •IBM FlashSystem

- Réduira l'I/O wait time sur les instances Oracle critiques.
- Fournira les données avec **la plus faible latence** possible sur le SAN.
- **Accélérera les applications avec un profil I/O-bound**, incluant le transactionnel, le traitement batch et l'analytique.



Delivers **Extreme Performance, Macro Efficiency, and Enterprise Reliability** for



## Exemple de cible : Accélération SAP

Problème client – la grande majorité des bases de données SAP ne dépasse pas 2TB. SAP encourage ses clients à améliorer les performances en adoptant l'architecture HANA (base de données en mémoire).

Cela entraîne souvent une redéfinition de l'architecture SAP, avec les coûts et les interruptions de service liés.

### •IBM FlashSystem

- Eliminera les **goulots d'étranglement I/O**
- Fournira les données avec **la plus faible latence** possible sur le SAN.
- Hébergera la volumétrie de la base **SAP** dans un pool très rapide.
- Permettra de se conformer aux recommandations HANA sans re définition de l'architecture.



Delivers **Extreme Performance, Macro Efficiency, and Enterprise Reliability** for





## Exemple de cible : Serveurs virtualisés et VDI

Problème client – La virtualisation des serveurs et du poste de travail (VDI Virtual Desktop Infrastructure) poussent les baies de stockage conventionnelles aux limites de leurs performances. Les utilisateurs sont alors impactés par les temps de réponse dégradés.

L'insertion de IBM FlashSystem dans ces environnements permet d'améliorer jusqu'à **16 fois** les temps de réponse, et de ne plus craindre les 'log on storm' ou 'boot storm' qui sont gommés grâce aux performances Flash.

### •IBM FlashSystem

- **Éliminera les problèmes de densité d'I/O et de points chauds** qui sont courants dans les environnements virtualisés.
- **Élimine les problèmes de performances lors des 'log on' ou 'boot' storm.**
- **Permet la consolidation du stockage** sans pénaliser la performance, entraînant un ratio Virtuel/Physique plus important et donc une optimisation de l'utilisation des serveurs physiques.



Delivers **Extreme Performance, Macro Efficiency, and Enterprise Reliability** for



# 12 IBM “Flash Centers of Competency”

→ Proof of Concept

→ Client Workload Benchmarks

→ Tailored IT Optimization Assessments and TCO Studies



# Smarter Storage for Smarter Computing: What's Next. Ready Now

1	<i>La stratégie IBM sur la technologie Flash</i>
2	<i>La proposition de valeur de FlashSystem</i>
3	<i>Exemples de mise en œuvre</i>
4	<i>Gains FlashSystems et retour d'expériences</i>
5	<i>FlashSystems dans un environnement IBM i</i>
6	<i>Points techniques</i>



*IBM FlashSystem...  
Enterprise Storage in microseconds, not milliseconds*

# Les solutions de stockage

## Bandes magnétiques dans robotique



**Temps d'accès : 20s**

**Flux séquentiels : sauvegardes**

## Disques SAS dans baie de stockage



**Temps d'accès : 5-10ms**

**Flux aléatoires : Bases de données, ERP, envts virtualisés, etc**

## Disques SSD dans baie de stockage



**1ms**

## Baie de stockage FlashSystem



**0,1ms**

# Le stockage sur disque, goulot d'étranglement dans l'IT

Durant les 10 dernières années:

- Vitesse CPU: croissance **8-10x**
- Vitesse DRAM: croissance **7-9x**
- Vitesse BUS: croissance **20x**
- Vitesse RESEAU: croissance **100x**
- Vitesse DISQUES: croissance **1.2x seulement**



# Différentes classes de stockage en technologie Flash

## **Direct Attach FI**

(PCIe adapters, tiroir SSDs)

temps de réponse IO optimisés, hautes performances, dans le serveur

Direct Attach

0,02ms

## **Direct Attach Flash**

Inconvénients: cher, dédié, limité en capacité (nombre d'emplacements PCI), données difficiles à protéger, non mutualisable

## **Baie de stockage 'classique'**

(Disques durs et SSD)

Haut niveau de fonctionnalités, avec forte capacité

Hybrid Storage Systems

1 à 10 ms

## **Baie de stockage 'classique'**

Inconvénients: performances



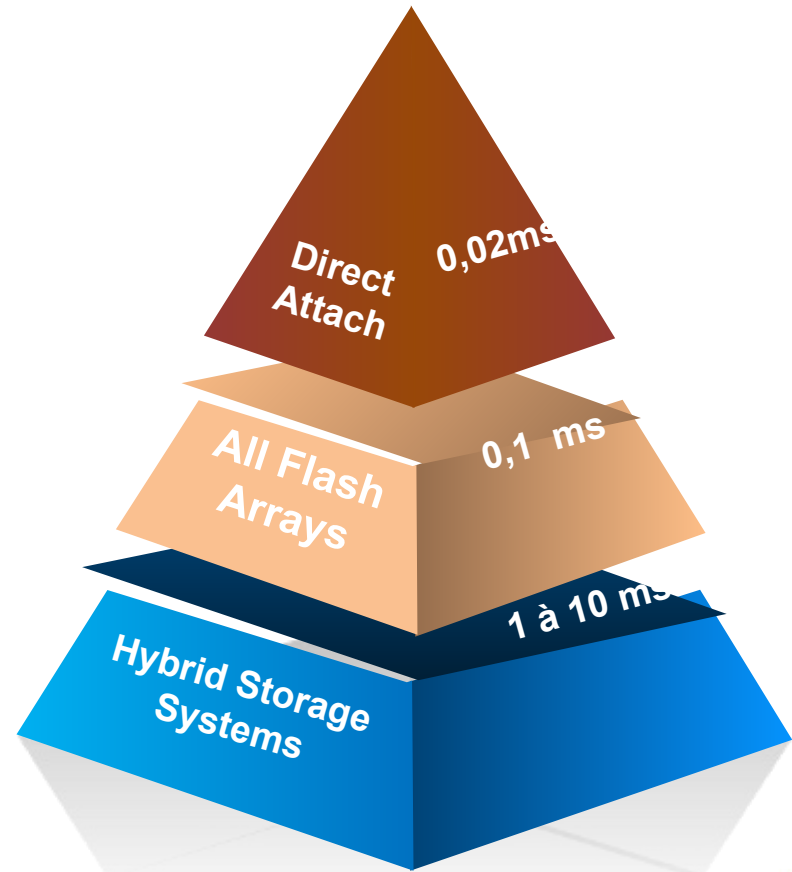
Comment répondre au besoin d'une technologie de stockage mutualisé, avec de très hautes performances, et une disponibilité de type 'Entreprise' ?

# Accélérer les applications critiques, Utiliser la pleine performance des serveurs

**IBM FlashSystem** apporte la performance extrême, de très faible temps de réponse, une disponibilité de type 'entreprise'



**150-200  $\mu$ s  
Latency**



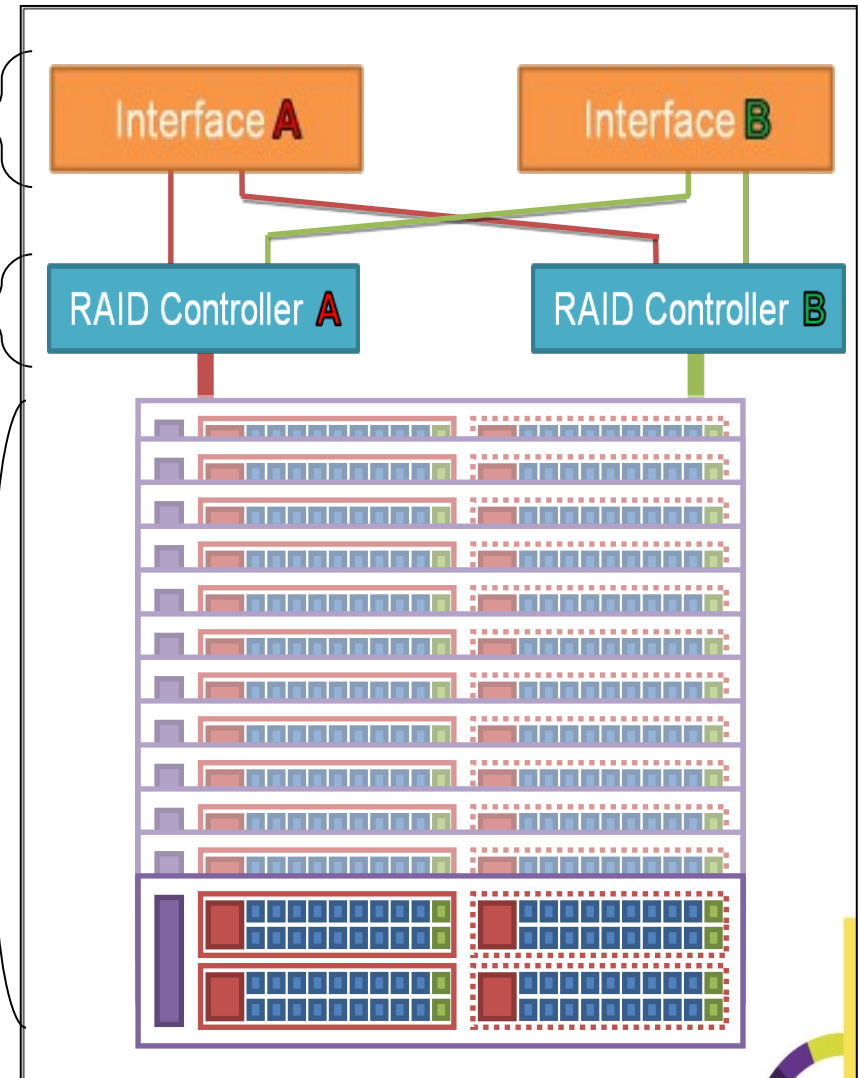
# FlashSystems en quelques mots

	FlashSystem 720 w/ RAID
<b>Sequential read</b> (256 KB) GB/s	3.3 GB/s (FC); 5 GB/s (IB)
<b>Sequential Write</b> (256 KB) GB/s	3.3 GB/s (FC); 4 GB/s (IB)
100% <b>Random Reads</b> (4 KB) IOPS	525k
<b>70-30 Random Read-Write</b> (4 KB) IOPS	450k
Max <b>random write</b> (4KB) IOPS	400k
Size / KW	1RU / 300W

External  
Interfaces  
(FC, IB)

RAID  
Controllers

Flash  
Modules



Une conception axée sur la performance :

- Fonctions microcodées.
- Absence de fonctions avancées « logicielles » :

Réplication, Snapshots, compression, Thin provisioning



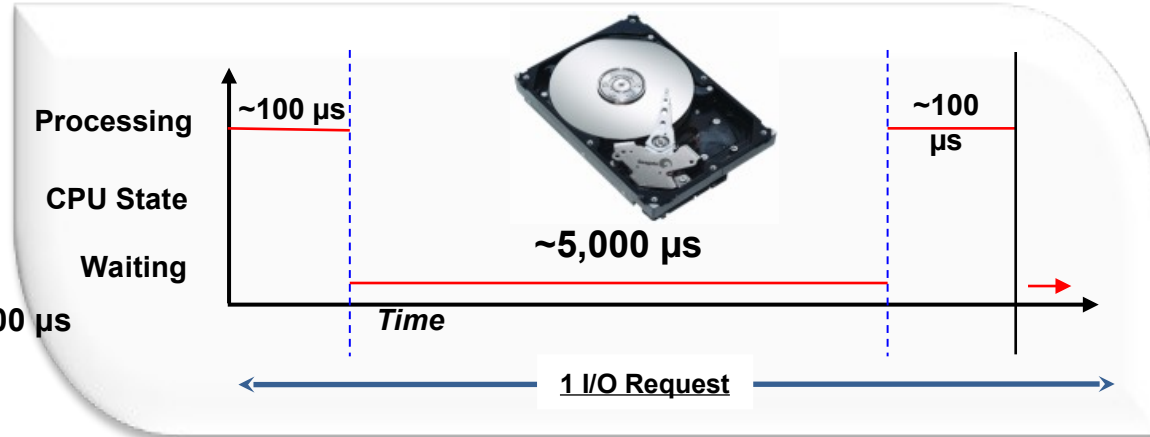
# Impact de la latence du stockage sur le rendement applicatif

## Temps de service d'un I/O sur **Disque** coté application

- 1. Demande d' I/O ~ 100 µs
- 2. Attente des données ~ 5,000 µs
- 3. Traite l' I/O ~ 100 µs

▪ Temps total coté application = 200 µs + 5,000 µs = 5,200 µs

▪ **“Rendement applicatif” = 200 / 5,200 = ~4%**



## Temps de service d'un I/O sur **SSD** coté application

Attente des données ~ 1,000 µs

• Temps total coté application = 200 µs + 1,000 µs = 1,200 µs

• **“Rendement applicatif” = 200 / 1,200 = ~17%**

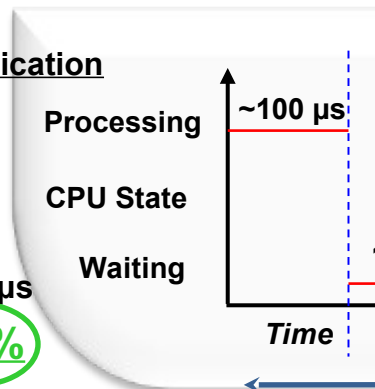


## Temps de service d'un I/O sur **FlashSystems** coté application

- 1. Demande d' I/O ~ 100 µs
- 2. Attente des données ~ 200 µs
- 3. Traite l' I/O ~ 100 µs

▪ Temps total coté application = 200 µs + 200 µs = 400 µs

▪ **“Rendement applicatif” = 200 / 400 = 50%**



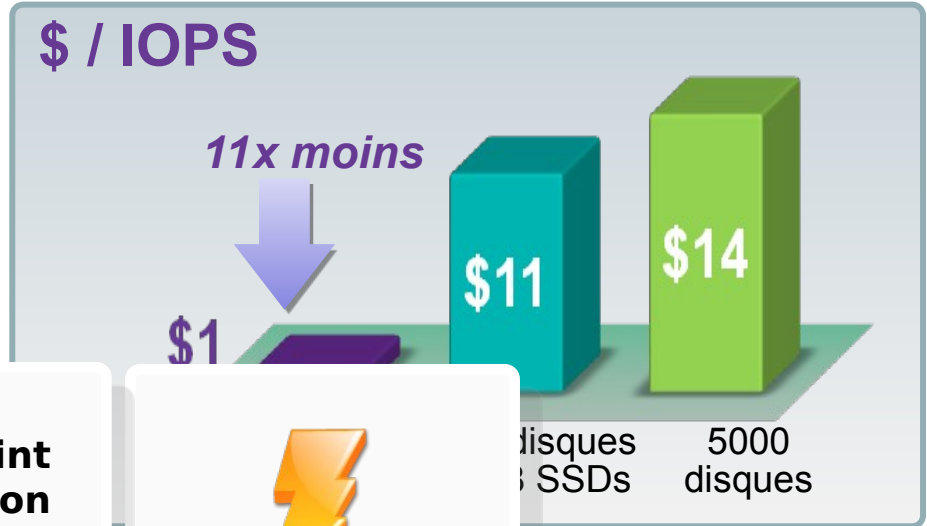
**85% Reduction**  
In batch processing times

**90% Reduction**  
In OLTP times

# Economie d'énergie et de place au sol

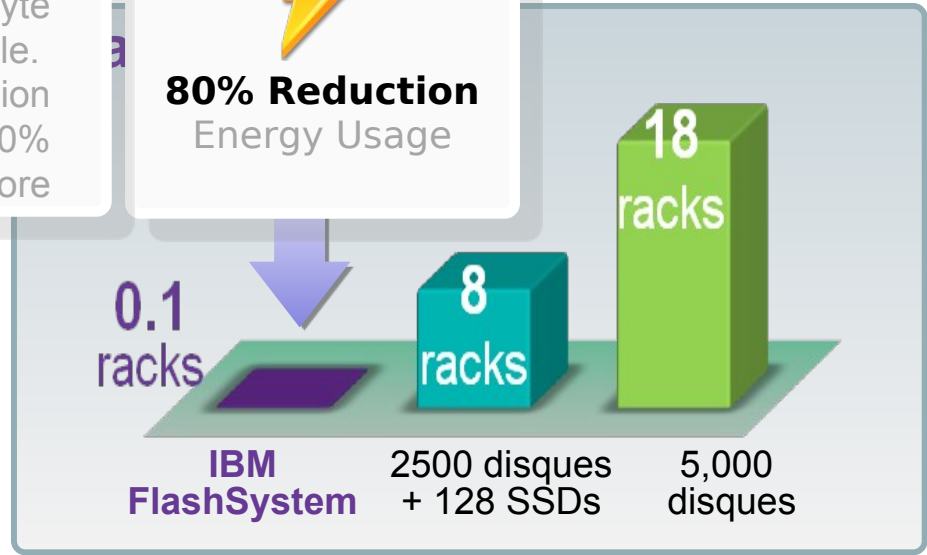
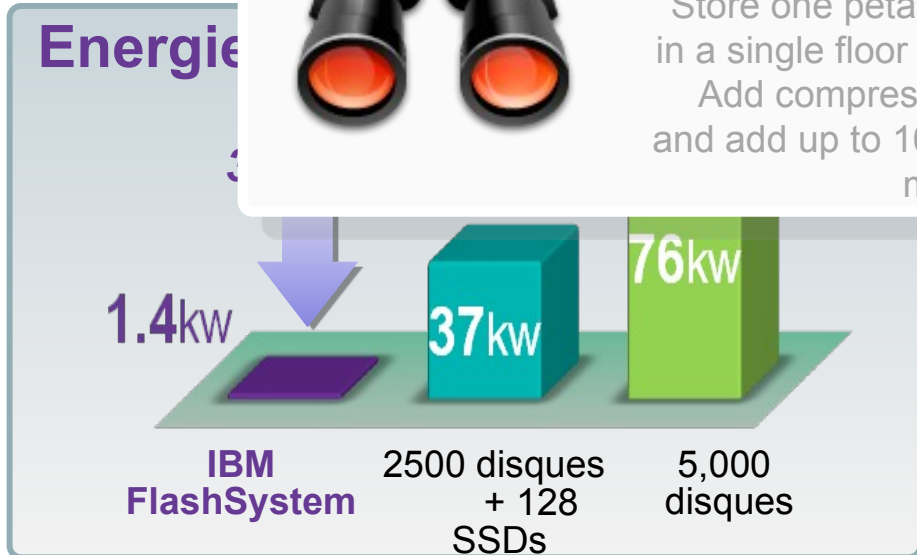
Pour un même besoin, 3 réponses très différentes en terme d'impact environnemental entre :

- FlashSystem
- SSD + disques
- Disques



**75% Footprint Reduction**

Store one petabyte in a single floor tile. Add compression and add up to 100% more

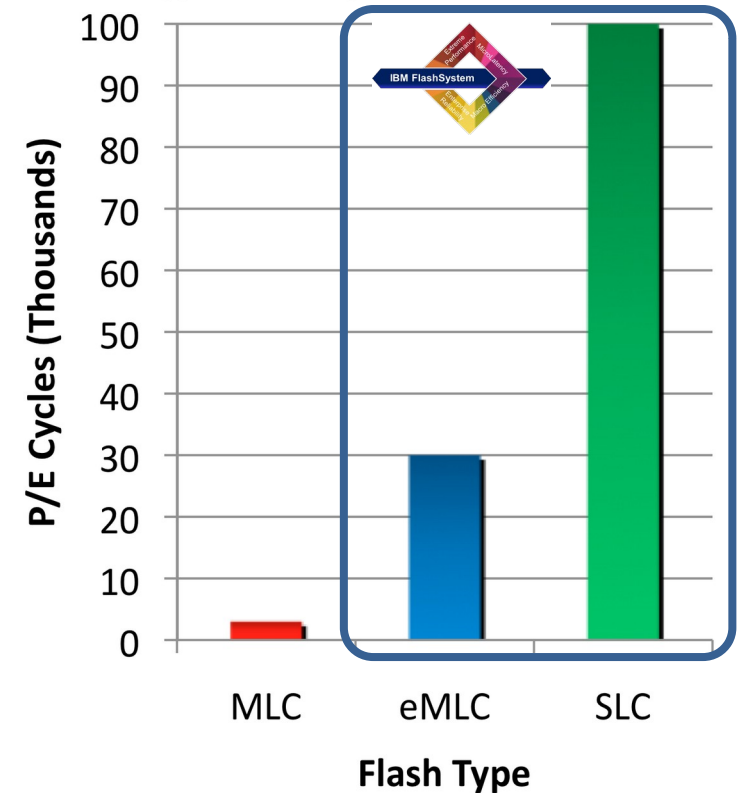


Test Power/DB2 : 1.3 Million IOPS, 43K+ Transactions per second, 13K Updates per second

# Qualité et durée de vie des Flash NAND

- Une mémoire Flash est un stockage électronique ,non volatile, qui peut être électriquement effacé (**E**rase) puis reprogrammé (**P**rogram).
- Un chip NAND ne s'use pas en lecture mais s'use en écriture. Sa durée de vie diminue donc au fur et à mesure des écritures.
- Le 1er cycle de **P/E** (**P**rogram/**E**rase) intervient dès le 2nd cycle d'écriture.
- La durée de vie des chips est liée à la technologie utilisée.
  - MLC (3 000 cycles P/E maximum par bit)
  - eMLC (30 000 cycles P/E maximum par bit)
  - SLC (100 000 cycles P/E maximum par bit)
- **IBM Flashsystem est de technologie eMLC (810/820) ou SLC (710/720).**
- Des techniques comme le wear leveling et l'overprovisioning permettent de "prolonger" la durée de vie des baies Flash.

## Typical Chip Endurance



# Durée de vie maximale (cycles d'écritures).

## 1 – Capacité de la baie Flash

Exemple avec **5 TB Flash System**

**5 TB =  $5 \times 10^{12}$  Octets = 40.000.000.000.000 bits.**

## 2 – Nombre de cycles P/E maximum par bit

**100.000** pour **SLC** (model 7xx)

**30.000** pour **eMLC** (model 8xx)

**3.000** pour **MLC** (non utilisé par IBM, mais par certains de ses concurrents).

$$\frac{\text{Flash Capacity} \times \text{Flash Quality}}{\text{Media Write Bandwidth}}$$

## 3 – Calcul du nombre maximum de cycles P/E de la baie Flash

Exemple pour **5 TB SLC** (model 710) :

$40.000.000.000.000 \times 100.000 = 4.000.000.000.000.000.000$  P/E cycles (ou writable Bits).

## 4 – Besoins du client en terme d'écritures / jour.

Exemple : **1 GB/sec =  $1 \times 8 \times 10^9 \times 3600 \times 24 = 86$  TB/jour = 691.200.000.000 bits par jour**

## 5 – Calcul de la durée de vie en fonction du workload client

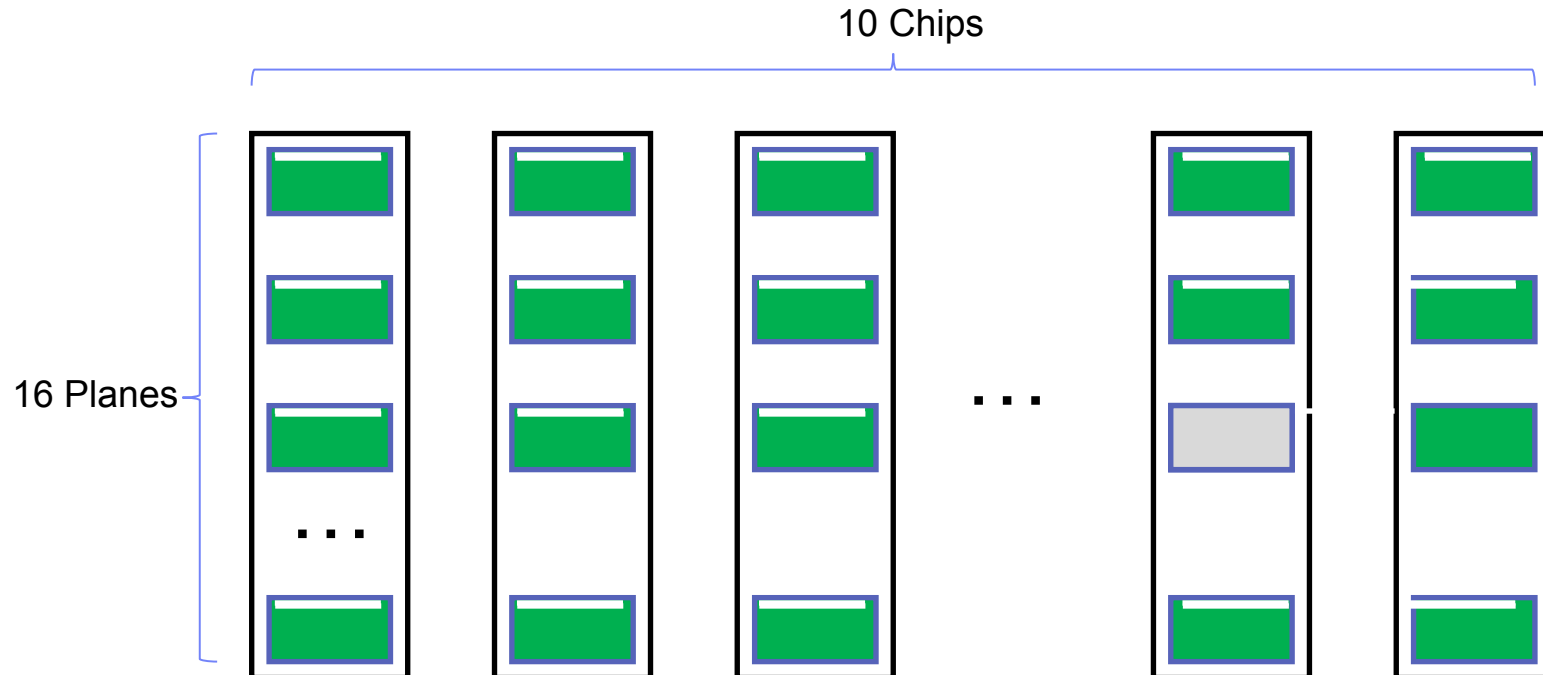
$4.000.000.000.000.000$  divisé par  $691.200.000.000.000$  donne **5 787 jours** soit **16 ans**

Baie Flash de 5TB	SLC	eMLC	MLC
Débit soutenu écritures GB/s maximum pour une durée de vie de 3 ans	5,3GB/s	1,6GB/s	0,2GB/s
Durée d'utilisation avant dégradation si 1GB/s soutenu en écriture.	Environ 16 ans	Environ 5 ans	173 jours

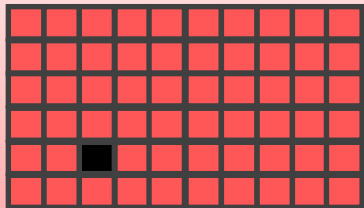
NB : La technologie MLC nécessite plus de capacité pour avoir une durée de vie acceptable

## Variable Stripe RAID™ (VSR)

- Patented Variable Stripe RAID allows RAID stripe sizes to vary.
- VSR uses Overprovisioning space to effectively handle failures and **avoid requiring maintenance**.
- If one die fails in a ten-chip stripe, only the failed die is bypassed, and then data is restriped across the remaining nine chips. **No system rebuild needed!**
- **Optional Flash module spare for 710/810.**



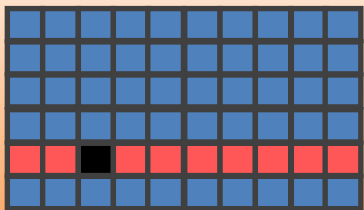
# Avantages de VSR : ne pas nécessiter d'arrêt pour maintenance et avoir toujours la performance optimale



No Parity

## Form Factor SSD

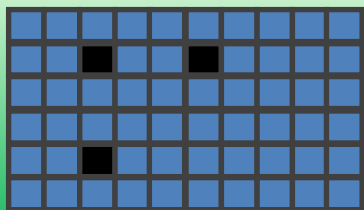
- Flash failure = Disk failure
- Requires top-level RAID
- Relatively frequent hot-swaps



Parity

## Enterprise Flash Drive

- Flash failure = Degraded state within module
- Performance impact on RAID set
- Hot-swap to resolve



Parity

## IBM FlashSystem with VSR

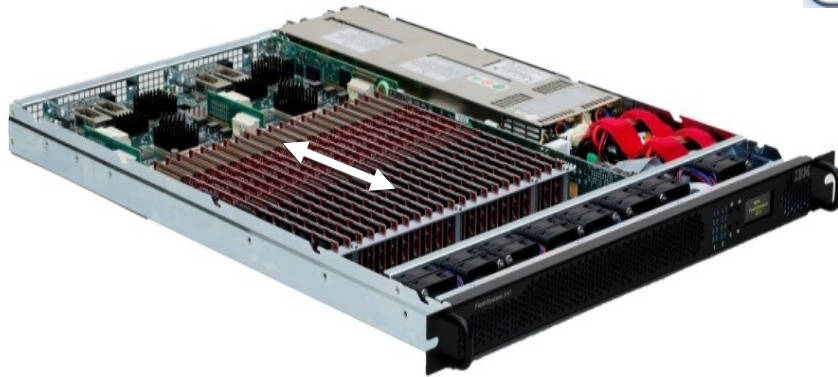
- Preserves Flash life
- Preserves performance
- Re-parity data in microseconds

Less maintenance touches while still preserving the **life**, **protection**, and **performance** of the Day-1 experience



# Différences entre les modèles FlashSystem x10 et x20

FlashSystem 710/810



1D RAID across Flash chips  
No Flash Hot-Swap

Flash Board capacity (**710**: 250 GB \* **810**: 500 GB)

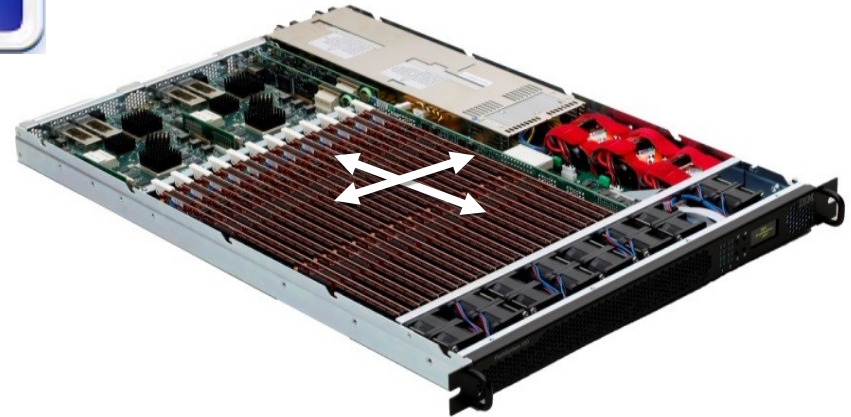
Flash Module capacity = 1 Flash Board

### Incremental Capacities

4-20 Flash Modules, upgrade 4 Modules  
(**710**: 1/2/3/4/5 TB \* **810**: 2/4/6/8/10 TB)



FlashSystem 720/820



**2D RAID** across Flash chips & Flash Modules  
**Flash Module Hot-Swap**

Flash Board capacity (**720**: 500 GB \* **820**: 1 TB)

Flash Module capacity = 1 or 2 Flash Board

### No capacity upgrade

(**720**: 5 ou 10 TB \* **820**: 10 ou 20 TB)



# Proposition de valeur d'IBM FlashSystem



**85% Reduction**  
In batch  
processing times



**90% Reduction**  
In OLTP times



**150-200  $\mu$ s**  
**Latency**



**80% Reduction**  
Energy Usage



**75% Footprint  
Reduction**  
Store one petabyte  
in a single floor tile.  
Add compression  
and add up to 100%  
more

**Enterprise  
Reliability**

High Availability,  
2D Flash RAID™  
and Variable  
Stripe RAID™





# Smarter Storage for Smarter Computing: What's Next. Ready Now

1	<i>La stratégie IBM sur la technologie Flash</i>
2	<i>La proposition de valeur de FlashSystem</i>
3	<i>Exemples de mise en œuvre</i>
4	<i>Gains FlashSystems et retour d'expériences</i>
5	<i>FlashSystems dans un environnement IBM i</i>
6	<i>Points techniques</i>



*IBM FlashSystem...*

*Enterprise Storage in microseconds, not milliseconds*

## 3 fonctionnements principaux pour FlashSystem

- Utilisée en parallèle d'une baie de stockage disque
- En remplacement de toute ou partie d'une baie de disques classique
- Utilisée sous un système de virtualisation (SVC/V7000)

### Avec les mêmes objectifs :

- ❖ **Améliorer de façon très importante les performances IOs.**
- ❖ **Réduire la place au sol et l'empreinte énergétique par rapport à une solution disque (\*)**
- ❖ **Réduire la facture logicielle sur les serveurs en leur délivrant les données beaucoup plus rapidement (\*\*)**

(\*) 500 000 I/O par sec dans une unité de 1U versus plus de 100U avec 2000 disques

(\*\*) L'application a ses performances multipliées par 12 avec la Flash !

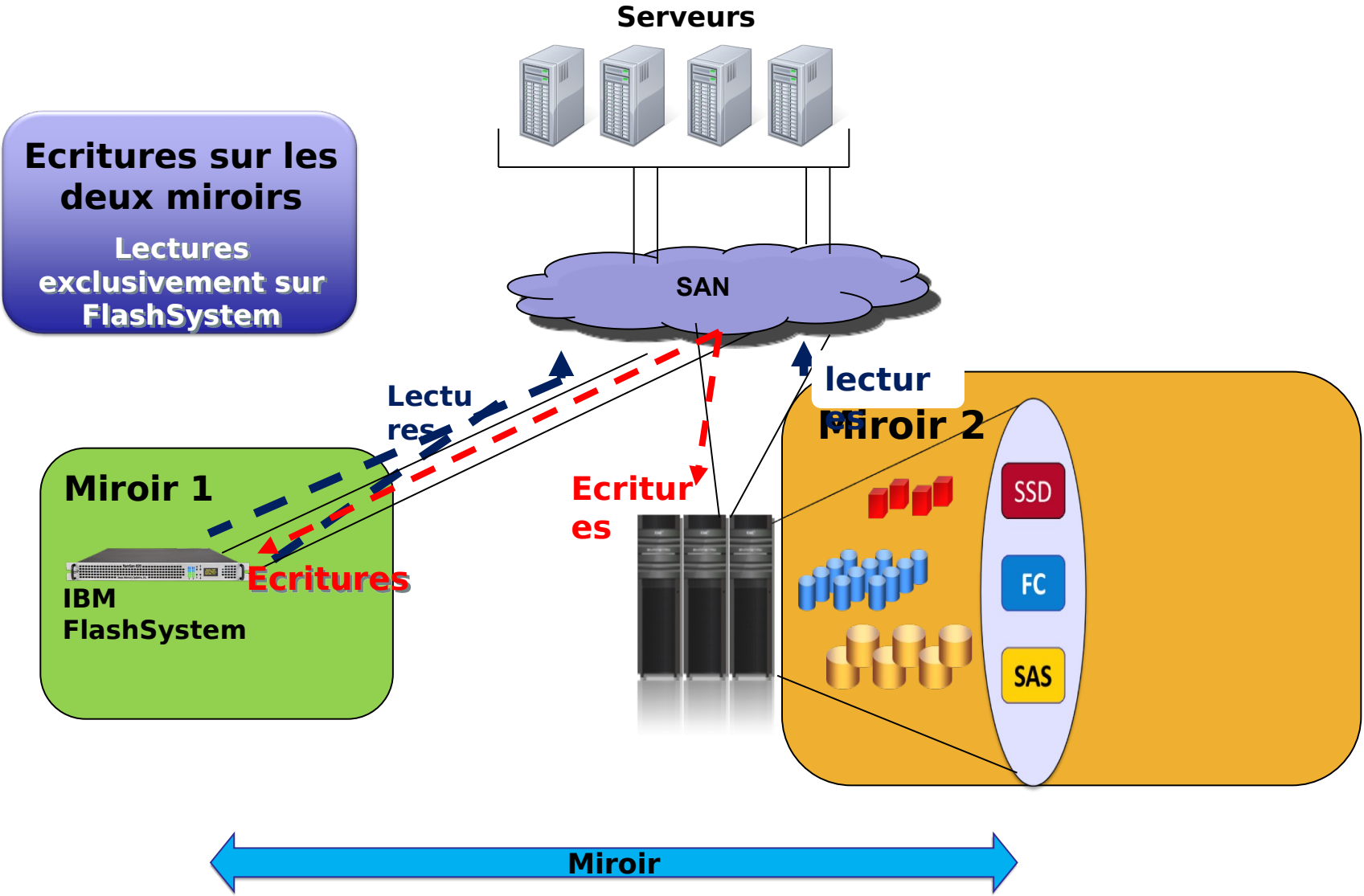
CPU plus utilisée → moins de coeurs → budget licences ISV réduit (Oracle, SQL)

# Smarter Storage for Smarter Computing: What's Next. Ready Now

1	<i>La stratégie IBM sur la technologie Flash</i>
2	<i>La proposition de valeur de FlashSystem</i>
3	<i>Exemples de mise en œuvre</i>
	<i>En parallèle du stockage existant</i>
	<i>Remplacer toute ou partie d'une baie de disque</i>
	<i>Intégration avec SVC</i>
4	<i>Gains FlashSystems et retour d'expériences</i>
5	<i>FlashSystems dans un environnement IBM i</i>
6	<i>Points techniques</i>

*IBM FlashSystem...*

*Enterprise Storage in microseconds, not milliseconds*



# Implementation Flash en parallèle du stockage disque existant



## Lectures “forcées” sur la baie Flash

Au niveau de l'OS (Volume Manager)

AIX LVM (native least queue read)

Veritas (Preferred Read Plex)

Solaris SVM

Au niveau Applicatif

Standby/Reporting Instance

SQL AlwaysOn

Oracle ASM / Oracle DataGuard



## Écritures sur la baie Flash et la baie de disque en miroir



## Cas client : Banco Azteca



### **Avant** Introduction de FlashSystems

Lectures : 8000/s  
Ecritures : 2000/s

Les performances ne sont pas suffisantes

### **Après** Introduction de FlashSystems en miroir de la baie de stockage disque primaire

Lectures : 40 000/s  
Ecritures : 10 000/s

La baie de stockage, libérée des lectures (qui se font exclusivement sur FlashSystems) accepte beaucoup plus d'écritures

**Multiplication des performances par 5**

# Différentes implémentations Flash

	Migration des données ?	Améliorations	Copie des données entre baies Flash	Nombre de baies FS	Type de baie FlashSystems
<b>En parallèle d'une baie primaire existante</b>	Non	<ul style="list-style-type: none"><li>•Lectures +++</li><li>•Ecritures (+)</li></ul>	Sans objet	Une	FSx20 avec RAID 2D



## Booste les performances applicatives et infra

- Tous les I/O 'Read' sur la baie Flash
- Augmentation des performances I/O 'Write' sur la baie existante



## Installation sans peine

- Flash en Mirroir de la baie existante
- Augmentation de la sécurité due au miroir
- Aucune introduction de risque dans l'infrastructure



## Performances améliorées de la baie en place

- Préservation des fonctions de la baie existante
- Les autres workloads du SAN inchangés
- Les fonctions de copie en place restent inchangées





# Smarter Storage for Smarter Computing: What's Next. Ready Now

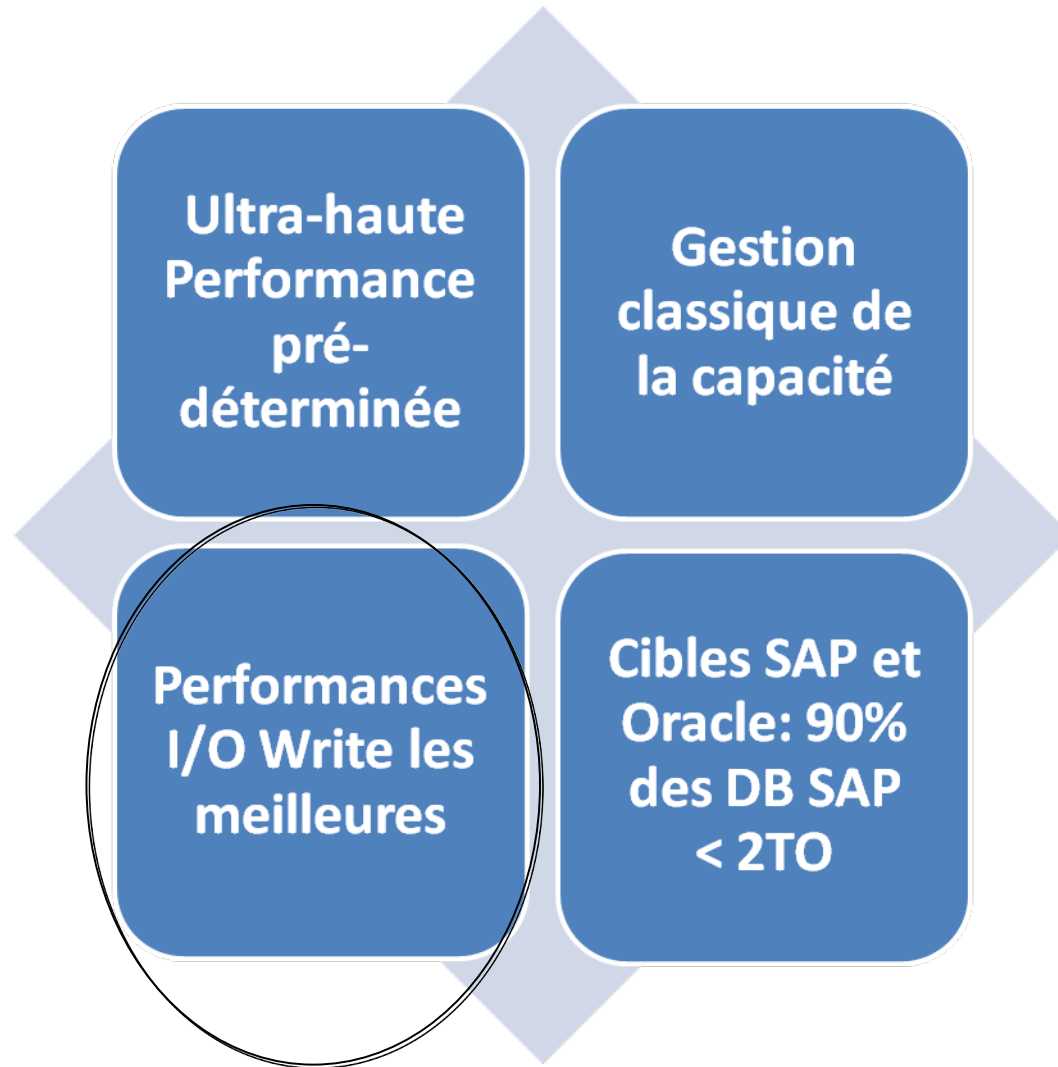
1	<i>La stratégie IBM sur la technologie Flash</i>
2	<i>La proposition de valeur de FlashSystem</i>
3	<i>Exemples de mise en œuvre</i>
	<i>En parallèle du stockage existant</i>
	<i>Remplacer toute ou partie d'une baie de disque</i>
	<i>Intégration avec SVC</i>
4	<i>Gains FlashSystems et retour d'expériences</i>
5	<i>FlashSystems dans un environnement IBM i</i>
6	<i>Points techniques</i>

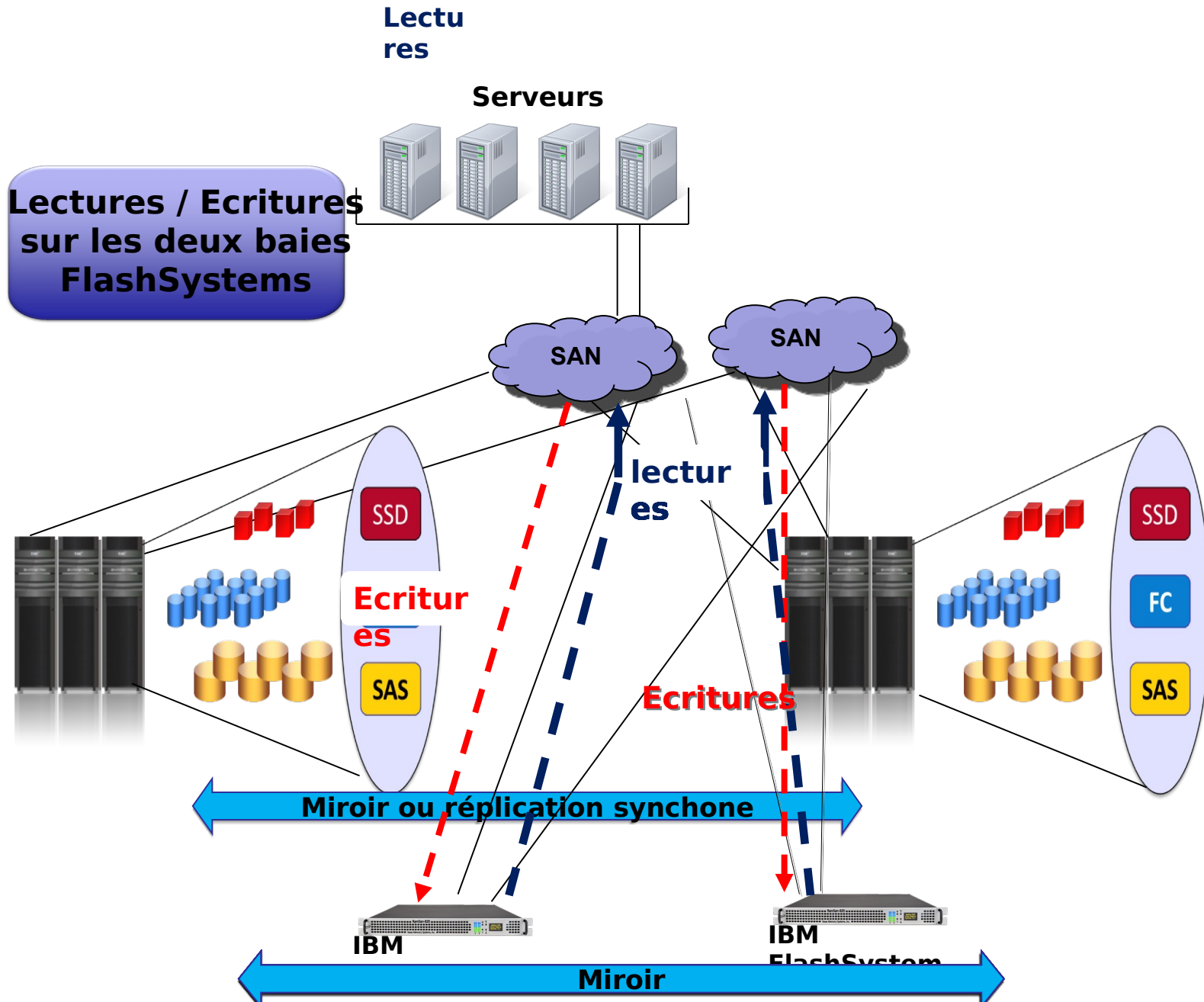


*IBM FlashSystem...*

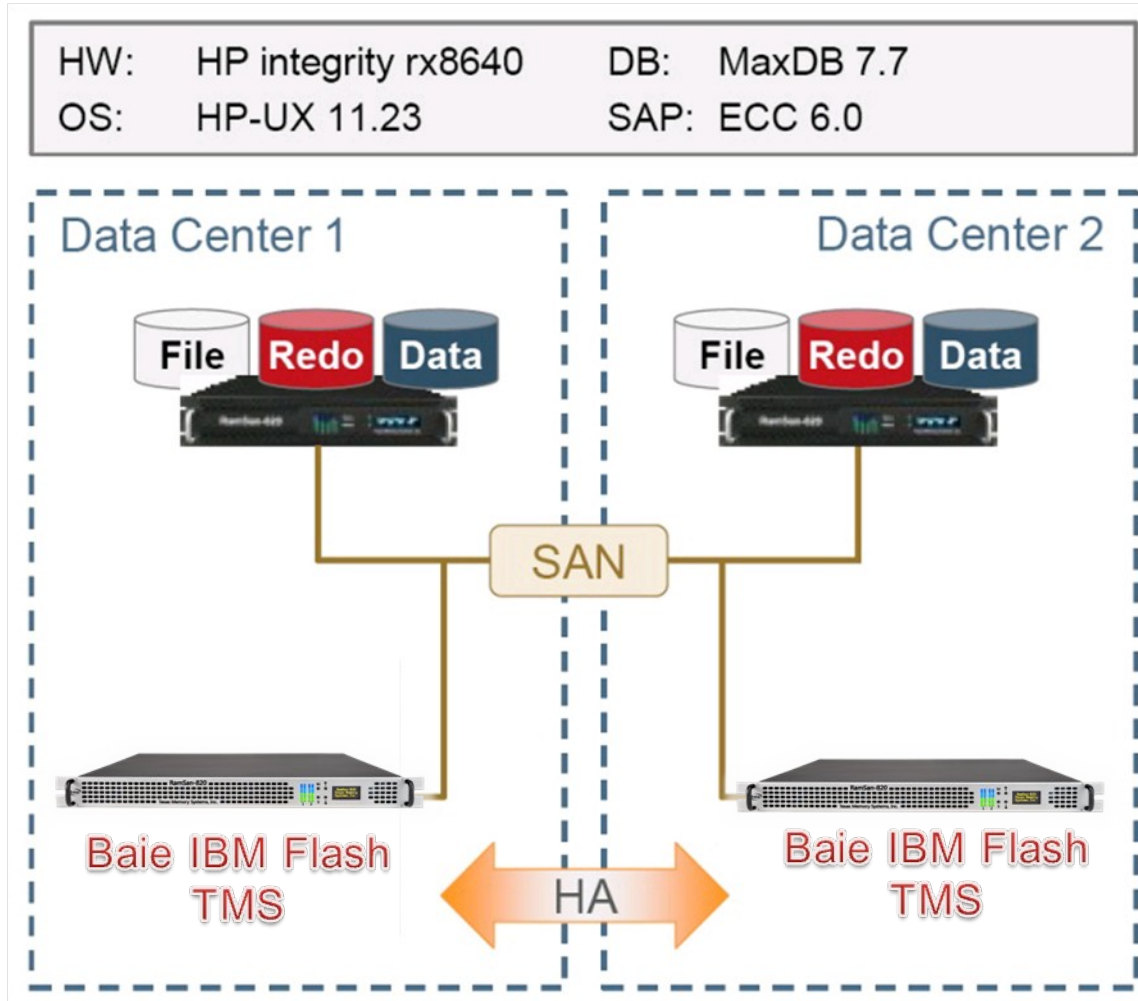
*Enterprise Storage in microseconds, not milliseconds*







# Cas client environnement SAP



L'administrateur de l'application SAP a identifié des problèmes de performances sur certains datasets de SAP

La partie des data SAP impactée a été déplacée sur 2 baies Flash TMS mirorrées

**12 x  
Performance  
increase**

# Différentes implémentations Flash

	Migration des données ?	Améliorations	Copie des données entre baies Flash	Nombre de baies FS	Type de baie FlashSystems
<b>En parallèle d'une baie primaire existante</b>	Non	<ul style="list-style-type: none"><li>•Lectures +++</li><li>•Ecritures (+)</li></ul>	Sans objet	Une	FSx20 avec RAID 2D
<b>En remplacement de disques</b>	Oui	<ul style="list-style-type: none"><li>•Lectures +++</li><li>•Ecritures +++</li></ul>	Miroir piloté par l'application ou l'operating system	Deux en miroir	FSx10 avec RAID 1D

# Smarter Storage for Smarter Computing: What's Next. Ready Now

1	<i>La stratégie IBM sur la technologie Flash</i>
2	<i>La proposition de valeur de FlashSystem</i>
3	<i>Exemples de mise en œuvre</i>
	<i>En parallèle du stockage existant</i>
	<i>Remplacer toute ou partie d'une baie de disque</i>
	<i>Intégration avec SVC</i>
4	<i>Gains FlashSystems et retour d'expériences</i>
5	<i>FlashSystems dans un environnement IBM i</i>
6	<i>Points techniques</i>

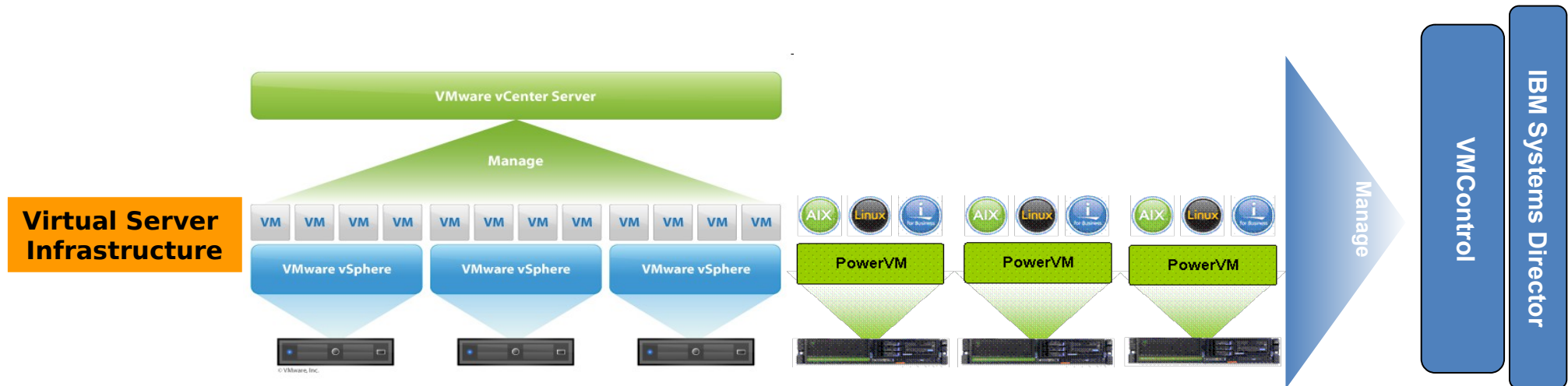


*IBM FlashSystem...*

*Enterprise Storage in microseconds, not milliseconds*



# Virtualisation : Pourquoi sur les serveurs et pas sur le stockage ?



- Notion de “**Server hypervisor**”
  - Les ressources physiques (serveurs) sont constituées en pool et sont utilisées par des machines virtuelles : **Utilisation optimisée du hardware.**
  - Les VMs sont mobiles : **Choix par l’IT des serveurs physiques les plus compétitifs.**
  - **Des fonctionnalités communes et des modes de gestion centralisée** sont disponibles pour les VMs, indépendamment du matériel.

Cela procure une réduction des coûts et une amélioration de la flexibilité

## Sur Intel

Plateforme de virtualisation : VMware vSphere, Red Hat EV, etc.

Gestion de la virtualisation : VMWare vCenter or Red Hat EV Manager, etc

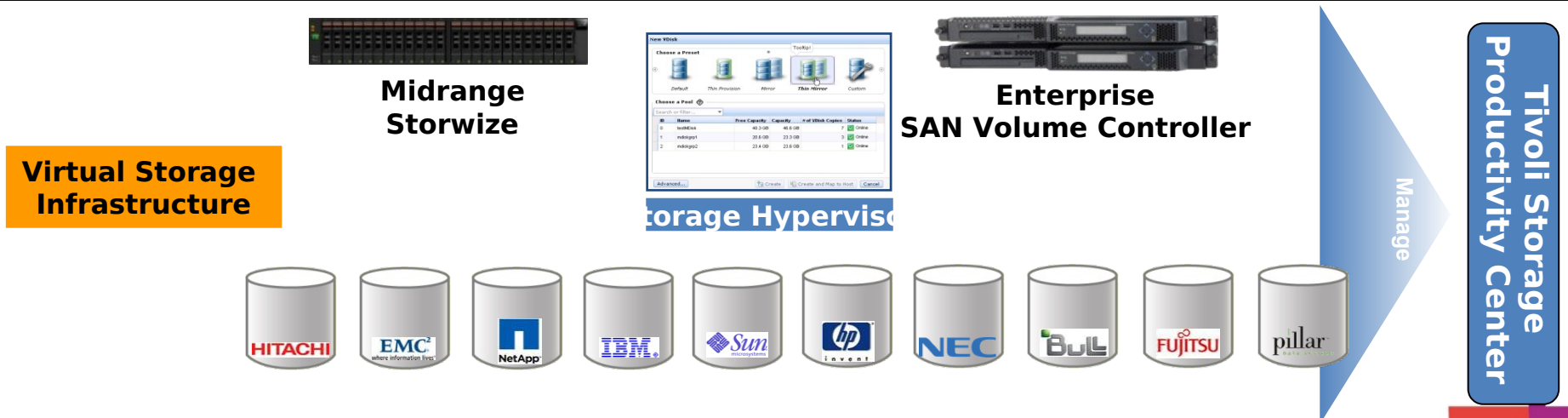
## Sur Power

Plateforme de virtualisation : IBM PowerVM, gestion de la virtualisation IBM Director VMControl

# La notion de “storage hypervisor”

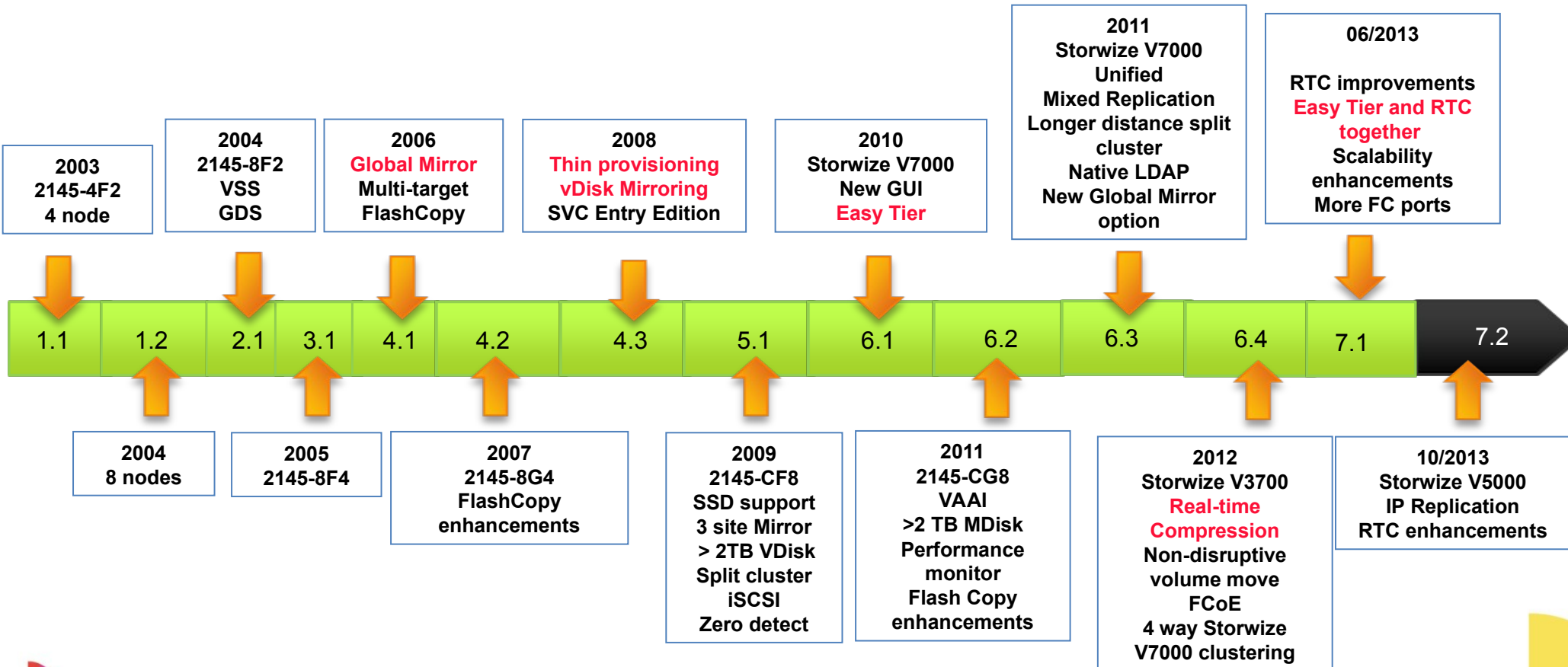
- Les ressources physiques (baies de disques) sont constituées en pool et sont utilisées par des disques virtuels : **Utilisation optimisée du hardware.**
- Les ressources disques sont banalisées : **Choix par l’IT du stockage le plus compétitif.**
- **Des fonctionnalités communes et des modes de gestion centralisée** sont disponibles pour les disques virtuels, indépendamment du matériel (Tiering, Migration online, copies locales et distantes, provisioning, supervision).

Cela procure une réduction des coûts et une amélioration de la flexibilité





# La virtualisation du stockage est au cœur de la stratégie IBM



# La famille Storwize



SAN Volume Controller



Flex System V7000



Storwize V7000 Unified



Storwize V7000



Storwize V5000

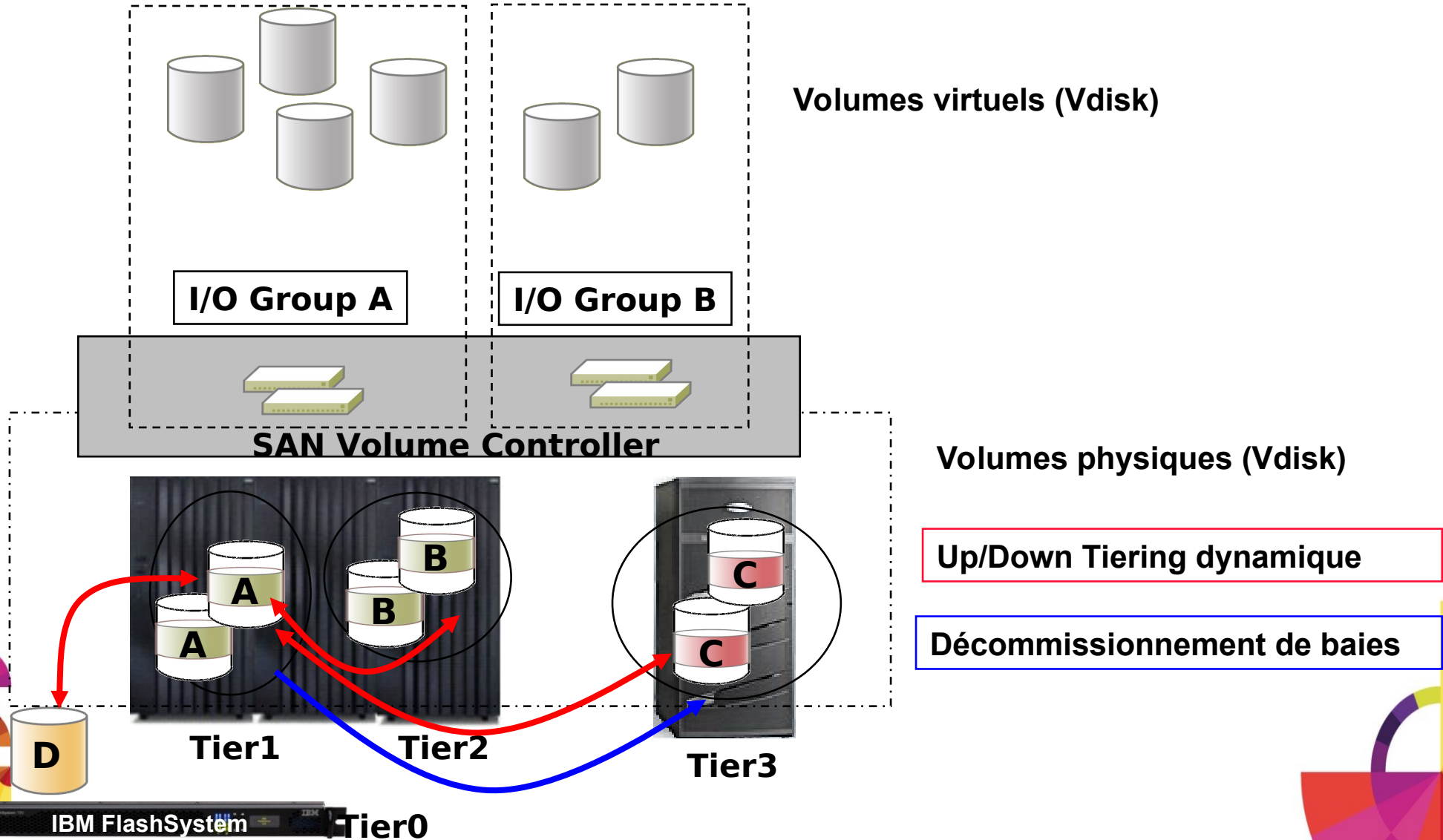


Storwize V3700



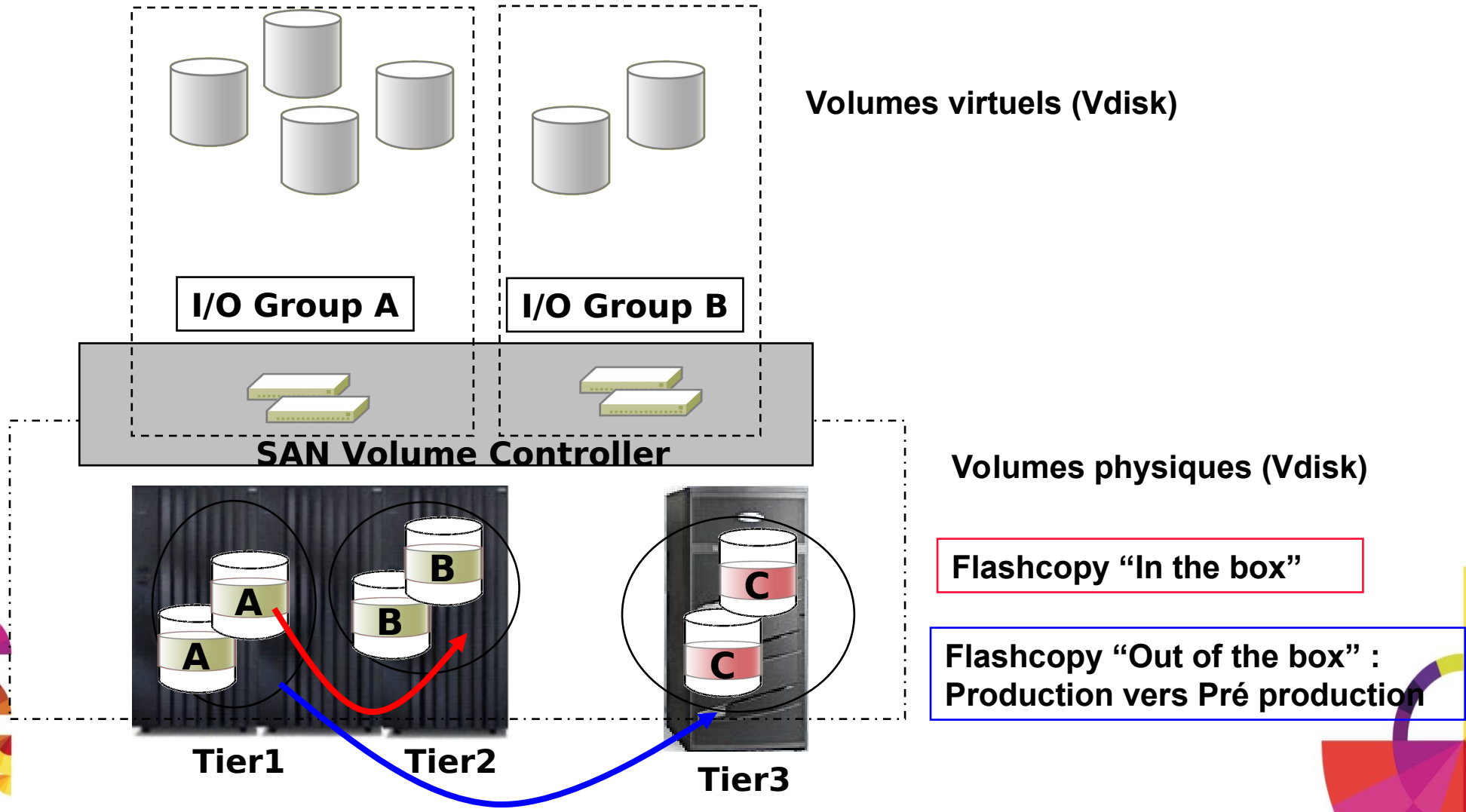
# Quelques avantages fondamentaux du SVC (1/5)

## Migration des données



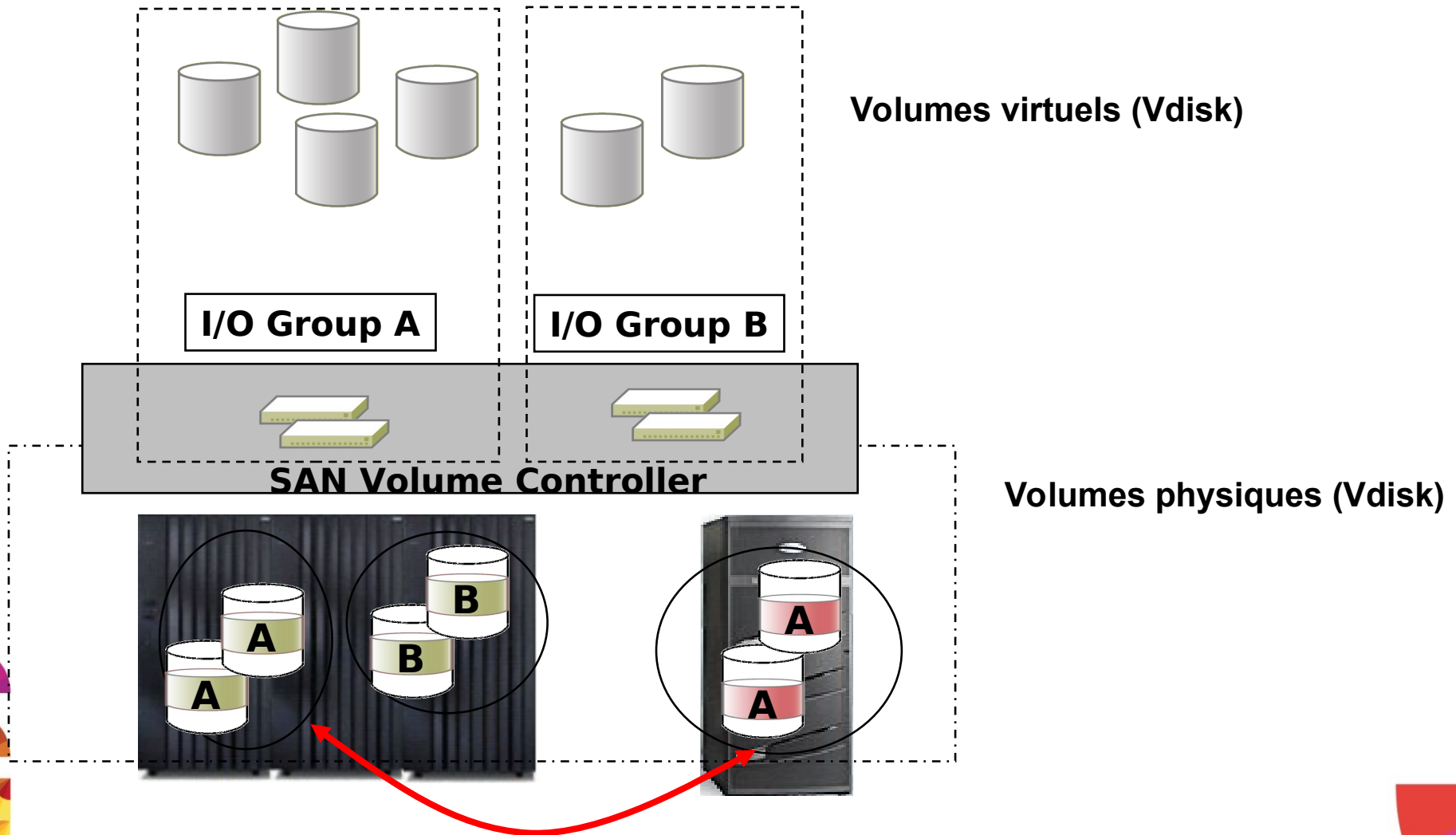
# Quelques avantages fondamentaux du SVC (2/5)

## Copies instantanées “in and out of the box”



# Quelques avantages fondamentaux du SVC (3/5)

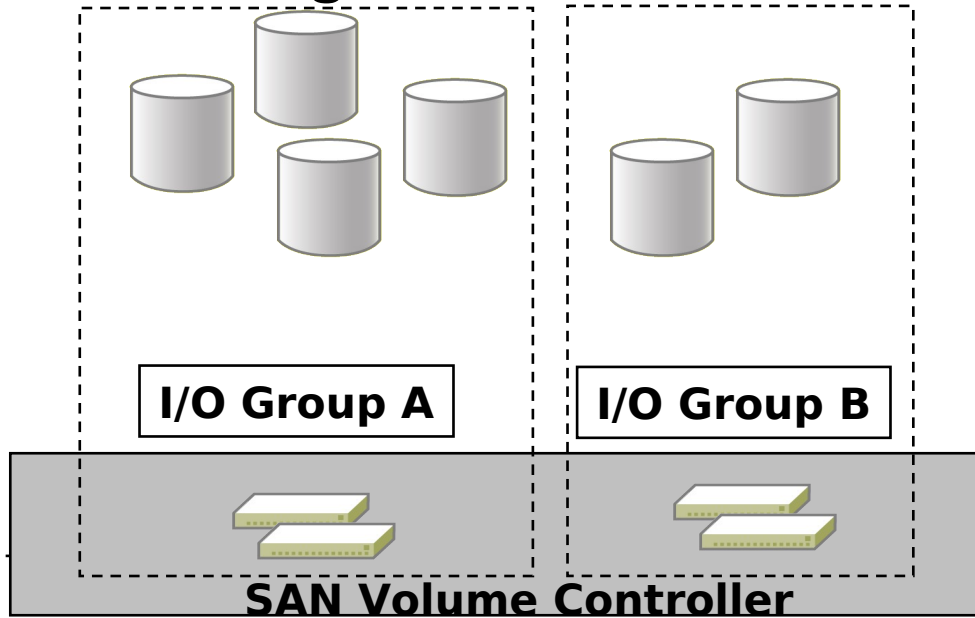
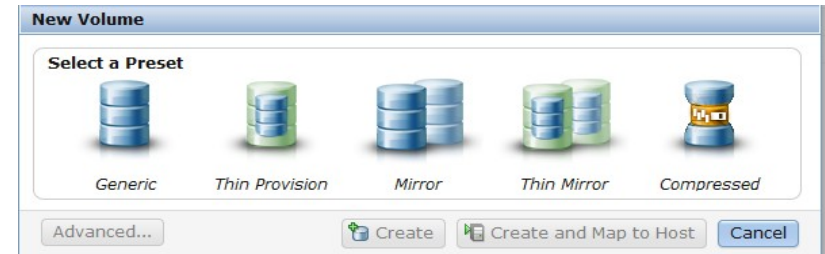
## Vdisk Mirroring



# Quelques avantages fondamentaux du SVC (4/5) Changement dynamique de type de volume basé sur Vdisk Mirroring

## Volumes virtuels (Vdisk) :

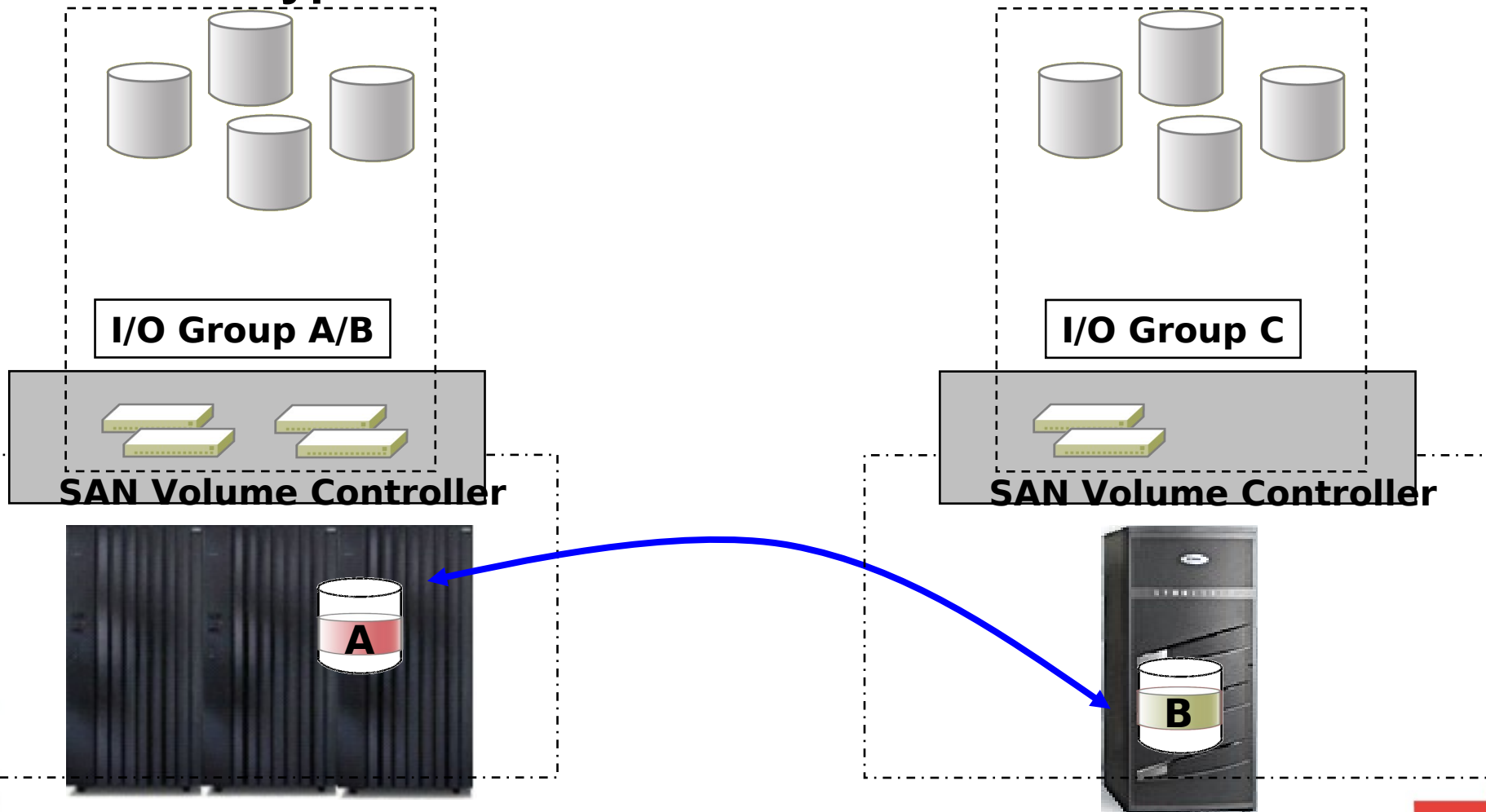
- Hard provisioning
- Thin provisioning
- Compressé



## Volumes physiques (Vdisk)



# Quelques avantages fondamentaux du SVC (5/5) Réplication (a)synchrone entre baies différentes, volumes de type différents.



# Intégrer un SVC dans une architecture existante

	Interruption de service	Effets clients
<b>Avec miroir OS (AIX LVM, ...)</b>	Non	Surcharge serveur. Synchronisation à faire hors charge nominale importante
<b>Avec migration depuis l'OS (VMware Storage vMotion, ...)</b>	Non	
<b>En mode IMAGE sur SVC</b>	Oui	Interruption de service minimale non liée à la volumétrie à migrer
<b>A froid (Robocopy Windows ou rsync Unix)</b>	Oui	Interruption de service liée à la volumétrie à migrer

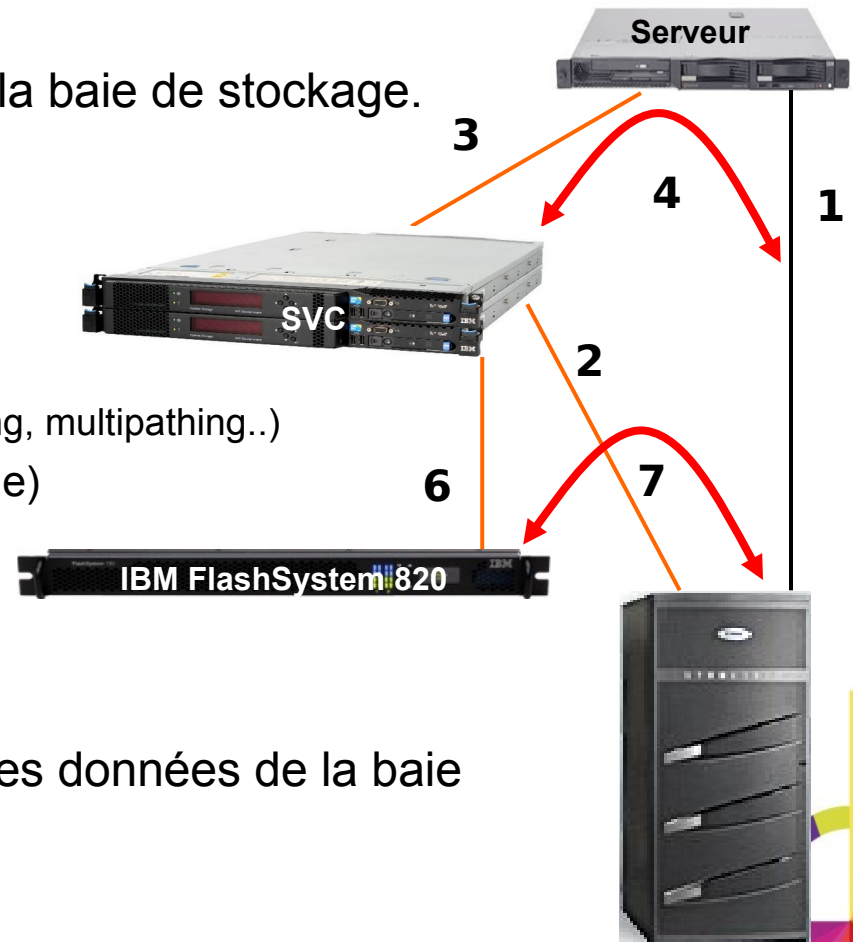




# Intégrer un SVC+Flash dans une architecture existante

## Mode Miroir OS

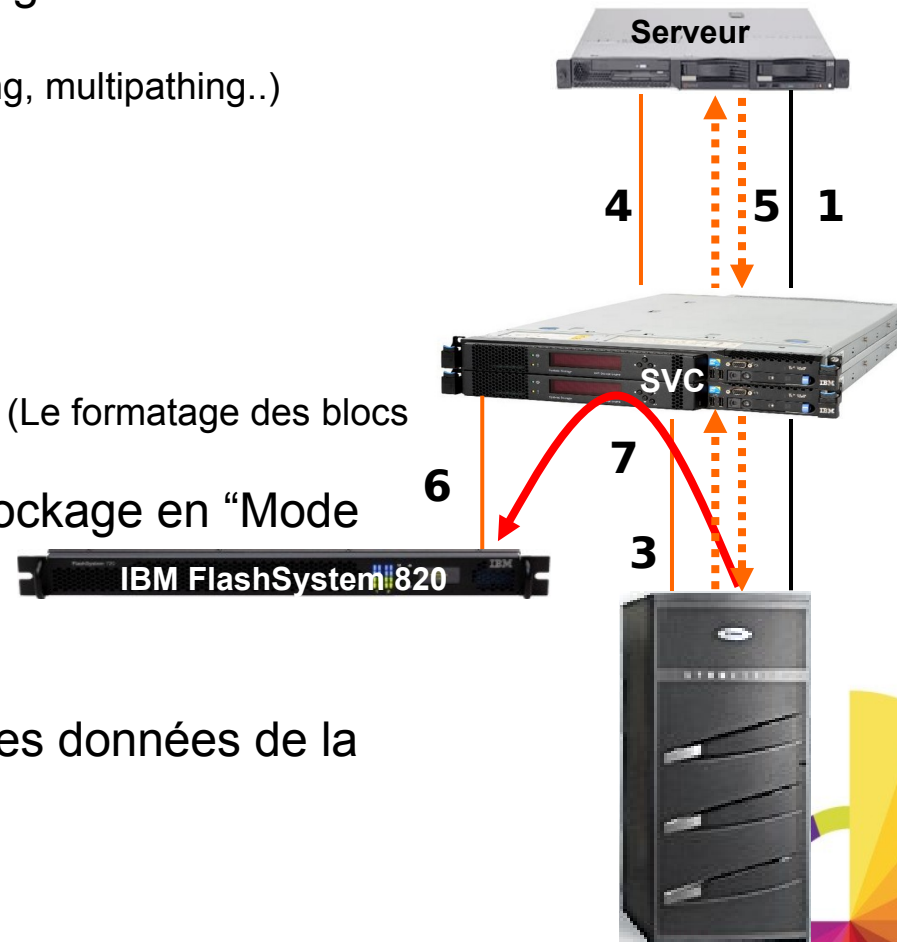
1. Le serveur accède “en direct” au LUN source de la baie de stockage.
2. Connexion de la baie de stockage au SVC.
  - Définition de SVC comme un ‘host’ pour la baie
  - Mapping de LUN cible de la baie au ‘host’ SVC
3. Connexion du serveur au SVC.
  - Carte FC séparée (si driver multipathing type Powerpath)
  - Préparation du serveur avant de le connecter au SVC (zoning, multipathing..)
4. Miroir initialisé par le serveur (AIX LVM, par exemple)
  - Lancement de la synchronisation inter volumes
5. Deconnexion du volume source de la baie
6. Connexion de la baie IBM FlashSystem au SVC.
  - Création de LUNs et mapping sur le SVC
7. Plusieurs options non-disruptives pour déplacer les données de la baie de stockage vers la baie FlashSystem
  - Migration (au niveau volume)
  - Volume Mirroring (au niveau volume)
  - Easy Tiering (au niveau extent)



# Intégrer un SVC+Flash dans une architecture existante

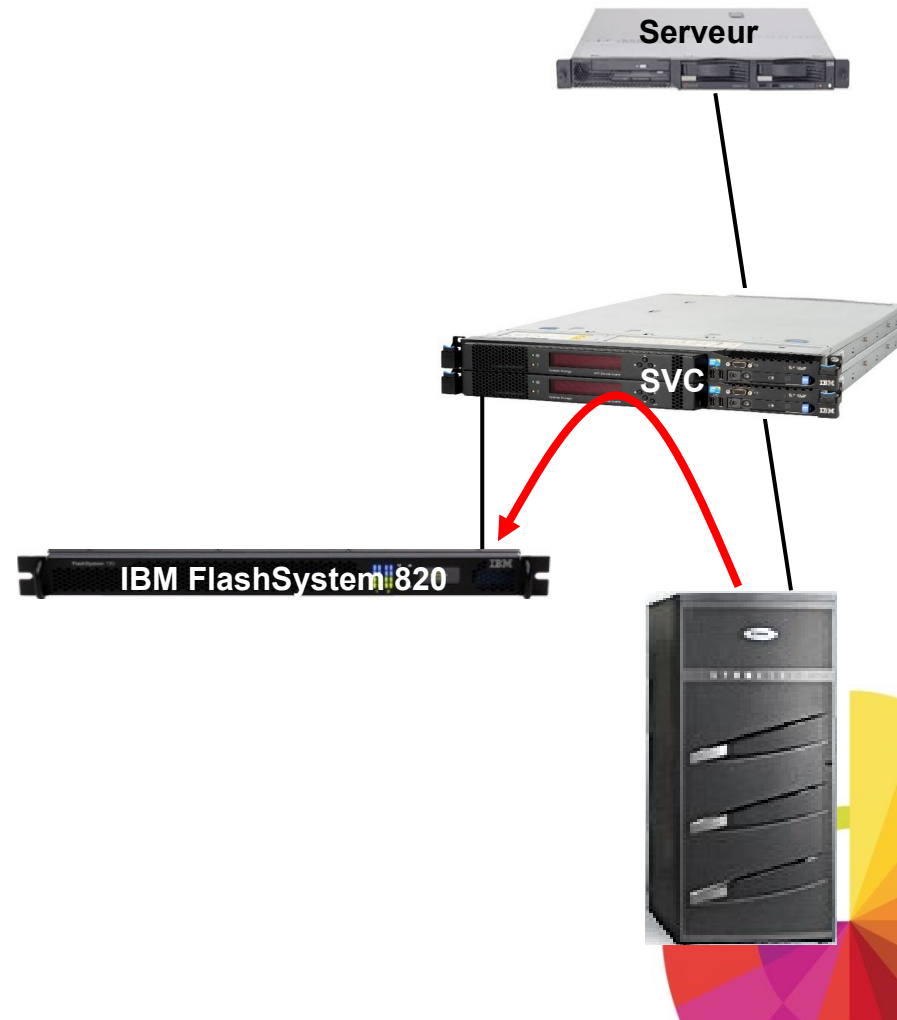
## Mode IMAGE

1. Le serveur accède “en direct” aux LUNs de la baie de stockage.
2. Déconnexion entre le serveur et la baie de stockage
  - UNMAP des LUNs
  - Préparation du serveur avant de le connecter au SVC (zoning, multipathing..)
3. Connexion de la baie de stockage au SVC.
  - Définition de SVC comme un ‘host’ pour la baie
  - Mapping des LUNs de la baie au ‘host’ SVC
  - Import des volumes en “Mode Image”
4. Connexion du serveur au SVC.
  - Mapping des volumes SVC en “Mode Image” sur le serveur (Le formatage des blocs reste le même)
5. Le serveur accède aux volumes sur la baie de stockage en “Mode Image” et peut repartir en production.
6. Connexion de la baie IBM FlashSystem au SVC.
  - Création de LUNs et mapping sur le SVC
7. Plusieurs options non-disruptives pour déplacer les données de la baie de stockage vers la baie FlashSystem
  - Migration (au niveau volume)
  - Volume Mirroring (au niveau volume)
  - Easy Tiering (au niveau extent)



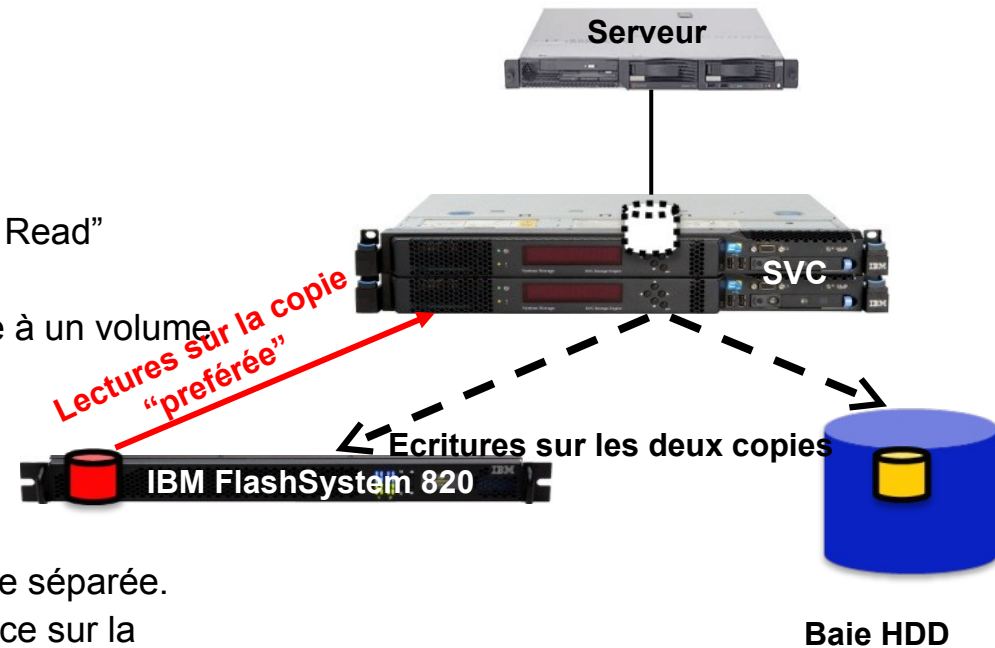
# Migration des données dans l'infrastructure virtualisée

- Une fois que le serveur est connecté au SVC, les données auxquelles il accède peuvent être déplacées de façon transparente d'un type de stockage à un autre, sous contrôle du virtualiseur :
  - Sans une intervention du serveur ou des applications.
  - Entre des systèmes de stockage hétérogènes.
- Les candidats à résider de façon permanente en Flash sont :
  - Logs
  - Temp-db
  - Indexes
  - Méta données
  - Base de données complète
- De façon temporaire et basée en général sur un calendrier:
  - Applications avec des échéances mensuelles/trimestrielles/...
  - Applications ayant un pic de charge à différents moments.



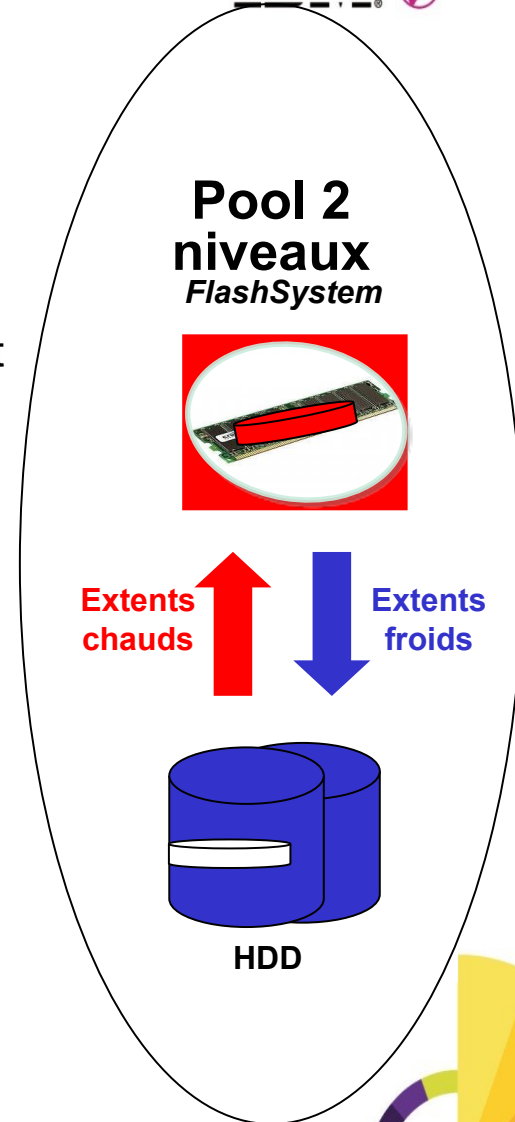
# SVC Volume Mirroring & Preferred Read

- Benefices : performance, haute disponibilité
- SVC gère deux copies du volume (copy0/copy1)
  - Lectures exclusivement depuis copy (c0) - "Preferred Read"
  - Ecritures sur les deux copies (c0 and c1)
  - Facilité de mise en oeuvre: il suffit d'ajouter une copie à un volume existant
- Un seul volume présenté au serveur
- Une fonction de haute disponibilité locale
  - Chaque copie du volume est sur une baie de stockage séparée.
  - Si une copie est indisponible, il y a continuité de service sur la seconde copie
  - Les copies sont automatiquement re synchronisées quand les disques/Flash sont à nouveau disponibles.
- Les copies peuvent être séparées.
  - Chacune peut fonctionner en production, de façon séparée.
- Les copies peuvent être en thin provisioning ou en mode compressé, de façon indépendante.
  - Utilisation optimisée de FlashSystem



## SVC Easy Tier

- Bénéfices : Performance et optimisation de la volumétrie FlashSystems nécessaire.
- Mouvement automatique, au niveau extent (un volume virtuel est constitué d'une concaténation d'extents) de données "chaudes" vers FlashSystems et de données "froides" vers HDD.
  - Réduction des coûts – FlashSystem utilisé seulement pour les données "chaudes".
  - La plupart des workloads ont moins de 10% de points chauds ("skew")
  - Facilité de Mise en Oeuvre – constitution d'un pool "hybride" constitué de LUNs FlashSystem et HDD.
- SVC Easy Tier
  - Peut être dynamiquement activé/désactivé pour un pool et/ou un volume
  - Un volume peut être copié/migré dedans/dehors du pool Easy Tier de façon non disruptive.
  - Evite aux administrateurs systèmes de passer du temps dans l'identification des points chauds.
  - Evite aux DBA de passer du temps dans l'identification des sources des problèmes de performance : Redo logs, Log buffers, etc.



# Historique et tendances de la compression chez IBM

Acquisition de la société Storwize et de la technologie Real Time Compression par IBM en juillet 2010



Sortie des appliances de compression NAS STN6500 et STN6800 en 2011

RTCa  
(Real Time Compression appliance)



LAN  
1Gbe-10Gbe



Filer NAS  
CIFS-NFS

Compression SAN dans V7K/SVC  
Compression NAS dans V7000U



Baie de stockage SAN

Acquisition de TMS en Septembre 2012



SAN

SAN

SAN



Pools hybrides EasyTier



# Taux de compression des données

- Les taux de compression maximum observés sont les suivants. Ils dépendent fortement du type de données.

<b>Databases</b>		<b>Up to 80%</b>
<b>Server Virtualization</b>	<b>Linux virtual OSes</b>	<b>Up to 70%</b>
	<b>Windows virtual OSes</b>	<b>Up to 55%</b>
<b>Collaboration</b>	<b>Office 2003</b>	<b>Up to 75%</b>
	<b>Office 2007 or later</b>	<b>Up to 25%</b>
<b>CAD/CAM</b>		<b>Up to 75%</b>

- L'outil *Comprestimator* permet d'obtenir de façon fiable le taux de compression que l'on pourra obtenir avec SVC.
- Download : <http://www-01.ibm.com/support/docview.wss?uid=ssg1S4001012>

## Différents cas d'usage de FlashSystem sous SVC

**Avec ou sans utilisation des fonctions de réplication inter SVC (Metro/Global Mirror)**

**Pool hybride Flash+HDD avec Easy Tier**

**En pool 100% Flash très haute performance**

**En miroir entre Flash et disques avec "preferred read" et SVC Vdisk Mirroring**

**Avec ou sans activation de la compression temps réel (RTC), ou des autres fonctions avancées (Thin Provisioning, Flashcopy**

**g, Flashcopy**





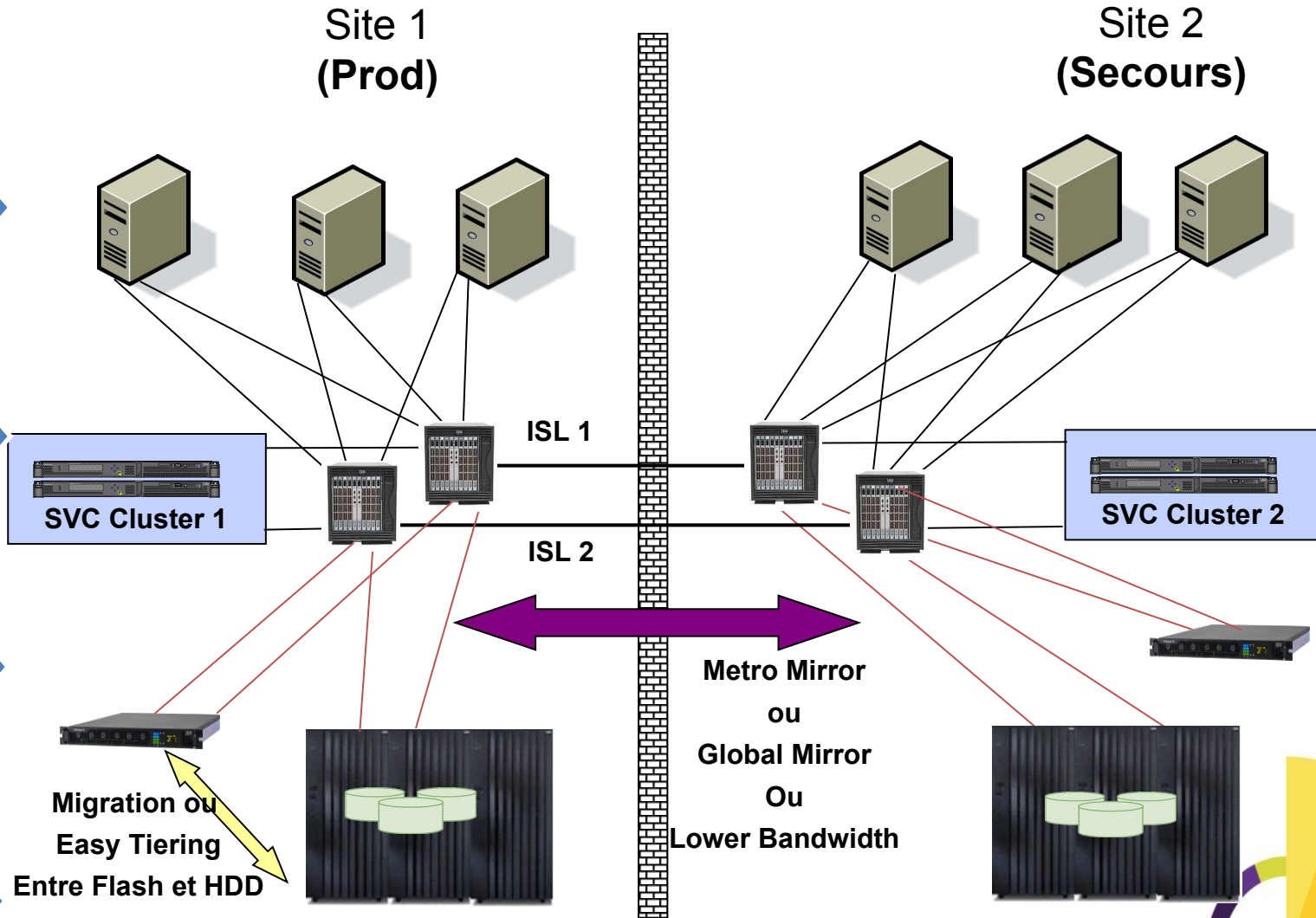
# Fonctionnement de deux SVCs en Failover / Failback

EasyTier déplace les données chaudes automatiquement

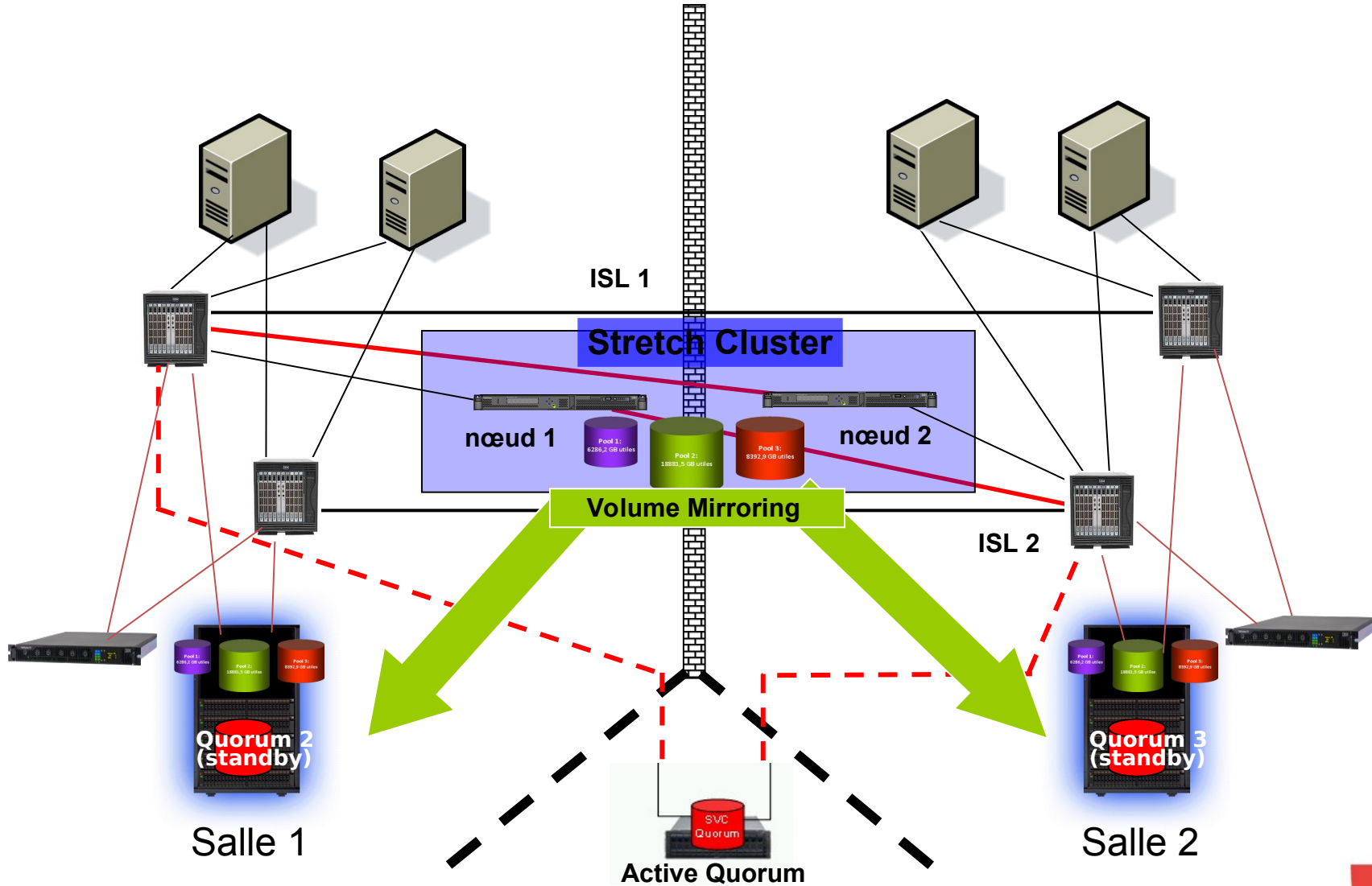
Accélération des applications ayant les volumes dans le même mdisk pool

Performances inégalées et 'scalable' par cluster SVC

Fonctions avancées (Copy Services, Thin provisioning, compression...)



# Fonctionnement en PCA avec un SVC en « Stretch cluster »



## Différentes implémentations Flash

	Migration des données ?	Améliorations	Copie des données entre baies Flash	Nombre de baies FS	Type de baie FlashSystems
<b>En parallèle d'une baie primaire existante</b>	Non	<ul style="list-style-type: none"> <li>• Lectures +++</li> <li>• Ecritures (+)</li> </ul>	Sans objet	Une	FSx20 avec RAID 2D
<b>En remplacement de disques</b>	Oui	<ul style="list-style-type: none"> <li>• Lectures +++</li> <li>• Ecritures +++</li> </ul>	Miroir piloté par l'application ou l'operating system	Deux en miroir	FSx10 avec RAID 1D
<b>Sous SVC</b>	Oui si SVC absent initialement	<ul style="list-style-type: none"> <li>• Lectures ++</li> <li>• Ecritures ++</li> <li>• Place occupée (si RTC ou thin provisioning)</li> <li>• Easy Tiering</li> <li>• Fonctions avancées du SVC</li> </ul>	Miroir ou réplication (a)synchrone du SVC	Deux en miroir ou en réplication	FSx10 avec RAID 1D

# Matrice de compatibilité

## Connexion de FlashSystems sur le SAN

- AIX 6.1 TL7 & TL8 + IV38225 (support par MPIO des volumes Flash)
- AIX 7.1 TL2 +IV38226 (support par MPIO des volumes Flash)
- VIOS 2.2.2 (w/ client AIX 6.1, AIX 7.1, RedHat 5, RedHat 6)
- Windows Server 2008 R2 SP1+
- Windows Server 2012
- SUSE ES 10 SP4
- SUSE ES 11 SP2
- RedHat EL 5.9
- RedHat EL 6.3 & 6.4
- VMware ESX 4.1 U3
- VMware ESX 5.1

## Connexion de FlashSystems sous SVC

- SVC 6.4.1.5 et versions ultérieures

<http://www-03.ibm.com/systems/support/storage/ssic/interoperability.wss>



# Smarter Storage for Smarter Computing: What's Next. Ready Now

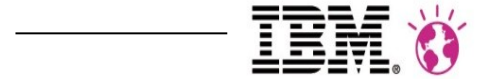
1	<i>La stratégie IBM sur la technologie Flash</i>
2	<i>La proposition de valeur de FlashSystem</i>
3	<i>Exemples de mise en œuvre</i>
4	<i>Gains FlashSystems et retour d'expériences</i>
5	<i>FlashSystems dans un environnement IBM i</i>
6	<i>Points techniques</i>



*IBM FlashSystem...*

*Enterprise Storage in microseconds, not milliseconds*

# Analyses AWR Oracle - Analyses DB2



Measures we can improve	Wait time accounted for	Current response time (microseconds)	Projected response time (microseconds)	New wait time
read by other session	23.00%	1557	400	5.91%
db file sequential read	17.59%	2899	400	2.43%
db file scattered read	7.42%	2764	400	1.07%
db file parallel write	0.89%	553	200	0.32%
log file parallel write	0.05%	1503	200	0.01%

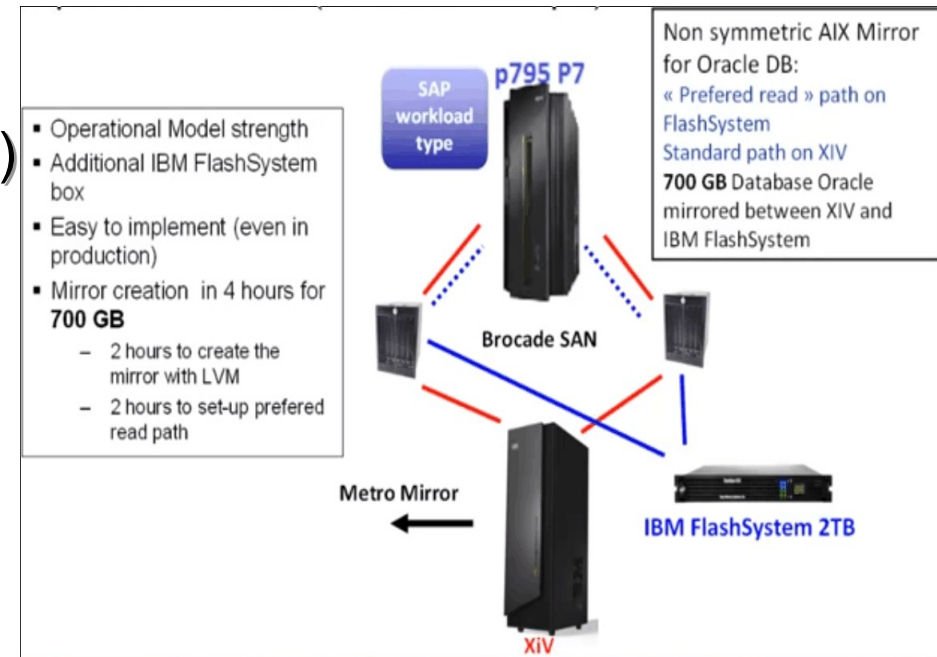


## Le POC sur site client ou à IBM Montpellier

- Parc de prêt du distributeur (Azlan/Arrow) ou du BP (DFI, SCC, ACMI, INETD, etc.)
- Parc de prêt IBM : 2 x FS720 5TB
- Montpellier : 1 x FS710 10TB



- Caractéristiques d'un POC :
  - Simple (1 baie en preferred read)
  - Court (7j-10j)
  - Non intrusif (preferred read)



# Les différentiateurs IBM

- **Performance : Faire un bench ou un POC, est la meilleure des réponses.**
  - Cela permet de mettre en évidence ce que l'on ne voit pas dans les brochures : qualité du Garbage Collection, limitation du Write Amplification, etc.
- **Haute disponibilité : Eviter les opérations de maintenance sans diminuer la performance.**
  - Panne d'un chip : VSR (Variable Stripe RAID) : RAID 5 adaptatif entre les chips d'un module.
  - Panne d'un module : RAID 5 avec spare entre les modules Flash.
- **Bundle avec SVC (SAN Volume Controller).**
  - Optimisation des achats de FlashSystems :
    - RTC (compression temps réel) des bases de données (70-80%), environnements virtualisés (60-70%) de production.
    - Easy Tiering : Ne sont montés en Flash que les blocs chauds.
    - La conjugaison de RTC et Easy Tiering.
  - Fonctions avancées du SVC :
    - Migration transparente d'un pool vers un autre
    - Réplication (a)synchrone



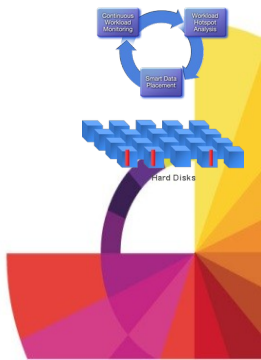
Compressed



Solid-State Storage



Hard Disk





En parallèle  
du stockage  
existant

# Synthèse de performances IBM Flash System en environnement Oracle

**3 indicateurs mesurés en PRODUCTION et significatifs pour apprécier l'impact de IBM Flash :**

- **Pour les workloads avec reads intensifs :**  
→ les temps de réponse sont 20 fois plus rapides avec IBM FlashSystem
- **Requêtes SQL Oracle (reads/writes)**  
→ elles s'exécutent **2,5 fois plus rapidement avec IBM FlashSystem**  
\*\*\*\* Mesures effectuées pendant 1h sur l'ensemble de requêtes Oracle exécutées sur serveur Unix 22 coeurs, quel que soit le profil (ETL, batch, interactif, etc.)
- **Batch de 1h30 dont 30% du temps passé en Wait IOs**  
→ ce batch dure seulement **63 minutes avec IBM FlashSystem**



# Performance Scenario: Oracle RAC, 4 Nodes

Enterprise Array, No Flash

IBM

## FlashSystem

2 million queries  
12.25 minutes to complete

16K Total IOPS  
4K per RAC Node

```
[oracle]$ time ./spawn_50.sh
```

```
real    12m15.434s
user    0m5.464s
sys     0m4.031s
```

2 million queries  
**1.3 minutes to complete**

**160K Total IOPS**  
**40K per RAC Node**

```
[oracle]$ time ./spawn_50.sh
```

```
real    1m19.838s
user    0m4.439s
sys     0m3.215s
```

**Un facteur 9x d'accélération!**



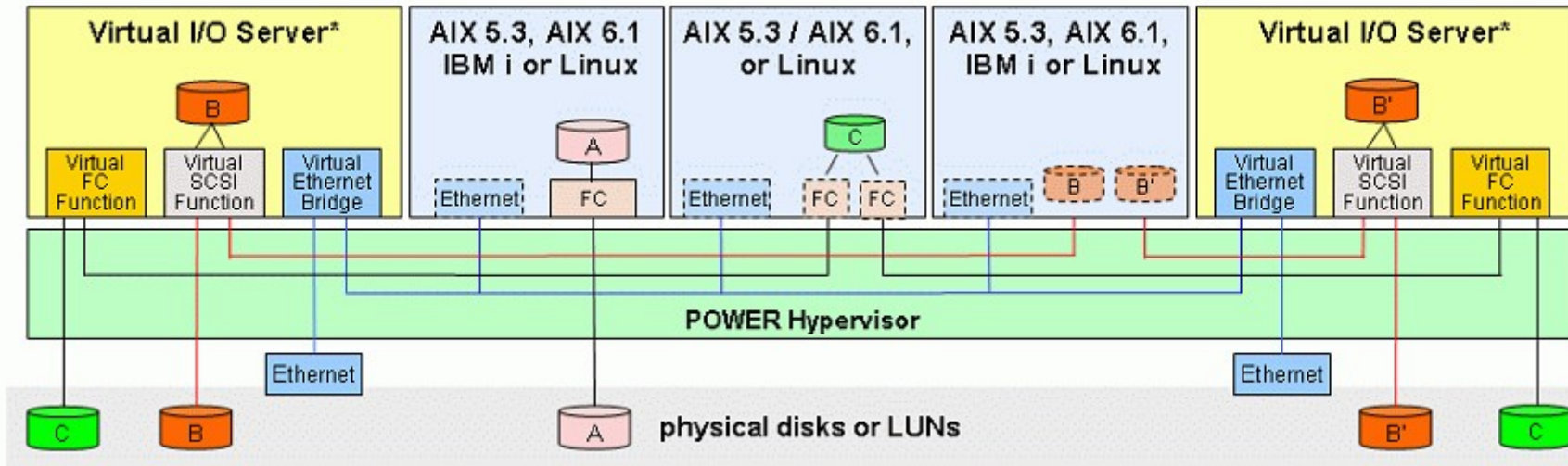
# Smarter Storage for Smarter Computing: What's Next. Ready Now

1	<i>La stratégie IBM sur la technologie Flash</i>
2	<i>La proposition de valeur de FlashSystem</i>
3	<i>Exemples de mise en œuvre</i>
4	<i>Gains FlashSystems et retour d'expériences</i>
5	<i>FlashSystems dans un environnement IBM i</i>
6	<i>Points techniques</i>



*IBM FlashSystem...  
Enterprise Storage in microseconds, not milliseconds*

## Spécificités IBM i: la virtualisation à travers le VIOS (1/2)

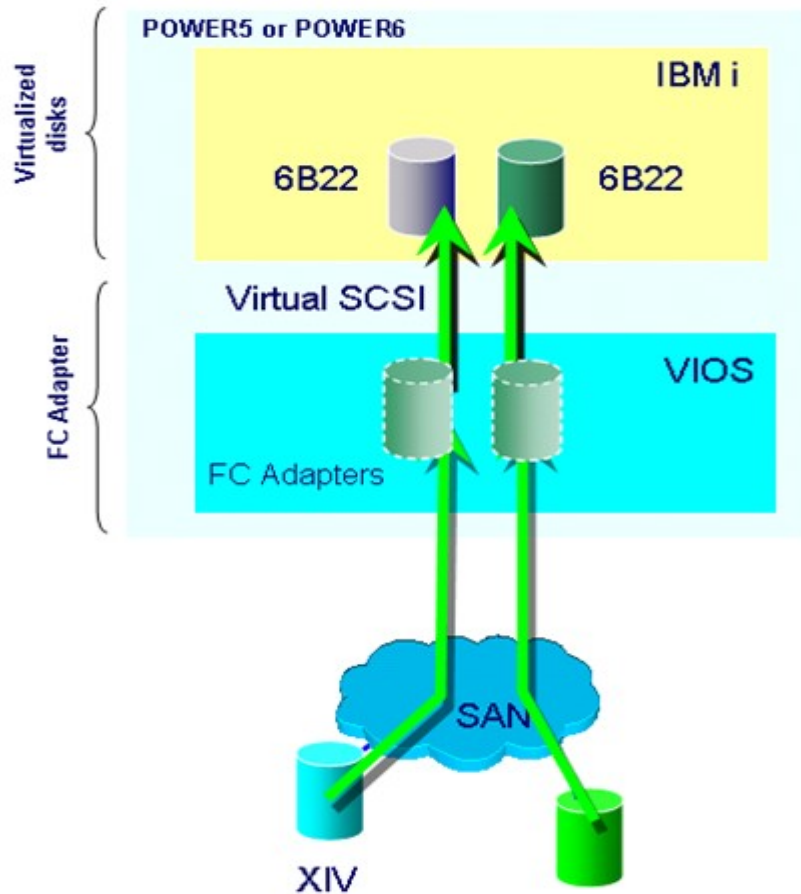


Connexion aux périphériques externes :

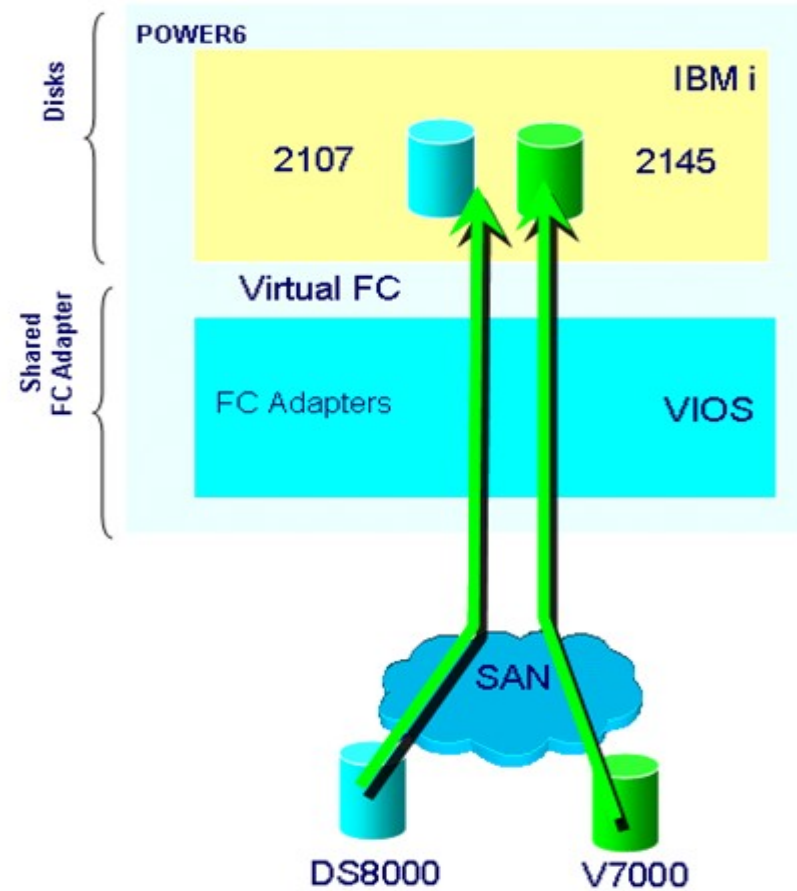
- En attachements direct FC ou via switches.
- A travers un ou plusieurs VIO servers :
  - Virtual SCSI ou vSCSI
  - Virtual Fibre Channel using NPIV (Node Port ID Virtualization)
  - Virtual Ethernet bridge using Shared Ethernet Adapter (SEA)

# Spécificités IBM i : la virtualisation à travers le VIOS (2/2)

## Virtual SCSI

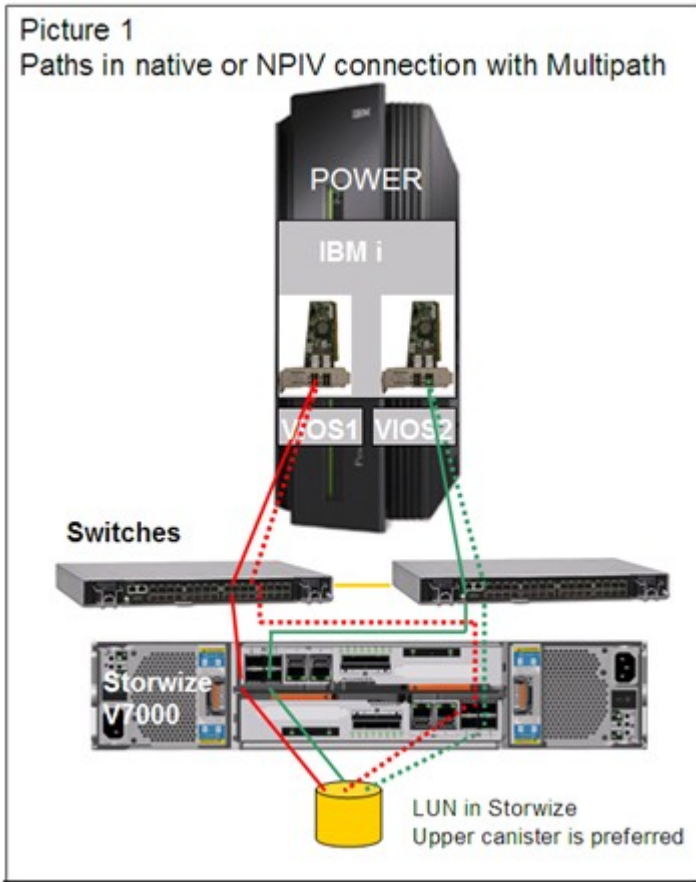


## N-Port ID Virtualization

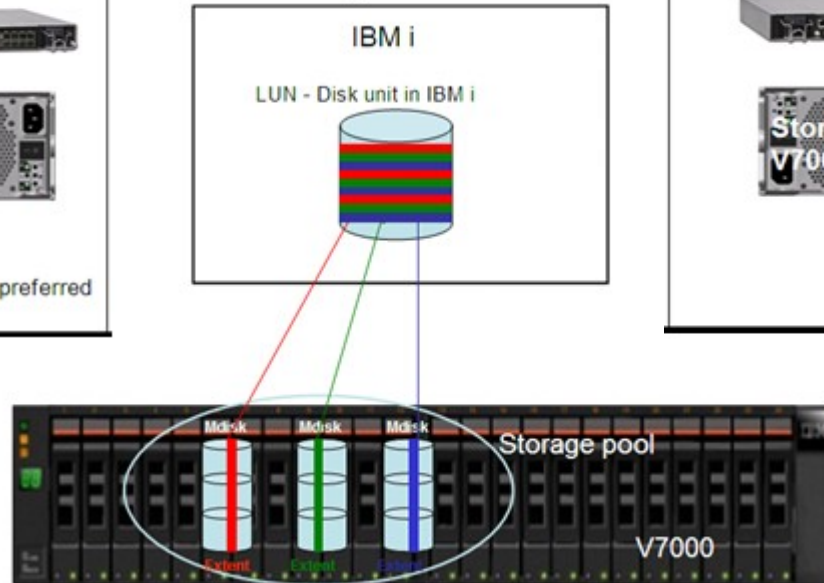
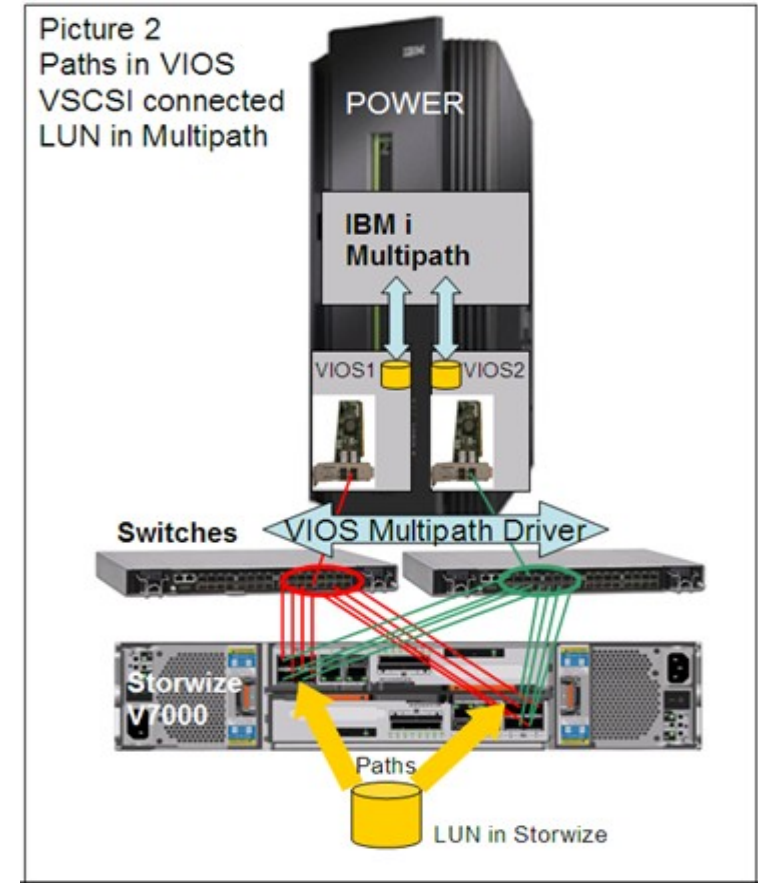


# La virtualisation de bout en bout : Serveur et stockage

Picture 1  
Paths in native or NPIV connection with Multipath



Picture 2  
Paths in VIOS  
VSCSI connected  
LUN in Multipath



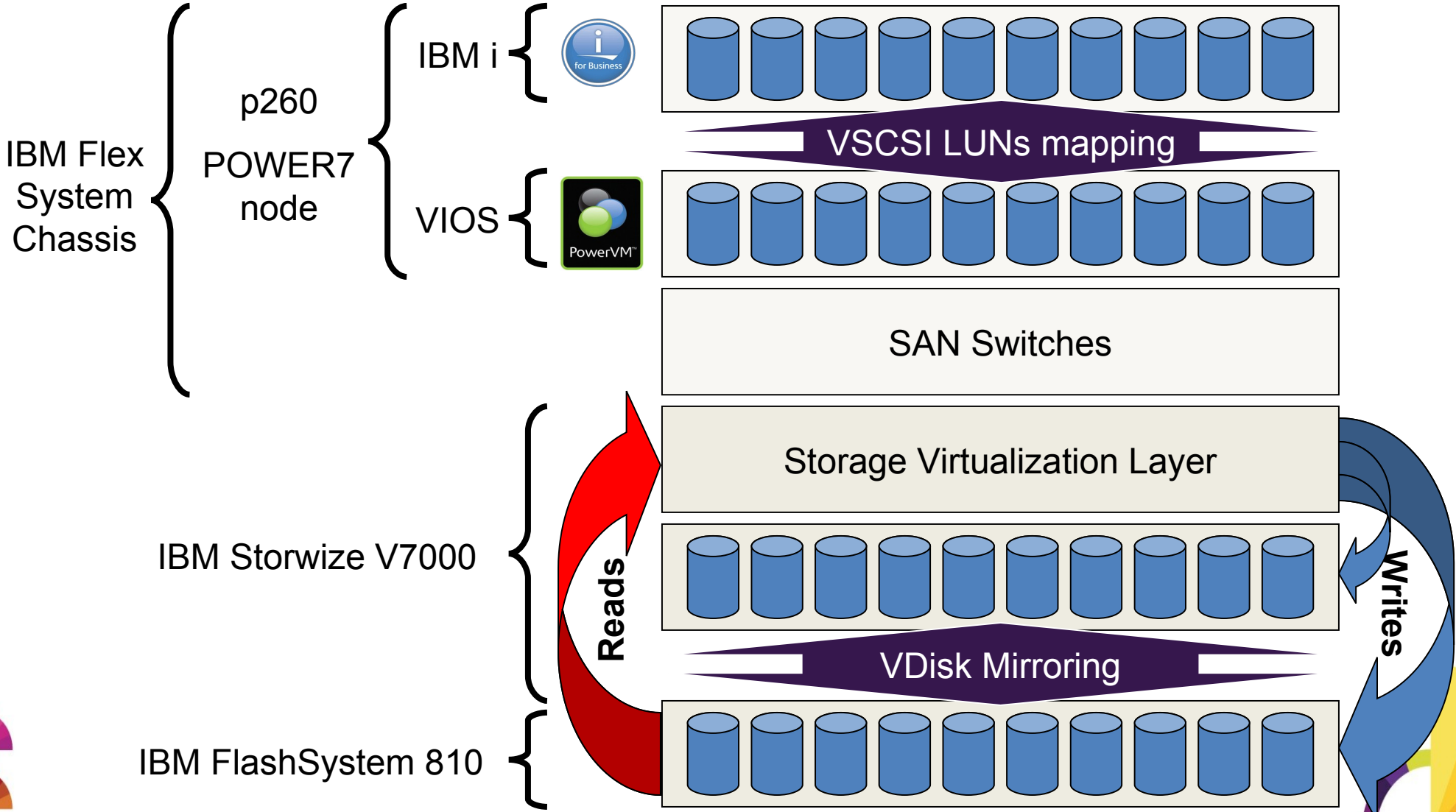
# IBM i matrice de compatibilité

Last updated September 23, 2013		Rack / Tower Systems								Notes	
		Server model			IBM i version		IBM i attach				
Storage Family		POWER5	POWER6	POWER7	6.1	7.1	Direct	Native	VIO S vS CSI	VIO S NPI V	
Storwize	V3700	⊘	✓	✓	✓	✓	✓	✓	✓	✓	<p><b>NPIV support requires IBM i 7.1 TR6 or later</b></p> <p><b>NATIVE support requires</b></p> <ul style="list-style-type: none"> <li>- POWER7 or POWER7+ servers</li> <li>- IBM i 7.1 TR6 plus PTFs MF56600, MF56753, MF56854 or their supersedes</li> <li>- and 6.4.1.4 or later SVCN3700/V7000 firmware.</li> </ul>
	V7000	⊘	✓	✓	✓	✓	✓	✓	✓	✓	
SVC		⊘	✓	✓	✓	✓	✓	✓	✓	✓	<p><b>DIRECT attach available for <u>POWER7</u> systems when utilizing PCIe 4Gb fibre channel adapters (feature 5774 or 5276)</b></p>
Flash Systems		⊘	✓	✓	✓	✓	✓	✓	✓	✓	<p><b>Flash Systems require either SVC or Storwize.</b> SVC/Storwize firmware level 6.4.1.5 or higher is required. See support documentation at the URLs listed on the notes page for additional information. VIOS VSCSI virtualization without SVC or Storwize is <u>not</u> supported.</p>

These tables do not include detailed considerations. For example, server firmware levels or required PTFs are not included.

- ✓ **POWER7** servers require IBM i licensed internal code (LIC) level 6.1.1 or later.
- ✓ **VIOS** support requires IBM i 6.1 or later and POWER6 or later server technology.
- ✓ **NPIV** requires IBM i LIC 6.1.1 or later with exceptions as noted below. NPIV capable fibre channel adapters and SAN switches are also required.
- ✓ **VIOS vSCSI** support requires a SAN switch.

# PoC #1 - Configuration

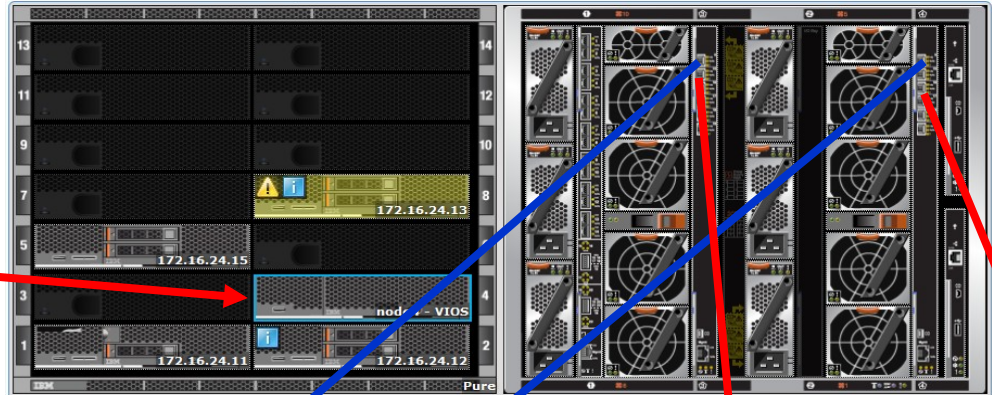




# PoC #1 – Composants matériels



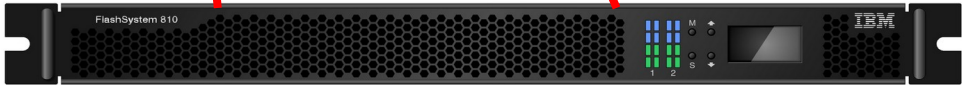
p260  
POWER7  
node



FC3171  
8Gb SAN  
Switches



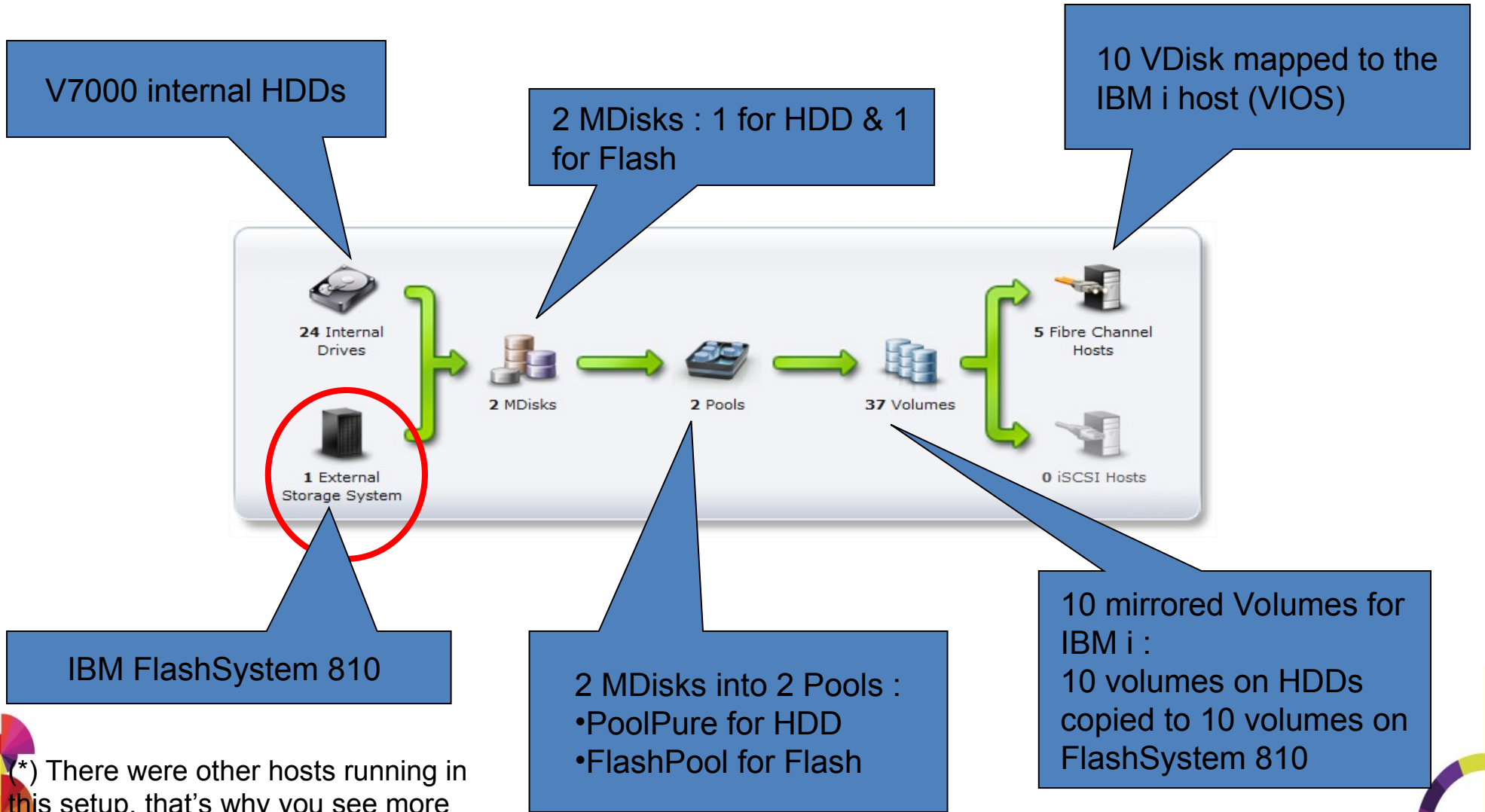
IBM Storwize V7000



IBM FlashSystem 810



# PoC #1 – Configuration Storwize V7000



(\*) There were other hosts running in this setup, that's why you see more hosts and volumes.



## PoC #1 - Tests descriptions (1/2)

- **Test 1** : A complete **power down restart cycle of IBM i** – self explaining, a good mix of CPU, memory, read and write IOs.
- **Test 2** : The goal was to **test streaming IOs** (large block size). I used a backup to a SAVF of a library containing 5 large files for a total of 13GB (2,6 GB average file size).
- **Test 3** : Here, the objective was to test an **analytic query like workload performing a lot of random reads**. In order to do this I used the GO DISKTASKS option 1. It submits a job that collects information about the disk space, this function captures information on all inactive objects.
- **Test 4** : This time, the goal was to test a **large set of relatively small file sizes**, so less streaming when compared with Test 2. I used a backup to a SAVF of a large IFS directory (/QIBM) containing a mix of 190.713 files of miscellaneous sizes for a total of 18 GB (99KB average file size).



# PoC #1 - Tests descriptions (2/2)

## Testing procedure :

After an IPL → sequence Test 2, Test 4, Test 3, then I issued a PWRDWNSYS RESTART(\*YES) for TEST 1 and cycled the tests sequence

## What was measured / compared ?

1. Elapsed time
2. IO/sec
3. Response time

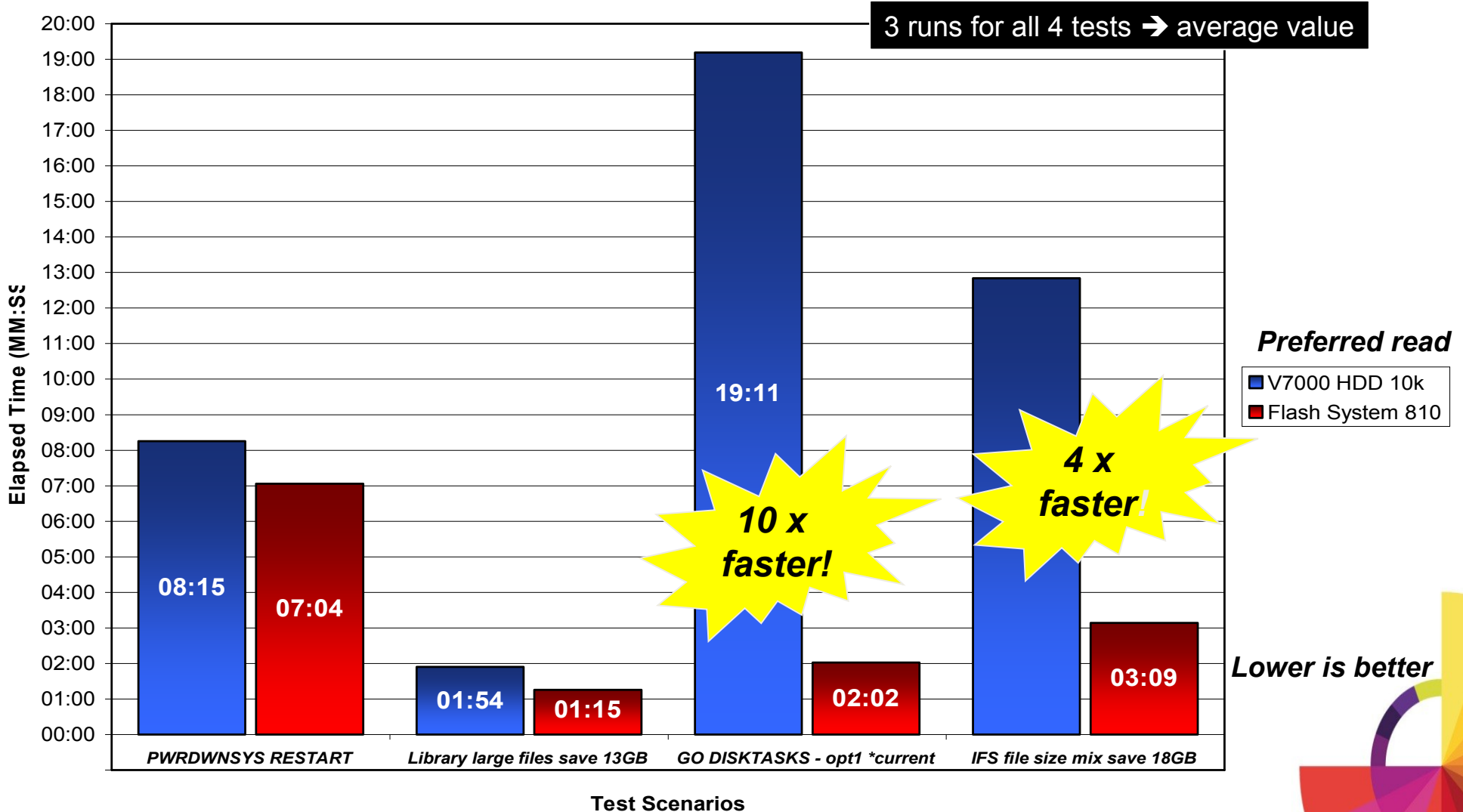
## What scenarios were tested ?

1. The **baseline**, using V7000 HDD LUNs (primary/preferred LUN on HDD)
2. Then, setting the **primary (preferred LUN) for each vDisk to the FlashSystem LUNs**. This in order to redirect all READS operations to Flash memory only while WRITES operations remain in sync on both copies.
3. The impact of **Real Time Compression** (RTC)
4. The impact of **disabling the IBM Storwize V7000 cache**
5. Splitting the VDisk mirror link to use the IBM **FlashSystem only** (with V7000 cache disabled or enabled)



# PoC #1 – Results

IBM i on IBM FlashSystem



# PoC #1 - Comparing the different results



	Preferred read on HDDs RTC OFF	Preferred read on Flash RTC OFF	Preferred read on HDDs RTC ON	Preferred read on Flash RTC ON	Preferred read on Flash RTC ON and CACHE OFF	Flash System only with CACHE OFF	Flash System only with CACHE ON
<b>Test 1</b> <i>PWRDWN SYS RESTART(*YES)</i>	8 m 15 s	7 m 4 s	4 m 33 s	3 m 4 s	9 m 49 s	2 m 54 s	2 m 50 s
<b>Test 2</b> <i>SAVLIB to SAVF 13GB</i>	1 m 54 s	1 m 15 s	4 m 33 s	5 m 47 s	12 m 11 s	54 s	58 s
<b>Test 3</b> <i>GO DISKTASKS OPTION 1</i>	19 m 11 s	2 m 2 s	16 m 31 s	2 m 25 s	7 m 31 s	1 m 43 s	1 m 59 s
<b>Test 4</b> <i>IFS file size mix save 18GB</i>	12 m 50 s	3 m 9 s	<i>no time left to complete this test ☹️</i>				

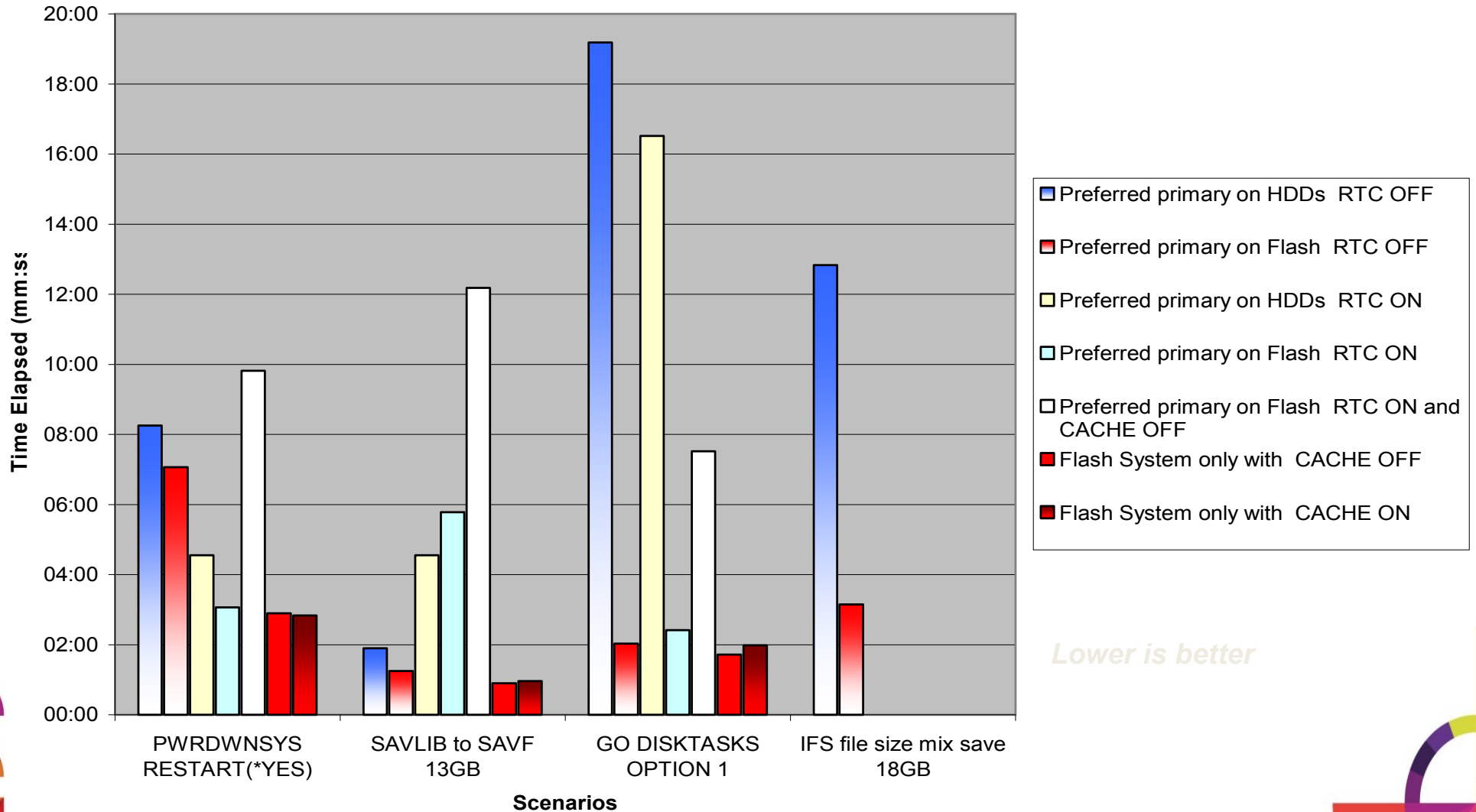
VDisk mirroring with preferred read on HDD vs preferred read on FlashSystem

VDisk mirror with Flash as primary is close to "Flash only" performance

# PoC #1 – Chart (all figures)



All results compared



Lower is better



# PoC #1 - Conclusions

- As expected, IBM FlashSystem is always boosting performance :
  - Modest improvement for a complete shutdown – restart cycle (1,17x)
  - Good improvement for save operations (1,5x in streaming), especially with IFS large number of small files (4x)
  - Impressive improvement for analytics and queries like workload (up to 10x)
  - The save tests were certainly slowed down by the synchronous writes to HDD
- With Real Time Compression turned on :
  - The addition of a VDisk mirror on IBM FlashSystems is improving DB performance compared to HDD only
  - But please, notice that the activation of RTC is in general negatively affecting the performance compared to the run without RTC (except for the shutdown-restart of the system where RTC seems to have a positive impact)
- With the IBM Storwize V7000 cache de-activated :
  - The runtime of the FlashSystem was also slightly better, **this confirms the very low latency of Flash memory compared to cache memory**





## PoC #1 - Remarks

- In the tests I ran, a “dedicated IBM FlashSystems” means an IBM **FlashSystem virtualized** by an IBM Storwize V7000.
- According to the interoperability matrix, native **direct attached IBM FlashSystems** (without an IBM Storwize V7000 or SAN Volume Controller) is not supported by IBM i at the time being. As such, I could not test it and could not demonstrate if IBM FlashSystems would possibly perform better than in a virtualized configuration.
- Only **one FlashSystem volume** was defined → performance may be improved by using **4 volumes** (as Mdisks).



# Smarter Storage for Smarter Computing: What's Next. Ready Now

1	<i>La stratégie IBM sur la technologie Flash</i>
2	<i>La proposition de valeur de FlashSystem</i>
3	<i>Exemples de mise en œuvre</i>
4	<i>Gains FlashSystems et retour d'expériences</i>
5	<i>FlashSystems dans un environnement IBM i</i>
6	<i>Points techniques</i>
	<i>Durée de vie</i>
	<i>Architecture des FlashSystems</i>
	<i>Protection des données</i>
	<i>Performances</i>

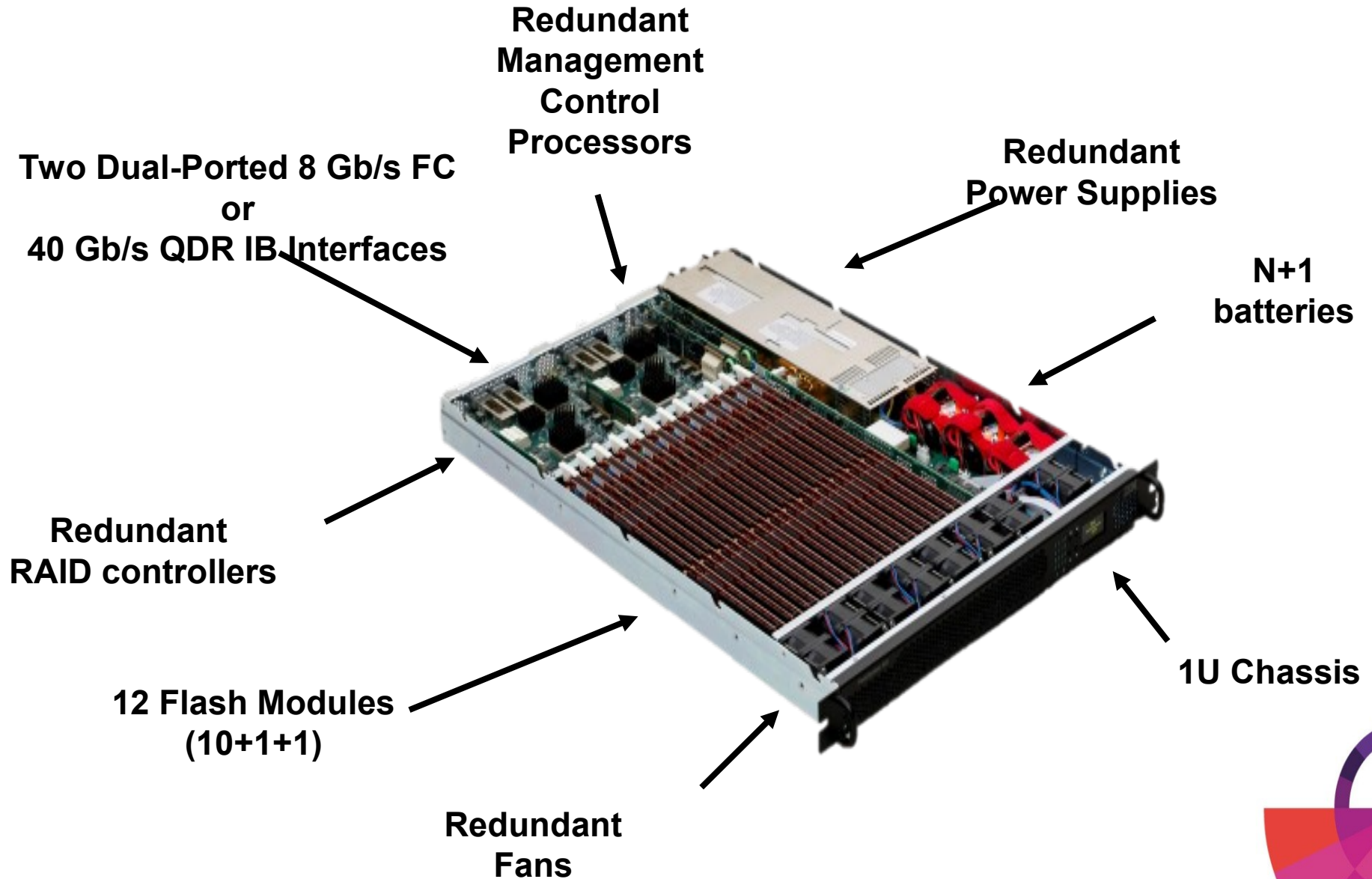


*IBM FlashSystem...*

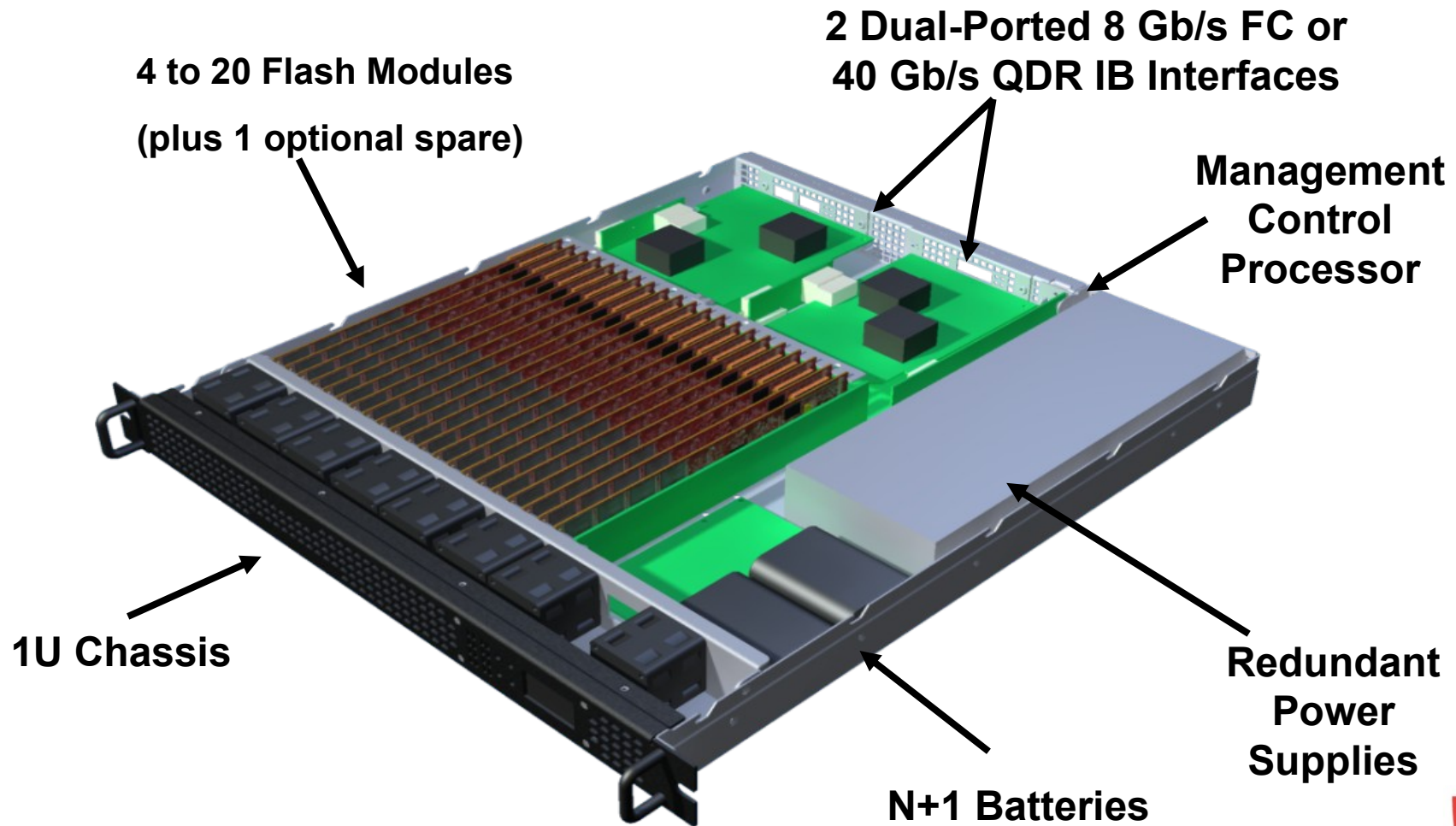
*Enterprise Storage in microseconds, not milliseconds*



# IBM FlashSystem 720 / FlashSystem 820 Architecture

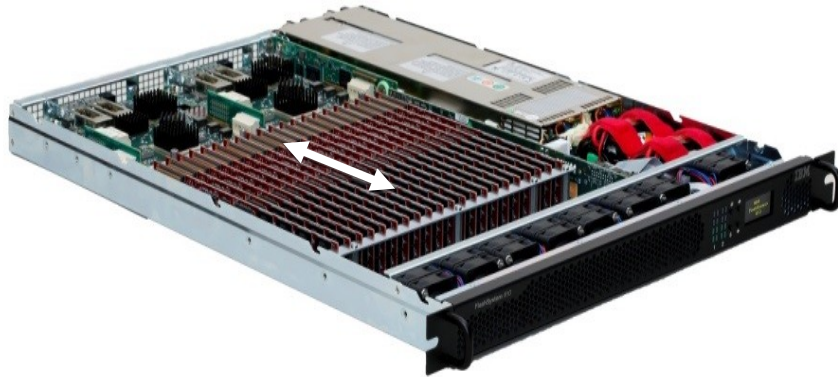


# IBM FlashSystem 710 / FlashSystem 810 Architecture



# Différences entre les modèles FlashSystem x10 et x20

## FlashSystem 710/810



1D RAID across Flash chips  
No Flash Hot-Swap

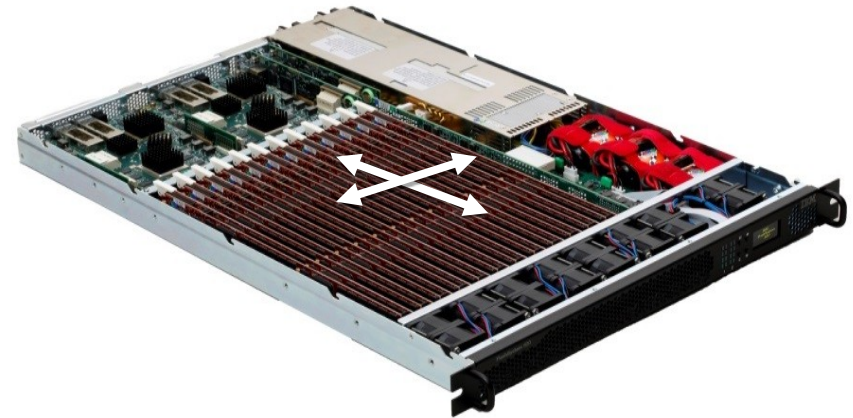
Flash Board capacity (**710**: 250 GB \* **810**: 500 GB)

Flash Module capacity = 1 Flash Board

### Incremental Capacities

4-20 Flash Modules, upgrade 4 Modules  
(**710**: 1/2/3/4/5 TB \* **810**: 2/4/6/8/10 TB)

## FlashSystem 720/820



**2D RAID** across Flash chips & Flash Modules  
**Flash Module Hot-Swap**

Flash Board capacity (**720**: 500 GB \* **820**: 1 TB)

Flash Module capacity = 1 or 2 Flash Board

### No capacity upgrade

(**720**: 5 ou 10 TB \* **820**: 10 ou 20 TB)

# FlashSystem 720 ou 820

	720	820
Technologie des chips	SLC	eMLC
Protection des données 2D Flash RAID	<ul style="list-style-type: none"> <li>• RAID 5 entre les 12 Flash Modules en 10+1+1 (data+parité+hot spare) ou RAID 0 en 11+1 (data+hot spare)</li> <li>• VSR RAID 5 dans chaque Flash Module en 9+1 (data+parité)</li> </ul>	<ul style="list-style-type: none"> <li>• RAID 5 entre les 12 Flash Modules en 10+1+1 (data+parité+ hot spare) ou RAID 0 en 11+1 (data+hot spare)</li> <li>• VSR RAID 5 dans chaque Flash Module en 9+1 (data+parité)</li> </ul>
Capacité d'un Flash Module Nombre maxi de Flash Module	0,5TB ou 1TB brut 12	1TB ou 2TB brut 12
Upgrade	Non, proposé en 6TB (12x0,5) ou 12TB (12x1TB), mais pas d'upgrade possible	Non, proposé en 12TB (12x1) ou 24TB (12x2TB), mais pas d'upgrade possible
Capacités	<ul style="list-style-type: none"> <li>• RAID 0 : 5,6 à 11,3 TiB</li> <li>• RAID 5 : 4,7 à 9,4 TiB</li> </ul>	<ul style="list-style-type: none"> <li>• RAID 0 : 11,3 ou 22,5 TiB</li> <li>• RAID 5 : 9,4 ou 18,8 TiB</li> </ul>
Latence	<ul style="list-style-type: none"> <li>• Write : 25 µs</li> <li>• Read : 100 µs</li> </ul>	<ul style="list-style-type: none"> <li>• Write : 25 µs</li> <li>• Read : 110 µs</li> </ul>
Performances (blocs de 4KB)	<ul style="list-style-type: none"> <li>• 100% Read : 525k IOPS</li> <li>• 70/30 R/W : 450k IOPS</li> <li>• 100% Write : 400k IOPS</li> </ul>	<ul style="list-style-type: none"> <li>• 100% Read : 525k IOPS</li> <li>• 70/30 R/W : 430k IOPS</li> <li>• 100% Write : 280k IOPS</li> </ul>
Hot swap	Flash Modules et Power	Flash Modules et Power
Watts	350	300

# FlashSystem 710 ou 810

	710	810
Technologie des chips	SLC	eMLC
Protection des données	<ul style="list-style-type: none"> <li>•VSR RAID 5 dans chaque Flash Module en 9+1 (data+parité)</li> <li>• 1 Flash Module de spare (option)</li> </ul>	<ul style="list-style-type: none"> <li>•VSR RAID 5 dans chaque Flash Module en 9+1 (data+parité)</li> <li>•1 Flash Module de spare (option)</li> </ul>
Capacité d'un Flash Module Nombre maxi de Flash Module	0,25TB brut donc 1TB pour 4 modules 20	0,5TB brut donc 2TB pour 4 modules 20
Upgrade	Oui, par groupe de 4 modules	Oui, par groupe de 4 modules
Capacités	<ul style="list-style-type: none"> <li>• RAID 5 : 4,7 TiB maxi par upgrade de 0,9 TiB (4 modules)</li> </ul>	<ul style="list-style-type: none"> <li>• RAID 5 : 9,4 TiB maxi par upgrade de 1,9 TiB (4 modules)</li> </ul>
Latence	<ul style="list-style-type: none"> <li>• Write : 60 µs</li> <li>• Read : 100 µs</li> </ul>	<ul style="list-style-type: none"> <li>• Write : 60 µs</li> <li>• Read : 110 µs</li> </ul>
Performances (blocs de 4KB)	<ul style="list-style-type: none"> <li>• 100% Read : 570k IOPS</li> <li>• 70/30 R/W : 490k IOPS</li> <li>• 100% Write : 400k IOPS</li> </ul>	<ul style="list-style-type: none"> <li>•100% Read : 550k IOPS</li> <li>• 70/30 R/W : 430k IOPS</li> <li>• 100% Write : 400k IOPS</li> </ul>
Hot swap	Power	Power
Watts	280	350

# Performances IBM FlashSystem en “natif” sur le SAN

	FlashSystem 820 w/ RAID	FlashSystem 720 w/ RAID	FlashSystem 810	FlashSystem 710
<b>Sequential</b> read (256 KB) GB/s	3.3 GB/s ( <b>FC</b> ); 5 GB/s ( <b>IB</b> )	3.3 GB/s ( <b>FC</b> ); 5 GB/s ( <b>IB</b> )	3.3 GB/s ( <b>FC</b> ); 5 ( <b>IB</b> )	3.3 GB/s ( <b>FC</b> ); 5 ( <b>IB</b> )
<b>Sequential</b> Write (256 KB) GB/s	2.8 GB/s	3.3 GB/s ( <b>FC</b> ); 4 GB/s ( <b>IB</b> )	2.8 GB/s	3.3 GB/s ( <b>FC</b> ); 4.5 GB/s ( <b>IB</b> )
100% <b>Random Reads</b> (4 KB) IOPS	525k	525k	550k	570k
<b>70-30 Random Read- Write (4 KB) IOPS</b>	<b>430k</b>	<b>450k</b>	<b>430k</b>	<b>490k</b>
Max <b>random write</b> (4KB) IOPS	280k	400k	400k ( <b>FC</b> ); 470 K ( <b>IB</b> )	400k ( <b>FC</b> ); 470k ( <b>IB</b> )

The table above shows each system's **maximum throughput capabilities**, summarized and rounded.



## Plus d'informations

go to <http://www.redbooks.ibm.com> and search for IBM FlashSystem

**IBM FlashSystem 710 and IBM FlashSystem 810:**

<http://www.redbooks.ibm.com/technotes/tips1002.pdf>

**IBM FlashSystem 720 and IBM FlashSystem 820:**

<http://www.redbooks.ibm.com/technotes/tips1003.pdf>

**IBM FlashSystem in OLTP Database Environments:**

<http://www.redbooks.ibm.com/technotes/tips0973.pdf>

**IBM FlashSystem in OLAP Database Environments:**

<http://www.redbooks.ibm.com/technotes/tips0974.pdf>

**IBM FlashSystem and SVC (and RTC) :**

<http://www.redbooks.ibm.com/Redbooks.nsf/RedpieceAbstracts/redp5027.html?Open>

**Matrice de compatibilité**

<http://www-03.ibm.com/systems/support/storage/ssic/interoperability.wss>

**Benchmark: IBM Montpellier**

