

IBM Power Systems



Virtualisation et IBM i

COMMON, Genève le 03 mai 2011

Alain PLU - IBM France

Agenda

- Les enjeux de la virtualisation
- Les partitions virtuelles IBM i
- PowerVM et la virtualisation de l'IBM i
- IBM i et la virtualisation du Fibre Channel

Les enjeux de la virtualisation

Evolution des technologies

- Processeurs
 - 32 threads par puce
- Bus PCI Express
 - 16 GO/s
- Ethernet
 - 10 Gb
- SAN
 - 8 Gb

La densification du matériel va permettre la mutualisation de la bande passante

Evolution des serveurs Power

Processeur/ Modèle	Power 5-520	Power 6-520	Power 7-720
Nombre de coeurs	2	2	8
Puissance	6 000 CPW	8 300 CPW	23 800 CPW
Bande passante I/O	5 GO/s	14 GO/s	30 GO/s
Nombre de tiroirs I/O	6 PCI-X	4 PCI-X DDR	2 PCI-E

■

Evolution du bus PCI

	PCI	PCI-X	PCI-E V1	PCI-E V2
Fréquence carte	33 Mhz	66-133-266 Mhz	2,5 Ghz	5 Ghz
Connectique	Parallèle	Parallèle	Série	Série
Débit de données	133 MO/s	0,5 - 1 GO/s	2,5 - 6,4 GO/s	5 - 16 GO/s
Mode de transfert	Uni-directionnel	Uni-directionnel	Bi-directionnel	Bi-directionnel

■

Evolution du SAN

	2002	2005	2008	201X
Bande passante	2 Gb	4 Gb	8 Gb	16 Gb
Type de carte	PCI	PCI-X	PCI-E	PCI-E V2
IO/s	1500	2600	21000	ND
Ports	Physique	Physique	Physiques et Virtuels	Physiques et Virtuels

■

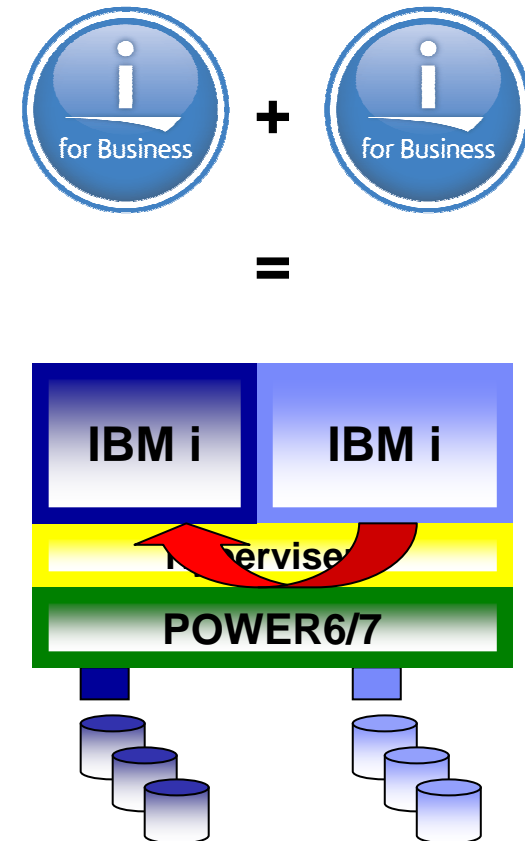
Evolution des technologies

- La virtualisation permet de mutualiser la capacité croissante de bande passante
- L'utilisation de connectique dédiée va devenir de moins en moins justifiable sauf pour les gros environnements de production
- Les nouvelles technologies reposent sur la virtualisation des I/O
 - Advanced Memory Sharing
 - Partition Hibernation
 - Live Partition Mobility

Les partitions virtuelles IBM i

Partitions virtuelles IBM i

- **Virtualisation à partir d'un autre IBM i**
- **La partition IBM i utilise des ressources d'entrée/sortie fournies par une autre partition IBM i**
 - Moins de cartes et de disques à acheter pour l'ensemble des partitions IBM i
 - Ajout facile de partitions dans le cadre de tests ou de développement
 - Mixage possible d'I/O virtuelles ou directes pour la partition cliente
- **Plateformes supportées**
 - Tous les Power Systems à base de POWER6 et POWER7
- **Gestion des partitions**
 - HMC



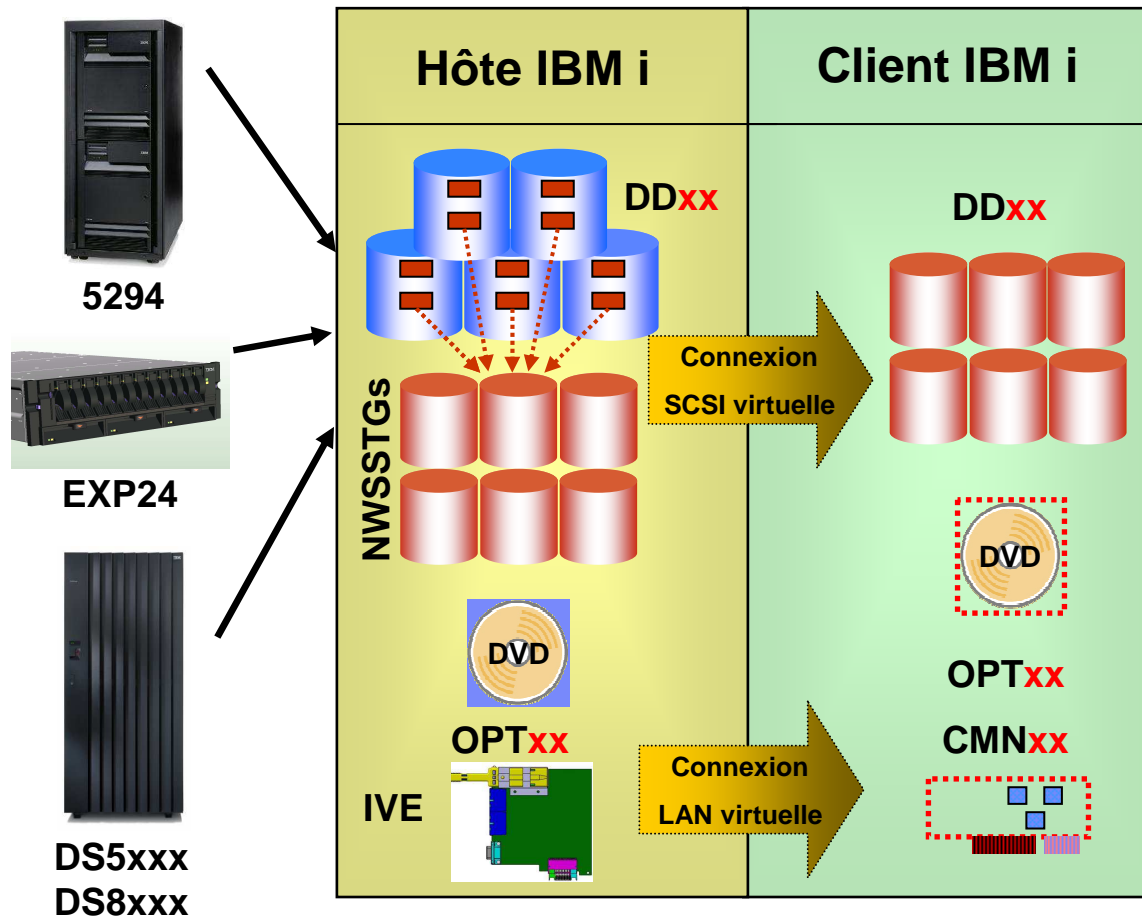
Prérequis et support

Pour la solution IBM i hôte virtualisant les ressources pour des partitions i clientes

- L'hôte IBM i hôte peut utiliser des disques internes, ou les baies supportées nativement
 - Il est possible d'avoir plusieurs niveaux de virtualisation, mais semble de peu d'intérêt si ce n'est que pour des problèmes de compatibilité d'environnement
- Il faut disposer de processeurs POWER6 ou POWER7
- IBM i 6.1 minimum

Les DS5100, 5300 et DS8xxx sont supportés sur tous les environnements POWER6 / POWER7

Partitions IBM i client et hôte



Prérequis

- Système POWER6 ou POWER7
- IBM i 6.1 pour l'hôte et le client

Disques

- Matériel associé à la partition hôte par la HMC
- Les disques virtuels sont des objets espace de stockage (NWSSTG)

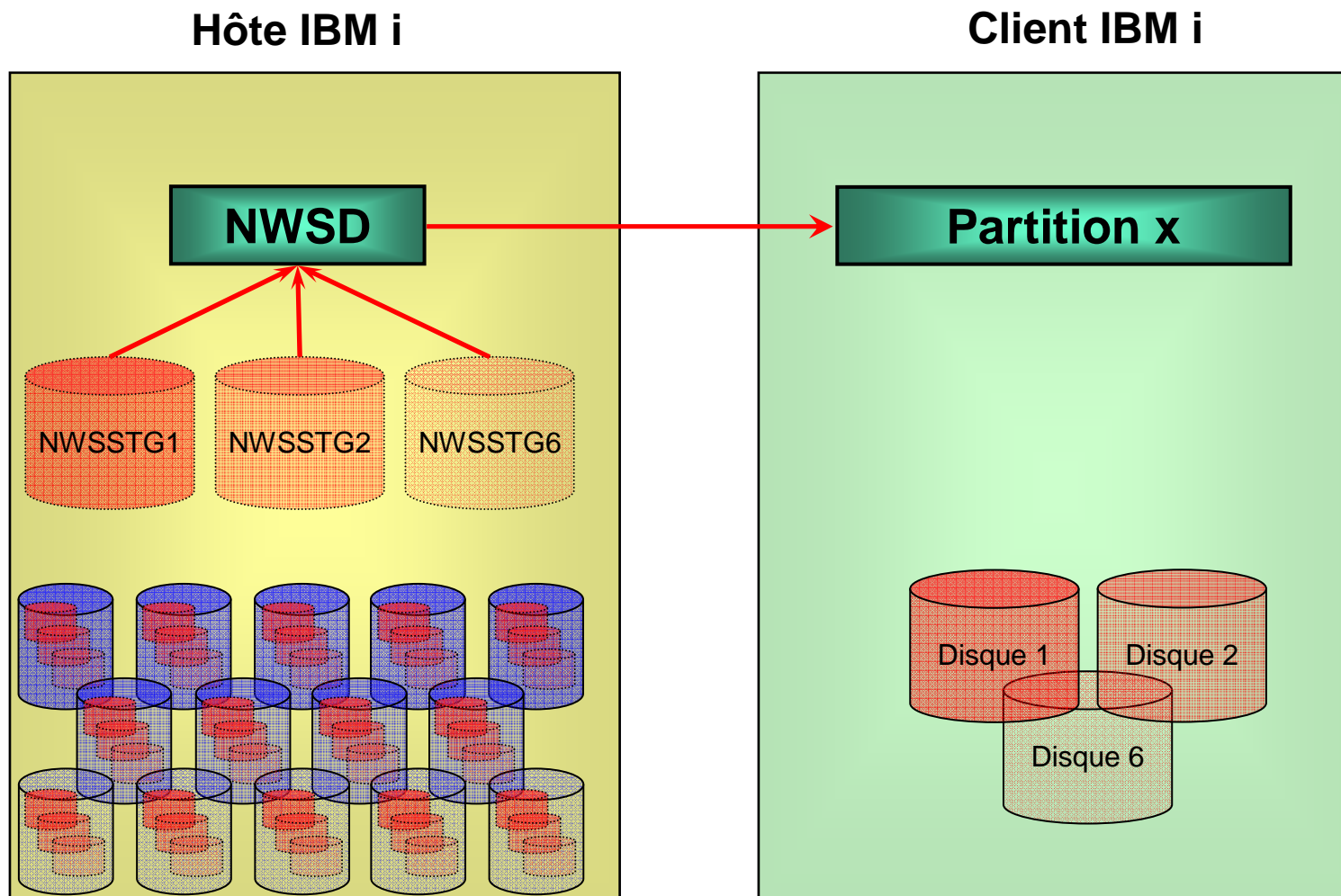
Unités optique

- Les unités optique de la partitions hôte sont accessibles dans la partition client

Réseau

- Routage entre une carte Ethernet et une carte virtuelle côté hôte, carte virtuelle côté client
- Carte IVE

Objets de configuration



Création de 6 espaces de stockage NWSSTG associés au serveur réseau NWSD

Objets de configuration

- **Unités de disques**
 - On peut associer un maximum de 16 NWSSTG à un NWSD
 - Si besoin de plus, créer d'autres NWSD ...

- **Unités optiques**
 - Les lecteurs de CD-ROM et DVD sont vus de toutes les partitions clientes
 - Les catalogues d'images sont vus de toutes les partitions clientes
 - Il est possible d'interdire l'accès à une ou plusieurs unités
Ressources unités restreintes (RSTDDEVRSC)

- **Unités de bandes**
 - les unités de bandes sont vues au travers du support virtuel SCSI
 - LTO3/4/5 et TS2900, TS3100, TS3200 (hors bandothèque)

- **Réseau**
 - Ports Ethernet virtuels

Sécurisation de la partition cliente

- **Il est possible de faire dépendre une partition cliente de deux partitions hôtes**
 - Permet de s'affranchir de l'arrêt pour maintenance de la partition hôte

- **On créera des configurations symétriques pour les deux hôtes**
 - Même nombre de cartes SCSI virtuelles
 - Même nombre de NWSD
 - Même nombre et même taille de NWSSTG

- **Les disques virtuels de la partition clientes seront mirrorés entre eux**
 - Miroir au niveau contrôleur
 - L'arrêt d'une partition hôte arrêtera la fonction miroir qui sera reprise au redémarrage de celle-ci

PowerVM et la virtualisation de l'IBM i

Qu'est-ce que PowerVM ?

Composants matériels et logiciels qui permettent la virtualisation des systèmes d'exploitation IBM i, AIX et Linux sur les serveurs à base de processeurs IBM Power



Dispositifs de PowerVM

Micro-Partitionnement

Virtual I/O Server

Integrated Virtualization Manager

Live Partition Mobility

Lx86



Partitionnement logique

Offre PowerVM*

Dispositif / Fonction	Express Edition	Standard Edition	Enterprise Edition
Serveurs supportés	520 / 550 POWER6 720 / 750 POWER7	JS12, JS21, JS22, PS7x Power Systems	JS12, JS22, POWER6/POWER7
Nombre de partitions maximum	2 LPARs + 1 VIOS par serveur	10 par processeur	10 par processeur
Gestion	IVM	IVM & HMC	IVM & HMC
Virtuel I/O Server	✓	✓	✓
Mobilité des partitions			AIX et Linux seulement
Pool de Processeurs partagés		✓ (POWER6 & HMC requis)	✓ (HMC Requisite)
Partage des processeurs dédiés inutilisés	✓	✓ POWER6	✓
Système d'exploitation	AIX / Linux / i	AIX / Linux / i	AIX / Linux / i
PowerVM Lx86	✓	✓	✓
Active Memory Sharing			✓

Certaines fonctions PowerVM ne sont pas disponibles sur tous les modèles

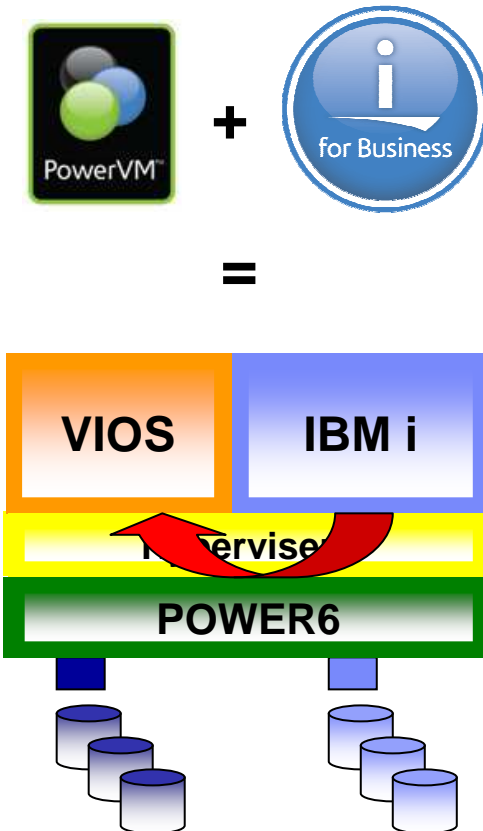
(<http://www.ibm.com/systems/power/software/virtualization/index.html>)

Qu'est-ce que VIOS ?

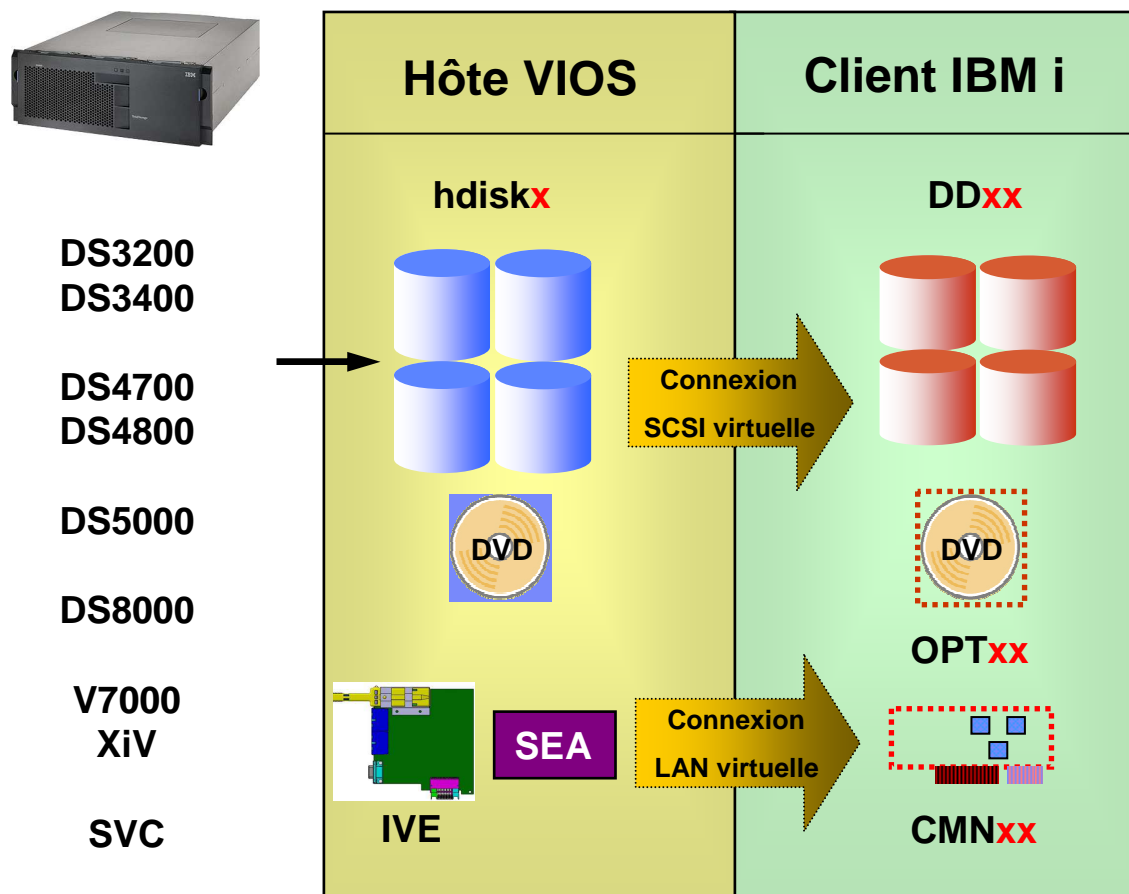
- Une appliance sous LPAR qui réside dans les serveurs et lames POWER5/6/7
- Facilite le partage des ressources physique I/O entre LPARs
- Fonction de base : I/O virtuels : virtual SCSI, NPIV, et SEA (layer-2 bridge, virtual Ethernet bridge)
- Fonctions évoluées : (LPAR Mobility, Advanced Memory Sharing, etc)
- Basée sur une version restreinte d'AIX
- VIOS sert les OS AIX, Linux et IBM i
- VIOS est packagé avec le micro-partitioning dans PowerVM

Virtualisation avec PowerVM

- **Virtualisation basée sur VIOS**
- **La partition IBM i utilise des ressources d'entrée/sortie fournies par une partition VIOS (Virtual I/O Server)**
 - VIOS est un composant de PowerVM
 - Nécessite niveau minimum OS IBM i 6.1
- **Plateformes supportées**
 - Tous les machines à base de POWER6 et POWER7
- **Stockage**
 - Supporte les baies DS3200, DS3400, DS4700, DS4800, DS5000, DS8000 V7000, XiV et SVC
- **Gestion des partitions**
 - HMC (requis à partir du 570 ou 770)
 - IVM



VIOS, IBM i 6.1 et stockage



Disques

- Matériel affecté à la partition VIOS
- Chaque unité logique est virtualisée directement
- Les LUNs sont au format 512 octets *open*

Unités optiques

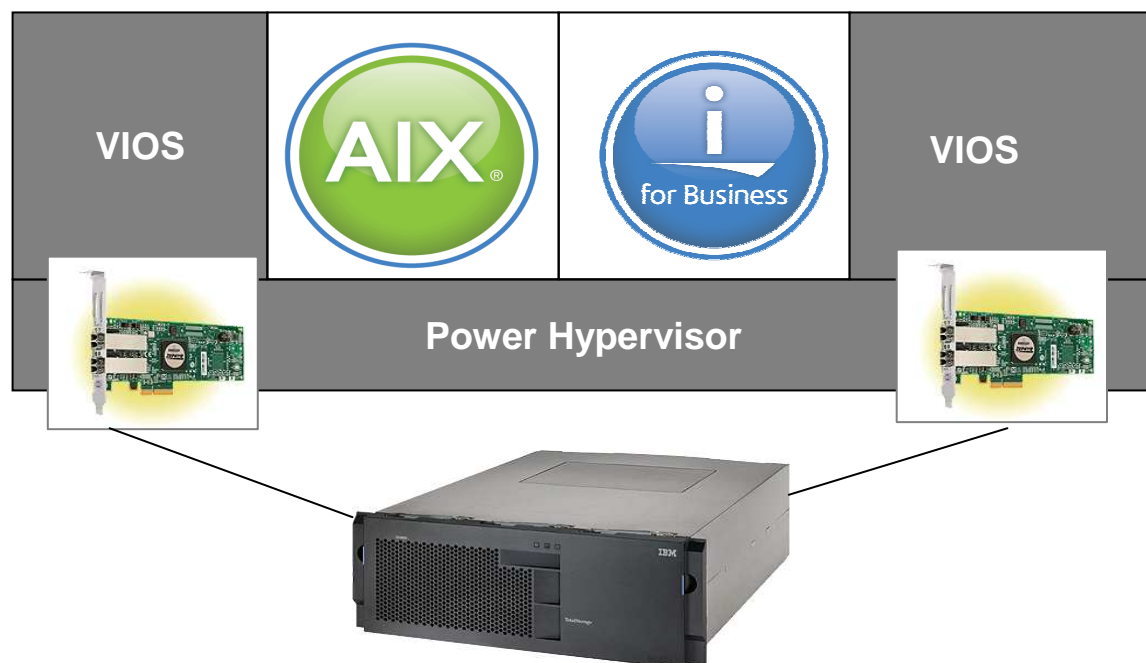
- Le lecteur DVD du VIOS est virtualisé directement

Réseau

- Le client IBM i dispose d'une carte Ethernet virtuelle
- Le VIOS utilise les fonctions de bridge virtuel pour interconnecter le réseau virtuel et le réseau physique

Support VIOS redondant avec MPIO client

- Permet aux partitions IBM i 6.1 d'utiliser deux partitions VIOS redondantes pour accéder à une baie de stockage.
- Améliore la disponibilité de l'ensemble des environnements virtualisés



Configuration IBM i et lames Power

- **Châssis BladeCenter H**

- DVD, Fibre, Ethernet, SAS switchs
- Stockage
 - DS3400, DS4700, DS4800, DS5000, DS8000, V7000, XiV ou SVC
 - Bande SAS externe et bandothèque

- **Châssis BladeCenter S**

- DVD, Ethernet, SAS switchs ou RAID SAS switchs
- Stockage
 - Disques SAS internes, DS3200
 - Bande SAS externe

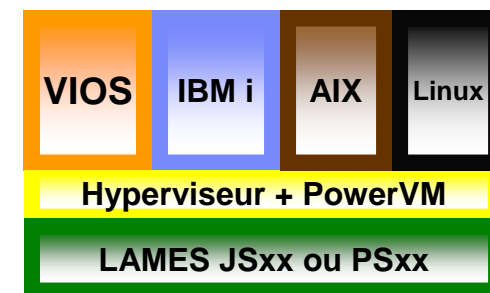
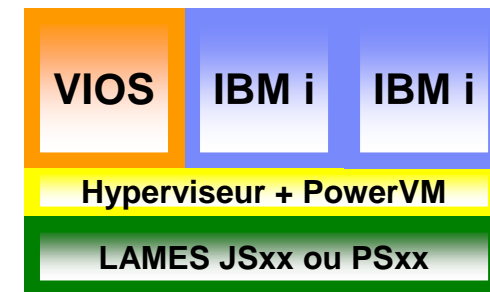
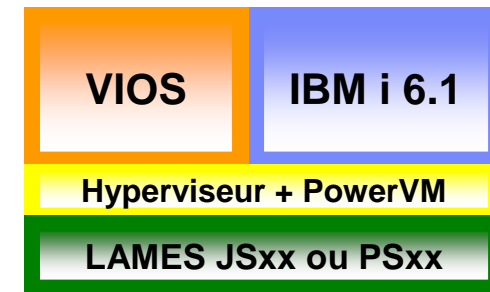
- **Lames JS12, JS22, JS23, JS43**

- **Lames PS700, PS701, PS702, PS703, PS704**

- **La partition VIOS gère les ressources des lames**

- IVM est utilisé pour créer et gérer les autres partitions

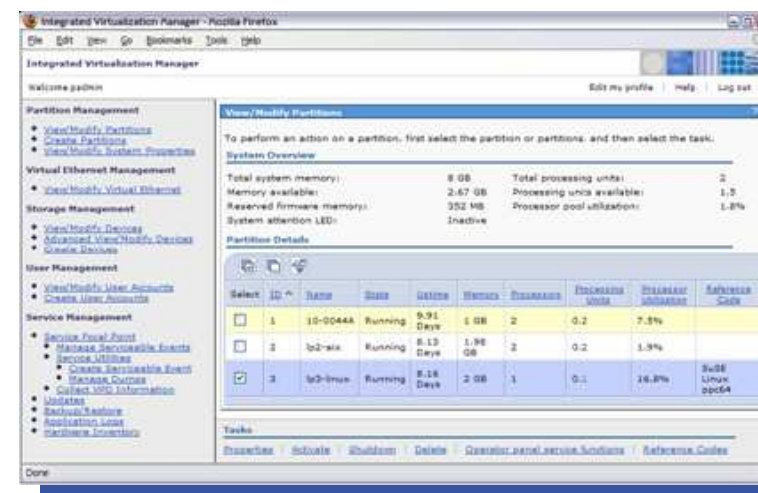
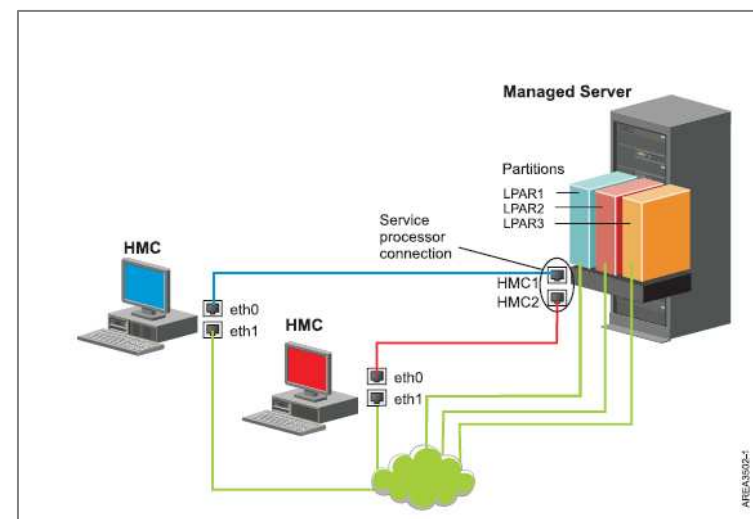
Exemples de lames partitionnées



Les consoles HMC et IVM

- **Hardware Management Console**
 - Matériel dédié
 - Console dédiée à la gestion des serveurs et de la virtualisation

- **Integrated Virtualization Manager**
 - Outil de gestion sous navigateur web pour les serveurs Power Express et les lames Power



Options de gestion avec IVM ou HMC



Lames Power

- **IVM est la seule possibilité**
 - Interface graphique WEB
 - Fonction de partitionnement limitée pour VIOS, IBM i 6.1, AIX, Linux
 - Gestion des disques virtuels, optiques et bibliothèque de supports virtuels
 - Simple et rapide pour le partitionnement et la virtualisation
 - Quelques fonctions de service
- **IVM fait partie du VIOS**
 - Inclus dans les versions PowerVM Standard et Enterprise

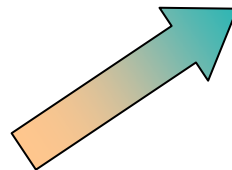
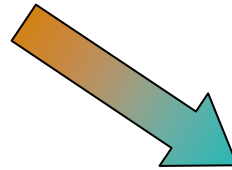
Serveurs POWER6/7 bas et milieu de gamme

- **HMC ou IVM sont utilisables**
- **IVM :**
 - Simple et rapide pour le partitionnement et la virtualisation
 - Quelques fonctions de service
- **HMC :**
 - Fonction de partitionnement étendue pour VIOS, IBM i 6.1, AIX, Linux
 - Nombreuses autres fonctionnalités (Service, Ras, etc.)

Serveurs POWER6/7 haut de gamme

- **HMC est la seule possibilité**
 - Interface graphique WEB
 - Fonction de partitionnement étendue pour VIOS, IBM i 6.1, AIX, Linux
 - Gestion des disques virtuels, optiques (besoin de configuration côté VIOS et IBM i)
 - Nombreuses autres fonctionnalités (Service, Ras, etc.)

Systems Director Management Console



SDMC:

- Browser Interface
- Manage Multiple Systems
- New User Interface
- Enhanced Functionality
- Support 1000 LPARs
- Supports P6 & P7

Systems Director Management Options

Hardware Appliance



IBM Provided Hardware

Plus Software Appliance:

- SDMC Management Base
- Systems Director Express Edition

Turn Key Solution

Software Appliance

Software Appliance:

- SDMC Management Base
- Systems Director Express Edition

Non Turn Key Solution

Utilize existing x86 virtualization infrastructure

Customer provided hardware

- Finite set of server options



Console SDMC : les fonctions

- Toutes les fonctions de configuration, de maintenance et de gestion fournies par la console HMC et IVM sont intégrées dans IBM Systems Director.
 - ▶ SDMC fournit la gestion des serveurs et de la virtualisation sous PowerVM
 - ▶ La gestion de la console est transférée dans Systems Director.

- Interface CLI Interface
 - ▶ Identique à la console HMC
 - ▶ Aucune différence de syntaxe
 - Des alias sont fournis pour une compatibilité presque totale

- Interface graphique
 - ▶ Fonctionnalités identiques
 - ▶ Mais nouvelle interface basée sur Systems Director
 - ▶ Simplification de la navigation et des tâches utilisateur.

Support de la bande virtuelle pour IBM i

Le support des unités de bande virtuelles permet aux partitions IBM i d'avoir accès directement aux unités de bande gérées par le Virtual I/O Server.



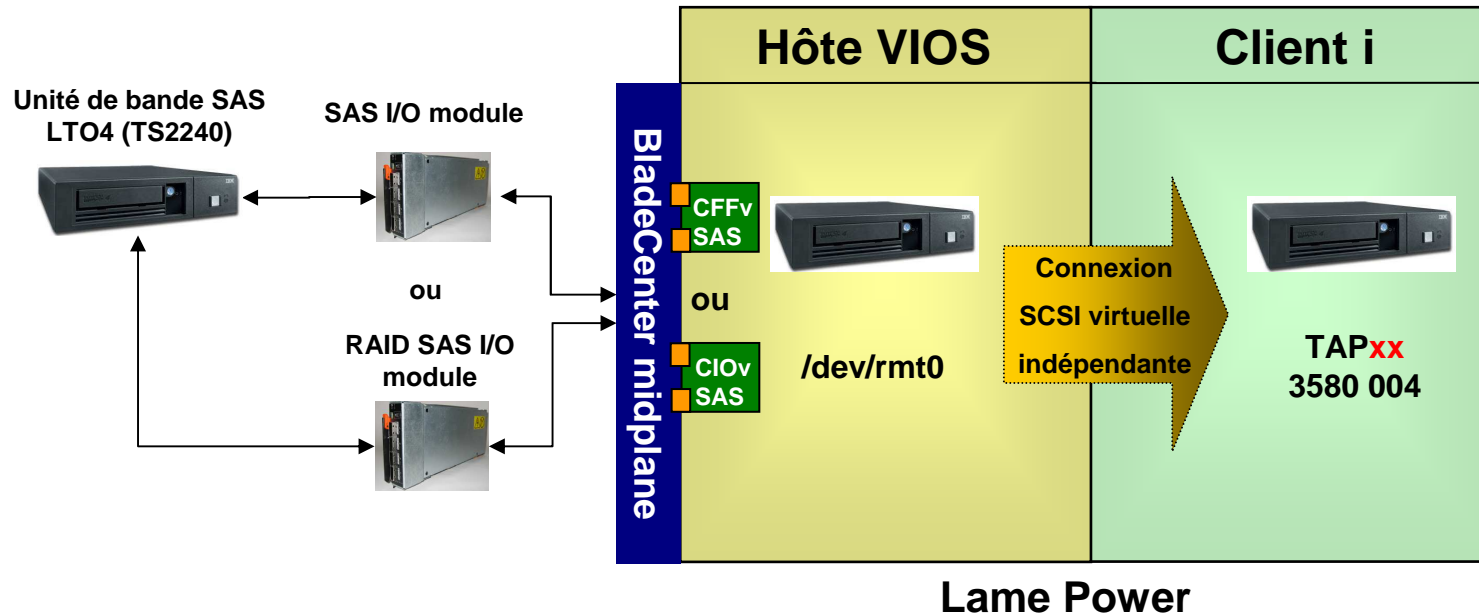
- Simplification des procédures de sauvegarde et de restauration dans le cadre des implémentations de IBM i sur des lames.
 - Partitions IBM i 6.1 sur lames JS12, JS22, JS23, JS43 et PS70x
 - Support des commandes SAV/RST et de BRMS
 - Supporté dans les châssis H et S

- Simplification des processus de migration des environnements traditionnels vers les lames



Support de la bande virtuelle pour IBM i

2

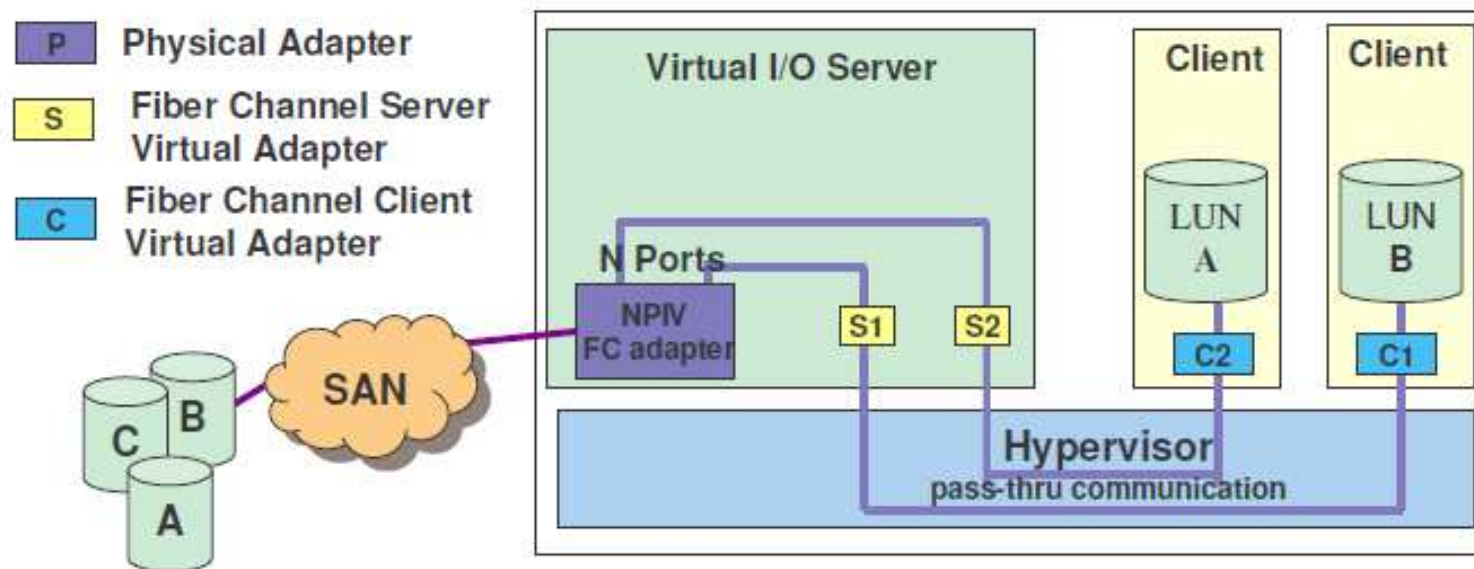


- Unité de bande TS2240 LTO4 ou TS2250 LTO5 SAS connectée à un switch SAS dans le châssis
 - **SAS Connectivity Module** ou **RAID SAS Switch Module** dans un châssis **BCS**
 - **SAS Connectivity Module** dans un châssis **BCH**
- Virtualisation de l'unité de bande par le Virtual I/O Server
- L'unité de bande est assignée à la partition IBM i avec IVM
- L'unité de bande est disponible en tant que TAPxx, type 3580 model 004

IBM i et la virtualisation du Fibre Channel

Qu'est ce que NPIV ?

- NPIV est une fonctionnalité Fibre Channel permettant le partage d'un N_Port physique par plusieurs N_Port IDs
 - **C'est un Passthru FC dans l'hyperviseur, sans besoin de VSCSI**
- Fournit aux partitions clientes des connexions FC directes aux ressources SAN
- Sans définir aucune unité virtuelle cible dans le VIOS



Virtualisation des lecteurs bande via NPIV et VIOS

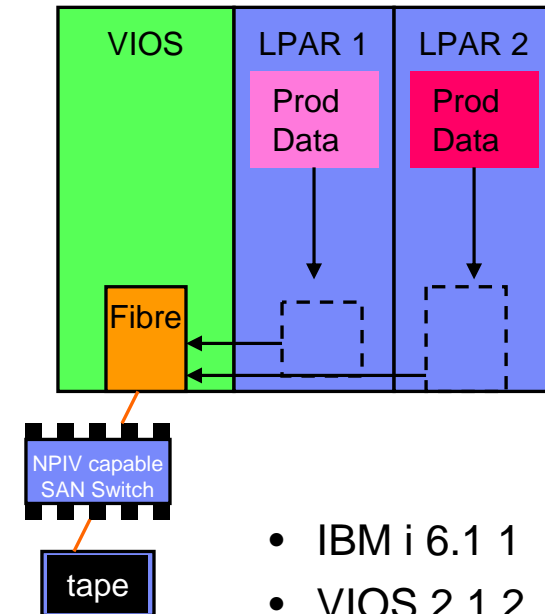
- NPIV = N-Port ID Virtualization
- Virtualise le port FC pour partage de carte FC du VIOS entre LPARs
- Carte FC 5735 (8 Gbit) requise ou équivalent Blade
- Bandothèques et lecteurs pris en charge :

3573 (TS3100/TS3200) avec LTO3/LTO4/LTO5

3576 (TS3310) avec LTO3/LTO4/LTO5

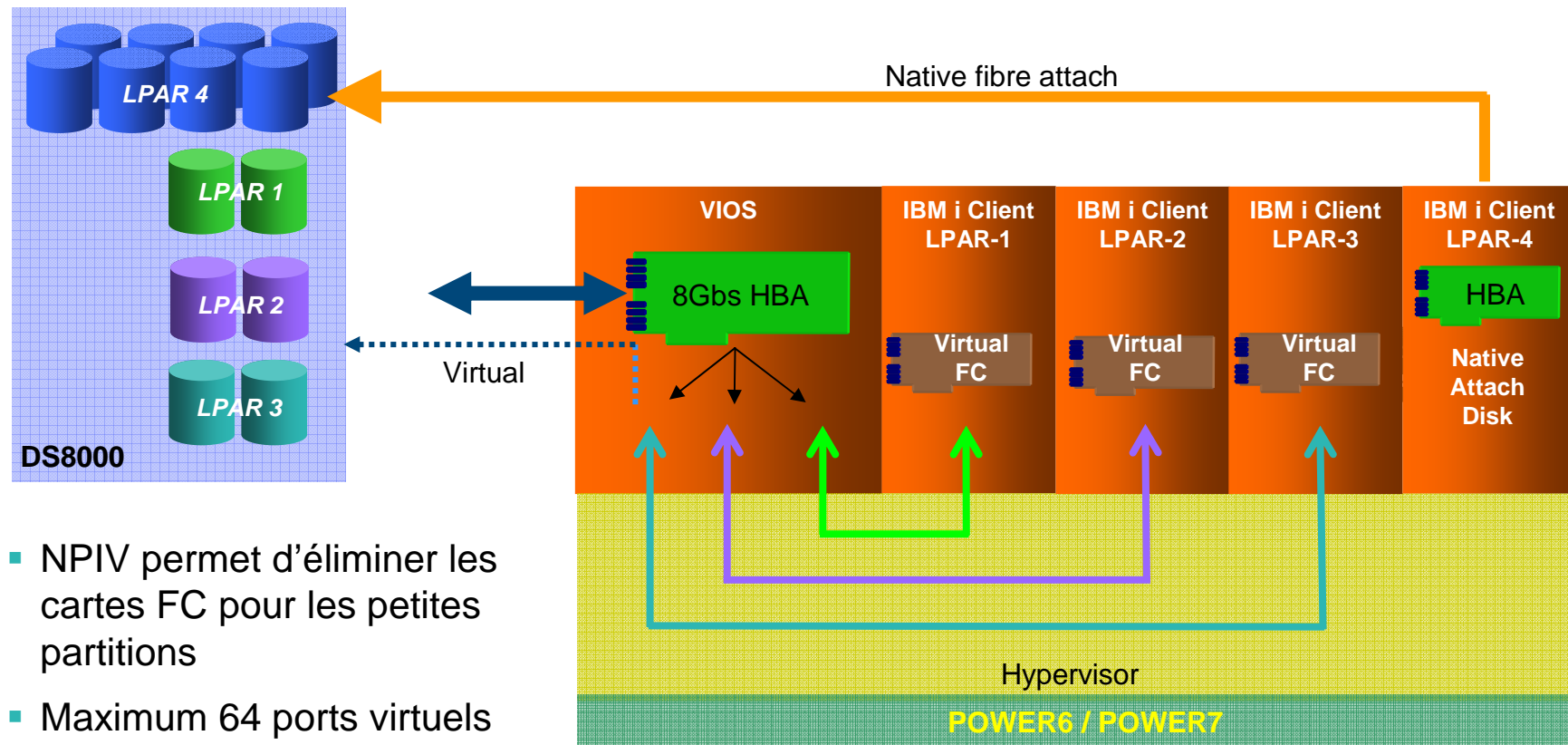
3577 (TS3400) avec TS1120/TS1130

3584 (TS3500) avec LTO3/LTO4/LTO5/TS1120/TS1130/3592-J1A



- IBM i 6.1 1
- VIOS 2.1.2

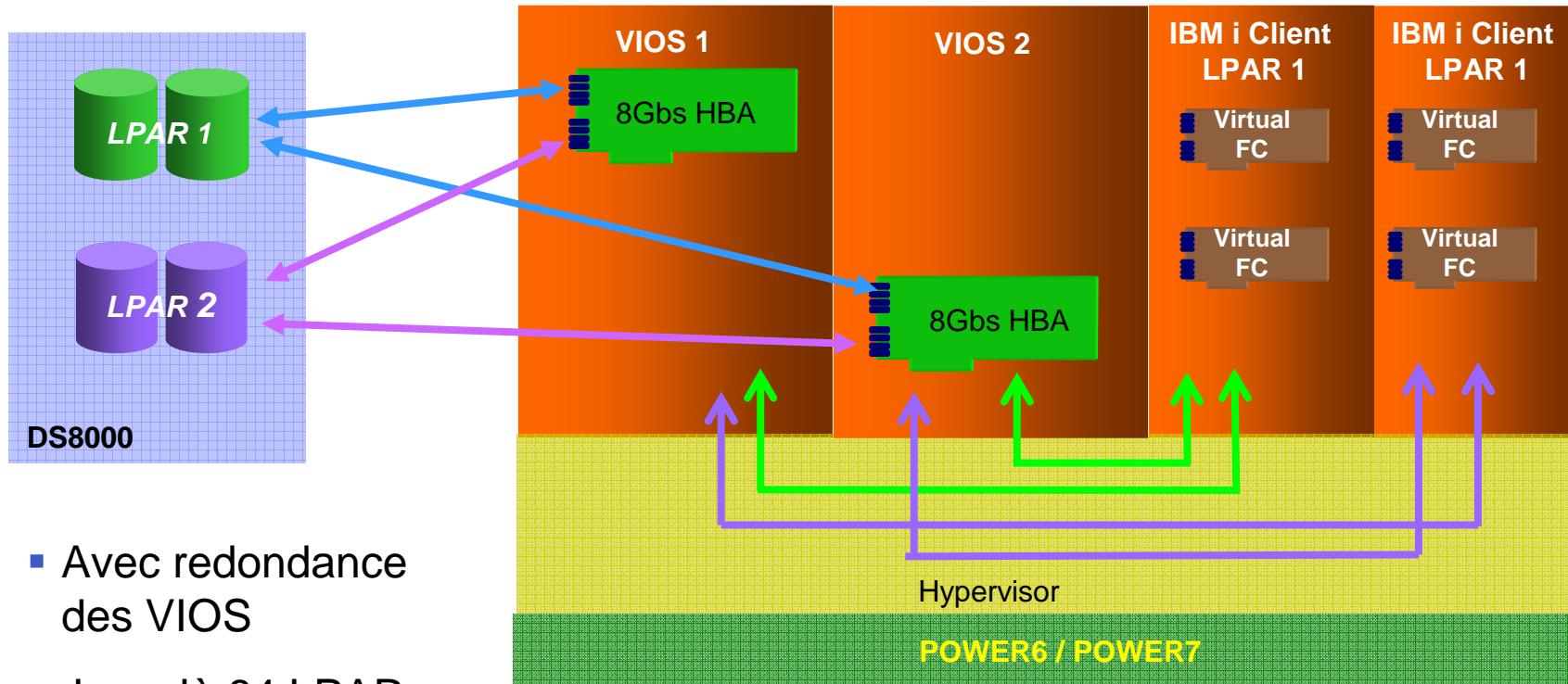
PowerVM VIOS avec NPIV et DS8000 (vue théorique)



- NPIV permet d'éliminer les cartes FC pour les petites partitions
- Maximum 64 ports virtuels (VFC) par carte FC
- Nécessite Switch SAN compatible NPIV

- IBM i 6.1.1
- VIOS 2.1.2

PowerVM VIOS avec NPIV et DS8000 (vue pratique)



- Avec redondance des VIOS
- Jusqu'à 64 LPARs clients IBM I
- DS8000 et DS5000

- IBM i 6.1.1
- VIOS 2.1.2

Synthèse technique

Le Virtual I/O Server fournit : la virtualisation du réseau, la virtualisation des disques et des unités optiques, le partage de cartes Ethernet et de cartes fibre (NPIV).

- Echange mémoire à mémoire = très efficace
- Exécuter les fonctions de pilote disque que fait normalement la partition cliente est un peu plus consommateur de ressources.
- Les I/O disques utilisent les "*Remote DMA*" = très efficace
- Le support NPIV utilise un mode passthru = très efficace
- L'Ethernet virtuel ne peut pas utiliser les fonctions d'accélération matérielle = efficace
- Réaffectation des buffers Ethernet = très efficace
- Le Virtual I/O Server peut avoir besoin de bufferiser le trafic réseau et dans ce cas avoir besoin de mémoire supplémentaire.

Conclusions :

- Toutes les fonctionnalités sont très efficaces, sauf l'Ethernet virtuel qui est seulement efficace
- Si les charges disque et réseau sont équivalentes, se focaliser sur la partie réseau qui est trois fois plus consommatrice en ressource processeur.