

Module INF 180 – INFORMATIQUE / TELCO's

SI / Systèmes d'Information
« IT / Information Technology »

Architecture & Concepts

INF_180

VIRTUALISATION

(Concepts et Techniques de la Virtualisation)

Résumé des Objectifs du Module :

- Être capable de : Comprendre les principes de la VIRTUALISATION (IT).
- Répondre à la question : Pourquoi et Comment ça marche (vulgarisation).

Cours magistral sur 2 jours

Michel MESTRALLET

Table des matières

1	Préambule et « avertissements »	6
2	Objectif du Module	6
2.1	Être capable de :	6
2.2	Répondre à la question :	6
3	Descriptif et résumé du contenu	7
4	Définition du concept de Virtualisation	8
5	Historique de la Virtualisation (UC et Stockage)	10
5.1	Historique de la Virtualisation des UC / Serveurs	10
5.1.1	Introduction (Historique de la Virtualisation)	10
5.1.2	Détails (Historique de la Virtualisation)	10
5.1.3	Utilisations	13
5.1.4	Définitions (Virtualisation, Machine Virtuelle, Hyperviseur)	13
5.1.5	Système de base	14
5.2	Historique de la Virtualisation du Stockage et de la Sauvegarde	14
6	Nouvelle génération des centres informatiques (DataCenter 3.0)	16
6.1	Data Center 1.0	16
6.2	Data Center 2.0	16
6.3	Data Center 3.0	16
6.4	La Consolidation	17
6.5	La Virtualisation	18
6.6	L'Automatisation	18
6.7	Le Green IT / Les Datacenters éco-énergétiques	19
7	La virtualisation des réseaux (Services réseaux virtualisés - VPN)	20
7.1	Introduction aux VPN	20
7.2	Principe de fonctionnement des VPN	20
7.2.1	Concept du VPN	20
7.2.2	Principe général VPN	21
7.2.3	Fonctionnalités des VPN	22
7.2.3.1	Le VPN d'accès	22
7.2.3.2	L'intranet VPN	22
7.2.3.3	L'extranet VPN	23
8	La virtualisation des serveurs	24
8.1	Concepts	24
8.2	Les différents types de virtualisation (Serveurs)	24
8.2.1	L'émulation	24
8.2.2	La virtualisation totale	25
8.2.3	La paravirtualisation	26
8.2.4	La virtualisation matérielle	26
8.3	Virtualisation matérielle	28
8.3.1	Virtualisation matérielle type x86	28
8.3.1.1	Cas d'INTEL	28
8.3.1.2	Cas d'AMD	29
8.3.2	Virtualisation matérielle IBM type p5 / p6 (« Partitionnement »)	31
8.3.2.1	Introduction	31
8.3.2.2	Ressources virtuelles	32
8.3.2.3	Gestion virtuelle	33
8.3.2.4	Accès virtuel	33

8.3.3	Serveurs UNIX - System p d'IBM / Hypervision, Micro-Partitionnement, Virtual I/O Server	34
8.3.3.1	Introduction à l'Hypervision « matérielle ».....	34
8.3.3.2	Virtualisation (Rappel) : Différentes implémentations possibles.....	34
8.3.3.3	Virtualisation / IBM : 40 ans de développement IBM	35
8.3.3.4	Micro-partitioning : CE Capacité d'Exécution.....	36
8.3.3.5	Micro-partitions: Ajustement des puissances	36
8.3.3.6	Illustration de la « Mobilité » des partitions - Live Partition Mobility / p6	37
8.3.3.7	En exemple : Gamme Serveurs IBM Power5 et Power6.....	39
8.4	Les acteurs du marché	40
8.4.1	Présentation « très évolutive » des acteurs X86 (marché « en consolidation ») !	40
8.4.2	Marché du contrôle des infrastructures.....	41
8.5	Architectures de virtualisation de serveurs.....	41
8.5.1	Introduction.....	41
8.5.2	Terminologies / Technologies	41
8.5.3	Architectures de virtualisation de serveurs.....	42
8.5.4	Principe de la Virtualisation	42
8.5.5	Quelles sont les intérêts de la virtualisation ?.....	44
8.5.6	Alléger l'administration système.....	45
8.5.7	Optimiser les ressources	46
8.5.8	Résumé des avantages d'une virtualisation des serveurs :.....	47
8.6	Mobilité des Machines Virtuelles.....	48
8.7	Hyperviseur sans aide du Hardware et HYPERVISEUR avec l'aide du Hardware « Micro Partitionnement ».....	49
8.7.1	Architectures de virtualisation de Serveurs :.....	49
8.7.2	Virtualisation supportée par HYPERVISEUR avec l'aide du Hardware « Micro Partitionnement ».....	49
8.7.3	Granularité des micro-partitions allant jusqu'à 1%.....	49
8.7.4	La technologie des processeurs p5 & p6 d'IBM	50
9	Mainframes & Virtualisation.....	51
9.1	Mainframes, le retour et la Virtualisation / Généralités	51
9.2	Le marché des Serveurs et des « Blades Centers » / Serveurs Lames.....	52
9.2.1	Marché Français des Serveurs	52
9.2.2	Marché Mondial des Serveurs Lames.....	53
10	La virtualisation du stockage (stockage virtualisé et le Thin provisioning, sauvegarde virtualisée / VTL et la technologie de déduplication, archivage et de son réseau spécialisé SAN).....	54
10.1	Introduction / Historique.....	54
10.2	Concepts	55
10.3	Le marché et les acteurs du marché du Stockage	56
10.3.1	Les Constructeurs du Marché du Stockage Informatique.....	56
10.3.2	Quelques Produits RAID / Baies de Stockage « High-End ».....	57
10.4	Le stockage sécurisé / Technologie RAID	58
10.4.1	Notion de RAID / Technologies RAID	58
10.4.2	Les différents niveaux de RAID / Technologie RAID :.....	58
10.4.3	RAID / Niveau 0 ➔ « Array 0 » / Mode « Striping ».....	59
10.4.4	RAID / Niveau 1 - Mode « Mirroring » / « Shadowing » et « Duplexing ».....	59
10.4.5	RAID / Niveau 3 - Transfert parallèle avec gestion de parité en temps réel	60
10.4.6	RAID / Niveau 4 - Disques de données indépendants avec gestion de parité partagée	61
10.4.7	RAID / Niveau 5 - Disques de données indépendants avec gestion de parité partagée et distribuée.....	62

10.4.8	RAID / Niveau 6 - Disques de données indépendants avec gestion de parité partagée, distribuée et répartie	63
10.4.9	RAID / COMBINAISONS (RAID 10, RAID 31).....	64
10.4.10	Le JBOD "Just a Bunch Of Disks" / NRAID	65
10.5	Les Architectures de virtualisation du stockage	68
10.5.1	HDS - Hitachi Data Systems	68
10.5.1.1	USP V – Universal Storage Platform (Business Continuity & Virtualization).....	68
10.5.1.2	Présentation résumée de l’USP V	69
10.5.1.3	Architecture interne USP V	70
10.5.1.4	L’unité de contrôle de l’USP V – Universal Storage Platform	71
10.5.2	IBM - SVC / SAN Volume Controller – « In Band »	72
10.5.2.1	Philosophie des SVC d’IBM	72
10.5.2.2	« SVC – Terminologie » / Architecture logique du SAN.....	72
10.5.2.3	Architecture logique du SAN Volume Controller	73
10.5.2.4	SAN Volume Controller - Environnements supportés	73
10.5.2.5	Avance IBM sur la Virtualisation... par IBM ! (3Q2007)	74
10.5.2.6	Stockage IBM : DS6000 & DS8000.....	74
10.5.2.7	Stockage IBM : DS6000 & DS8000.....	75
10.5.3	EMC (EMC ²)	76
10.5.3.1	VIRTUALISATION EMC ² / INVISTA (« Out of Band »)	76
10.5.3.2	VIRTUALISATION EMC ² basée sur le SAN / INVISTA (« Out of Band »).....	76
10.5.3.3	Le modèle objet d’Invista	77
10.5.3.4	Mobilité dynamique des volumes.....	77
10.5.3.5	Déploiement simplifié d’EMC Invista dans un SAN.	78
10.5.3.6	EMC ² (DMX-3) : Vision par HDS !	78
10.5.3.7	STOCKAGE EMC ² - Gamme DMX.....	79
10.6	Clones, Snapshots, Réplication à distance (PCA, PRA, PRCA).....	80
10.6.1	Remote Replication with Hitachi Universal Replicator Software (HUR).....	80
10.6.2	HUR: Détails de la technologie (“Key Technology”)	81
10.6.3	Hitachi Universal Replicator Software: Asynchronous Remote Copy.....	81
10.6.4	Avantages avec HUR.....	82
10.6.5	Continuité de service avec HUR (Virtualisation du stockage).....	83
10.7	Technologie de « Thin provisioning ».....	84
10.7.1	Hitachi Dynamic Provisioning (HDP) sur USP V	84
10.7.2	Hitachi Dynamic Provisioning (HDP) / Gestion d’approvisionnements de volumétrie.....	85
10.7.3	Hitachi Dynamic Provisioning (HDP) / – Seuils d’alerte.....	86
10.8	La Sauvegarde virtualisée / VTL / Technique de déduplication	87
10.8.1	Data Center / La Sauvegarde	87
10.8.2	Les challenges de la gestion du Stockage et de sa Sauvegarde	87
10.8.3	Les acteurs majeurs (Constructeurs et OEM) VTL	88
10.8.4	Le Marché du Disque Dur et de la Bande dans les Sauvegardes.....	89
10.8.5	Les chiffres du marché VTL :	89
10.8.6	Le Marché du VTL et de la Déduplication : « Carré magique »	90
10.8.7	Sauvegarde / Data Center	90
10.8.8	Architecture VTL (indépendance ou non / aux librairies).....	92
10.8.8.1	Architecture avec dépendance du type réel de la librairie gérée	92
10.8.8.2	Architecture avec indépendance du type réel de la librairie « gérée »	92
10.8.9	La technologie de Déduplication (SIS, SIR, CAS).....	93
10.8.9.1	Introduction à la technologie de Déduplication (SIS, SIR, CAS)	93
10.8.9.2	La technologie de Déduplication / Concept général.....	93

10.8.9.3	La technologie de Déduplication / Illustration (Phase 1)	94
10.8.9.4	La technologie de Déduplication / Illustration (Phase 2)	95
10.8.9.5	La technologie de Déduplication / « Hashing »	96
10.8.9.6	Déduplication / Sauvegarde sur disques	96
10.8.10	La technologie de Déduplication « Temps réel »	97
10.8.11	Technologie de Déduplication / Reconstitution des données	98
10.8.12	Exemple d'implémentation : Bloc diagramme (Diligent Technologies / IBM)	99
10.8.13	La technologie de Déduplication / Résumé	100
10.8.14	La Technologie de Déduplication en TELCO's	100
10.8.15	Les Acteurs majeurs avec « De-Dupe » (Constructeurs / Éditeurs et OEM)	101
11	Concept de la virtualisation du poste de travail	103
11.1	Le « Bureau Virtuel »	103
11.2	Le « Streaming Applicatif »	103
11.3	La Virtualisation du Poste de Travail	103
11.3.1	Introduction	103
11.3.2	Présentation	103
11.3.3	Concept VDI (« voix, données, images » Câblage et connectique normalisée) :	104
11.3.4	Architecture VDI	104
11.3.5	Expérience Utilisateur / Architecture WAN, LAN et Local	105
11.3.6	Virtualisation du Poste de Travail : Principales technologies	105
12	Concept de la virtualisation applicative	106
12.1	Généralités de la virtualisation applicative	106
12.2	Principe de la virtualisation applicative	106
12.3	Intérêt de la virtualisation applicative	107

1 Préambule et « avertissements »

Ce document est un support de cours et un complément d'information pour le stagiaire, il sera à compléter par les explications et commentaires au cours des 2 jours de ce module.

Remarque :

Beaucoup d'illustrations sont issues de « la toile », et une partie de la prose de certains paragraphes est également issue de la même origine. Ces documents sources sont donc du domaine public. Ces pages Web peuvent être retrouvées par le moteur de recherche favori du lecteur en utilisant certains mots clés associés aux pages recherchées. Pour éviter de donner des adresses de liens obsolètes, aucune bibliographie ne sera intégrée à ce support de cours. D'autre part, en faisant ses propres recherches, le stagiaire peut donc trouver d'autres illustrations appropriées à ce support de cours. Bonne navigation !

2 Objectif du Module

2.1 Être capable de :

- Comprendre les principes de la VIRTUALISATION (IT).
- Comprendre ce qu'est la virtualisation, les étapes d'une démarche de virtualisation, ses enjeux, ses contraintes.

2.2 Répondre à la question :

- Répondre à la question : Pourquoi et Comment ça marche (vulgarisation).

3 Descriptif et résumé du contenu

- **Définition du concept de Virtualisation.**
- **Historique de la Virtualisation (UC et Stockage).**
- **Nouvelle génération des centres informatiques (DataCenter 3.0) :**
 - **La consolidation,**
 - **La virtualisation,**
 - **L'automatisation.**
- **La virtualisation des réseaux (services réseaux virtualisés / VPN).**
- **La virtualisation des serveurs :**
 - **Concepts,**
 - **Les acteurs du marché,**
 - **Architectures de virtualisation de serveurs,**
 - **Mobilité des Machines Virtuelles,**
 - **Hyperviseur sans aide du Hardware et HYPERVISEUR avec l'aide du Hardware « Micro Partitionnement ».**
- **La virtualisation du stockage (stockage virtualisé et le Thin provisioning, sauvegarde virtualisée / VTL et la technologie de déduplication, archivage et de son réseau spécialisé SAN) :**
 - **Concepts,**
 - **Le marché et les acteurs du marché du Stockage,**
 - **Le stockage sécurisé / Technologie RAID,**
 - **La technologie des disques (FC et SATA II) et les SSD,**
 - **Les Architectures de virtualisation du stockage,**
 - **Classes de stockage et Mobilité / Migration « on line » des Données,**
 - **Clones, Snapshots, Réplication à distance (PCA, PRA, PRCA),**
 - **Technologie de « Thin provisioning »,**
 - **La Sauvegarde virtualisée / VTL / Technique de déduplication.**
- **Concept de la virtualisation du poste de travail.**
- **Concept de la virtualisation applicative.**

4 Définition du concept de Virtualisation

Virtualisation → Abstraction Physique / Logique → Leurre !

La virtualisation

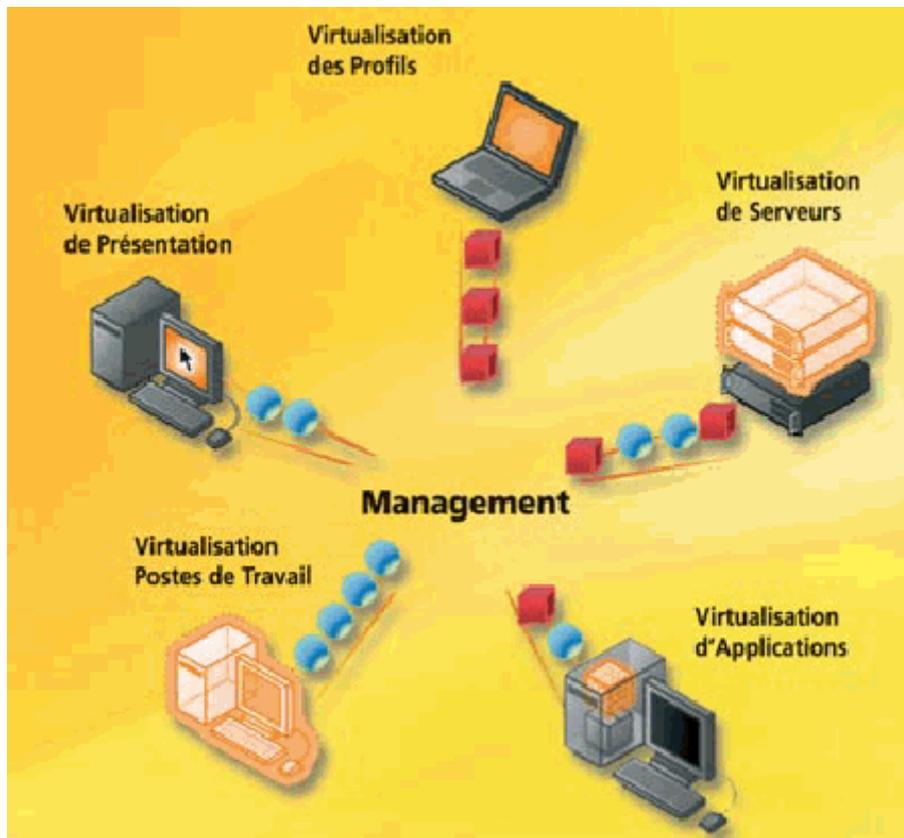
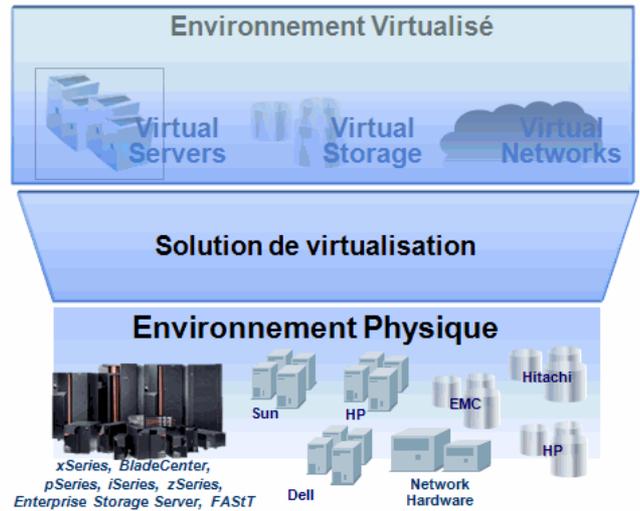
Une vue logique des ressources IT

Plutôt que de disposer de ressources de calcul et de stockage de manière traditionnelle, c'est à dire dédiées à des environnements physiques, la virtualisation permet de s'appuyer sur des pools de ressources logiques.

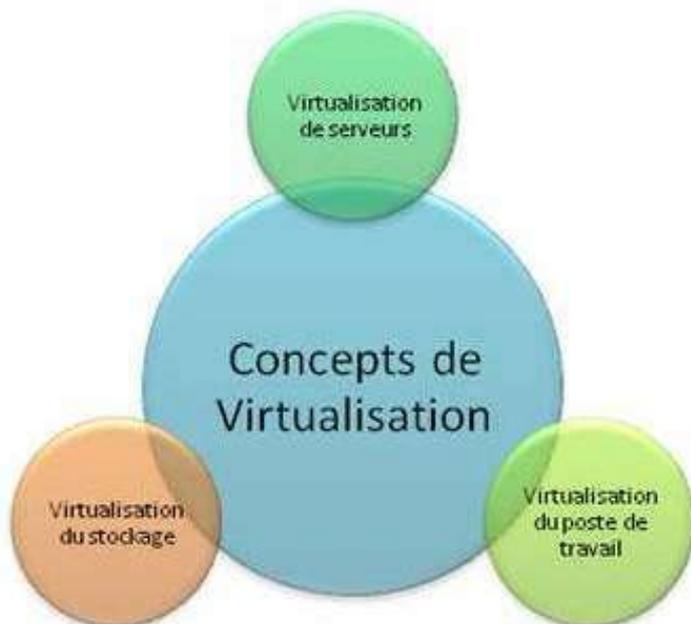
En d'autres mots, la **virtualisation** fournit une **vue logique** plutôt que **physique** des ressources.

La virtualisation permet de :

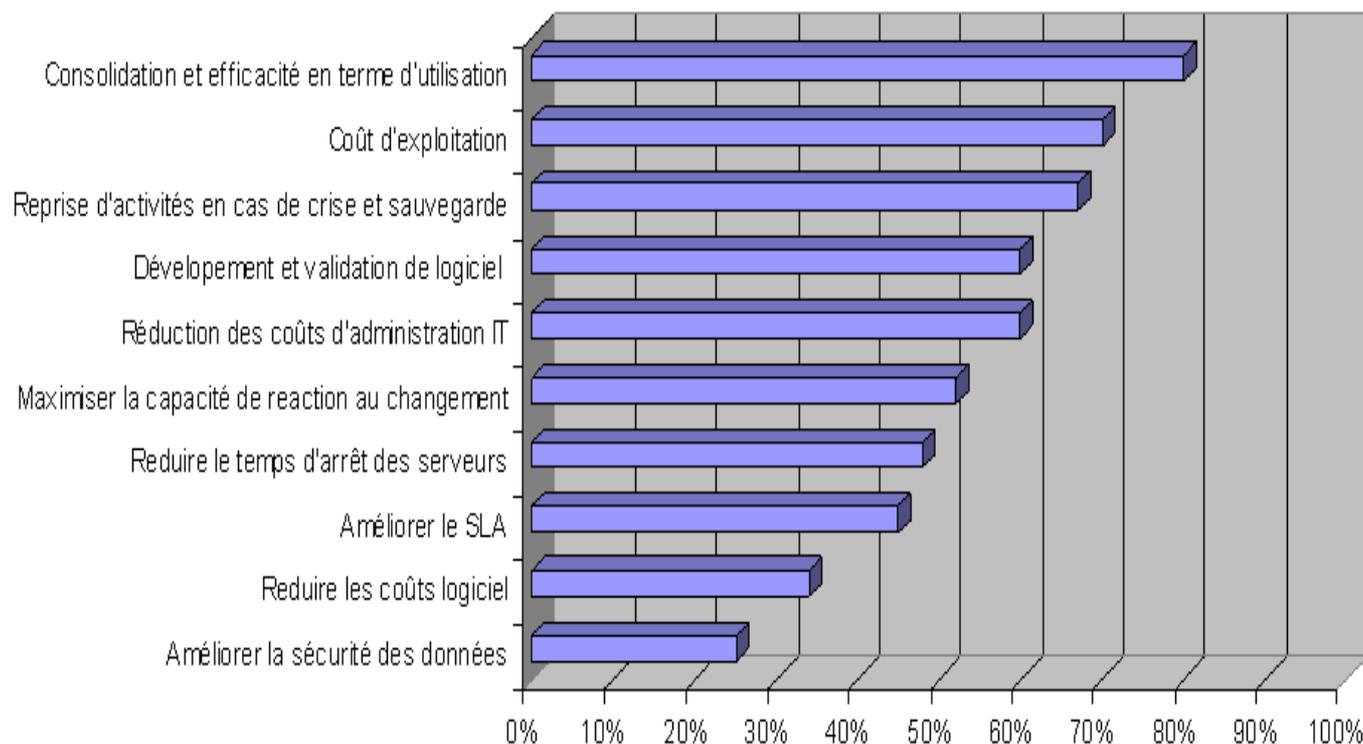
- simplifier l'infrastructure IT
- optimiser le taux d'utilisation des ressources IT.
- faire face aux variations de charge



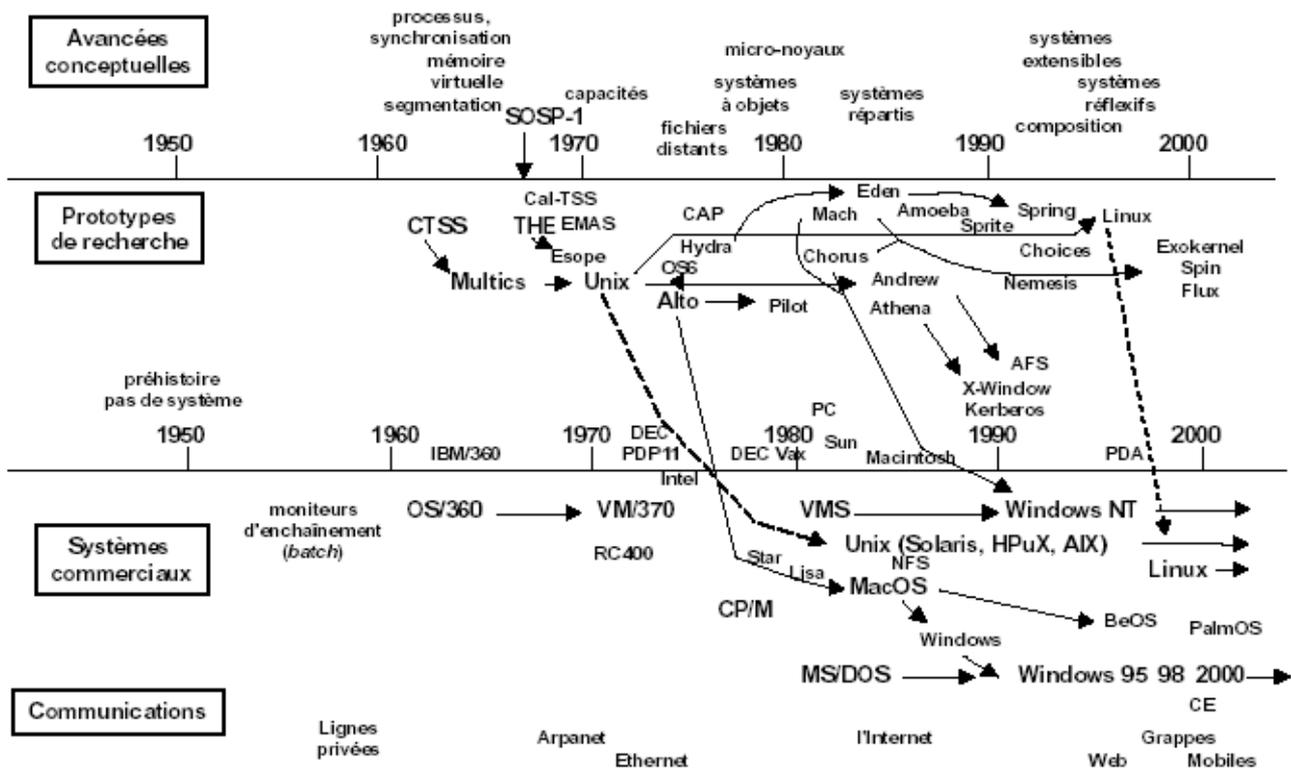
Les aspects de la virtualisation sont multiples et variés :



Les raisons évoquées par les DSI pour virtualiser sont multiples et variés :



5 Historique de la Virtualisation (UC et Stockage)



5.1 Historique de la Virtualisation des UC / Serveurs

5.1.1 Introduction (Historique de la Virtualisation)

Historiquement, la virtualisation est apparue dans les années 70 (IBM VM 370). Elle permettait de faire tourner sur un mainframe plusieurs OS incompatibles (DOS/VSE et OS/MVS...) et créait déjà des machines individuelles (CMS), préfigurant le fonctionnement des premiers PC.

5.1.2 Détails (Historique de la Virtualisation)

Juin 1959 : Christopher Strachey publie un papier intitulé « Time Sharing in Large Fast Computer » à la conférence internationale sur le traitement de l'information à l'UNESCO.

En 1974, Christopher clarifie son idée présenté en 1959 dans un email adressé à Donald Knuth. Il y explique que son papier précédent porte principalement sur la multiprogrammation. Cette idée de multiprogrammation se rapporte à l'idée de ne plus attendre le chargement d'un périphérique, et de pouvoir travailler sur une autre tâche sur son ordinateur. C'est en ce terme qu'il parle de « Time Sharing ».

1960 : La première utilisation de la multiprogrammation peut être attribuée à l'ordinateur Atlas. Le projet Atlas est développé conjointement par l'Université de Manchester et Ferrandi Ltd.

En 1961, les SER (Supervisor Extracode Routine) forment les « branches principales » du superviseur (terme employé pour décrire un hyperviseur). Ils sont activés aussi bien par des routines (partie d'un code informatique destinée à être utilisé plus d'une fois) d'interruption que par des instructions d'« extracode ». Le superviseur d'Atlas utilise alors une machine virtuelle. Une autre machine virtuelle est alors employée pour faire tourner les programmes des utilisateurs.

Dans la même année, au MIT (Massachusetts Institute of Technology), le CTSS (Compatible Time Sharing System) est développé sur des machines IBM 704.

Le programme de supervision du CTSS repose sur une console I/O :

- organisation via des jobs fg (« foreground » ou « premier plan ») et bg (« background » ou en tâche de fond)
- stockage temporaire
- etc...

Le superviseur a le contrôle direct de toutes les interruptions collectées.

En 1963, le projet MAC du MIT a pour but le design et l'implémentation du meilleur temps partagé (« Time Sharing ») d'un système basé sur le CTSS.

Vers 1965, le centre de recherche d'IBM est en charge du projet M44/44X. Son but étant d'évaluer les concepts naissant sur le « **partage du temps** ».

L'architecture est la suivante :

- une machine principale IBM 7044 (M44)
- et d'autres machines où chacune est une image expérimentale de la machine principale (M44)

L'espace d'adressage de 44X réside alors dans la hiérarchie de la mémoire de M44, implémenté via de la mémoire virtuelle et de la multiprogrammation.

En 1967, l'idée de virtualisation matérielle est développée via le modèle 67 (System/360) (Voir son pupitre ci-dessous) d'IBM. Ce modèle virtualise toutes les interfaces matérielles via la VMM (Virtual Machine Monitor). Le système d'exploitation est le TSS (« Time-Sharing System ») aussi appelé le superviseur. L'IBM 360/67 est équipé d'un « hyperviseur » appelé CP/67 capable de supporter plusieurs machines virtuelles 360/65 tournant sous des versions de système différentes. CP/67 a engendré VM/CMS (CMS correspond à la notion de système de temps partagé), qu'IBM a commercialisé pendant les années 1970-1980.



Pupitre d'un modèle IBM 360/67

La VMM tourne directement sur le « vrai » matériel, ceci permet l'existence de plusieurs machines virtuelles : VMS. Où chaque machine virtuelle tourne sur son propre système d'exploitation. Les machines virtuelles d'IBM de nos jours sont des offres très respectées et des plateformes de traitement robustes.

Durant les années 90, la notion de « processeur simulé » voit le jour. C'est la machine P-code (ou pseudo-code). C'est un langage machine qui est exécuté dans une machine virtuelle. Le langage Java™ a suivi le modèle de P-code pour sa machine virtuelle. Ceci a permis la distribution large des programmes de Java au-dessus des architectures innombrables en mettant en communication simplement la JVM (Java Virtual Machine).

A la fin des années 90 et au début des années 2000, la société VMware développe et popularise un système propriétaire de virtualisation logicielle des architectures du type x86 pour les architectures de type x86.

L'architecture IA-32 (X86) fournit alors quelques instructions pour la virtualisation. Il faut savoir que les instructions liées au mode privilégié peuvent renvoyer différents résultats en fonction du mode. Par exemple, l'instruction x86 STR s'occupe de rapporter l'état de sécurité.

Mais, la valeur qu'elle renvoie dépend du niveau de privilège. Ainsi, c'est pour cela que dès 2006 les fabricants de processeur x86 que sont AMD et Intel ont proposé dans leurs gammes de la virtualisation matérielle.

5.1.3 Utilisations

La virtualisation séduit autant les professionnels IT que les enseignants pour de multiples raisons :

1. Développement

- tracer les programmes : faire du « débogage »
- tester la portabilité : plusieurs systèmes d'exploitation sur un ordinateur
- déploiement et migration de machines virtuelles
- possibilité de créer des points de restauration de machines virtuelles par le biais de « Snapshots » : image d'une machine virtuelle à un instant donnée.

2. Enseignement

- isoler des machines entre elles : empêcher des étudiants de communiquer entre eux lors de devoirs, ou de casser des machines

3. Sécurité par l'isolation

- s'emprisonner dans un répertoire d'un système hôte, où l'accès à ce dernier est impossible depuis le répertoire
- restreindre l'accès à une partie de la mémoire
- jail (« prison ») : enfermement dans une machine virtuelle avec un accès restreint à la mémoire

5.1.4 Définitions (Virtualisation, Machine Virtuelle, Hyperviseur)

Définition 1 : **Virtualisation** (définition générale) elle a pour but de donner un environnement système au programme afin de le faire croire être dans un environnement matériel.

Définition 2 : **Machine virtuelle** (VM - Virtual Machine) le sens originel de machine virtuelle est la création de plusieurs environnements d'exécution sur un seul ordinateur, dont chacun émule l'ordinateur hôte. Cela fournit à chaque utilisateur l'illusion de disposer d'un ordinateur complet alors que chaque machine virtuelle est isolée des autres. Le logiciel hôte qui fournit cette fonctionnalité est souvent dénommée superviseur ou hyperviseur. Ce concept va plus loin que celui des simples temps partagées où chaque

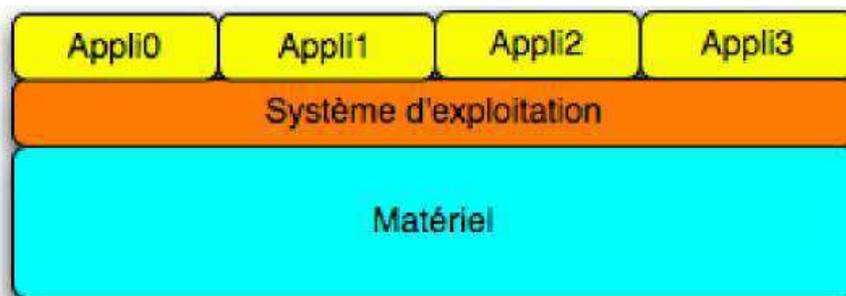
utilisateur dispose seulement d'un espace de développement personnel, et non d'une machine simulée entière.

Définition 3 : **Hyperviseur** (ou VMM - Virtual Machine Monitor) c'est le logiciel hôte qui fournit la fonctionnalité d'exécuter une machine virtuelle sur un système. Le système d'exploitation communique directement avec le matériel.

5.1.5 Systeme de base

Un tel système est basé sur un système d'exploitation prévu pour fonctionner avec le matériel utilisé.

Tout d'abord, l'application envoie des instructions au système d'exploitation qui permet au matériel de l'exécuter. La figure ci-dessous décrit ce système de base.



Systeme de base

5.2 Historique de la Virtualisation du Stockage et de la Sauvegarde

Dès 1980, la start-up française COPERNIQUE développe une baie de stockage, le DIRAM 32, gérant jusqu'à 16 disques en double accès (en interface SMD de l'époque) et pouvant être connecté jusqu'à 4 ordinateurs d'un même constructeur.

Ce produit « virtualisait » :

- des couples de disques en Mirroring pour assurer la continuité d'activité en cas de panne d'un disque (concept que l'on désigne actuellement par technologie RAID-1,

- un disque « virtuel » simulé en mémoire électronique dans une partie du cache d'E/S
- un processeur programmable de traitement sur données des disques pour des recherches associatives sur critères via un canal standard d'E/S disque. Il faut noter que ce concept novateur, qui a permis, entre autres gros projets, de développer l'Annuaire Électronique du téléphone de France Télécom (renseignement téléphonique sur Minitel via le 11 / 3611 de FT), n'a jamais été ré-implémenté dans les baies de stockage d'IBM, de HDS, ni d'EMC².

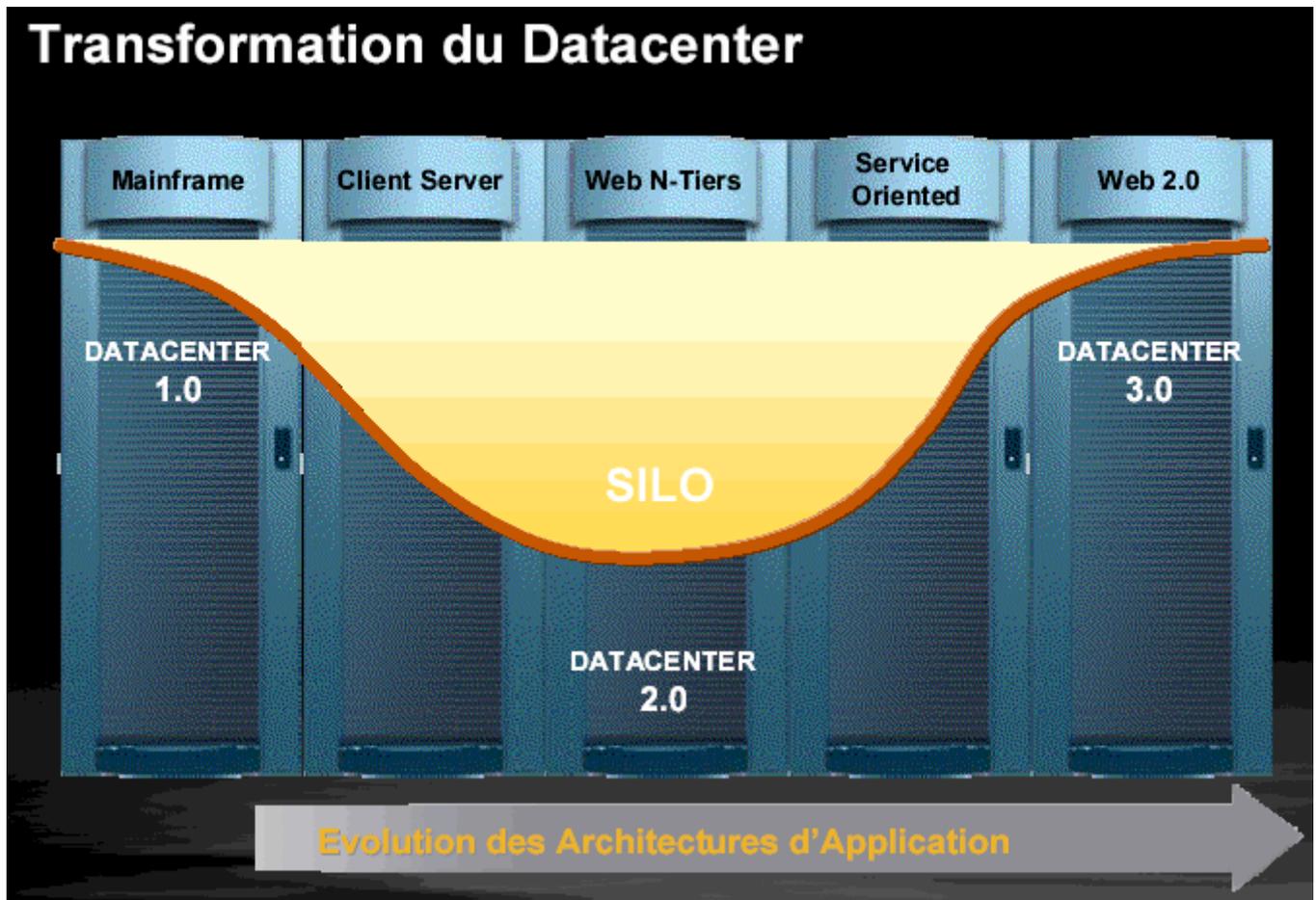
La virtualisation s'est invitée dans le stockage en tout début de période (2000) dès le déploiement d'environnements SAN / NAS avec la création d'unités logiques virtuelles LUN (« Logical Unit Number ») qui ont permis un premier niveau d'abstraction entre les volumes de stockage vus par les applications et le stockage physique mono ou multi-disques. Du côté technologie, on peut donc estimer qu'il y a déjà une bonne maturité.

L'éditeur américain DataCore a été, dès 1999, l'un des premiers à proposer une offre « mature » de Virtualisation « In-Band » du Stockage au travers du SAN avec sa suite « SanSymphony ». Son Appliance « In-Band » était basée sur des serveurs « multiprocesseurs » à base d'X86 sous NT. Depuis, vers 2003, cette solution a été reprise « à l'identique » dans son architecture et ses fonctionnalités par IBM avec ses « SVC - SAN Volume Controller » à base de 2 serveurs PowerPC montés en cluster (nœud SVC).

L'implémentation technique de l'offre de Virtualisation de HDS et d'IBM est du type « In-Band » :

La virtualisation est directement intégrée aux baies de stockage chez HDS (USP – Universal Storage Platform / gamme TagmaStore). Pour IBM, la virtualisation est gérée par le SVC (nœud de cluster) qui est une Appliance spécialisée indépendante des baies de stockage, et qui gèrent la virtualisation pour les baies de stockage qui y sont connectés. Les SVC sont proposés par IBM avec les baies DS4000, DS6000 et même les DS8000. Pour ces derniers, IBM introduit progressivement des couches de virtualisation directement dans les baies DS8000, mais l'offre est très en retard par rapport à toutes les fonctionnalités des SVC. D'autre part, l'introduction de la virtualisation au niveau des DS6000 pose probablement un problème de stratégie par rapport à ses SVC. L'implémentation technique de l'offre de Virtualisation d'EMC² pour ses baies DMX est du type « Out of Band », la virtualisation est traitée au niveau des directeurs du SAN gérant les baies de stockage. C'est l'offre de virtualisation la plus en retard par rapport à HDS et IBM.

6 Nouvelle génération des centres informatiques (DataCenter 3.0)



DataCenter V3.0 → Nouvelle génération des centres informatiques

6.1 Data Center 1.0

Centralisation à l'extrême autour de Mainframes accessibles par les utilisateurs au travers de différents terminaux.

6.2 Data Center 2.0

L'arrivée des technologies clients / serveurs nous a amené à l'étape Data center 2.0 évoluant ainsi d'une « Architecture » hautement centralisée à une autre hautement distribuée.

6.3 Data Center 3.0

Par exemple Vision de CISCO (24/07/2007)

Le Data Center 3.0 va combiner le meilleur de ces deux mondes en permettant de « recentraliser » les serveurs et le stockage autour d'une plateforme réseau(x) hautement disponible et évolutive en temps réel, tout en préservant les avantages et la souplesse d'une architecture distribuée.

➔ Les acteurs concernés sont les Acteurs majeurs dans le stockage, les serveurs et les systèmes d'exploitation.

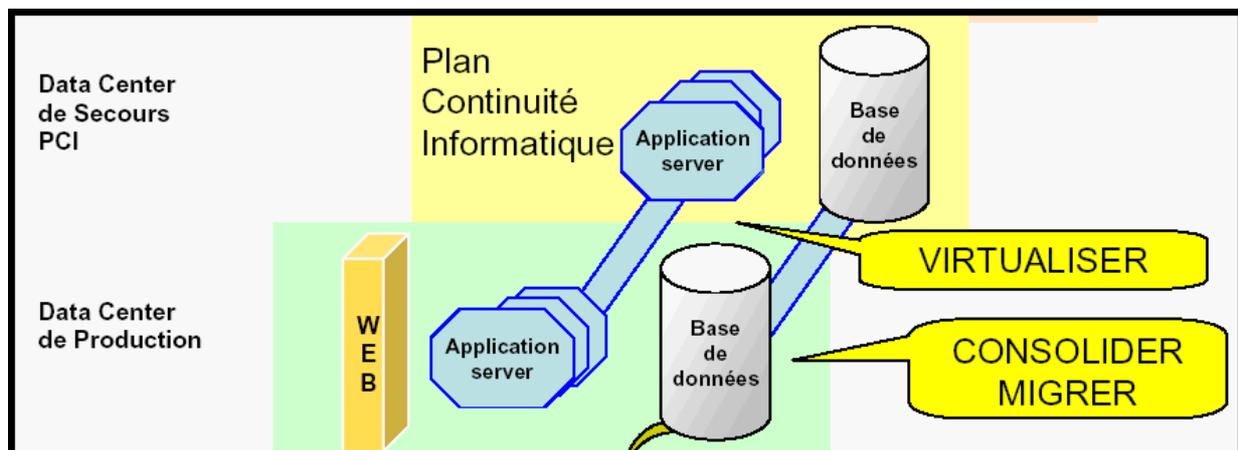
La réalisation du Data Center 3.0 va reposer sur 3 piliers :

- La consolidation,
- La virtualisation,
- L'automatisation.

6.4 La Consolidation

La consolidation de serveurs permet de réduire le TCO (« Total Cost of Ownership » ou coût total de possession), le temps de maintenance, simplifier l'administration de l'infrastructure, sécuriser l'architecture systèmes et réseaux et mettre en œuvre des PRA.

Démarche Consolidation / Virtualisation du DataCenter



La virtualisation est une technologie clé pour les projets de consolidation.

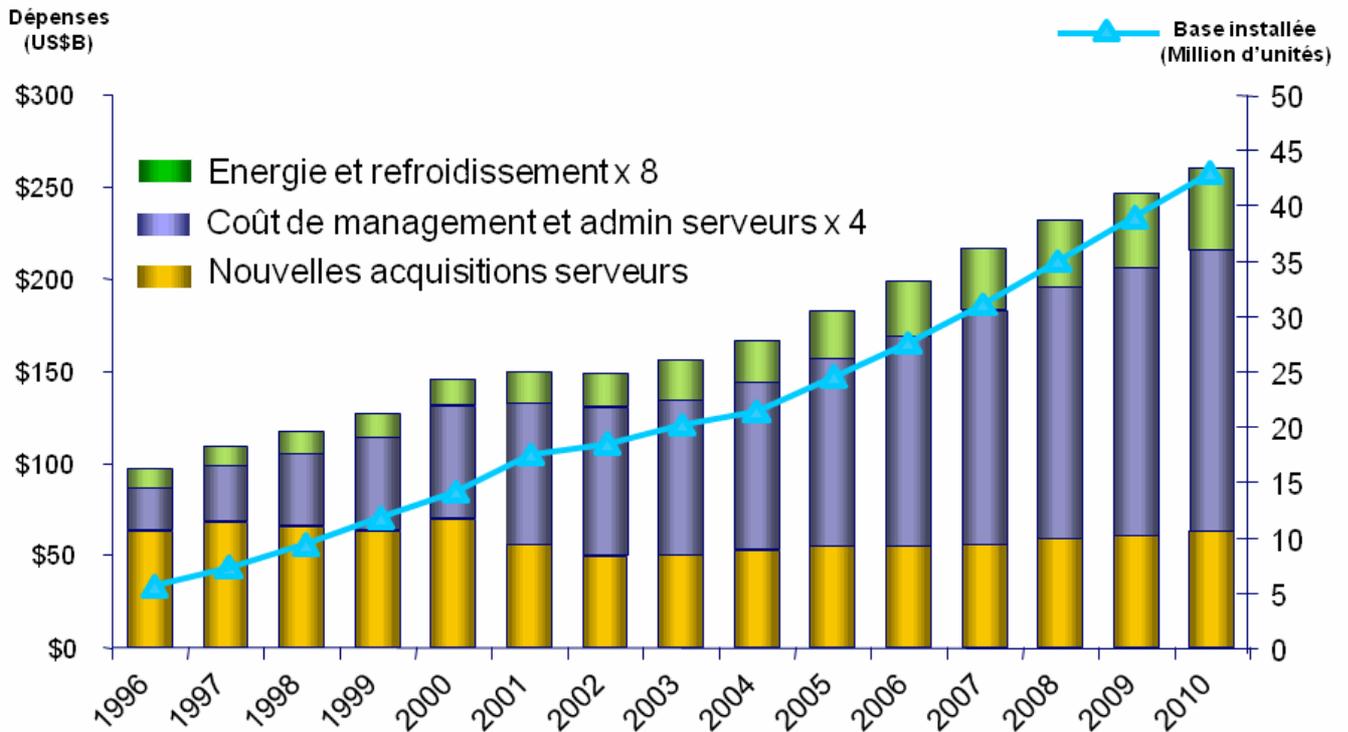
Déploiement informatique : trop de serveurs ➔ Démarche

- Comment réduire la complexité et simplifier l'administration des systèmes d'information ?
- Mes serveurs sont utilisés à 10-20 ou 30% de leur puissance, comment optimiser ?
- Coûts en électricité & Climatisation (Développement Durable / DSI - Initiatives « Bio Data Center »), et Occupation au sol, comment améliorer mon ROI ?

➔ Voir également le « Green IT » ci-après.

6.5 La Virtualisation

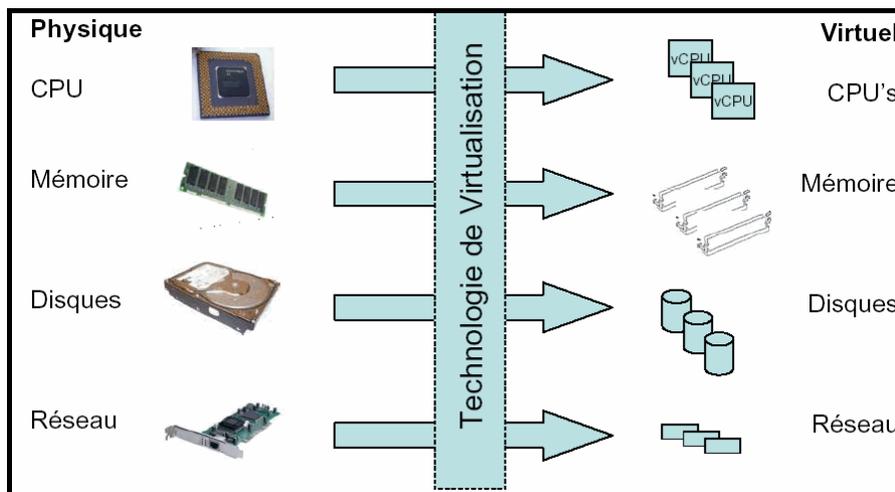
Une nécessité : Optimiser les ressources IT / investissements.



6.6 L'Automatisation

L'efficacité de l'automatisation IT améliorera l'utilisation des ressources tout en libérant les administrateurs IT pour qu'ils puissent innover sur l'aspect des métiers de leur organisation.

Infrastructure informatique virtualisée – même principe

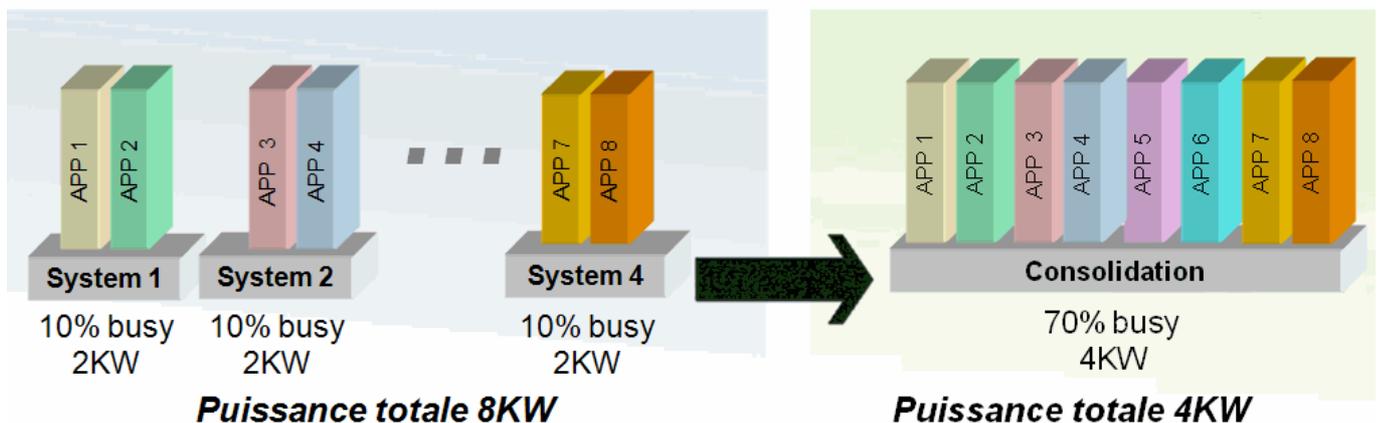


L'infrastructure virtualisée apporte :

- plus de flexibilité
 - une meilleure utilisation des ressources
 - une meilleure disponibilité à coût réduit
 - des coûts d'administration réduits
 - une meilleure interopérabilité
 - la protection de vos investissements.
- ... et offre des coûts réduits et une meilleure agilité.

6.7 Le Green IT / Les Datacenters éco-énergétiques

Consolidation de serveurs et du stockage : vers une meilleure gestion de l'énergie



La consolidation de serveurs et du stockage, combinée à la virtualisation est une approche très optimisée pour répondre à la réduction des coûts énergétiques

7 La virtualisation des réseaux (Services réseaux virtualisés - VPN)

→ **Réseau Privé Virtuel** ou RPV (« Virtual Private Network » en anglais, abrégé en VPN)

7.1 Introduction aux VPN

→ **La Virtualisation des Réseaux est essentiellement implémentées par Services.**

Le VPN (pour Virtual Private Network) est une technologie de "Réseau Privé Virtuel". Il permet à un ordinateur distant d'avoir, via Internet, un accès direct et totalement sécurisé à un autre ordinateur ou à un réseau local. Cette technologie dispense d'avoir recours à de coûteuses solutions de location de connexions privées et spécifiques.

Les applications et les systèmes distribués font de plus en plus partie intégrante du paysage d'un grand nombre d'entreprises. Ces technologies ont pu se développer grâce aux performances toujours plus importantes des réseaux locaux. Mais le succès de ces applications a fait aussi apparaître un de leur écueil. En effet si les applications distribuées deviennent le principal outil du système d'information de l'entreprise, comment assurer leur accès sécurisé au sein de structures parfois réparties sur de grandes distances géographiques ? Concrètement comment une succursale d'une entreprise peut-elle accéder aux données situées sur un serveur de la maison mère distant de plusieurs milliers de kilomètres ? Les VPN ont commencé à être mis en place pour répondre à ce type de problématique. Mais d'autres problématiques sont apparues et les VPN ont aujourd'hui pris une place importante dans les réseaux informatique et l'informatique distribuées.

7.2 Principe de fonctionnement des VPN

7.2.1 Concept du VPN

Les réseaux locaux d'entreprise (**LAN** ou RLE) sont des réseaux internes à une organisation, c'est-à-dire que les liaisons entre machines appartiennent à l'organisation. Ces réseaux sont de plus en plus souvent reliés à Internet par l'intermédiaire d'équipements d'interconnexion. Il arrive ainsi souvent que des entreprises éprouvent le besoin de communiquer avec des filiales, des clients ou même du personnel géographiquement éloignées via internet.

Pour autant, les données transmises sur Internet sont beaucoup plus vulnérables que lorsqu'elles circulent sur un réseau interne à une organisation car le chemin emprunté n'est pas défini à l'avance, ce qui signifie que les données empruntent une infrastructure réseau publique appartenant à différents opérateurs. Ainsi il n'est pas impossible que sur le chemin parcouru, le réseau soit écouté par un utilisateur indiscret ou même détourné. Il n'est donc pas concevable de transmettre dans de telles conditions des informations sensibles pour l'organisation ou l'entreprise.

La première solution pour répondre à ce besoin de communication sécurisé consiste à relier les réseaux distants à l'aide de liaisons spécialisées. Toutefois la plupart des entreprises ne peuvent pas se permettre de relier deux réseaux locaux distants par une ligne spécialisée, il est parfois nécessaire d'utiliser Internet comme support de transmission.

Un bon compromis consiste à utiliser Internet comme support de transmission en utilisant un protocole d'encapsulation" (en anglais « *tunneling* », d'où l'utilisation impropre parfois du terme "tunnelisation"), c'est-à-dire encapsulant les données à transmettre de façon chiffrée. On parle alors de **réseau privé virtuel** (noté *RPV* ou **VPN**, acronyme de *Virtual Private Network*) pour désigner le réseau ainsi artificiellement créé.

Ce réseau est dit *virtuel* car il relie deux réseaux "physiques" (réseaux locaux) par une liaison non fiable (Internet), et *privé* car seuls les ordinateurs des réseaux locaux de part et d'autre du VPN peuvent "voir" les données.

Le système de *VPN* permet donc d'obtenir une liaison sécurisée à moindre coût, si ce n'est la mise en œuvre des équipements terminaux. En contrepartie il ne permet pas d'assurer une « qualité de service » comparable à une ligne louée dans la mesure où le réseau physique est public et donc non garanti.

7.2.2 Principe général VPN

Un réseau VPN repose sur un protocole appelé "protocole de tunneling". Ce protocole permet de faire circuler les informations de l'entreprise de façon cryptée d'un bout à l'autre du tunnel. Ainsi, les utilisateurs ont l'impression de se connecter directement sur le réseau de leur entreprise.

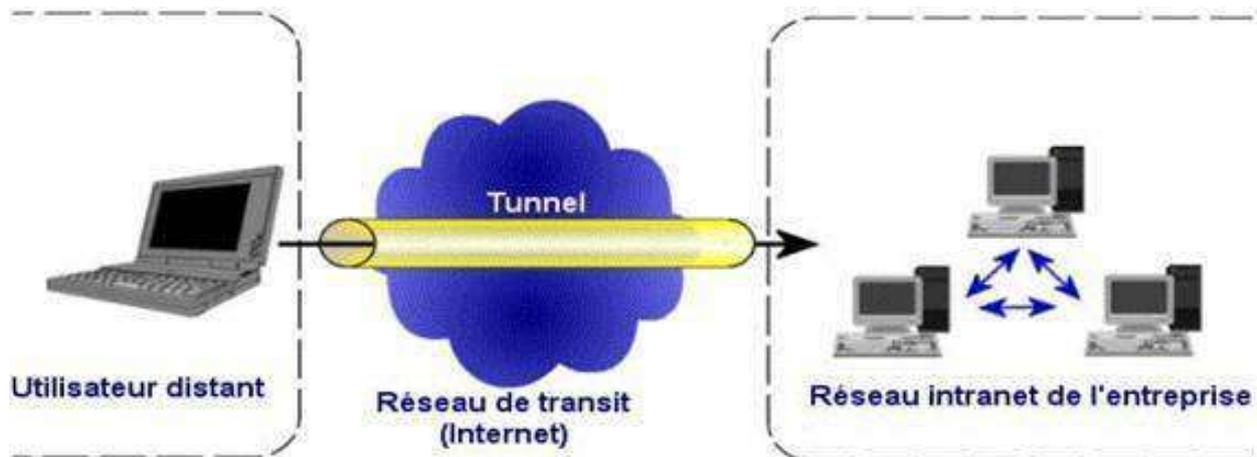
Le principe de tunneling consiste à construire un chemin virtuel après avoir identifié l'émetteur et le destinataire. Par la suite, la source chiffre les données et les achemine en empruntant Ce chemin virtuel. Afin d'assurer un accès aisé et peu coûteux aux intranets ou aux extranets d'entreprise, les réseaux privés virtuels d'accès simulent un réseau privé, alors qu'ils utilisent en réalité une infrastructure d'accès partagée, comme Internet.

Les données à transmettre peuvent être prises en charge par un protocole différent d'IP. Dans Ce cas, le protocole de tunneling encapsule les données en ajoutant une en-tête. Le tunneling est l'ensemble des processus d'encapsulation, de transmission et de désencapsulation.

7.2.3 Fonctionnalités des VPN

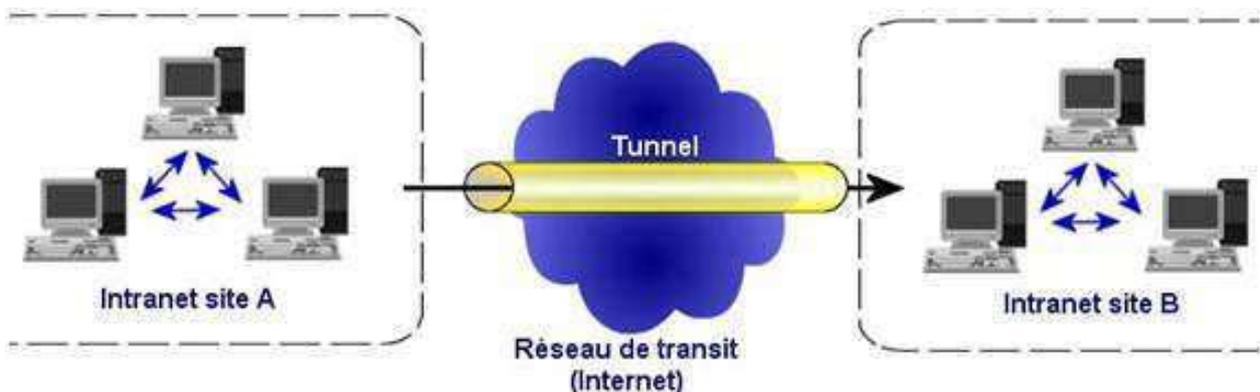
➔ Il existe 3 types standards d'utilisation des VPN. En étudiant les schémas types d'utilisation, il est possible d'isoler les fonctionnalités indispensables des VPN.

7.2.3.1 Le VPN d'accès



Le VPN d'accès est utilisé pour permettre à des utilisateurs itinérants d'accéder au réseau privé. L'utilisateur se sert d'une connexion Internet pour établir la connexion VPN.

7.2.3.2 L'intranet VPN



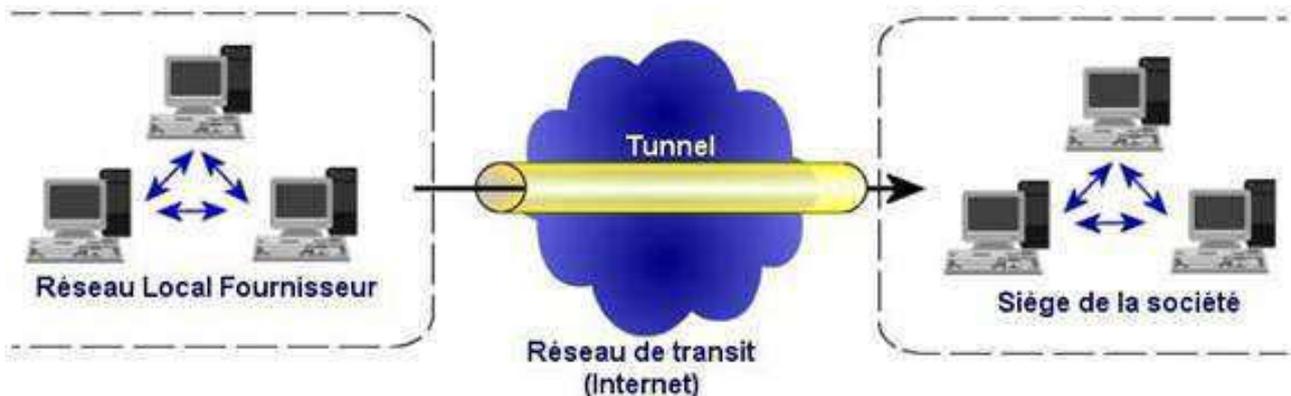
➔ L'intranet VPN est utilisé pour relier au moins deux intranets entre eux.

Ce type de réseau est particulièrement utile au sein d'une entreprise possédant plusieurs sites distants.

Le plus important dans ce type de réseau est de garantir la sécurité et l'intégrité des données. Certaines données très sensibles peuvent être amenées à transiter sur le VPN (base de données clients, informations financières...). Des techniques de cryptographie sont mises en œuvre pour vérifier que les

données n'ont pas été altérées. Il s'agit d'une authentification au niveau paquet pour assurer la validité des données, de l'identification de leur source ainsi que leur non-répudiation. La plupart des algorithmes utilisés font appel à des signatures numériques qui sont ajoutées aux paquets. La confidentialité des données est, elle aussi, basée sur des algorithmes de cryptographie. La technologie en la matière est suffisamment avancée pour permettre une sécurité quasi parfaite. Le coût matériel des équipements de cryptage et décryptage ainsi que les limites légales interdisent l'utilisation d'un codage "infaillible". Généralement pour la confidentialité, le codage en lui-même pourra être moyen à faible, mais sera combiné avec d'autres techniques comme l'encapsulation IP dans IP pour assurer une sécurité raisonnable.

7.2.3.3 L'extranet VPN



Une entreprise peut utiliser le VPN pour communiquer avec ses clients et ses partenaires. Elle ouvre alors son réseau local à ces derniers. Dans ce cadre, il est fondamental que l'administrateur du VPN puisse tracer les clients sur le réseau et gérer les droits de chacun sur celui-ci.

8 La virtualisation des serveurs

8.1 Concepts

Il s'agit de l'ensemble des techniques matérielles et/ou logicielles qui permettent de faire fonctionner sur une seule machine physique plusieurs systèmes d'exploitation et ou plusieurs applications, isolés les uns des autres, comme s'ils fonctionnaient sur des machines physiques distinctes.

La consolidation de serveurs permet de réduire le TCO (« Total Cost of Ownership » ou coût total de possession), le temps de maintenance, simplifier l'administration de l'infrastructure, sécuriser l'architecture systèmes et réseaux et mettre en œuvre des PRA.

8.2 Les différents types de virtualisation (Serveurs)

8.2.1 L'émulation

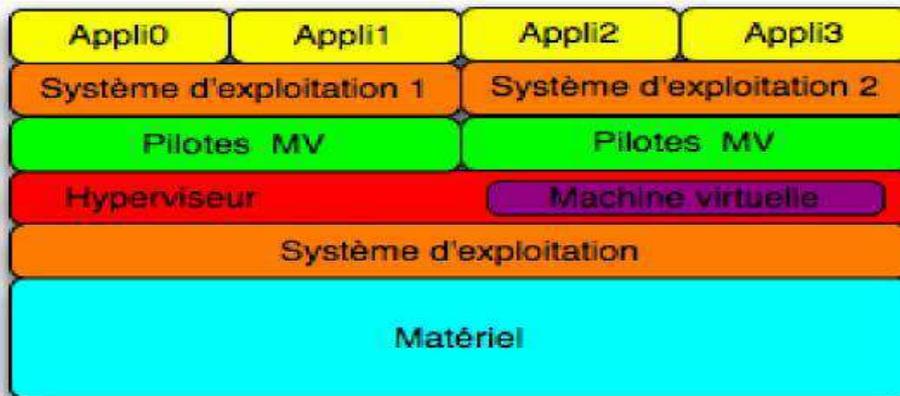
Dans ce cas, un environnement est créé où l'on simule les périphériques d'E/S (« hardware »).

Chaque application donne des instructions via des pilotes à l'hyperviseur qui doit passer par le système d'exploitation pour exécuter celles-ci (voir la figure ci-après).

Avantage : ce type de virtualisation permet de travailler avec des systèmes prévus pour un matériel totalement différent que celui sur lequel il tourne grâce à l'émulation du matériel.

Inconvénient : il y a une grosse perte de performance du fait du nombre de couches traversées et de la conversion nécessaire pour passer d'un langage correspondant à un matériel à l'autre.

Utilisations : en premier lieu, il est possible alors de tester des applications multiplateformes sans avoir besoin d'acquérir le matériel nécessaire. Ou alors de lancer une application prévue pour un autre matériel (jouer à des jeux d'anciennes consoles sur un ordinateur personnel). L'émulation est très utilisée par les constructeurs de calculateurs numériques (UC/Processeurs) pour étudier, pour développer et commencer la mise au point du logiciel d'une machine en cours de développement, donc dont le matériel n'existe pas encore ou n'est pas au point.



L'émulation

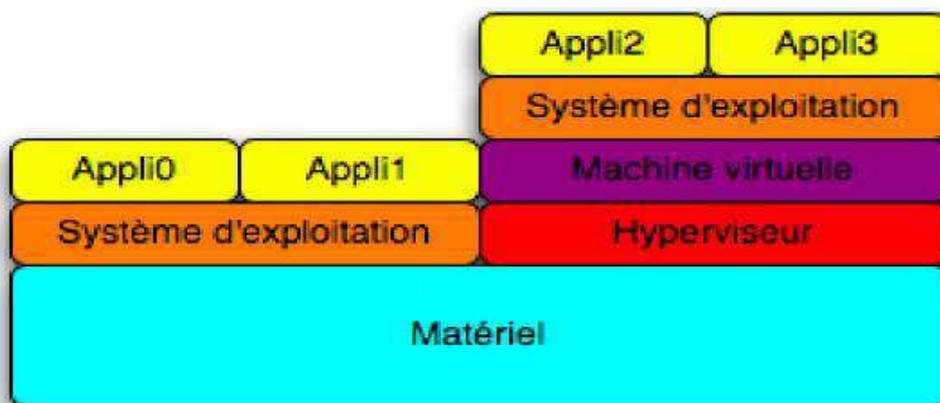
8.2.2 La virtualisation totale

Il s'agit d'une émulation plus performante mais plus restrictive car le processeur de la machine virtuelle doit être du même type que celui de la machine hôte. Ainsi on peut placer l'hyperviseur au niveau du système d'exploitation pour avoir un accès plus rapide au matériel (via les pilotes du système d'exploitation hôte). Les applications du système d'exploitation invité ont une couche de moins à traverser pour atteindre le matériel (Voir la figure ci-après).

Avantages : meilleures performances que l'émulation du fait de l'accès plus rapide au matériel.

Inconvénients : les performances ne sont pas optimales dans l'usage de certains périphériques : type carte accélératrice 3D. De plus, il y a réduction des systèmes d'exploitation virtualisables.

Utilisations : faire tourner des systèmes d'exploitation virtuellement en les laissant accéder aux différents périphériques au travers des pilotes installés sur le système hôte. Des logiciels tels que VMWare Workstation ou VirtualPC de Microsoft l'utilisent.



La virtualisation totale

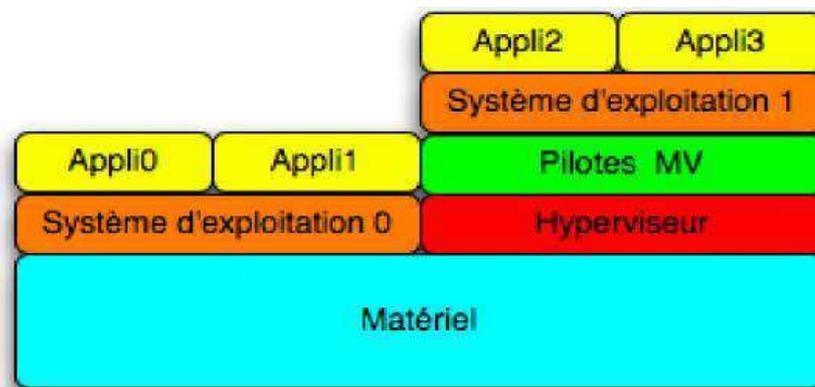
8.2.3 La paravirtualisation

On fait tourner l'hyperviseur sur le matériel et les systèmes d'exploitation invités au dessus de ce dernier (Voir. la figure ci-après).

Avantages : meilleures performances que la virtualisation totale.

Inconvénients : les systèmes d'exploitation invités doivent être modifiés afin de tourner avec l'hyperviseur. La machine virtuelle doit pouvoir tourner sur le processeur (physique) de la machine hôte.

Utilisations : Des logiciels comme Xen ou UML l'utilisent.



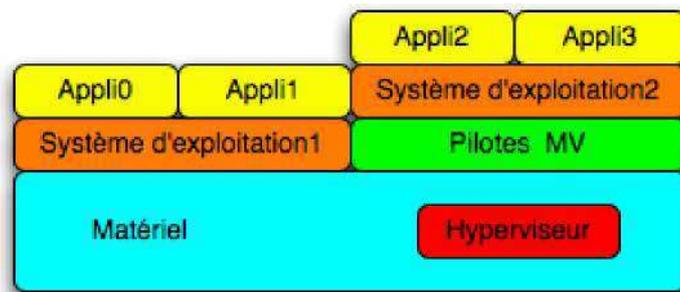
La paravirtualisation

8.2.4 La virtualisation matérielle

Des instructions sont ajoutées au processeur pour qu'il serve d'hyperviseur à l'aide du « HAL » (Hardware Abstraction Layer). Les systèmes d'exploitation invités sont au même niveau que ceux hôtes (Voir la figure ci-après).

Avantages : certains processeurs permettent un accès direct à la mémoire des invités. Les performances sont optimales. Les processeurs ne sont pas émulés et les systèmes d'exploitation inchangés.

Inconvénient : il faut un processeur spécifique qui supporte les nouvelles instructions (HAL). Ainsi, si une machine ne comporte pas ce type de processeur, alors il est nécessaire de s'équiper d'une machine récente qui comportera un processeur compatible.



La virtualisation matérielle

Ce type de virtualisation, qui est la plus sûre et la plus performante, est détaillé dans un chapitre spécifique.

Remarque : IBM, avec ses plateformes de serveurs à base des processeurs p5 et p6 supportent une virtualisation matérielle de « micro-partitionnement » sous UNIX (AIX) et sous LINUX.

8.3 Virtualisation matérielle

8.3.1 Virtualisation matérielle type x86

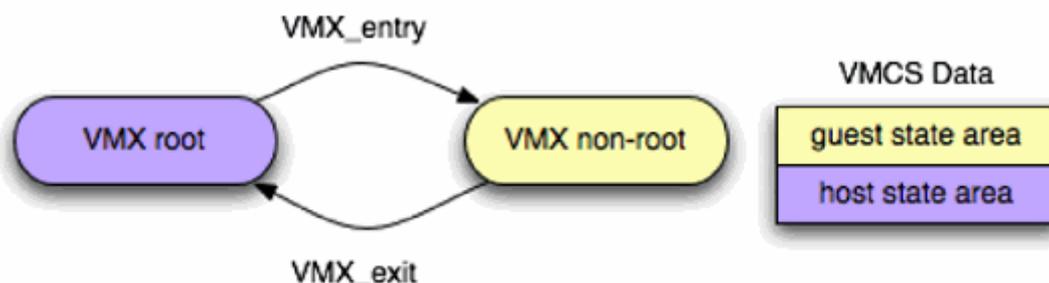
Ces technologies peuvent être aussi bien utilisées par Xen, VMWare ou encore User-mode Linux.

8.3.1.1 Cas d'INTEL

La technologie de virtualisation développée par Intel (Intel-VT - anciennement Vanderpool) supporte aussi bien l'architecture x86 (VT-x) que celle ItaniumR (VT-i). **Un nouveau mode d'exécution apparaît alors : VMX.**

Ce mode comporte un niveau racine (root) pour l'hyperviseur, correspondant à des rings inférieurs à 0, et un niveau normal pour les systèmes d'exploitation invités (machine virtuelle) (non-root), correspondant aux anciens rings de 0 à 3. Le niveau root est entièrement privilégié, alors que l'autre ne possède aucun privilège, même au « ring 0 ».

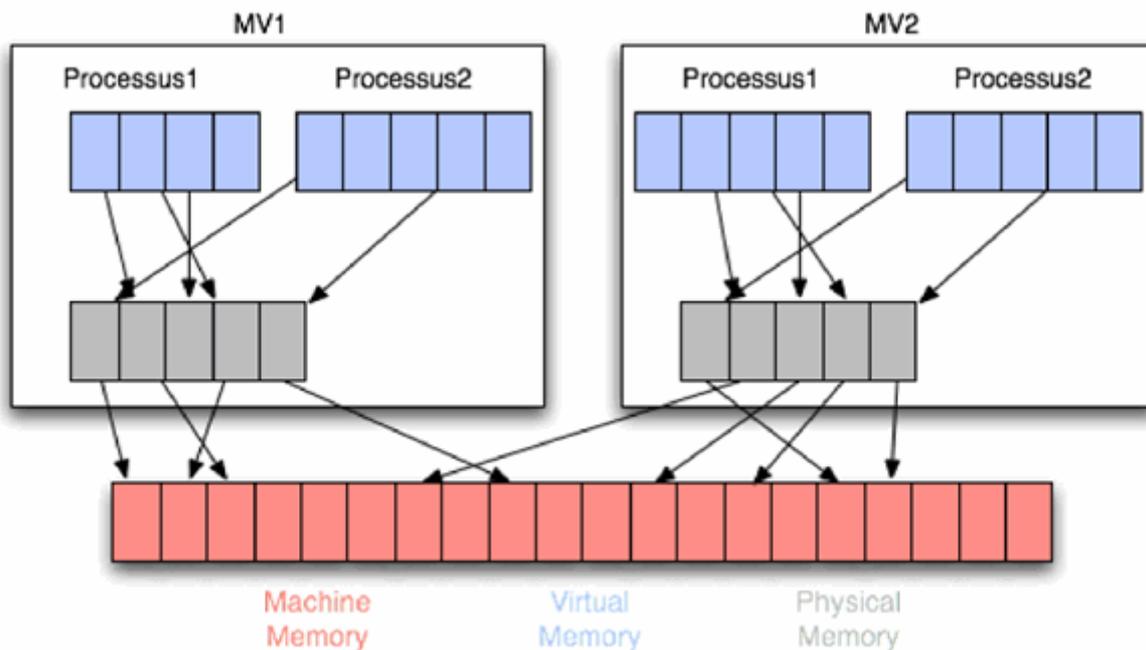
L'instruction VMX entry permet de passer de VMX root à VMX non-root, tandis que l'instruction VMX exit permet l'inverse. Cette technologie permet une certaine flexibilité dans la définition des instructions permettant à la machine virtuelle (système d'exploitation invité) de sortir de l'emprise de la VMM (VM exit). La figure ci-dessous montre le passage d'un niveau à l'autre via les deux instructions expliquées précédemment.



La technologie INTEL-VT

En mode root l'hyperviseur stocke dans une structure de données les informations relatives au matériel associé à un système invité (processeur(s), blocks mémoire(s), entrées/sorties). Sous la technologie VT-x, celle-ci se nomme Virtual Machine Control Structure (VMCS).

Depuis 2007, une 2ème génération de cette technologie est apparue. Elle inclut en plus un support matériel pour la gestion de la mémoire virtuelle : Nested Page Tables (NPT). Deux sortes de table de page existent : Virtual Memory -> Physical Memory et Physical Memory -> Virtual Memory. Cette nouveauté permet des gains significatifs de performance pour les applications faisant un usage intensif de la mémoire ou générant un grand nombre de processus système. La figure ci-après décrit ce nouveau support.



Gestion de la mémoire virtuelle

8.3.1.2 Cas d'AMD

La technologie de virtualisation développée par AMD se nomme AMD-V (anciennement Pacifica).

Il faut savoir que lors du boot un processeur AMD 64 démarre en mode x86 32 afin de maintenir la compatibilité avec les systèmes d'exploitation 32 bits. Le boot-loader d'un système d'exploitation 64 bits exécute une instruction qui bascule le processeur en mode x86 64. De manière similaire, un processeur fonctionnant avec la nouvelle technologie démarre en mode invité (toutes les fonctions de virtualisation sont alors inactives) jusqu'à ce qu'un hyperviseur compatible ne soit démarré.

La structure Virtual Machine Control Block (VMCB) a la même fonction que celle de VT-x.

Une fois la structure VMCB définie, l'hyperviseur bascule le processeur en mode invité, puis passe la main à l'OS invité associé qui démarre (se considérant seul au monde sur le matériel) exécutant son code privilégié en ring 0 et son code applicatif en ring 3. L'hyperviseur s'occupe alors de l'exécution et intercepte les instructions privilégiées. Quand une instruction demande un accès à une ressource définie dans la VMCB le processeur bascule en mode root. L'hyperviseur gère alors la requête.

L'instruction VMCALL est une instruction qui permet aux systèmes d'exploitation invités de négocier des ressources avec l'hyperviseur.

La Translation Look-aside Buffer (TLB) est une table qui contient les références entre adresses réelles et virtuelles des pages mémoire récemment accédées. Le Device Exclusion Vector (DEV), quant à lui, permet à l'hyperviseur de savoir immédiatement si un accès à une page mémoire est légitime ou non.

Sous AMD-V, le support matériel pour la gestion de la mémoire se nomme : Nested Pages ou Nested Page Tables (NPT).

8.3.2 Virtualisation matérielle IBM type p5 / p6 (« Partitionnement »)

8.3.2.1 Introduction

Le but de ce chapitre est de donner une introduction à l'alternative de la virtualisation sur les machines x86 (VMware par exemple) sur les plateformes « haut de gamme » des serveurs UNIX d'IBM :

AIX (UNIX d'IBM) → Plateformes prises en charge : IBM eServer p5 / p6 et pSeries, iSeries

Pour beaucoup, la virtualisation est assimilée au partitionnement. C'est là une réaction naturelle de la part d'une génération d'utilisateurs d'informatique professionnelle qui ont créé et partagé des charges de travail et des unités virtuelles avec les hyperviseurs de machine virtuelle (VM) et les PR/SM des grands systèmes IBM et/ou les stations de travail VMware.

La plateforme IBM Virtualization Engine (IBM Virtualization Engine) étend les possibilités de la virtualisation bien au-delà du domaine du simple partitionnement sur une seule plateforme ou au sein d'une seule technologie. Composant standard des systèmes IBM eServer et Systems Storage (auparavant appelé TotalStorage) basés sur la technologie POWER5, la gamme VE permet aux entreprises de toutes les tailles de déployer virtuellement les applications sur des plateformes IBM et non IBM, et de gérer en temps réel du stockage, des unités en réseau et des systèmes informatiques hétérogènes comme s'il s'agissait d'une seule entité.

Pour IBM, la plateforme du Virtualization Engine correspond à une nouvelle étape dans la virtualisation, la technologie évoluant vers une stratégie informatique globale. En consolidant les ressources et en regroupant dans un pool les ressources de stockage sous-utilisées, les entreprises réduisent les coûts et les tâches de gestion, et deviennent capables de fonctionner et de définir des objectifs « Métier » sans se heurter aux limites de leurs infrastructures informatiques. Le processus de la virtualisation elle-même est conçu en adéquation avec les besoins métier : vous pouvez virtualiser des environnements partitionnés individuellement, par exemple le test d'application et la production, au cours d'un déploiement progressif en plusieurs phases.

Dans le cadre de la stratégie IBM Systems pour l'On Demand Business, celui-ci incite les clients à passer au "tout-virtuel". La plateforme IBM Virtualization Engine permet de virtualiser l'infrastructure informatique d'un client.

"Nous essayons de changer le débat sur la virtualisation", explique Steve Cole, IBM eServer worldwide marketing, Systems and Technology Group. "Tout virtualiser ne se limite pas à créer des partitions, c'est virtualiser un environnement composé de serveurs IBM, HP, Dell, d'unités de bande, de fichiers et de stockage en tout genre. C'est transformer toute la structure en un seul grand pool de ressources logiques. Cette approche vous permet vraiment de définir vos priorités métier : l'informatique ne vous dicte plus ce que vous pouvez faire et ne pas faire".

Nous allons balayer quelques-unes des principales technologies composant la plateforme Virtualization Engine. Vous verrez qu'IBM a non seulement amélioré la virtualisation, mais qu'il a transformé la technologie en une solution métier complète. Le tout est suivi d'une présentation des autres options IBM de virtualisation. Dans "Au service de la ville", vous pourrez lire l'étude de cas de la ville de Richmond (Virginie, États-Unis) qui a fait confiance au contrôleur de volume du réseau de stockage SAN IBM TotalStorage pour optimiser ses ressources de stockage et sécuriser et mieux contrôler son environnement informatique. Dans "Le tout-virtuel", vous découvrirez comment le tandem formé par la technologie IBM POWER5 / POWER6 et AIX tire parti de la virtualisation.

8.3.2.2 Ressources virtuelles

- **Le partitionnement logique dynamique (LPAR dynamique)** permet à certaines ressources système telles que les processeurs, la mémoire et les composants d'E/S d'être ajoutées et supprimées dans les partitions dédiées alors qu'elles sont en cours d'utilisation.

Plateformes prises en charge : IBM eServer p5 et pSeries, iSeries, zSeries et z9.

- **Le micro-partitionnement**, introduit avec les systèmes à microprocesseur POWER5, permet aux administrateurs d'allouer plus efficacement les ressources de serveur en affectant les ressources à des fractions de processeurs et non à des processeurs entiers. Les ressources peuvent facilement être partagées et les modifications d'allocation des ressources se déroulent de façon transparente pour les utilisateurs.

Plateformes prises en charge : IBM eServer p5 et pSeries, iSeries.

- **Les E/S virtuelles** désignent une partition réservée à un usage spécial qui fournit des ressources réseau d'E/S virtuelles aux partitions clientes. Le serveur d'E/S virtuelles possède les ressources partagées avec les clients. Une carte physique affectée à une partition peut être partagée par une ou plusieurs autres partitions, et permet aux administrateurs de réduire le nombre de cartes physiques requises pour les clients individuels. Le serveur d'E/S virtuelles réduit les coûts en évitant de devoir utiliser des cartes réseau, des cartes d'unité de disque et des unités de disque dédiées.

Plateformes prises en charge : IBM eServer p5 et pSeries, iSeries, z9, zSeries et xSeries.

- **Les réseaux LAN virtuels (VLAN)**, permettent aux ordinateurs de se comporter comme s'ils résidaient sur le même segment de réseau LAN, même s'ils sont situés physiquement sur des segments différents. Des réseaux LAN virtuels extrêmement flexibles fournissent des possibilités de virtualisation en vous permettant de prioriser le trafic sur les réseaux partagés. Un réseau LAN virtuel est organisé en fonction de schémas de trafic et non d'emplacements physiques.

Plateformes prises en charge : z9, zSeries, iSeries et pSeries.

- Les chemins de communication **Ethernet virtuels** sont mis en place entre plusieurs systèmes d'exploitation (par exemple AIX 5L et Linux) sans qu'il soit nécessaire d'ajouter des matériels. Des segments Ethernet virtuels peuvent être créés dynamiquement et l'accès à un segment LAN virtuel peut être limité pour des motifs de sécurité ou de séparation du trafic.

Plateformes prises en charge : IBM eServer p5 et pSeries, iSeries.

- **Les HyperSockets fournissent** une connexion réseau TCP/IP virtuelle intégrée entre plusieurs combinaisons de LPAR et serveurs virtuels exécutant Linux on zSeries, Linux on S/390, z/OS V1R2, z/VM V4R2 et VSE/ESA V2R7 ou une version supérieure sur le zSeries.

Plateformes prise en charge : IBM System z9.

- **IVM (Integrated Virtualisation Manager)**, qui s'adresse aux petites et moyennes entreprises, fournit une interface de gestion de système s'exécutant dans un navigateur qui permet de gérer un seul système sans utiliser de console HMC (Hardware Management Console). IVM vous permet de créer et de gérer des partitions logiques, de configurer des réseaux Ethernet virtuels, de gérer du stockage sur le serveur d'E/S virtuel, et ainsi de suite.

Plateformes prise en charge : IBM eServer p5, pSeries et OpenPower.

- **SAN Volume Controller (SVC)** permet d'apporter des modifications au stockage physique moyennant une interruption minimale, voire nulle, des applications. Il combine la capacité de plusieurs systèmes de stockage sur disque dans un seul pool de stockage, qui peut être géré depuis un point central. Il vous permet en outre d'utiliser des services de copie avancés sur les systèmes de stockage de différents fabricants.

Plateformes prise en charge : systèmes de stockage IBM et non IBM.

- **SAN File System** a pour but de simplifier la gestion des fichiers dans les réseaux SAN et de réduire leurs coûts de gestion, en fournissant un stockage et une gestion de données centralisés et basés sur des règles à des serveurs, des systèmes de stockage et des systèmes d'exploitation hétérogènes.

Plateformes prise en charge : systèmes de stockage IBM et non IBM.

8.3.2.3 Gestion virtuelle

- **IBM Director** fournit une interface utilisateur Java unique qui permet de surveiller l'utilisation et les performances de ressources système hétérogènes (notamment les configurations Windows, AIX 5L v5.3, Linux on Power, Linux on zSeries et VMware, ainsi que les unités de stockage) et d'effectuer le suivi de l'inventaire et des événements. Les administrateurs peuvent gérer les tâches, prendre des mesures correctives de base et émettre des commandes distribuées.
- **Enterprise Workload Manager (EWLM)** fournit une vue complète et consolidée des transactions et de leurs performances dans une infrastructure hétérogène, définie en fonction des objectifs « Métier ». Certaines ressources système telles que les partitions POWER5 peuvent aussi être gérées avec EWLM.
- L'outil **RDS (Resource Dependency Services)** détecte automatiquement les ressources IBM et non IBM, physiques et logiques, et mappe la topologie virtualisée, les associations et les dépendances de tout environnement défini par le client.

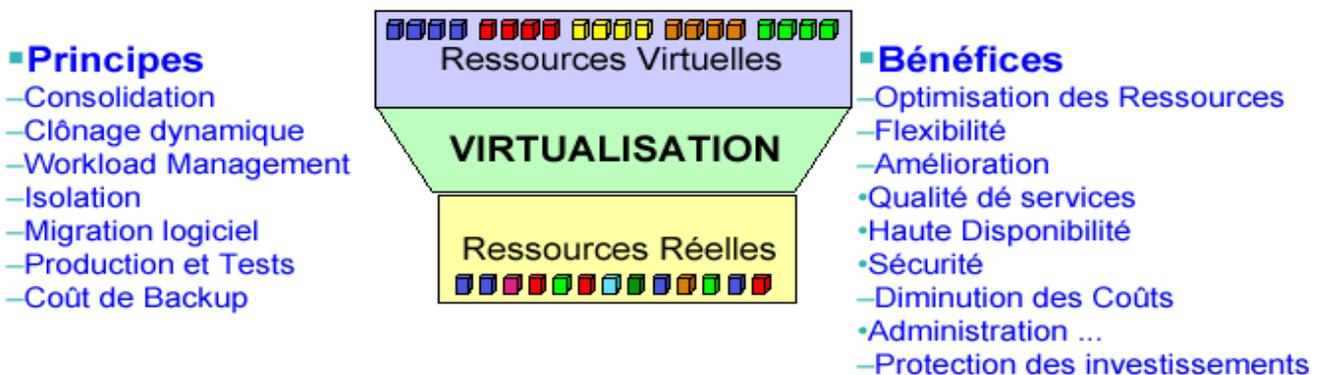
8.3.2.4 Accès virtuel

- **Virtualization Engine Console** comprend un ensemble complet de tâches de gestion et fournit une seule vue graphique et consolidée de la santé des ressources et des fonctions de surveillance dans plusieurs systèmes d'exploitation et plusieurs domaines/disciplines de gestion (par exemple IBM Director, CSM, MC, Tivoli, EWLM).

- **IBM TotalStorage Productivity Center** est l'équivalent de la console VE pour le stockage. Il centralise la gestion des unités SAN, des disques, de la réplication, de la planification de la capacité et de l'utilisation du stockage.

8.3.3 Serveurs UNIX - System p d'IBM / Hypervision, Micro-Partitionnement, Virtual I/O Server

8.3.3.1 Introduction à l'Hypervision « matérielle »



• **En Résumé la Virtualisation**

Aide à la réduction des coûts Matériel et Logiciel

Optimisation des ressources + réduction des licences logiciels

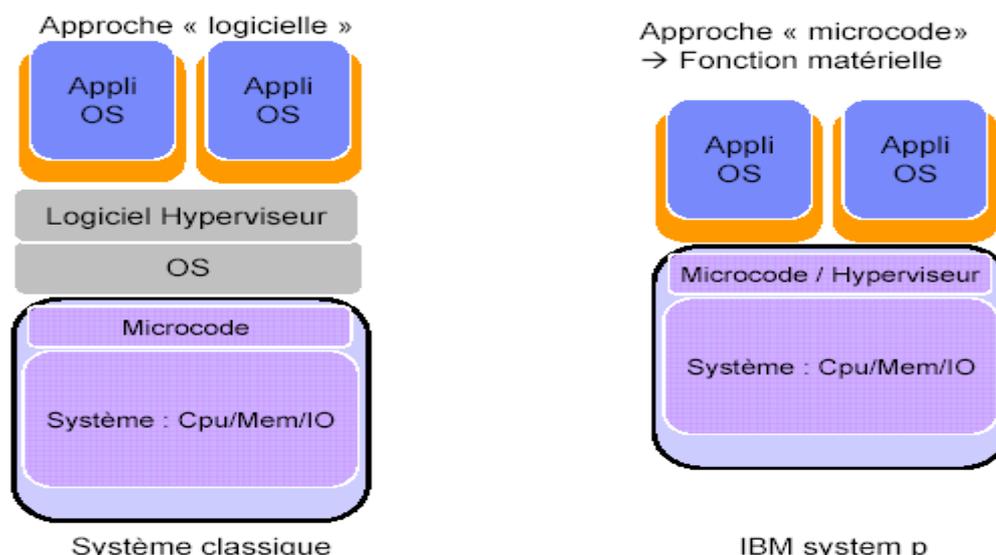
Amélioration de la flexibilité et réponses rapides à la demande

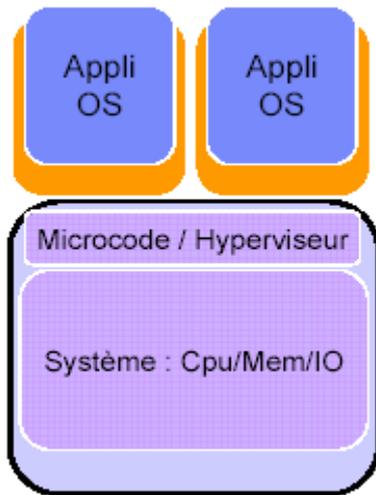
Ressources virtuelles s'ajustent plus facilement (On Demand)

Réduction des coûts d'administration

Moins de serveurs physiques, industrialisation des tâches

8.3.3.2 Virtualisation (Rappel) : Différentes implémentations possibles

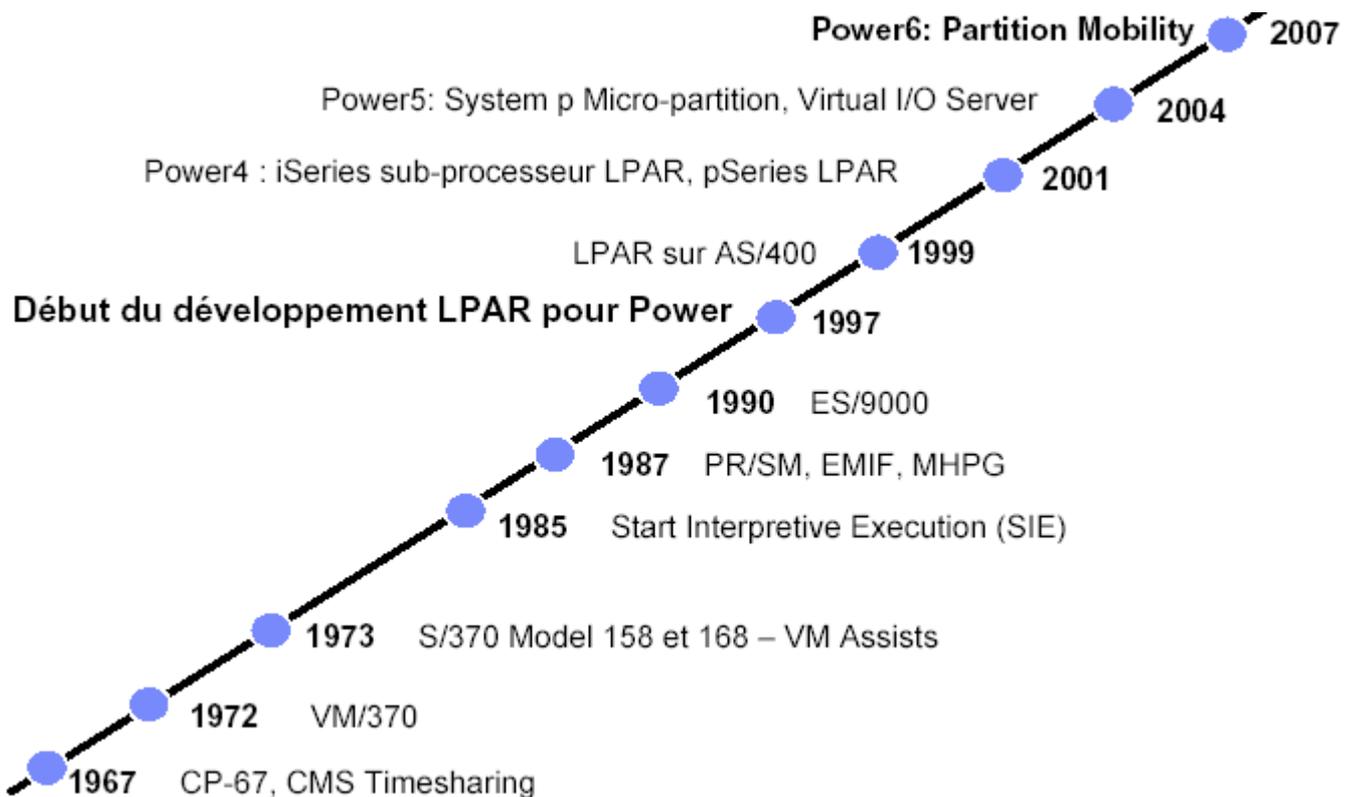




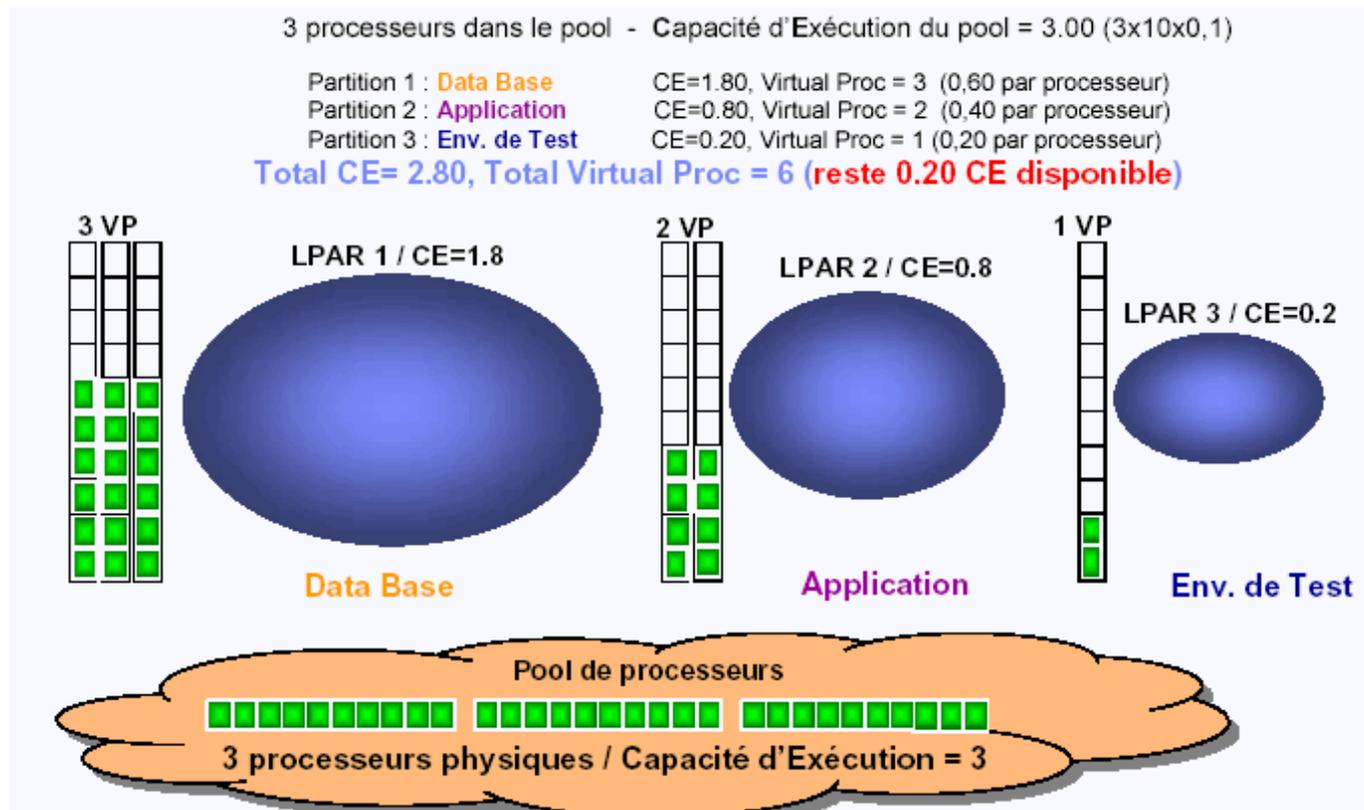
IBM system z9 (Mainframe) : 1989 PR/SM
 IBM system p5 (Unix/Linux): 2001 (hyperviseur)
 2004 (virtualisation)
 IBM power6 (Unix/Linux): 2007 (mobilité)

→ Bien noter la présence d'un Hyperviseur « Embedded », c'est-à-dire en microcode dans le hardware

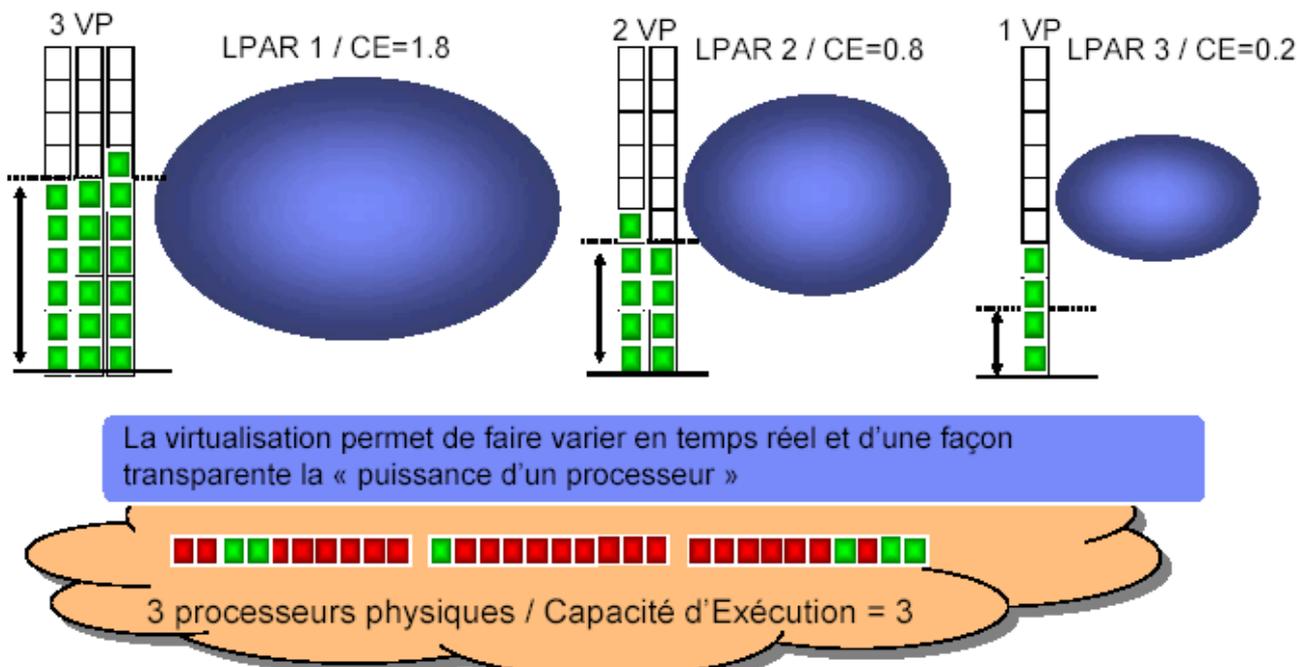
8.3.3.3 Virtualisation / IBM : 40 ans de développement IBM



8.3.3.4 Micro-partitioning : CE Capacité d’Exécution



8.3.3.5 Micro-partitions: Ajustement des puissances



8.3.3.6 Illustration de la « Mobilité » des partitions - Live Partition Mobility / p6

Live Partition Mobility (POWER6)

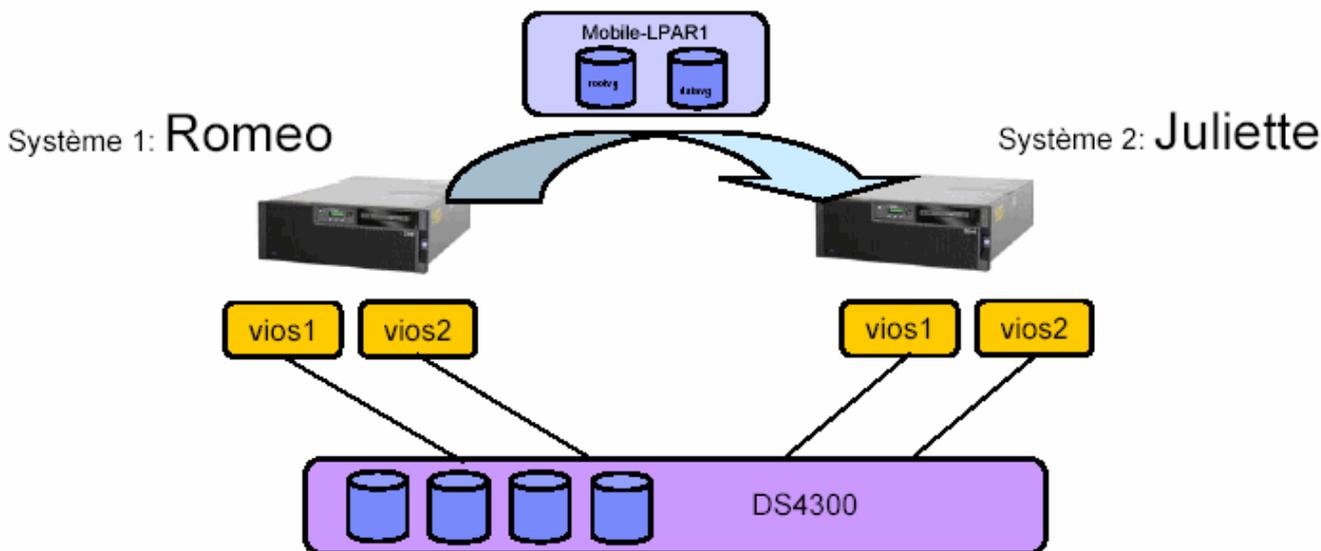
Migration de partition active

La migration de partition active permet de déplacer une partition d'une machine physique à une autre sans interrompre le fonctionnement du système d'exploitation et sans perturber les applications

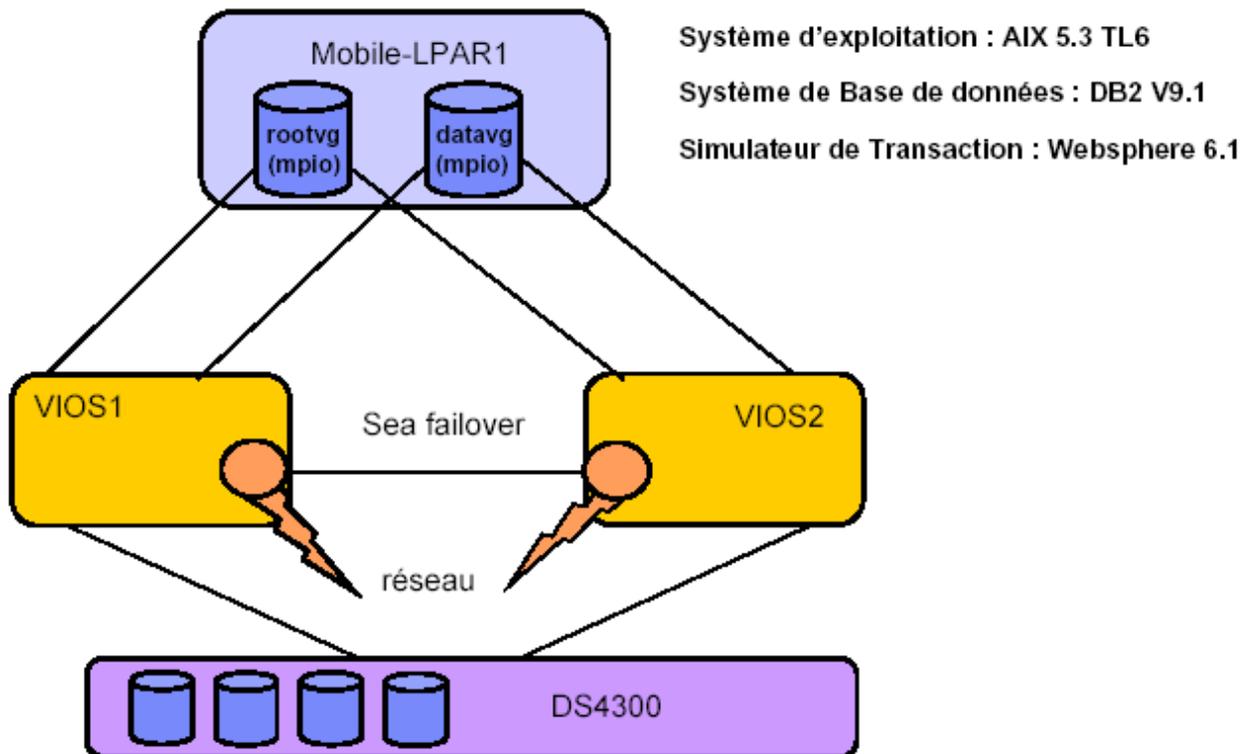
The diagram illustrates the Live Partition Mobility architecture. At the top is the **HMC** (Hardware Management Console). Below it are two server racks, each containing a **LPAR** (Logical Partition) and a **VIOS** (Virtual I/O Server). These are connected to a central **LAN** and **SAN** (Storage Area Network). A **Boot Paging Application** is connected to the SAN. A small image of a POWER6 processor is shown in the top right corner.

La "Partition Mobility" est supportée sur POWER6™ en AIX 5.3, AIX 6.1 et Linux

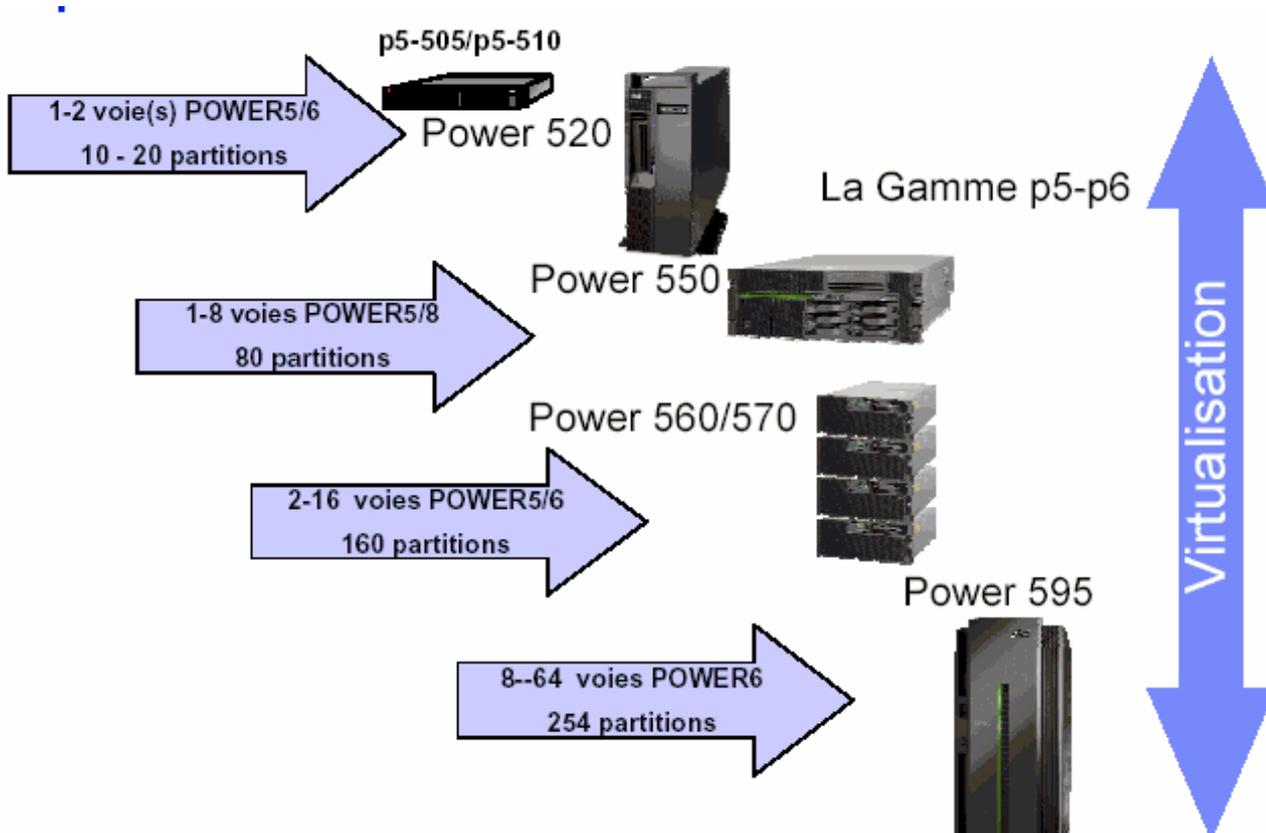
Illustration (dém.) :



Ci-dessous, en exemple, détails de la configuration de la partition :



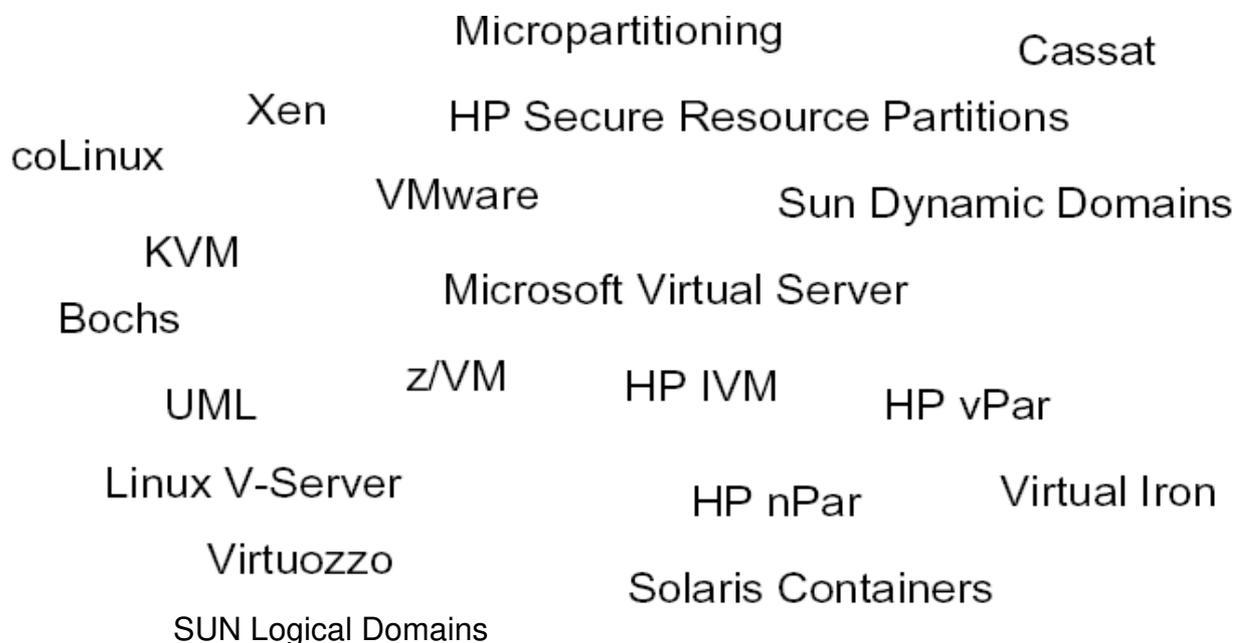
8.3.3.7 En exemple : Gamme Serveurs IBM Power5 et Power6



8.4 Les acteurs du marché

8.4.1 Présentation « très évolutive » des acteurs X86 (marché « en consolidation ») !

Hyperviseur	Editeur	Haute disponibilité	Base (type)
XenServer	Citrix	XenMotion	Xen (paravirtualisation)
Virtual Server 2005 Hyper-V (lancé à l'été 2008 mais en bêta dans Windows Server 2008)	Microsoft	Couplage avec Microsoft Cluster	Propriétaire (virtualisation complète)
Parallels Virtuozzo Containers	Parallels (ex-SWsoft)	Support du clustering intégré depuis la dernière version	Isolation
Intégré à Red Hat Enterprise Linux	Red Hat	Couplage avec Red Hat Cluster Suite	Xen (paravirtualisation)
Intégré à Suse Linux Enterprise Server	Novell	Intégrée à ZENworks Orchestrator	Xen (paravirtualisation)
VMware infrastructure	VMware	VMotion	Propriétaire (virtualisation complète)



8.4.2 Marché du contrôle des infrastructures



- Les acteurs s'affrontent pour le contrôle du marché de l'Infrastructure
- VMware / Groupe EMC²
- XenSource / Groupe CITRIX

8.5 Architectures de virtualisation de serveurs

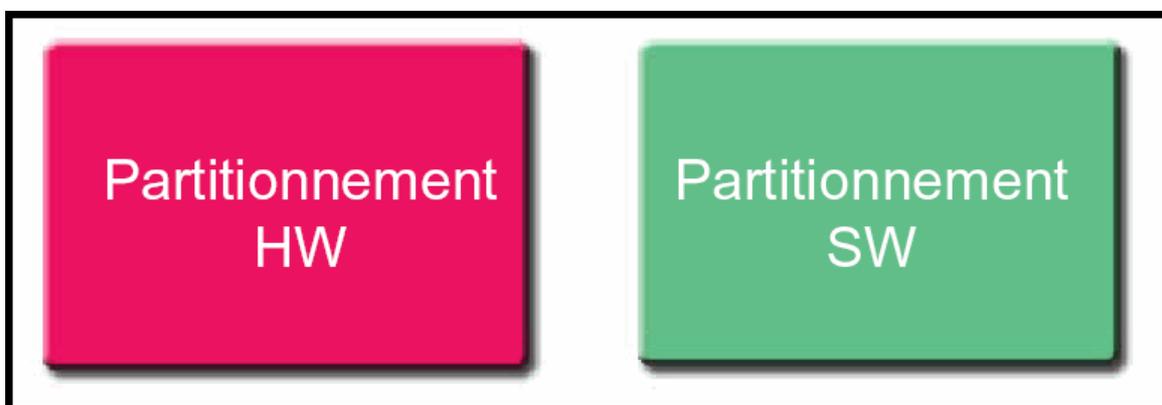
8.5.1 Introduction

Le cas standard d'un « Serveur », c'est un ordinateur (UC) géré par un système d'exploitation (OS).

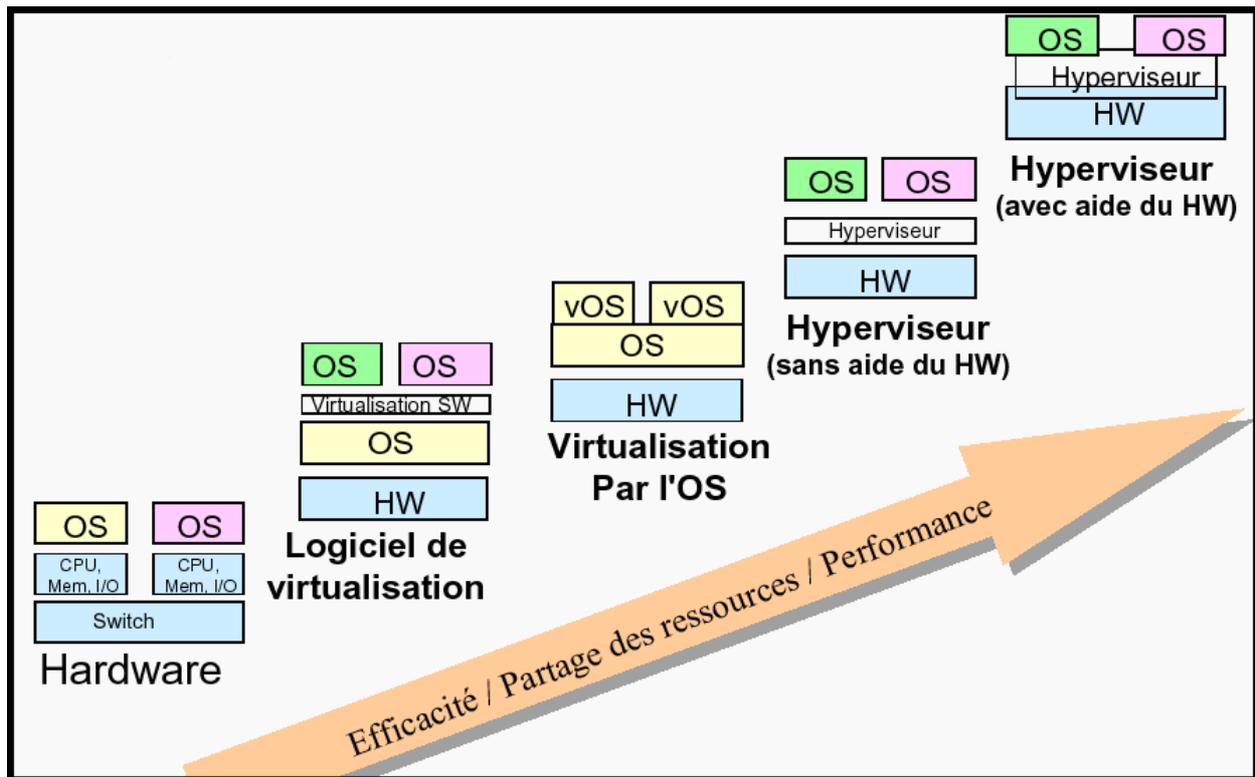
La virtualisation, elle, fait croire à un système d'exploitation qu'il est seul à gérer une machine unique, bien que ce ne soit pas le cas.

Pour quoi faire ?

8.5.2 Terminologies / Technologies



8.5.3 Architectures de virtualisation de serveurs

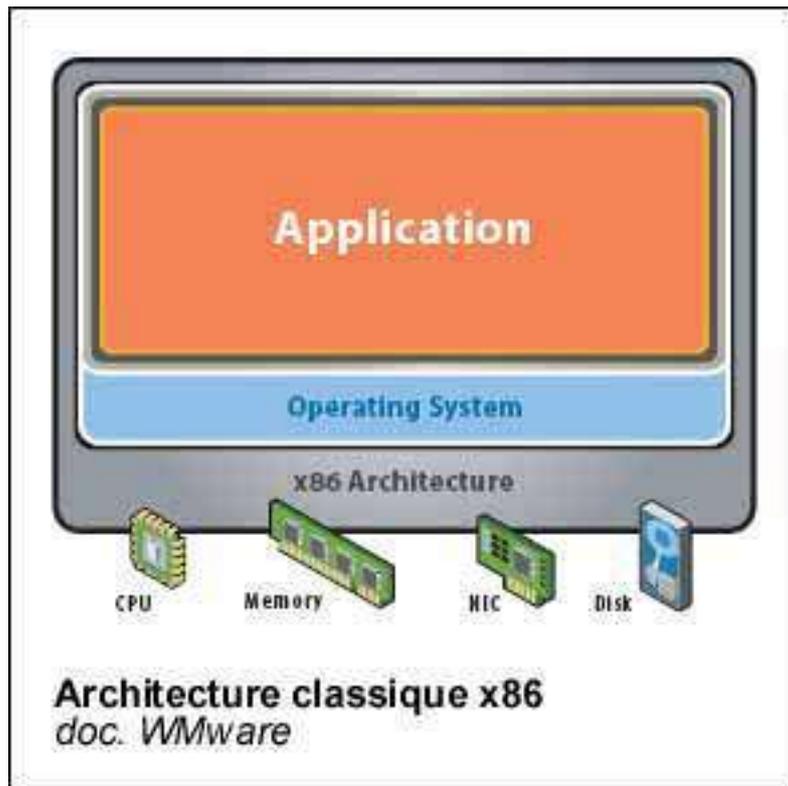


8.5.4 Principe de la Virtualisation

La raison d'être d'un ordinateur est d'exécuter des tâches, développées sous forme d'applications.

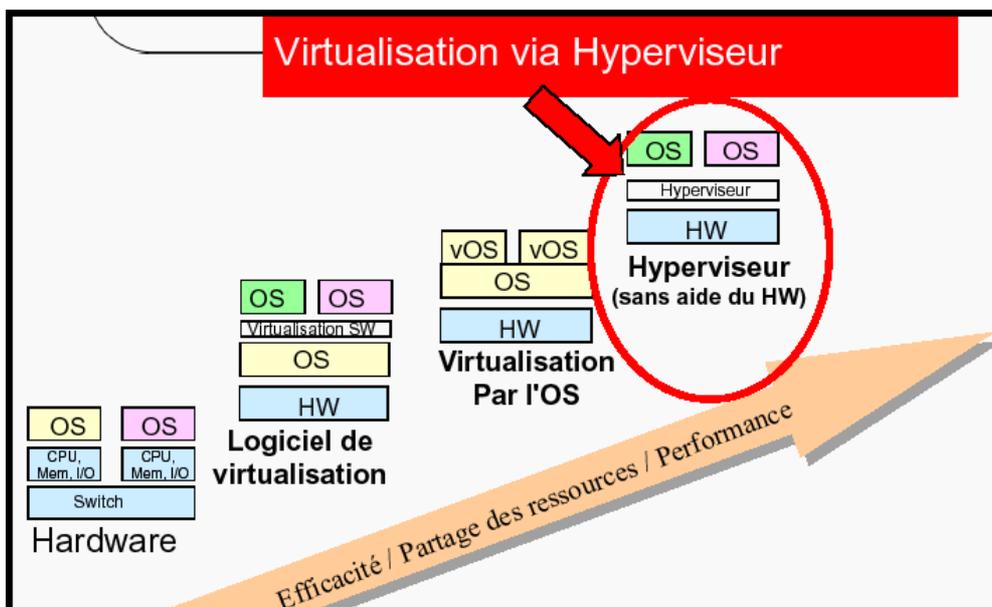
Dans le cas classique, une machine est gérée par un système d'exploitation (le superviseur), qui est chargé de 2 responsabilités :

- organiser le passage des applications (lancement, mise en attente, swap, arrêt... principalement),
- fournir les ressources à ces applications (mémoire, accès aux données stockées, accès au réseau...).

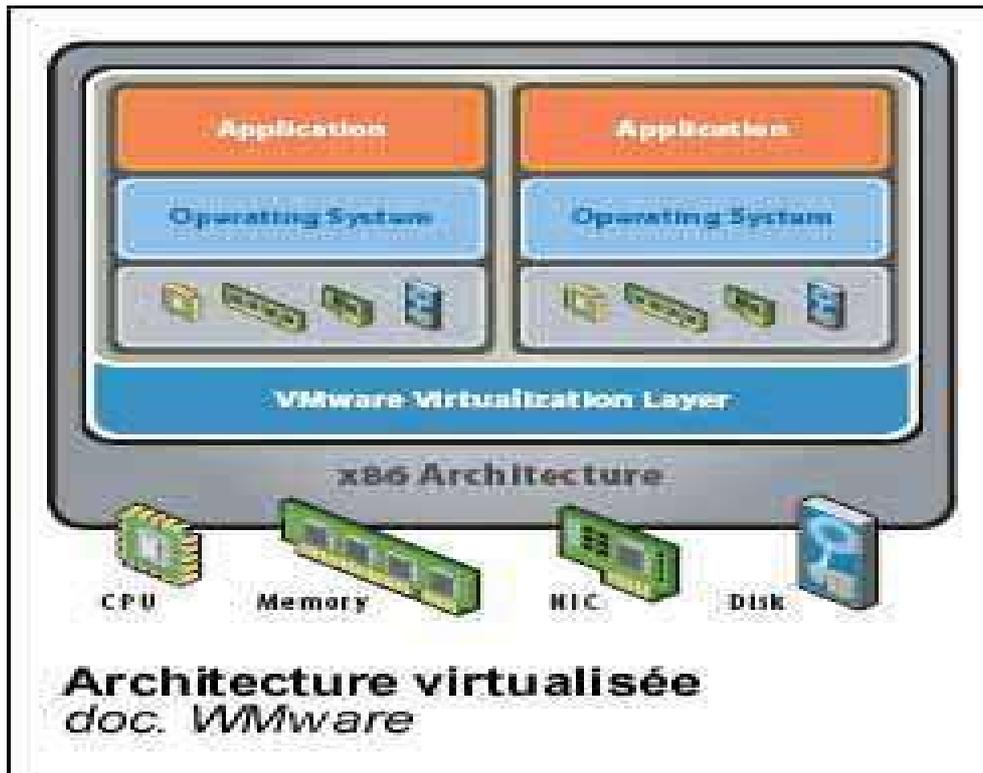


La virtualisation ajoute une couche de supervision supplémentaire (l'hyperviseur), qui :

- d'une part, assure la gestion de toutes les ressources physiques,
- d'autre part, propose plusieurs environnements étanches où un OS autonome pourra tourner dans chacun d'entre eux.



Virtualisation supportée par HYPERVISEUR (Ex VMware ESX et Xen)



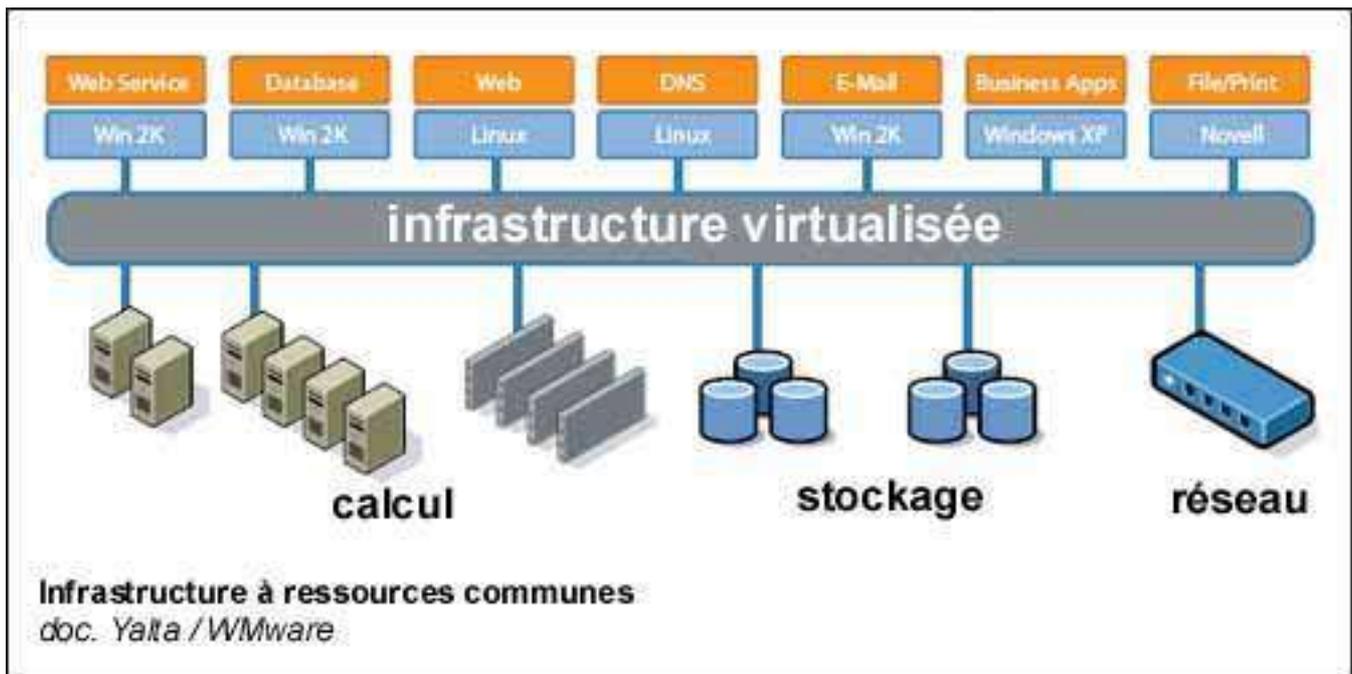
La couche de virtualisation fournira à chaque OS les mêmes types de ressources que si celui-ci était directement aux commandes de la machine.

8.5.5 Quelles sont les intérêts de la virtualisation ?

Les deux principaux intérêts de la virtualisation sont :

- optimiser l'utilisation des ressources, spécialement celle du calculateur (un serveur classique n'utilise la puissance disponible qu'une faible partie du temps),
- alléger la charge de l'administration système (en regroupant dans une même machine physique plusieurs machines existantes).

Historiquement, la virtualisation est apparue dans les années 70 (IBM VM 370). Elle permettait de faire tourner sur un mainframe plusieurs OS incompatibles (DOS/VSE et OS/MVS...) et créait déjà des machines individuelles (CMS), préfigurant le fonctionnement des premiers PC.



Depuis, et en dehors des 2 principales raisons (optimisation des ressources et simplification de l'administration), la virtualisation permet donc de faire cohabiter des environnements habituellement incompatibles :

- Windows et Linux, par exemple,
- des environnements de production et des environnements de tests en situation,
- des nouvelles versions d'environnement en cours de basculement,
- des environnements complets issus de l'intégration de filiales ou de sociétés nouvellement achetées (consolidation).

Dans ces cas cités, l'environnement pourra bénéficier de la puissance de la machine hôte et/ou de conditions de connexion semblables.

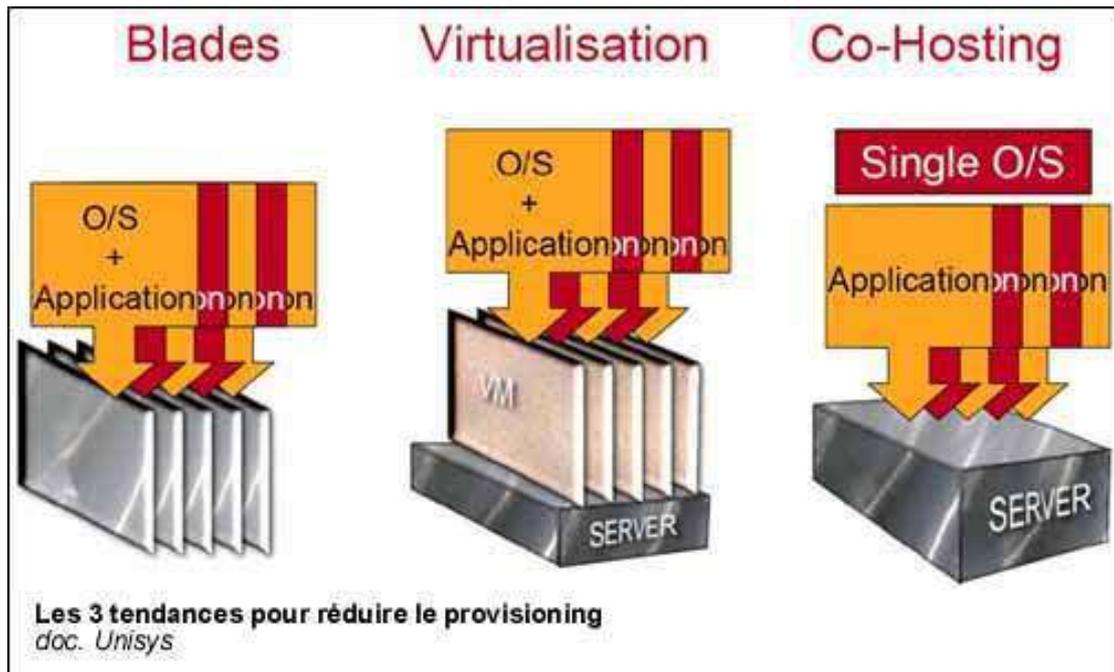
8.5.6 Alléger l'administration système

Si l'on ajoute une couche logicielle supplémentaire, potentiellement consommatrice de ressources et cause possible d'incidents, c'est paradoxalement souvent pour alléger la charge de l'administration système.

Les mainframes, système unique de l'entreprise, autrefois maîtres du terrain, ont cédé la place. Ils ont été remplacés par des serveurs de plus en plus nombreux (beaucoup de type x86). Les raisons, leur simplicité relative d'administration, le faible coût de licence, la scalabilité...

Administrer et maintenir ces nombreuses machines (parfois plusieurs milliers) est souvent devenu un vrai problème. Ils existent plusieurs solutions pour réduire cette charge :

- les « bladecenters », qui rassemblent à chaque niveau de rack les interfaces utiles (alimentation, réseau, supervision...);
- le mainframe,
- la virtualisation.



➔ Virtualiser, c'est donc rassembler dans une même machine, d'une part, un grand nombre de systèmes indépendants, d'autre part, un seul jeu de licences (« dans certains cas ») et d'interfaces.

8.5.7 Optimiser les ressources

On ne remplacera jamais 100 machines utilisées à 10% de leur temps par 10 machines utilisées à 100 %, mais peut-être 40 ou 50 machines.

Et diminuer par 2 ou 3 la puissance machine, c'est gagner sur tous les tableaux :

- moins de m² ultra-sécurisés,
- moins d'énergie électrique ultra-fiabilisée,
- moins de m³ ultra-climatisés...

...et aujourd'hui, on sait que l'on dépense plus pour les m² de la salle machine que pour l'achat des serveurs. Une bonne raison d'optimiser la puissance machine.

La virtualisation est promise comme une étape sur l'alignement de l'IT sur la stratégie business.

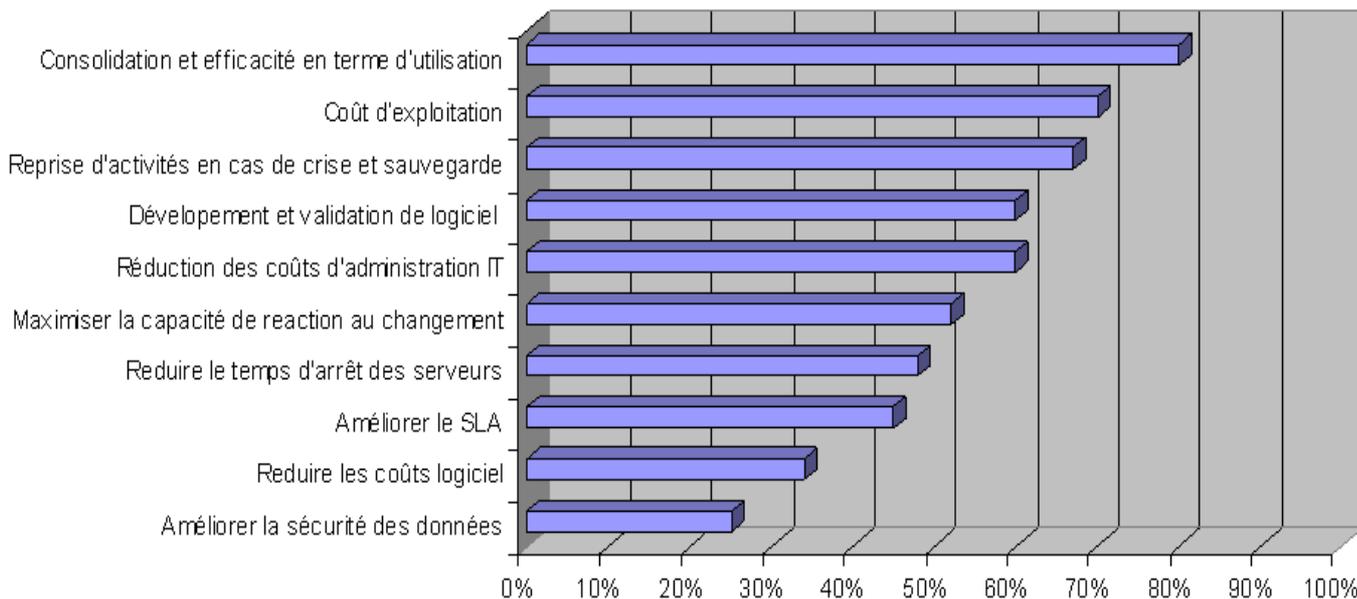
infrastructure	basique	standardisée	rationalisée	virtualisée	orientée services	orientée business
coûts	non évalués	coûts fixes	optimisés	partagés	fonction de l'utilisation	fonction du business
interface métier		SLA	SLA	SLA flexibles	SLA de bout en bout	SLA business
organisation		centralisée	consolidée	partagée	orientée service	orientée business
management	suivant les besoins	cycle de vie	maturité	dynamique	bout-en-bout	business
Modèle de maturité de l'infrastructure <i>doc. Yalta, d'après Unisys</i>						

→ **SLA** (« Service Level Agreement ») : Accords sur la qualité de service (Relations contractuelles entre la DSI, ses clients et ses fournisseurs).

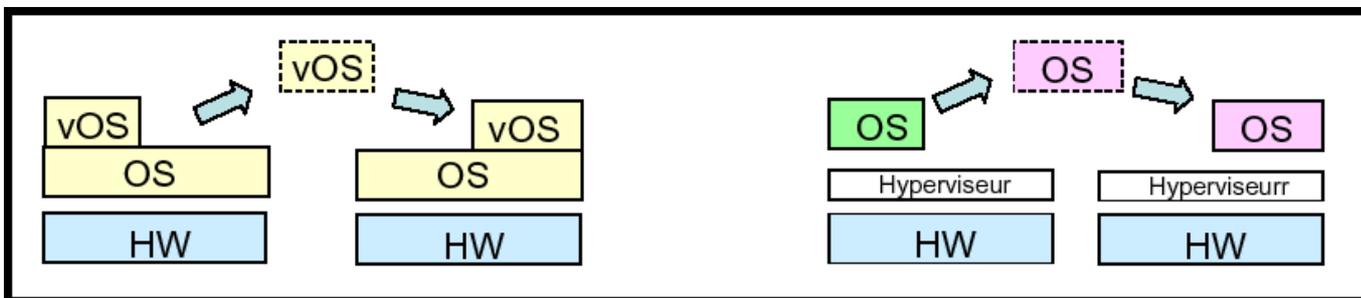
8.5.8 Résumé des avantages d'une virtualisation des serveurs :

- un taux d'utilisation du matériel augmenté de 50 à 70 %
- un délai de provisionnement des nouvelles applications de quelques minutes
- des coûts matériels et logiciels réduits de 40 %
- des coûts d'exploitation réduits de 50 à 70 %
- enfin, un ratio nombre de serveurs gérés / administrateur augmenté de 50 pour 1 à plus de 150 pour 1

Résumé : Pourquoi VIRTUALISER :



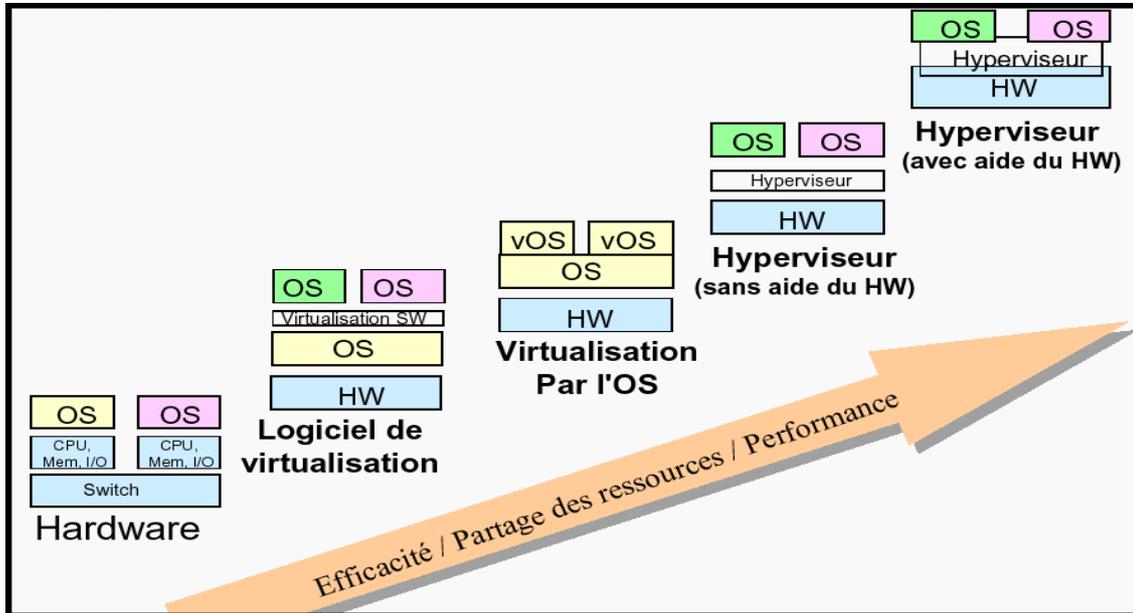
8.6 Mobilité des Machines Virtuelles



→ Voir également le chapitre : Virtualisation des Serveurs / IBM : Illustration de la « Mobilité » des partitions - Live Partition Mobility / p6

8.7 Hyperviseur sans aide du Hardware et HYPERVISEUR avec l'aide du Hardware « Micro Partitionnement »

8.7.1 Architectures de virtualisation de Serveurs :



8.7.2 Virtualisation supportée par HYPERVISEUR avec l'aide du Hardware « Micro Partitionnement »

8.7.3 Granularité des micro-partitions allant jusqu'à 1%

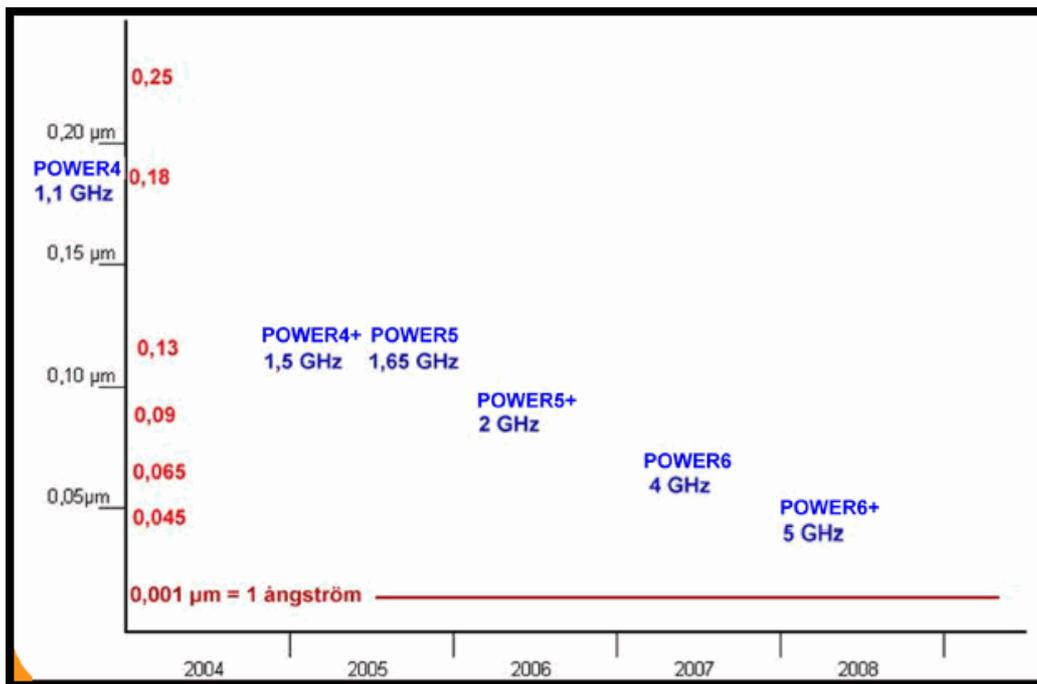
The diagram shows a physical CPU being partitioned by a hypervisor into virtual CPUs, which are then further divided into micro-partitions. Each micro-partition contains multiple virtual processors (labeled 'AIX L').

- Le micro-partitionnement se base sur les processeurs virtuels:
 - Plus petite partition possible: 10% d'un CPU physique
 - Granularité: 1%
- Une Micro-partition peut contenir une multitude de processeurs virtuels

Permet d'avoir des partitions <1 CPU physique.

→ Voir également le chapitre : Virtualisation des Serveurs / IBM : Serveurs UNIX - System p d'IBM / Hypervision, Micro-Partitionnement, Virtual I/O Server

8.7.4 La technologie des processeurs p5 & p6 d'IBM



Remarque :

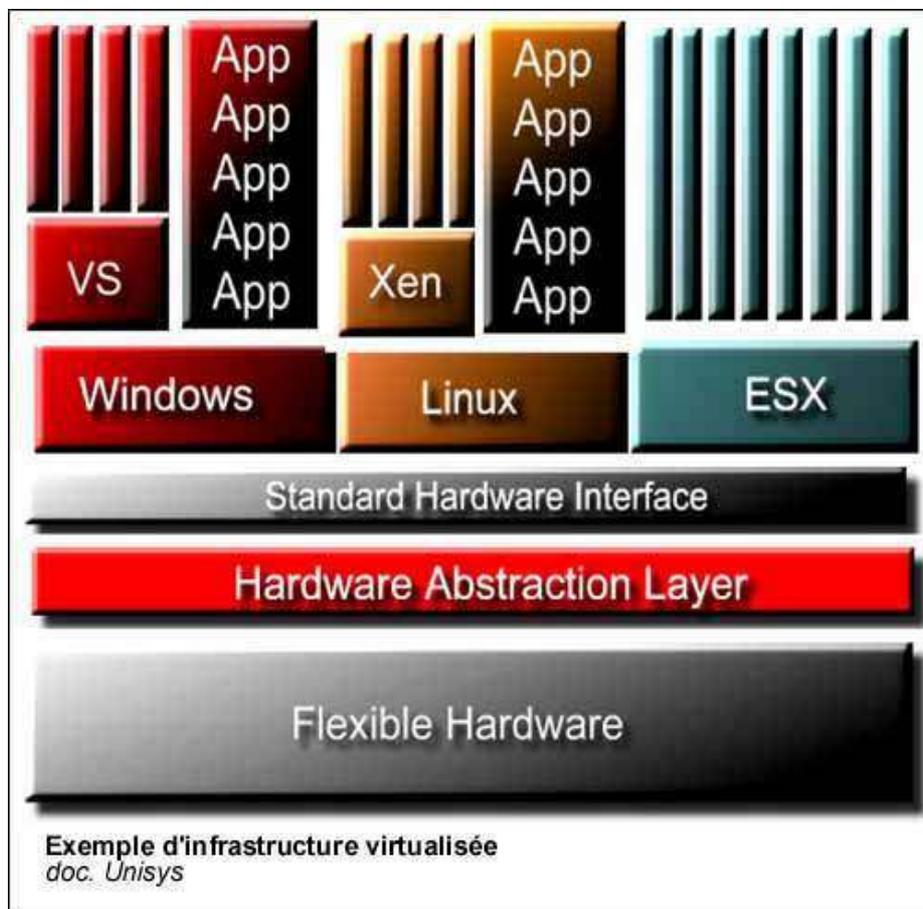
Serveurs p5 & p6 / IBM ↔ Serveurs ESCALA / BULL qui est un OEM d'IBM

9 Mainframes & Virtualisation

9.1 Mainframes, le retour et la Virtualisation / Généralités

Pour pouvoir faire tourner plusieurs systèmes d'exploitation avec leurs applications dans une même machine, il faut des machines puissantes. Le mainframe connaît donc un regain d'intérêt. Les constructeurs comme Bull, HP, IBM ou Unisys les proposent toujours à leur catalogue.

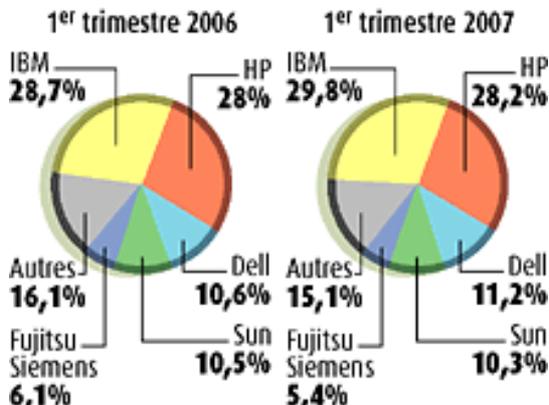
Mais attention, le mainframe s'est en grande partie débarrassé de ses vieux défauts (coûts élevés de l'hardware, coûts de licence, lourdeur de l'administration). Le mainframe est désormais construit à partir d'éléments courants du commerce, comme des processeurs x86 alignés en rang d'oignons et des systèmes d'exploitation open source.



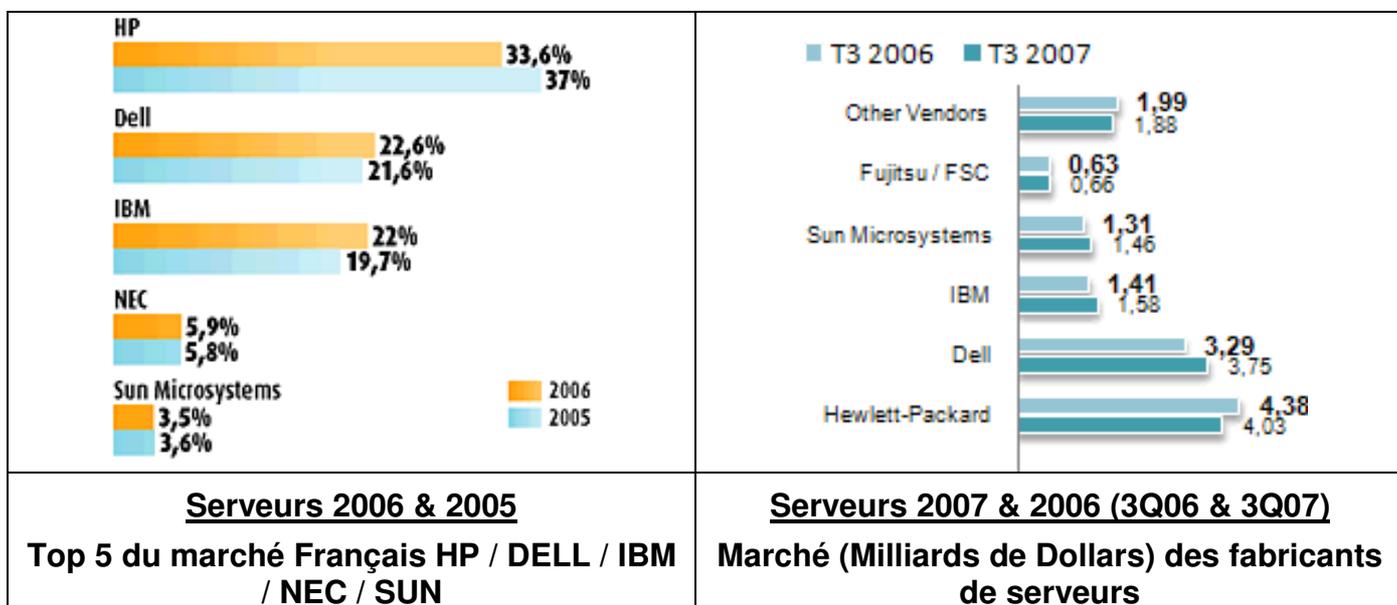
Comme on le voit sur le document d'Unisys, le mainframe est capable de faire tourner en même temps si nécessaire plusieurs environnements de virtualisation concurrents (Microsoft Virtual Server, Linux Xen, VMWare ESX).

9.2 Le marché des Serveurs et des « Blades Centers » / Serveurs Lames

9.2.1 Marché Français des Serveurs



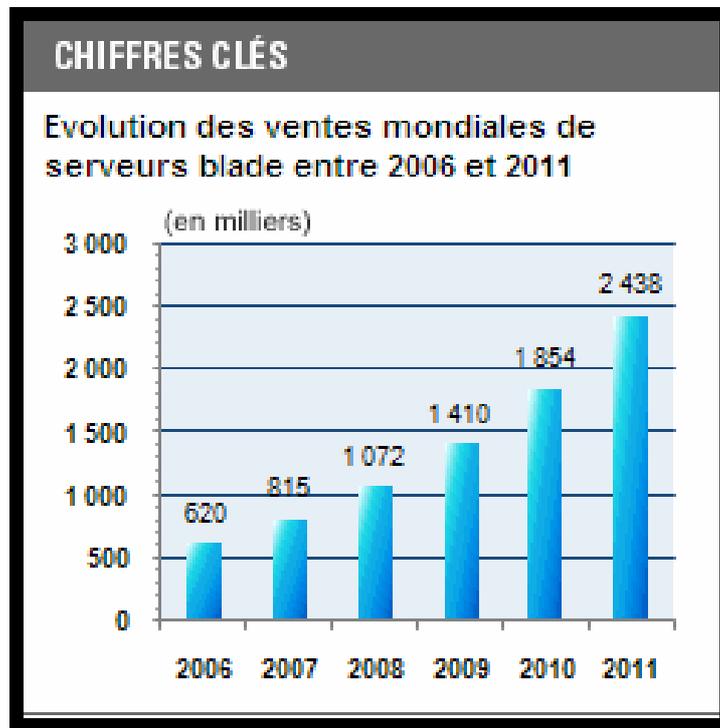
Marché Français en volume des 5 premiers fabricants de serveurs 1Q06 & 1Q07



➔ En France / 2006 : HP est le leader en volumes avec 95.700 unités livrées et 33,6% de parts de marché.

9.2.2 Marché Mondial des Serveurs Lames

Les ventes de serveurs lames auront quadruplé d'ici 2011



(Origine : Cabinet d'études iSupply)

Les ventes mondiales de serveurs lames devraient atteindre 2,4 millions d'unités en 2011.

→ Environ 31,5% de taux de croissance annuel moyen durant les quatre années à venir.

En 2011, les serveurs lames devraient représenter 21,6% de l'ensemble des serveurs vendus dans le monde.

Rappel : La technologie « lame » est apparue vers 2001 au niveau des Serveurs, mais a été développée au niveau des baies de stockage (RAID) dans la deuxième partie des années 90.

10 La virtualisation du stockage (stockage virtualisé et le Thin provisioning, sauvegarde virtualisée / VTL et la technologie de déduplication, archivage et de son réseau spécialisé SAN)

10.1 Introduction / Historique

Dès 1980, la start-up française COPERNIQUE développe une baie de stockage, le DIRAM 32, gérant jusqu'à 16 disques en double accès (en interface SMD de l'époque) et pouvant être connecté jusqu'à 4 ordinateurs d'un même constructeur.

Ce produit « virtualisait » :

- des couples de disques en Mirroring pour assurer la continuité d'activité en cas de panne d'un disque (concept que l'on désigne actuellement par technologie RAID-1,
- un disque « virtuel » simulé en mémoire électronique dans une partie du cache d'E/S
- un processeur programmable de traitement sur données des disques pour des recherches associatives sur critères via un canal standard d'E/S disque. Il faut noter que ce concept novateur, qui a permis, entre autres gros projets, de développer l'Annuaire Électronique du téléphone de France Télécom (renseignement téléphonique sur Minitel via le 11 / 3611 de FT), n'a jamais été ré-implémenté dans les baies de stockage d'IBM, de HDS, ni d'EMC².

La virtualisation s'est invitée dans le stockage en tout début de période (2000) dès le déploiement d'environnements SAN / NAS avec la création d'unités logiques virtuelles LUN (« *Logical Unit Number* ») qui ont permis un premier niveau d'abstraction entre les volumes de stockage vus par les applications et le stockage physique mono ou multi-disques. Du côté technologie, on peut donc estimer qu'il y a déjà une bonne maturité.

L'éditeur américain DataCore a été, dès 1999, l'un des premiers à proposer une offre « mature » de Virtualisation « In-Band » du Stockage au travers du SAN avec sa suite « SanSymphony ». Son Appliance « In-Band » était basée sur des serveurs « multiprocesseurs » à base d'X86 sous NT. Depuis, vers 2003, cette solution a été reprise « à l'identique » dans son architecture et ses fonctionnalités par IBM avec ses « SVC - SAN Volume Controller » à base de 2 serveurs PowerPC montés en cluster (nœud SVC).

L'implémentation technique de l'offre de Virtualisation de HDS et d'IBM est du type « In-Band » : La virtualisation est directement intégrée aux baies de stockage chez HDS (USP – Universal Storage Platform / gamme TagmaStore). Pour IBM, la virtualisation est gérée par le SVC (nœud de cluster) qui est une Appliance spécialisée indépendante des baies de stockage, et qui gèrent la virtualisation pour les baies de stockage qui y sont connectés. Les SVC sont proposés par IBM avec les baies DS4000, DS6000 et même les DS8000. Pour ces derniers, IBM introduit progressivement des couches de virtualisation directement dans les baies DS8000, mais l'offre est très en retard par rapport à toutes les fonctionnalités des SVC. D'autre part, l'introduction de la virtualisation au niveau des DS6000 pose probablement un problème de stratégie par rapport à ses SVC. L'implémentation technique de l'offre de Virtualisation d'EMC² pour ses baies DMX est

du type « Out of Band », la virtualisation est traitée au niveau des directeurs du SAN gérant les baies de stockage. C'est l'offre de virtualisation la plus en retard par rapport à HDS et IBM.

A ce jour, les acteurs principaux du stockage et de sa virtualisation, en classe « Enterprise / High-End » sont EMC², HDS et IBM. Pour HP et SUN, ce sont des OEM d'HDS, sachant qu'IBM, en « Middle-Range » (Famille DS4000), est un OEM de LSI Logic (Engenio). LSI Logic est un « géant » du stockage méconnu du public qui est troisième au niveau mondial en milieu de gamme / Middle-Range.

La maturité de l'offre « Virtualisation du Stockage » des trois leaders est directement transcrite dans les résultats d'une étude d'A.G. Edwards de juillet 2006 « High-End RAID Market Share Analysis », pour le marché mondial du stockage de la classe « Enterprise » / « High-End » : Depuis fin 2005, HDS est passé devant EMC², ce dernier perdant des parts de marché. De même pour IBM qui reste en troisième position. Dans la période 2002 / 2006, seul HDS a progressé sur ce marché.

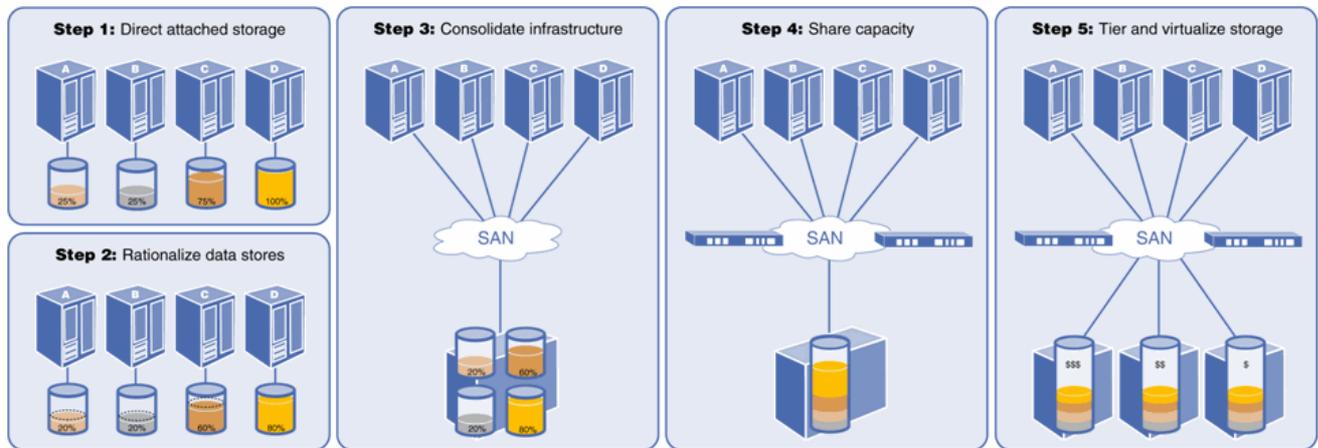
A ce jour, l'offre de virtualisation du stockage, devenue mature, permet d'évoluer « en douceur » vers un Système d'Information (« Data Centre ») ouvert et flexible, nécessaire aux impératifs de demain pour l'Entreprise. La capacité des entités « Métier » à s'adapter rapidement aux marchés de son Entreprise devient un facteur décisif de compétitivité. Ce défi fait peser de lourdes contraintes sur les systèmes d'information, et notamment sur le centre de données, dont les données stratégiques font partie du patrimoine de l'Entreprise . Une Entreprise (organisation Entreprise Cap 2020), ne pourra être flexible que si son système d'information l'est également. Le SI devra passer par la virtualisation du stockage, la virtualisation des serveurs et celle des infrastructures afin d'obtenir un SI à la demande des besoins de l'Entreprise.

10.2 Concepts

La virtualisation du stockage permet de masquer la disparité des ressources de stockage et de les présenter comme un volume logique homogène. La virtualisation est un middleware composé de la combinaison d'équipements (serveurs, proxy, commutateurs réseaux) et de logiciels éventuellement embarqués dans les Directeurs (SAN) et/ou dans les baies de stockage. Cette couche exploite les technologies de mapping de ressources physiques et intègre les méthodes propres à l'administration du stockage. Elle assure les relations entre les différents protocoles des environnements de stockage SAN et NAS (FC, iSCSI, SCSI, SATA, etc.).

Exemple d'Optimisation du Stockage

D'une infrastructure distribuée vers une infrastructure à étages



- Consolidation → Migration des données
- Virtualisation → Liée au choix de l'architecture (et donc du constructeur et/ou éditeur de la solution de stockage virtualisé)
- Optimisation → Classe de stockage / Thin provisioning

10.3 Le marché et les acteurs du marché du Stockage

10.3.1 Les Constructeurs du Marché du Stockage Informatique

Par exemple, dans le cadre d'un appel d'offres (AO), la consultation directe des constructeurs leaders du stockage comme HDS (Hitachi Data Systems), IBM et EMC² permet de couvrir l'ensemble des produits recherchés dans le cadre du renouvellement de baies de stockage / DataCenter.

Les Constructeurs et Éditeurs du domaine du stockage, dans la gamme des produits visés (« High-End »), sont :

- HDS (Hitachi Data Systems),
- IBM,
- EMC²,
- STK (StorageTek) – Société, très active dans les domaines de la sauvegarde sur librairie, a été rachetée par SUN en juin 2005. Elle n'est pas très représentative au niveau du stockage par rapport aux trois leaders listés ci-dessus.

→ Et les OEM's de ces constructeurs / éditeurs de logiciels :

- DELL (DELL est OEM d'EMC² pour la gamme CLARiiON),
- HP (HP est OEM de HDS)
- SUN (SUN est OEM de HDS)
- BULL (BULL est un intégrateur de solutions d'EMC²)

Au niveau des suites logicielles d'accompagnement des produits de stockage et du SAN, on trouve :

- **HDS (Hitachi Data Systems),**
- **IBM (1),**
- **EMC² (2),**

➔ Et les OEM's de VERITAS (SYMANTEC) / éditeurs de logiciels :

- SUN (SUN est OEM de VERITAS au niveau de « Storage Foundation » et NetBackup)
- HP (HP est OEM de VERITAS au niveau de « Storage Management Suite », Storage Foundation, Enterprise Security partnership)
- DELL (DELL est OEM de VERITAS au niveau de Backup-Exec)

(1) : Pour IBM, le produit du type « **Appliance** » **Hardware / Software** « **SVC** » (**SAN Volume Controller**) d'« IBM TotalStorage Virtualization », a remplacé chez IBM son ancien produit commercialisé en OEM de DataCore Software « SanSymphony ». Ce produit logiciel tournant sous NT à l'époque présentait, dans les années 2000 / 2001, les mêmes objectifs et la même architecture de virtualisation du SAN.

(2) : Pour EMC², la virtualisation (Invista) est « sous-traitée » au niveau des Directeurs du SAN.

10.3.2 Quelques Produits RAID / Baies de Stockage « High-End »

EMC Symmetrix

DMX800 / DMX1000 / DMX2000 / DMX3000 / DMX-3 / DMX-4

IBM TotalStorage

DS6800 / DS8100 / DS8300 / DS8000 Turbo Series

Hitachi / HDS

TagmaStore / NSC55 / USP 100-600-1100 / Hitachi Universal Storage Platform VM (USP VM)

➔ Pour mémoire :

HP StorageWorks (OEM d'HDS)

XP10000 / XP12000 / HP StorageWorks XP12000/10000

Sun StorEdge (OEM d'HDS)

9985 / 9990

10.4 Le stockage sécurisé / Technologie RAID

10.4.1 Notion de RAID / Technologies RAID

RAID est un acronyme signifiant « **R**edundant **A**rray of **I**nexpensive / **I**ndependent **D**isk » ou ensemble redondant de disques indépendants.

Un ensemble (« Array ») de type RAID (Baie de Stockage) est une collection de disques agissant comme une unité unique de stockage qui supporte la tolérance de panne de disque(s) sans perte de données et qui fonctionne de façon indépendante des autres sous-systèmes.

10.4.2 Les différents niveaux de RAID / Technologie RAID :

Un groupe de recherche de l'université Californienne de BERKELEY en 1989 a défini six niveaux de RAID. Les ARRAY et RAID (matrice de disques durs) étaient déjà utilisés dans l'industrie à la sortie de ce rapport.

Chaque niveau correspond à la manière dont les données sont stockées sur les différents disques (un compromis entre coût, sécurité et vitesse).

La compréhension des ces niveaux est importante car chaque niveau est optimisé pour une utilisation différente.

➔ La société Française COPERNIQUE avait développé en 1980 un RAID-1 avec un cache d'Entrée / Sortie sécurisé sur batteries, le DIRAM 32 qui pouvait gérer jusqu'à 8 couples indépendants de 2 disques en Miroirs. Cette baie de stockage contrôlait jusqu'à 16 disques en double accès en interface SMD / Storage Module Drive de CDC. C'était le premier RAID transactionnel Européen qui a été commercialisé et vendu au niveau Mondial pendant plus de 10 ans.

Extrait / U. C. Berkeley :

« In 1989, the RAID (Redundant Arrays of Inexpensive Disks) group at U. C. Berkeley built a prototype disk array called RAID-I. The bandwidth delivered to clients by RAID-I was severely limited by the memory system bandwidth of the disk arrays host workstation. We designed our second prototype, RAID-H, to deliver more of the disk array bandwidth to file server clients. A custom-built crossbar memory system called the XBUS board connects the disks directly to the high-speed network, allowing data for large requests to bypass the server workstation. RAID-II runs Log-Structured File System (LFS) software to optimize performance for bandwidth-intensive applications. The RAID-II hardware with a single XBUS controller board delivers 20 megabytes/second for large, random read operations and up to 31 megabytes/second for sequential read operations. A preliminary implementation of LFS on RAID-II delivers 21 megabytes/second on large read requests and 15 megabytes/second on large write operations. »

Remarque : Dans ce qui suit, seules les technologies qui sont principalement utilisées par les principaux constructeurs de RAID sont résumées ci-après.

10.4.3 RAID / Niveau 0 → « Array 0 » / Mode « Stripping »

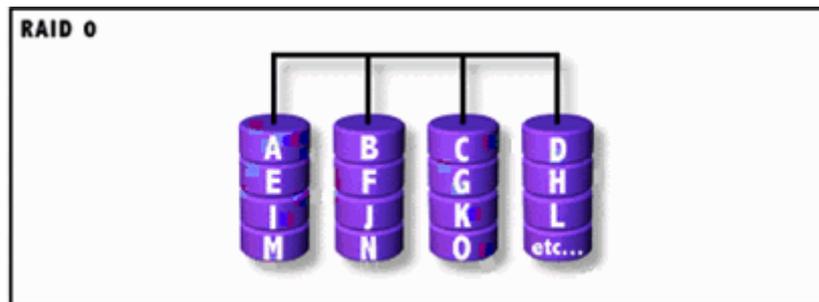
En mode **RAID 0 / ARRAY 0**, la donnée à stocker est répartie sur différents disques et aucune information de redondance n'est stockée, il en résulte une vitesse de transfert importante mais aucune sécurité par redondance d'information.

→ C'est la raison pour laquelle on notera ARRAY-0 et non RAID-0

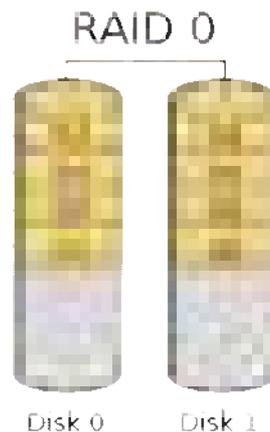
Donc, la moindre panne de disque entraîne la perte irrémédiable de données.

Le nombre minimum de disques requis, pour ce mode est de 2.

Ce niveau de RAID / ARRAY est aussi appelé mode « **STRIPPING** ».



→ Remarque : Le RAID-0 est également un RAID-3 (voir ci-après) dont le disque de parité est absent ou en panne.



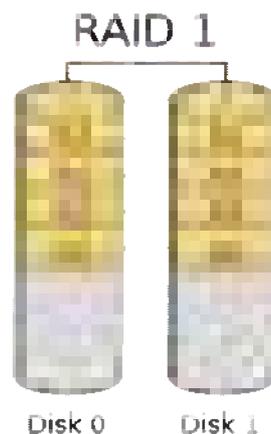
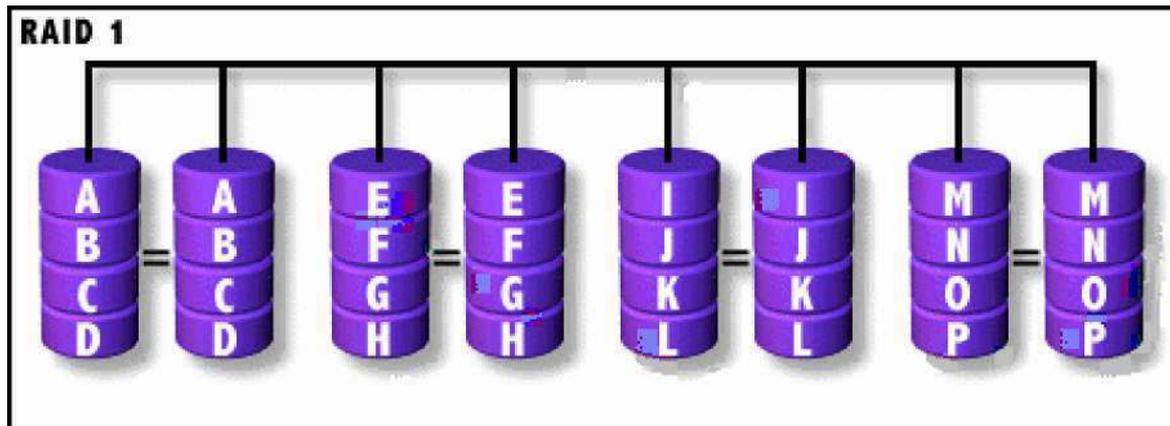
10.4.4 RAID / Niveau 1 - Mode « Mirroring » / « Shadowing » et « Duplexing »

En mode **RAID 1**, la donnée est intégralement dupliquée d'un disque sur un autre, d'où une redondance importante en cas de panne d'un disque.

Les performances sont supérieures à un disque seul mais le coût du Mégaoctet est très élevé (X2).

Ce mode peut être réalisé directement par le système d'exploitation de la machine « hôte » ou « serveur », entraînant ainsi une charge non négligeable du CPU.

Ce mode requiert au minimum 2 disques et est communément appelé « **MIRRORING** » lorsqu'il est réalisé sur le même canal d'interface Disques (SCSI, FC...), et « **DUPLEXING** » lorsque 2 canaux indépendants sont utilisés pour 2 disques en RAID-1.



10.4.5 RAID / Niveau 3 - Transfert parallèle avec gestion de parité en temps réel

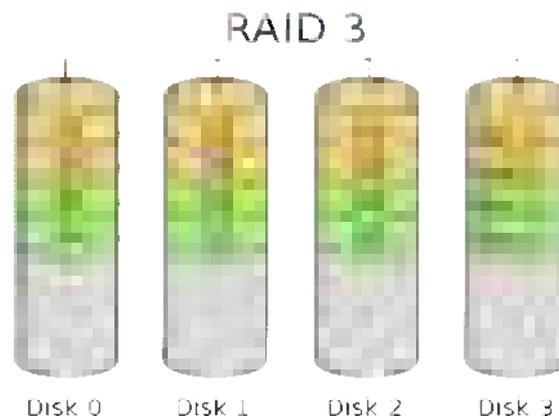
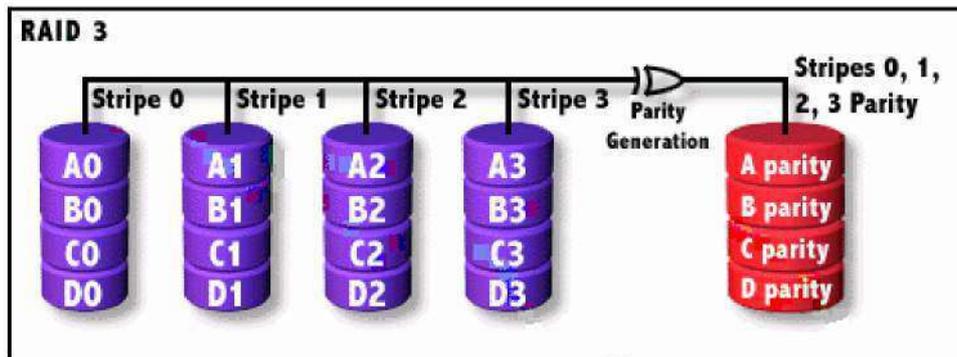
En mode **RAID 3**, la donnée à stocker est répartie en octets sur différents disques (en général 4 ou 8) («**STRIPPED**») éventuellement synchronisés en rotation et le système génère et stocke une parité sur un seul disque de parité.

Il en résulte des taux de transfert très importants en lecture mais aussi en écriture, en général taux de transfert d'un disque multiplié par le nombre de disques gérés en RAID-3 en parallèles (X4 ou X8).

La panne d'un disque de données a très peu d'impact sur la performance et le nombre minimum de disques pour l'utilisation du mode 3 est de 3 mais en général c'est 4 ou 8 disques plus le disque de parité, On dit RAID-3 4+1 ou 8+1.

La complexité du contrôleur est moyenne en mode RAID 3, mais il est très difficile de le réaliser uniquement par logiciel pour des raisons de performance !

En réalité, très peu de constructeurs implémentent un réel RAID niveau 3, car en général, ces RAID-3 sont spécialisés pour des applications de vidéo numérique, applications de postproduction (TV et cinéma) et des applications scientifiques et militaires nécessitant de très hautes performances en termes de débit d'E/S.



10.4.6 RAID / Niveau 4 - Disques de données indépendants avec gestion de parité partagée

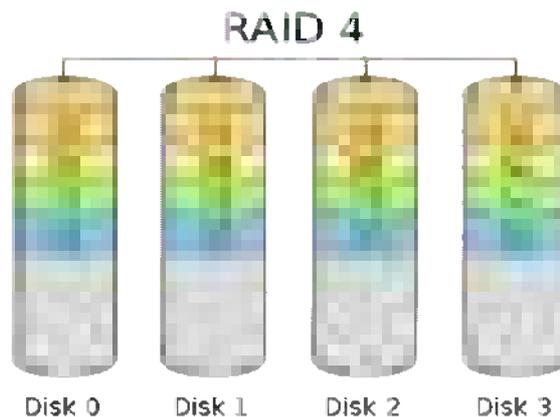
En mode **RAID 4**, un block entier de données est stocké sur un seul disque et le système génère et stocke une parité de plusieurs blocks provenant de différents disques de données sur un seul disque de parité.

Les taux de transferts sont très importants en lecture de larges fichiers mais ce mode est peu performant en écriture.

D'autre part, du fait du partage de la parité sur des blocks entiers de données, la reconstruction des données peut s'avérer difficile en cas de panne d'un disque. La panne d'un disque a un impact moyen sur les performances du système.

Dans ce mode, les disques de données sont indépendants et ne sont pas lus en parallèle sauf dans le cas où les données proviennent de disques différents.

La taille d'un block peut varier entre la taille d'un secteur et plusieurs Mégaoctets suivant le contrôleur utilisé et le coût du Mégaoctet est relativement faible en mode 4.



→ Ce mode n'est plus utilisé au profit du RAID-5 traité ci-après.

10.4.7 RAID / Niveau 5 - Disques de données indépendants avec gestion de parité partagée et distribuée

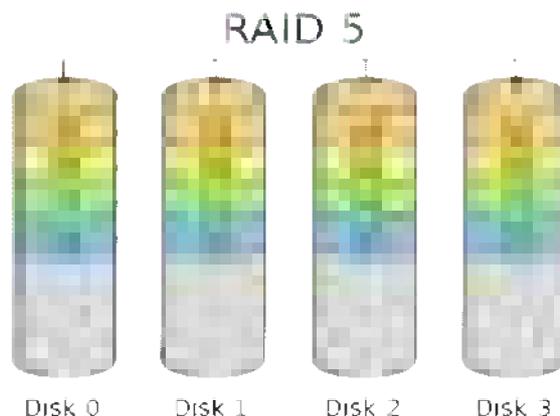
En mode **RAID 5**, les informations sont stockées de manière similaire au mode 4, par contre la parité est générée et est stockée de façon distribuée sur les disques de données.

Les performances sont importantes pour la lecture de larges fichiers et bonnes pour la lecture de petits fichiers et en mode écriture.

D'autre part, du fait du partage distribué de la parité sur des blocks entiers de données, la reconstruction des données peut s'avérer difficile en cas de panne d'un disque. La panne d'un disque a un impact moyen sur les performances du système.

Dans ce mode, les disques de données sont indépendants et ne sont pas lus en parallèle sauf dans le cas où les données proviennent de disques différents.

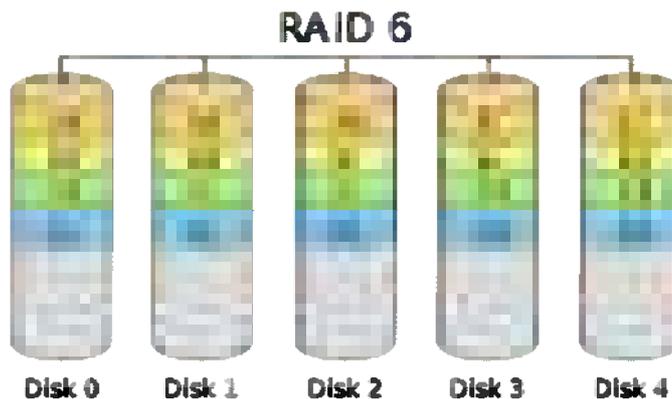
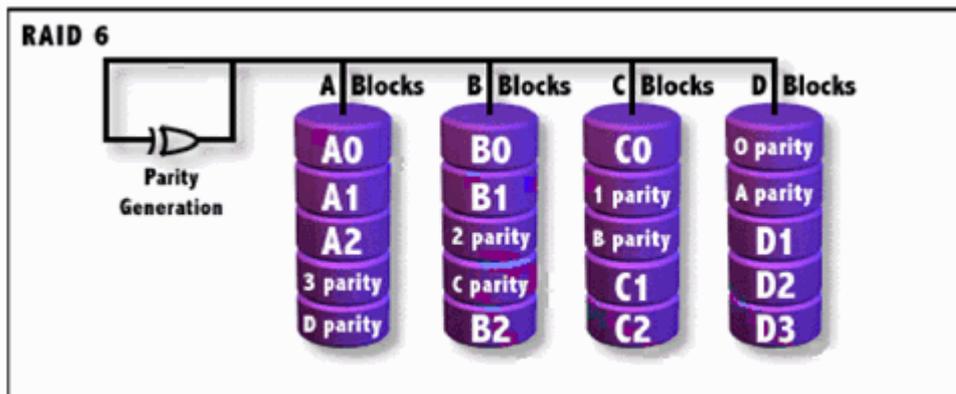
La taille d'un block peut varier entre la taille d'un secteur et plusieurs Méga octets suivant le contrôleur utilisé et le coût du Méga octet est relativement faible en mode 5. Ce mode requiert au minimum 3 disques.



10.4.8 RAID / Niveau 6 - Disques de données indépendants avec gestion de parité partagée, distribuée et répartie

Le mode **RAID 6** est similaire au mode 5 mais utilise plusieurs disques de parité (en général 2 disques de parité).

➔ De ce fait, le mode 6 admet de perdre plus d'un disque (jusqu'à deux disques) de données et de continuer à fonctionner en mode dégradé.

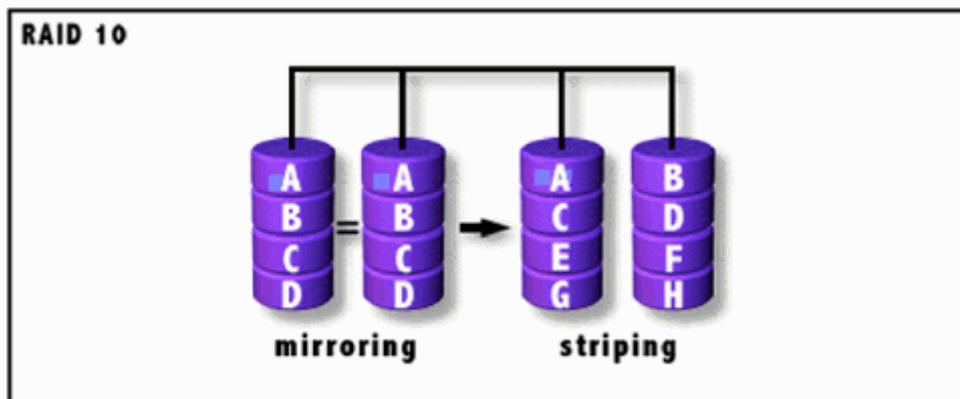


➔ Idem RAID-5 mais avec génération de deux parités pour accepter la perte de deux disques.

10.4.9 RAID / COMBINAISONS (RAID 10, RAID 31)

Certains constructeurs de contrôleur RAID propose des combinaisons de plusieurs niveaux de RAID, en voici quelques exemples : RAID 0+1 et RAID 3+1

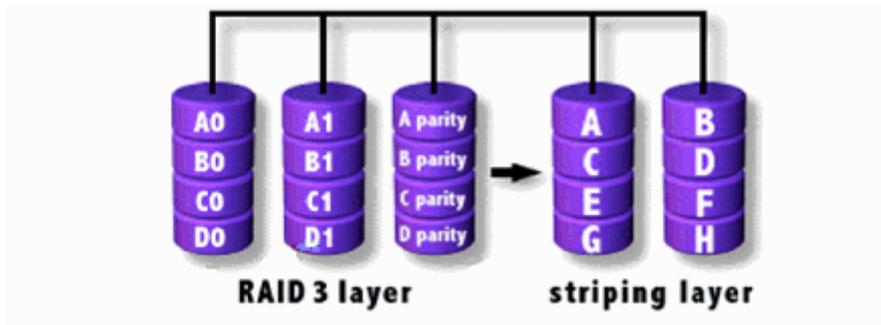
RAID 10 (« Miroir de Striping »)



RAID 31 (« Striping sécurisé avec parité type RAID-3 »)

Combinaison de la Technologie RAID-3 et du « striping » (3 disques minimum)

RAID 31



10.4.10 Le JBOD "Just a Bunch Of Disks" / NRAID

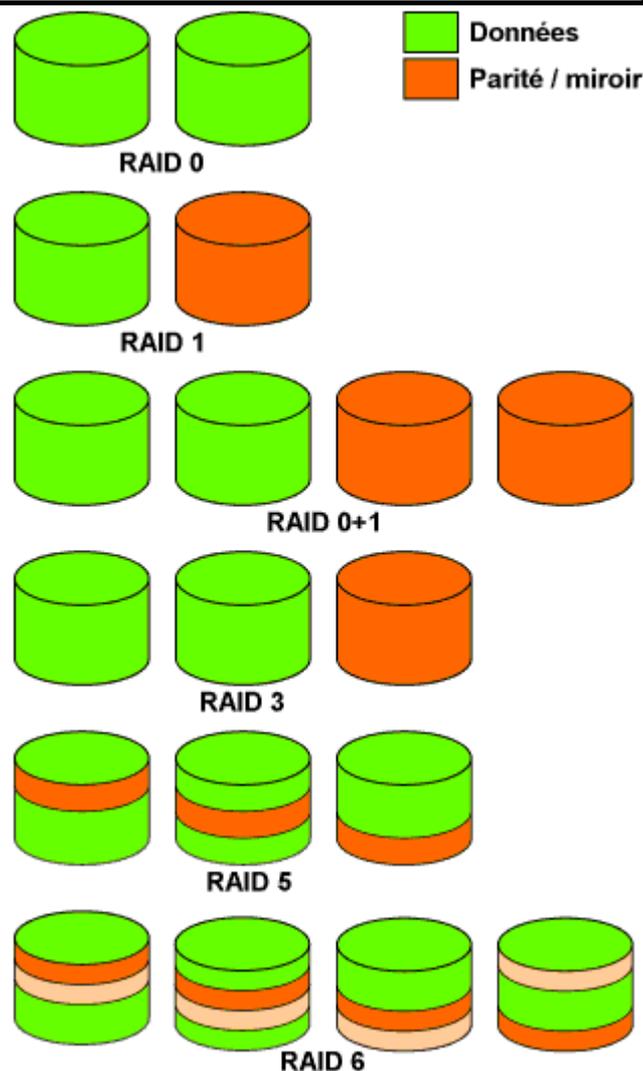
- Littéralement le **JBOD** veut dire "**Just a Bunch Of Disks**".
- Ce mode se nomme aussi **NRAID** ("Non-RAID") chez certains constructeurs.
- Celui ci correspond à un agrégat (« un de aspect de la Virtualisation ») de disques et n'apporte **aucune sécurité**.
- Celui ci permet simplement de créer un volume logique basé sur plusieurs disques physiques.

[Disque 1 - 15 Go] + [Disque 2 - 20 Go] = [Volume C - 35 Go]

➔ Le JBOD permet la plus part du temps d'augmenter un volume existant fonctionnant sans RAID par l'ajout d'un disque physique.

➔ L'augmentation de la capacité d'un RAID-1, RAID-5 ou RAID-6 est également possible « on line » sur certaines baies de stockage « High-End ».

Résumé des différents niveaux de RAID les plus utilisés :



RAID 0 : Les données sont réparties sur deux disques. Ainsi, l'espace de stockage est multiplié par 2, et les accès sont plus rapides. Ce niveau n'offre **aucune sécurité**.

RAID 1 : Les données sont écrites simultanément sur le disque et son miroir. La capacité de stockage n'est pas accrue, mais la sécurité est optimale.

RAID 0+1 : Ce niveau combine les deux niveaux 0 et 1. Espace de stockage plus important, sécurité optimale, rapidité d'écriture et de lecture, au détriment d'un coût plus élevé.

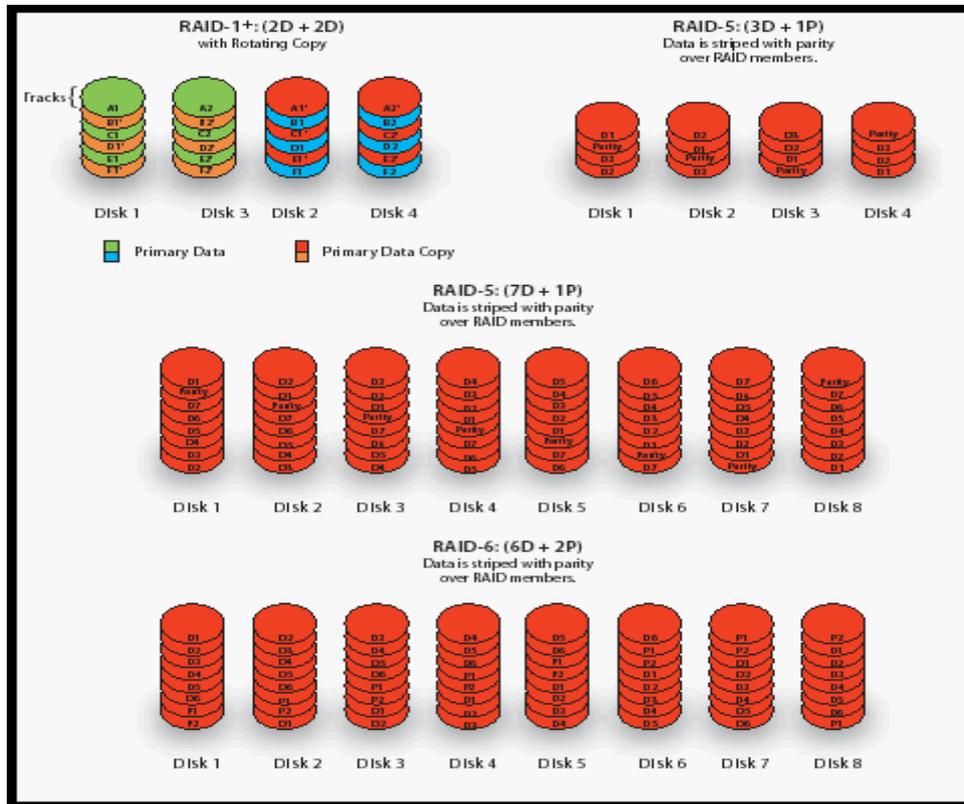
RAID 3 : L'un des disques est dédié à l'enregistrement des sommes de contrôle ; les données sont enregistrées sur les autres disques. L'écriture met en œuvre tous les disques, la lecture n'active que les disques de données.

RAID 5 : Ici, la parité est répartie de manière égale entre les disques de l'ensemble. Les opérations d'écriture sont plus lentes qu'en niveau 3, mais la lecture est plus rapide.

RAID 6 : Dans ce niveau, deux ensembles de parités sont calculés et répartis entre les disques

de l'ensemble. Ce niveau supporte une panne de *deux disques simultanément* (au lieu d'un seul en RAID 5).

Exemple avec HDS : RAID-1+, RAID-5, and/or RAID-6



RAID-1+ is available in 2 Data + 2 Data and 4 Data + 4 Data disk arrangements

RAID-6: two redundant sets of parity information for each stripe of data.

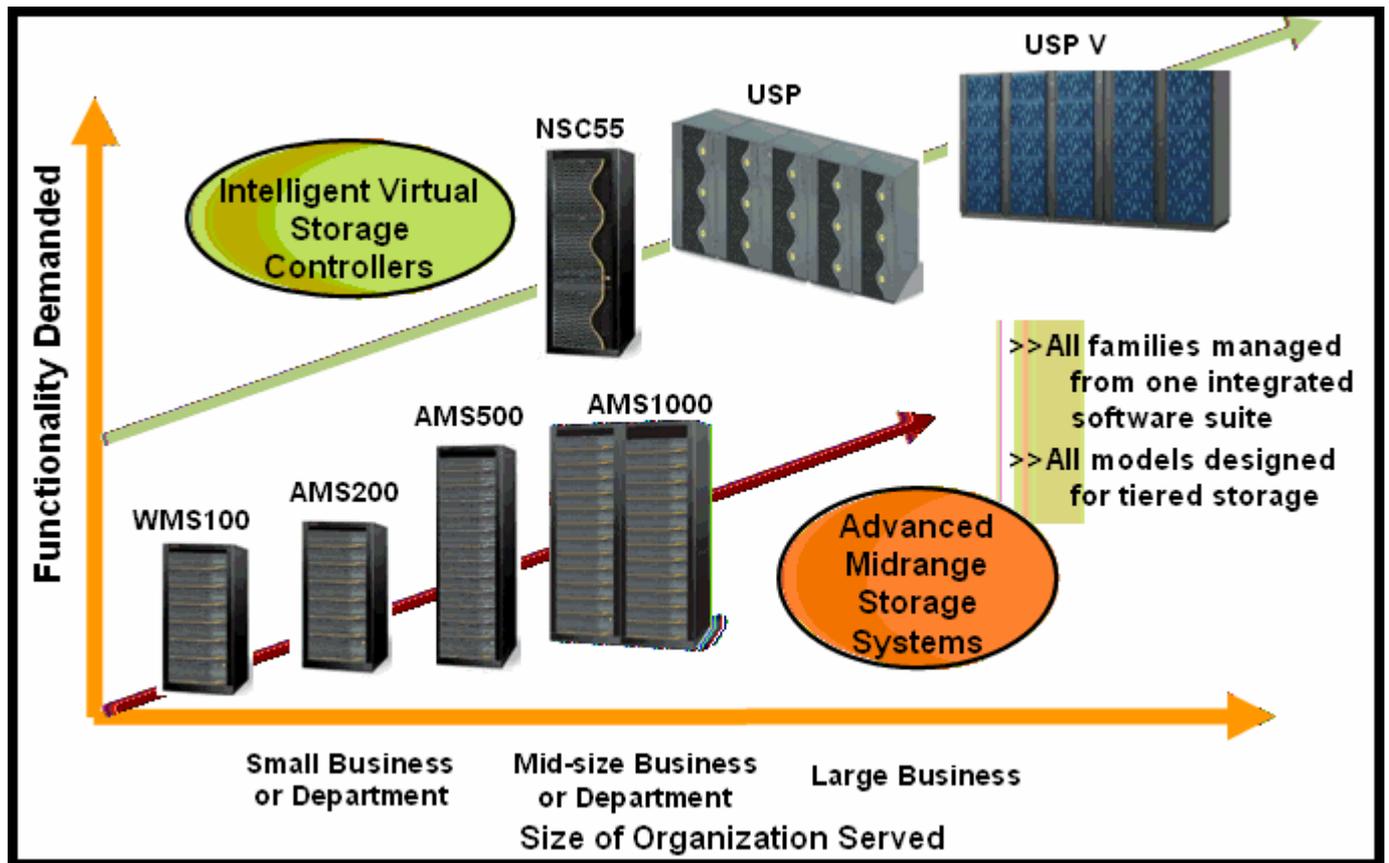
RAID-6 can recover from the failure of two drives in the array while other single RAID levels can handle at most one fault.

10.5 Les Architectures de virtualisation du stockage

10.5.1 HDS - Hitachi Data Systems

10.5.1.1 USP V – Universal Storage Platform (Business Continuity & Virtualization)

Hitachi Services Oriented Storage Solutions Product Portfolio



➔ Les « Suites » Software et le Logiciel d'Administration sont communs
 (« Common software tools » from a single management console)

Storage Management and Hitachi HiCommand Suite Software

10.5.1.2 Présentation résumée de l'USP V

- Baie de stockage **haut de gamme Hitachi**
 - Consolidation, performances et disponibilité
 - Systèmes ouverts
 - Systèmes mainframe

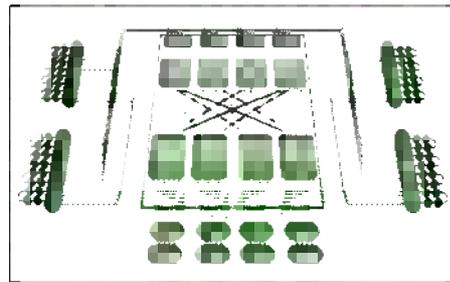
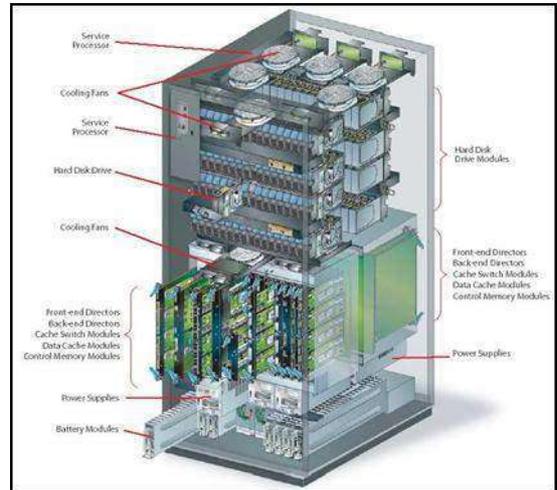
- **Fabriquée** pour l'Europe en France, à **Orléans**
 - Usine « **Hitachi Computer Products** », **Europe France**
 - Unité fabriquée depuis **juin 2007**
 - Un produit conforme aux normes sur l'environnement
 - Norme RoHS : Composants électroniques fabriqués sans utiliser certains éléments nocifs (plomb, cadmium, chrome...)

- **Commercialisée** par Hitachi Data Systems, Sun et HP
 - sous le nom d'**USP V** pour **Hitachi Data Systems**
 - sous le nom de ST9990V pour Sun Microsystems
 - sous le nom de XP24000 pour Hewlett Packard (HP)

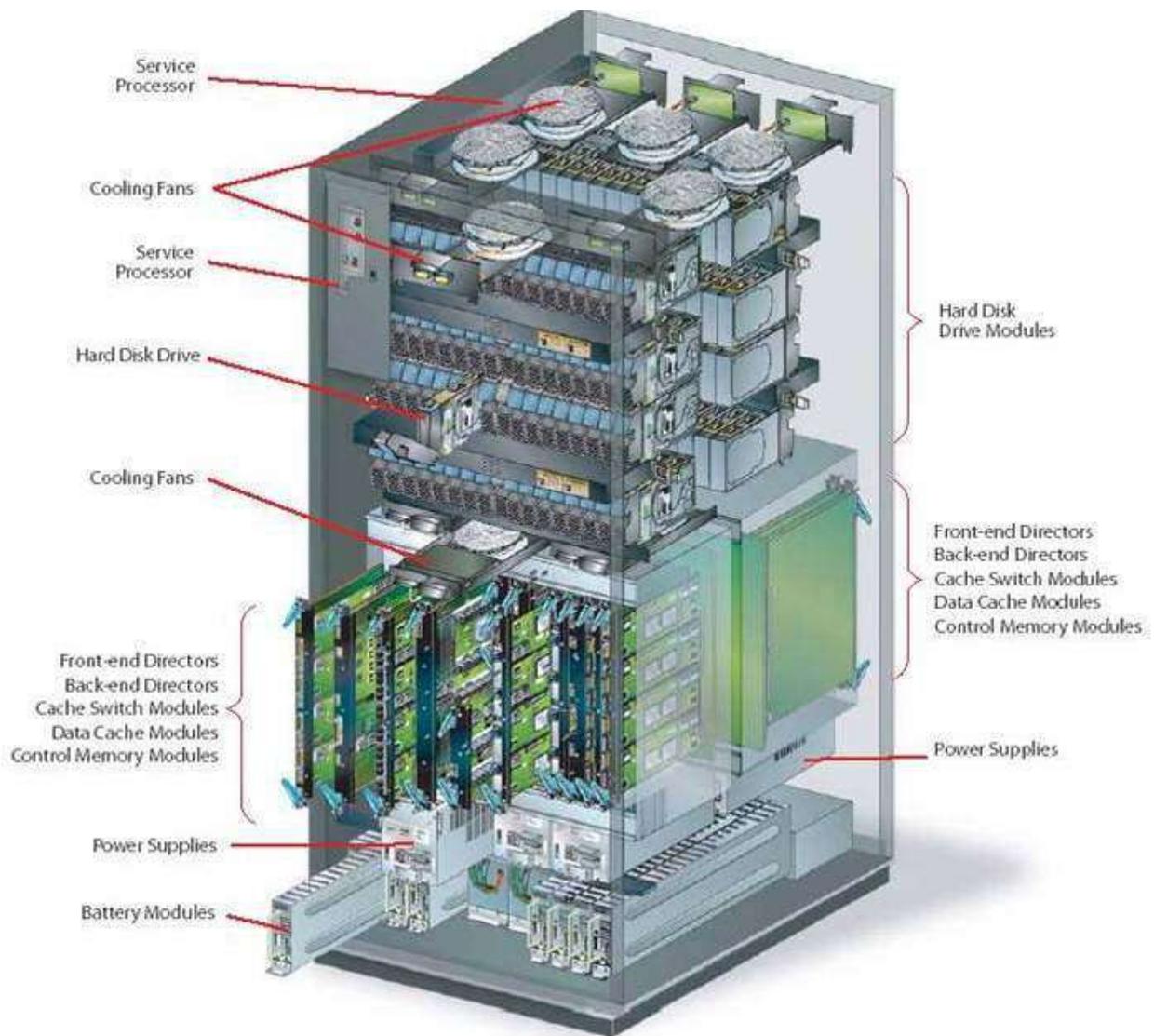


10.5.1.3 Architecture interne USP V

- Connectique maximale
 - SAN FC : 224 x Fibre Channel (1, 2 et/ou 4 Gbps)
 - Mainframe : 112 x FICON, 112 x ESCON
- Mémoire cache système
 - Jusqu'à **32 Go**
- Mémoire cache pour les données
 - Jusqu'à **256 Go utiles (mémoire cache dupliquée)**
 - 512 Go et 1.024 Go
- Matrice de commutation interne
 - Quatrième génération de Matrice HiStar™
 - Bande passante de commutation jusqu'à **106 Go/s**
- Volumétrie alliant performance et consolidation
 - 73 Go 15 krpm, 146 Go 15 krpm, 300 Go 10 krpm
 - Jusqu'à 1.152 disques FC (à 4 Gb/s) en RAID
- 65.536 adresses en open et en mainframe
- **Supporte la technologie RAID 6**
- **Hitachi Dynamic Provisioning (HDP)**



10.5.1.4 L'unité de contrôle de l'USP V – Universal Storage Platform

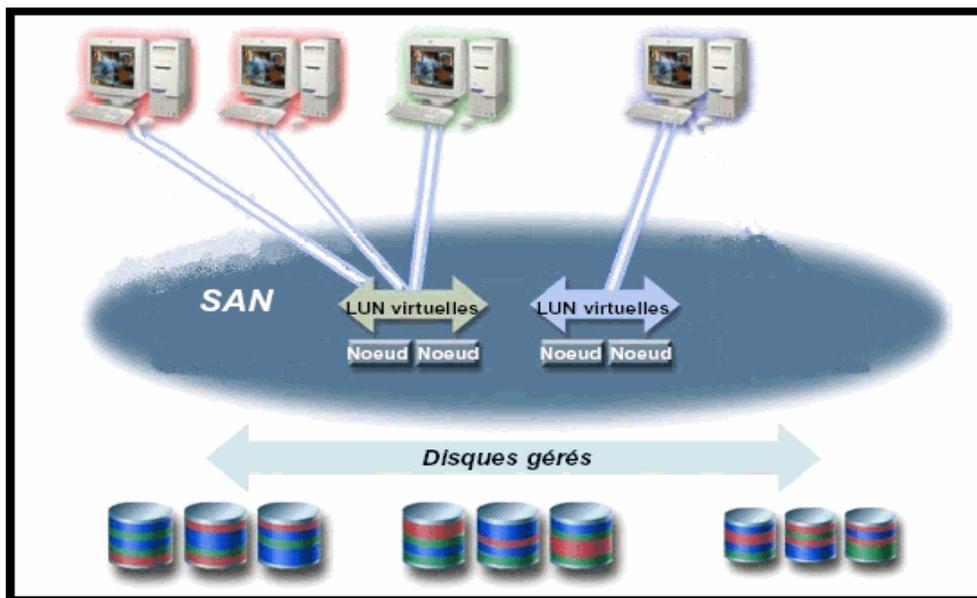


10.5.2 IBM - SVC / SAN Volume Controller – « In Band »

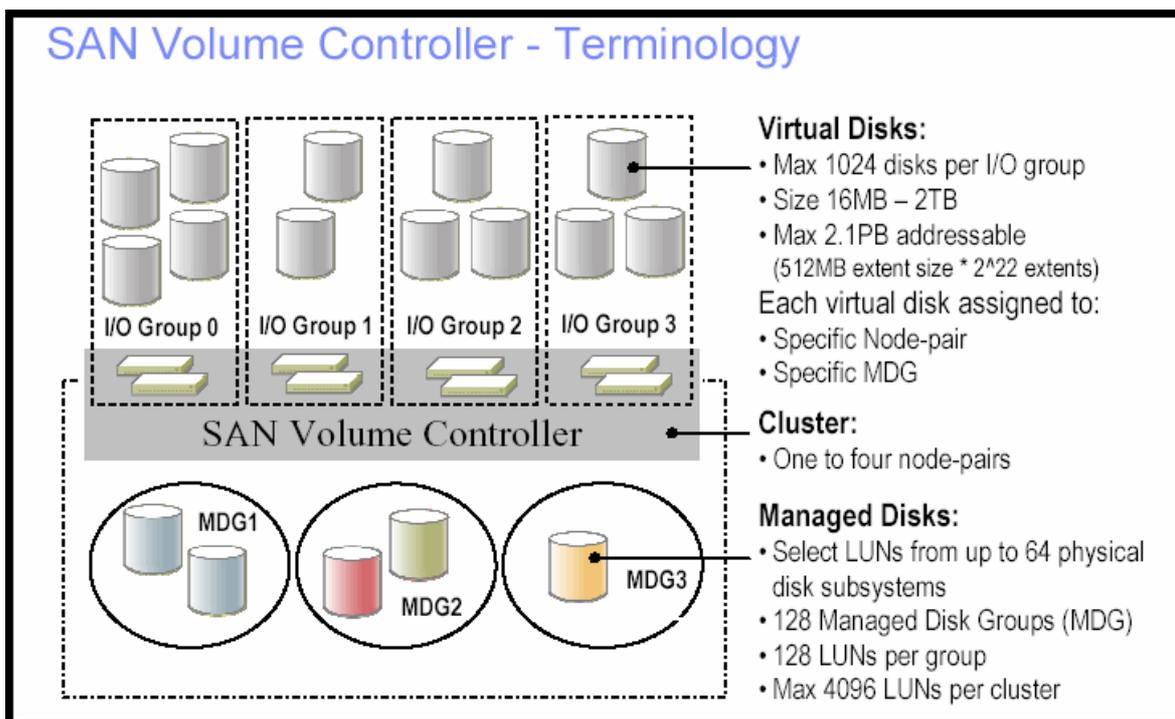
SAN → Storage Area Network

10.5.2.1 Philosophie des SVC d'IBM

SVC → Pools de disques gérés, répartis sur plusieurs sous-systèmes de stockage.

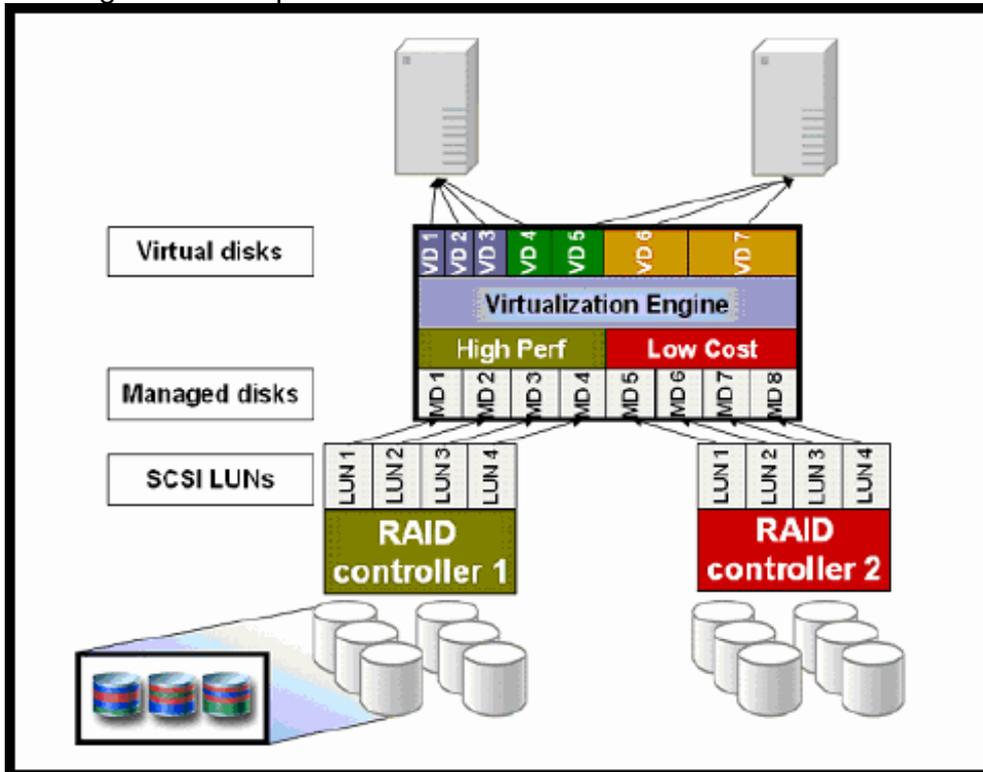


10.5.2.2 « SVC – Terminologie » / Architecture logique du SAN



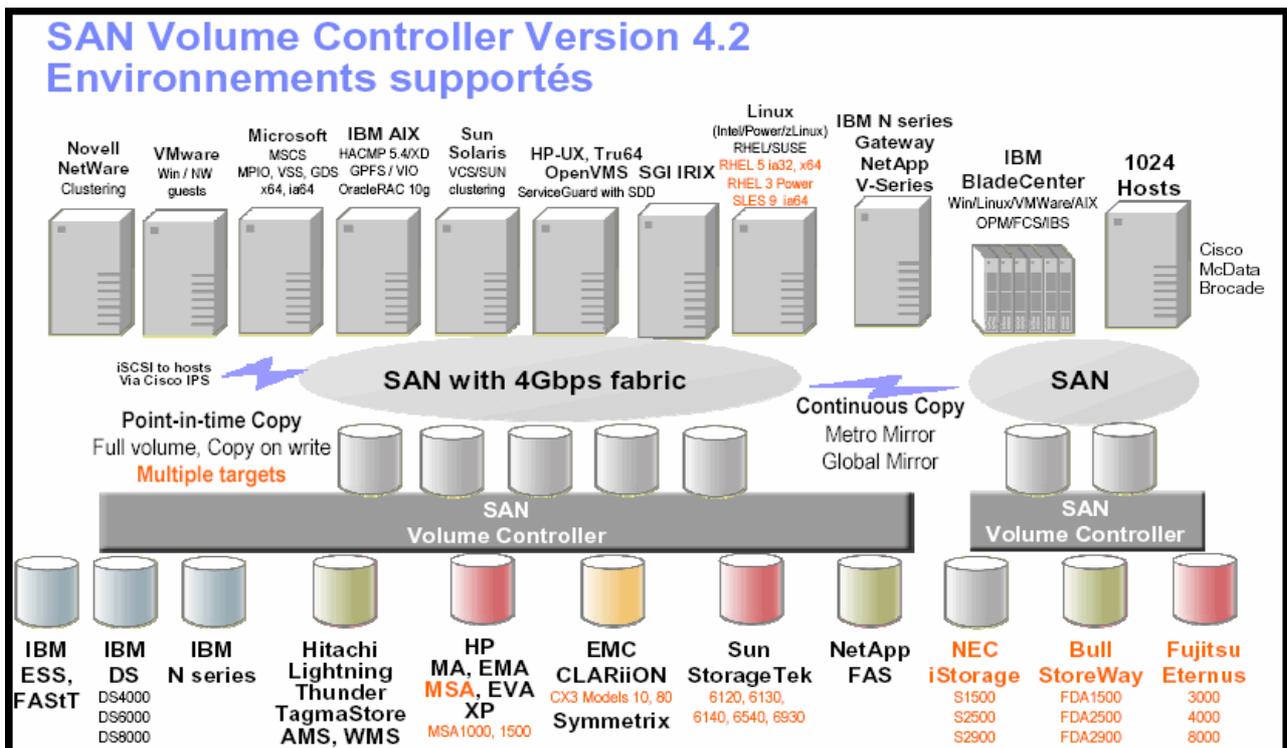
10.5.2.3 Architecture logique du SAN Volume Controller

Les LUN's (« Logical Unit Number ») des différentes baies de disques sur le SAN son vues comme des « Managed Disks » par le SVC.



➔ « Striping » sur plusieurs « Managed Disks » du groupe de disque.

10.5.2.4 SAN Volume Controller - Environnements supportés



10.5.2.5 Avance IBM sur la Virtualisation... par IBM ! (3Q2007)

- 2900 SVC* – 40 pays, 20 industries
- Meilleure utilisation stockage 20-30%
- Diminution des arrêts
- Support stockage hétérogène
- Forte implication des BPs – 75% des ventes
- Meilleur résultat SPC-1
- Seule solution certifiée par le SNIA SMI-S

	IBM	EMC	HDS	HP
Market Penetration	✓ 2900 sales ✓ 130+ ref.	Avail for Pilots	Few signs of virtualization yet	Same as HDS
Product Maturity	✓ SVC - 27 months, 8 th release ✓ SFS – 23 months, 3 rd release	None	1 year (TagmaStore) NSC55 6 months	Same as HDS
Portfolio	✓ 4 – SVC, SFS, DS8000 LPARs, Virtual Tape	1 Invista	2 (TagmaStore and NSC55)	2 (XP12000 & XP10000 OEM HDS)
Heterogeneity	✓ 6 (IBM, Dell, EMC, HP, HDS, Sun)	Some models of EMC, IBM, HDS & HP	3 (EMC, HDS, IBM)	Same as HDS
Scalability	✓ From small SANs (1TB) to large enterprises (>125TB)	Large enterprise > 50TB	TagmaStore Large enterprise NSC55 Midrange	Same as HDS
Entry List Price	✓ \$40K (SVC)	\$225K	\$700K Tag \$300K NSC55	> HDS



10.5.2.6 Stockage IBM : DS6000 & DS8000



10.5.2.7 Stockage IBM : DS6000 & DS8000

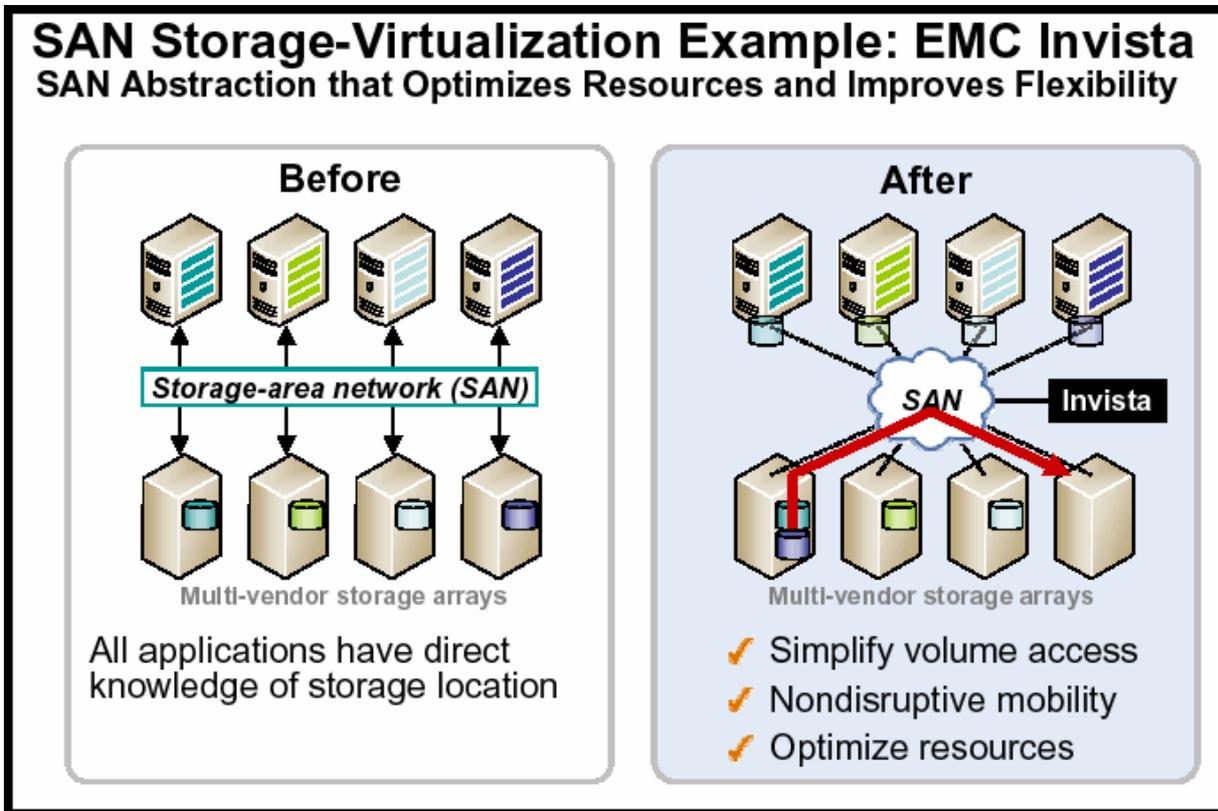
Attachment



	DS4800	DS6800	DS8100 / DS8300
Intel®	YES	YES	YES
Unix®	YES	YES	YES
iSeries™	NO	YES	YES
zSeries®	NO	YES	YES

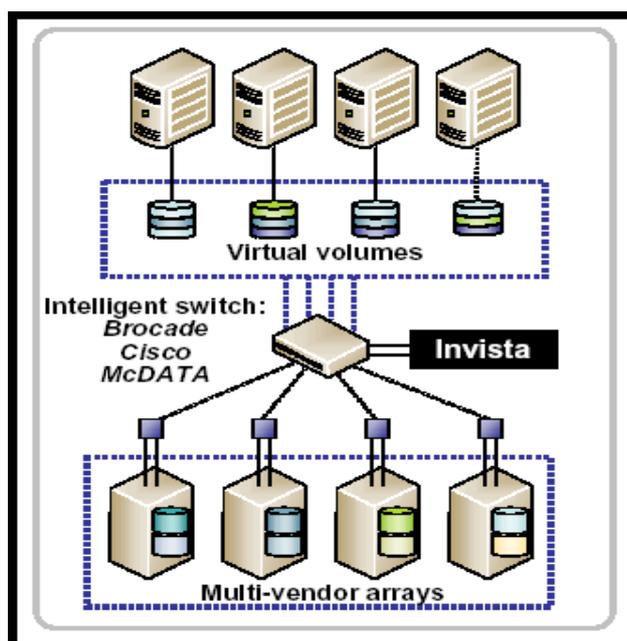
10.5.3 EMC (EMC²)

10.5.3.1 VIRTUALISATION EMC² / INVISTA (« Out of Band »)



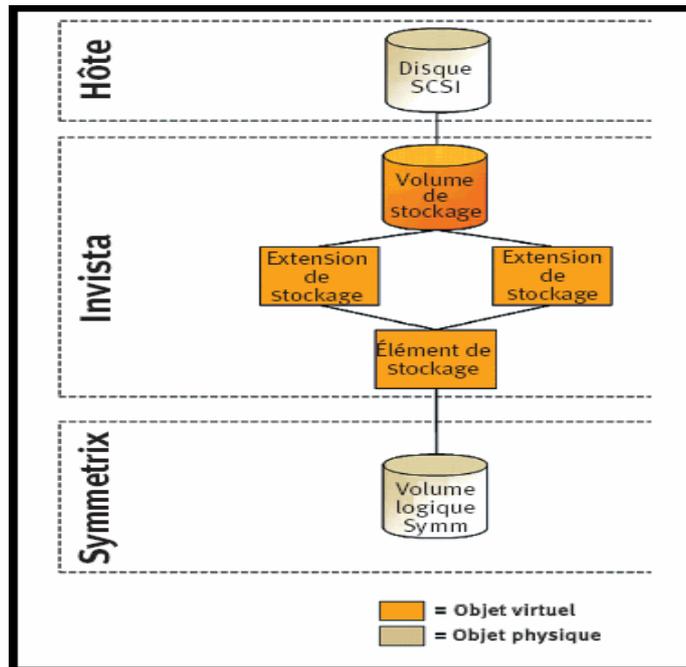
10.5.3.2 VIRTUALISATION EMC² basée sur le SAN / INVISTA (« Out of Band »)

→ SAN : BROCADE & McDATA / CISCO donc 3 développements spécifiques !

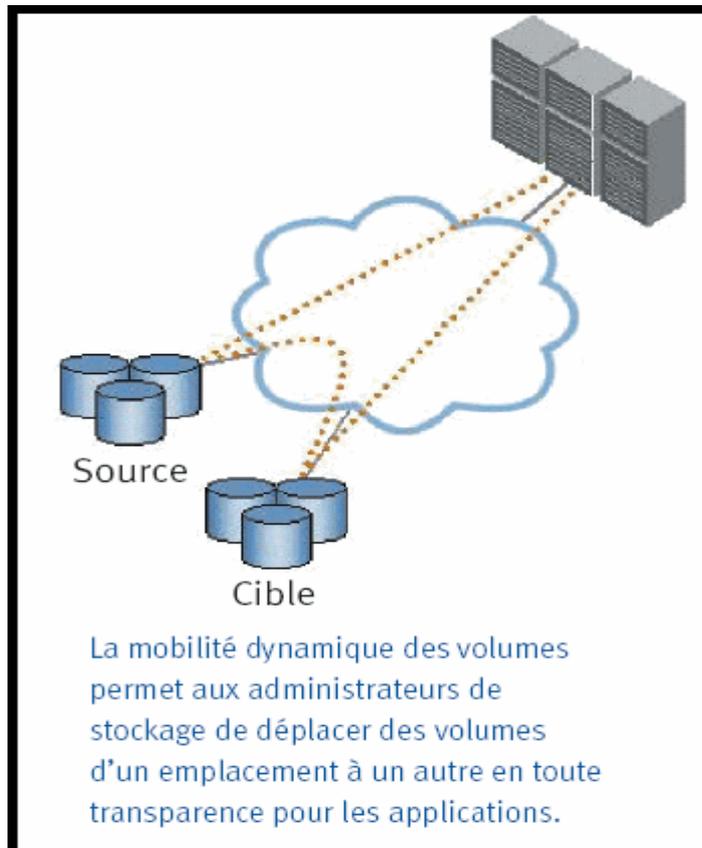


10.5.3.3 Le modèle objet d'Invista

Le modèle objet d'Invista est un modèle à plusieurs couches, il est fondé sur l'abstraction des objets physiques.

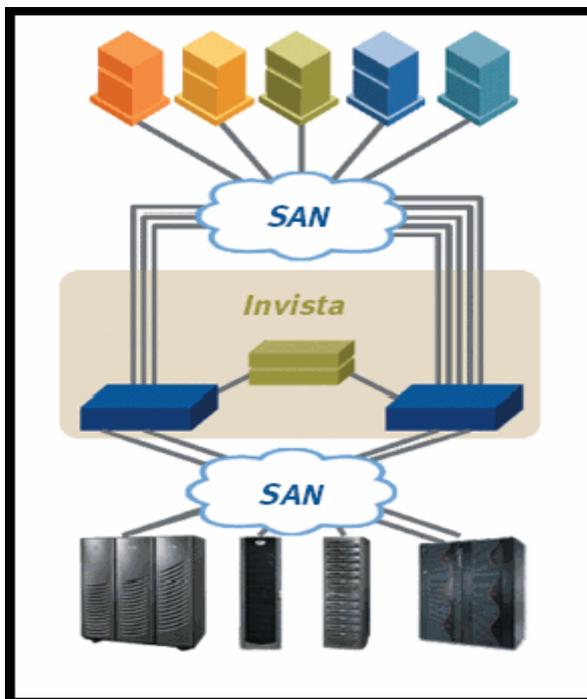


10.5.3.4 Mobilité dynamique des volumes



10.5.3.5 Déploiement simplifié d'EMC Invista dans un SAN.

Rappel : La Virtualisation du stockage d'EMC² est basée sur le SAN par INVISTA (« Out of Band »).



10.5.3.6 EMC² (DMX-3) : Vision par HDS !

**The Storage Industry:
The Archaic View**

HITACHI
Inspire the Next

**Monolithic High-end Disk Array
circa 2007**

**EMC DMX-3: 27-Foot Long,
Over 13 Tons of Stone Age Storage Iron**

Hitachi Data Systems © 2007 Hitachi Data Systems

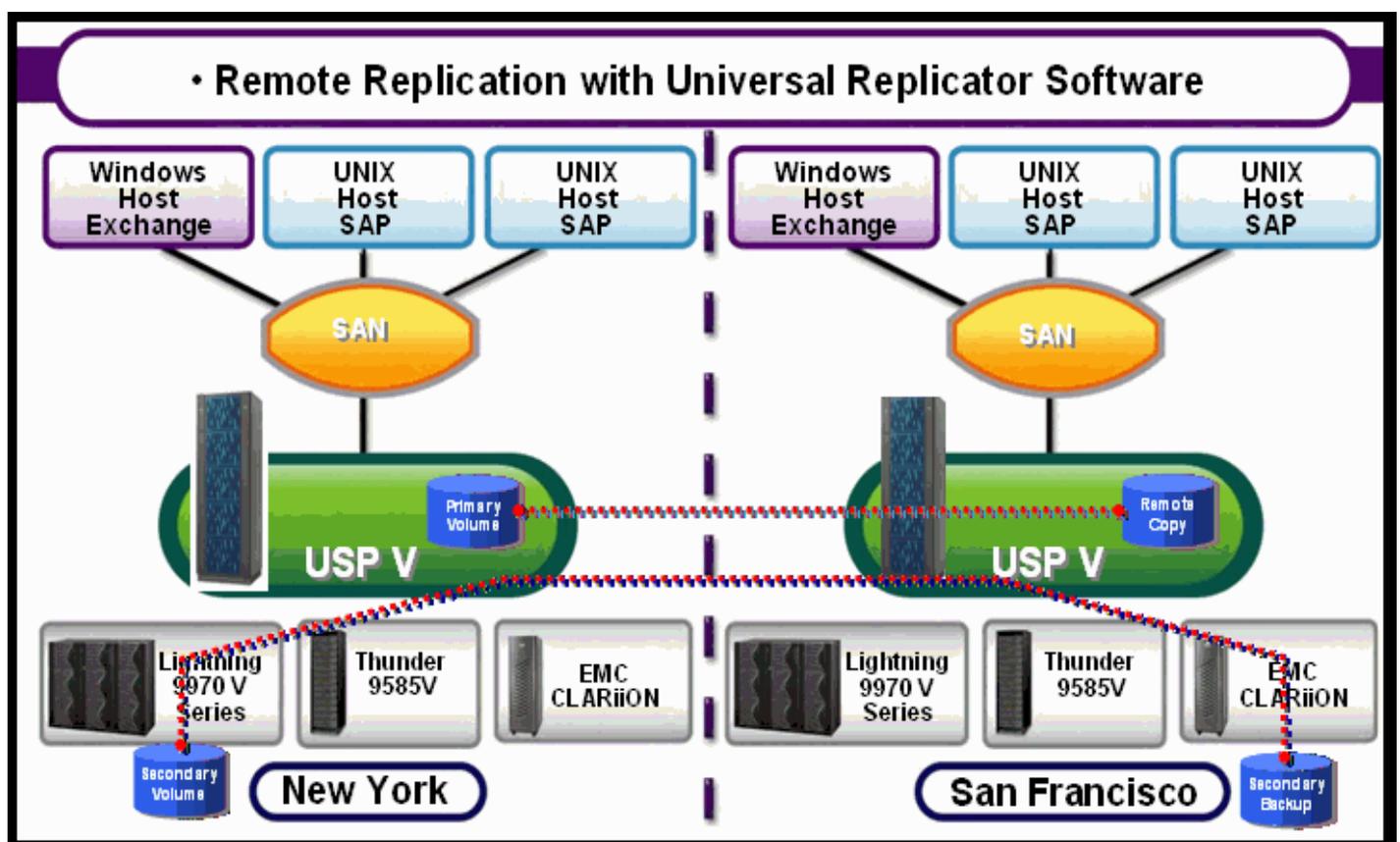
10.5.3.7 **STOCKAGE EMC² - Gamme DMX**

	Conception	Nb. max. de disques	Capacité (brute)	Caraux de disques	Directeur de mémoire globale	Mémoire globale maximale	Connectivité*
<p>DMX-4</p>  <p>» Notice tech.</p>	240/2400	1054 To	8/64 x Fibre Channel 4 Gbit	2 à 8	512 Go	256 Go	64 x FC 4 Gbit 64 x FC 2 Gbit 64 x ESCON 48 x FICON 2 Gbit 8 x GigE SRDF 48 x iSCSI
<p>DMX-4 950</p>  <p>» Notice tech.</p>	32/360	180 To	8 ou 16 x 4Gbit	2	128 Go	64 Go	16 x FC 4 Gbit 16 x FC 2 Gbit 8 x GigE SRDF 10 x GigE iSCSI
<p>DMX-3</p>  <p>» Notice tech.</p>	240/2400	1054 To	8/64 x Fibre Channel 2 Gbit	2 à 8	512 Go	256 Go	64 x FC 4 Gbit 64 x FC 2 Gbit 64 x ESCON 48 x FICON 2 Gbit 8 x GigE SRDF 48 x iSCSI
<p>DMX-3 950</p>  <p>» Notice tech.</p>	32/360	180 To	8 ou 16 x 4Go	2	128 Go	64 Go	16 x FC 4 Gbit 16 x FC 2 Gbit 8 x GigE SRDF 10 x GigE iSCSI

10.6 Clones, Snapshots, Réplication à distance (PCA, PRA, PRCA)

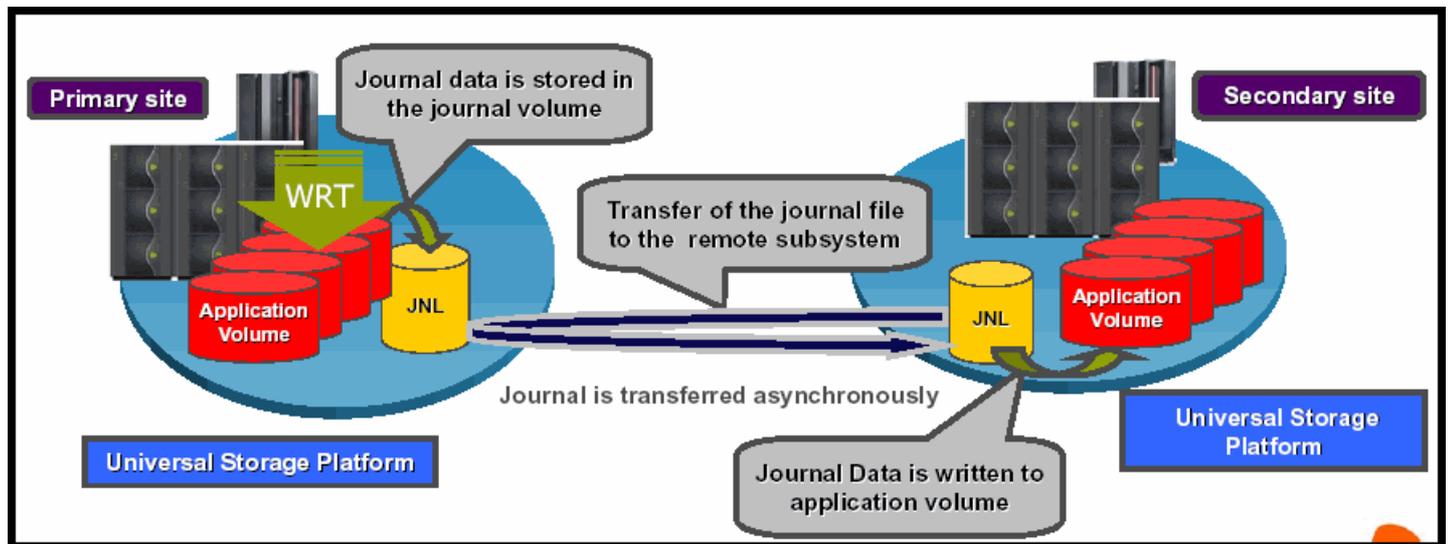
- L'USP d'HDS intègre de nombreux moyens de copier les données :
 - Clone (1 pour 1 interne) et/ou Snapshot (copie différentielle interne)
 - **ShadowImage**
 - Réplication synchrone (de baie à baie) ou asynchrone (de baie à baie)
 - **TrueCopy**
 - Réplication asynchrone journalisée (de baie à baie / grandes distances) **Hitachi Universal Replicator (HUR)**

10.6.1 Remote Replication with Hitachi Universal Replicator Software (HUR)

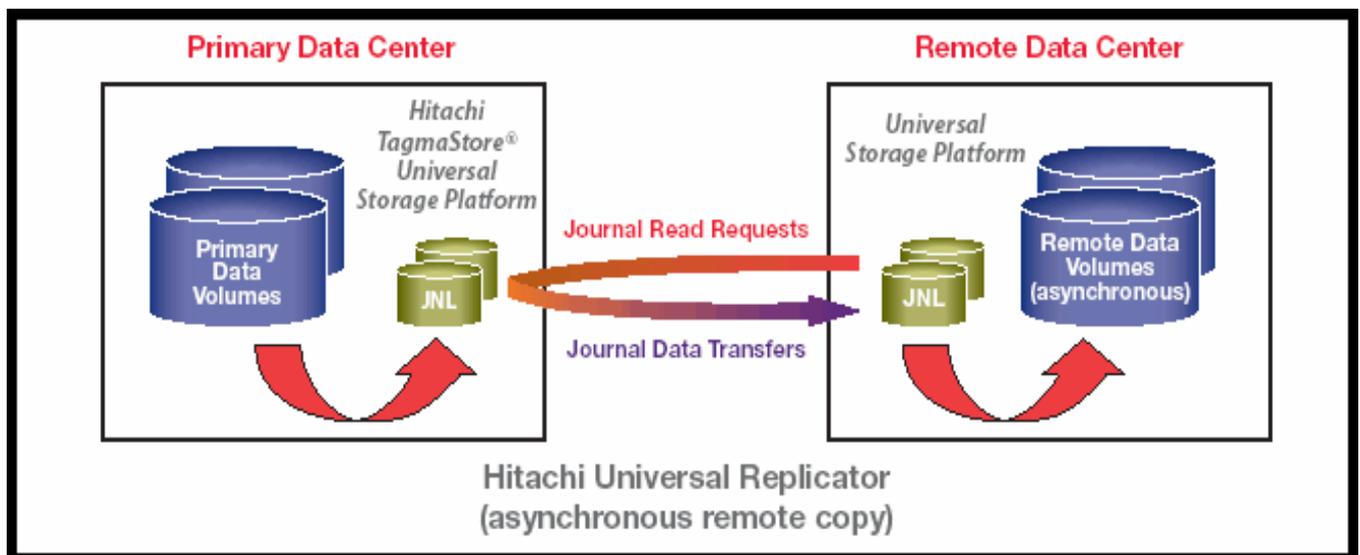


10.6.2 HUR: Détails de la technologie (“Key Technology”)

- Written data is chronologically reflected to the secondary site via journal volumes
- The journal supports consistency across multiple volumes
- The journal is copied by the remote system
- Sorting by time stamp, sequence number



10.6.3 Hitachi Universal Replicator Software: Asynchronous Remote Copy



Universal Replicator software uses disk-based journal volumes and a pull-style replication engine to move data from the primary site to the remote site.

10.6.4 Avantages avec HUR

- **Hitachi Universal Replicator** allows the replication of any volume hosted on any **Universal Storage Platform V** or external storage
- Lowers cache utilization through the use of performance-optimized disk-based journaling
- "Pull vs. Push": controls the replication processes from the target system thus improving performance of the production/primary system
- Maximizes the use of bandwidth by adapting to the available network resources and leveraging performance-optimized disk journals at the same time : better RPO control
- *Results in effective protection and significant optimization of storage resources for critical applications*

➔

➔

➔

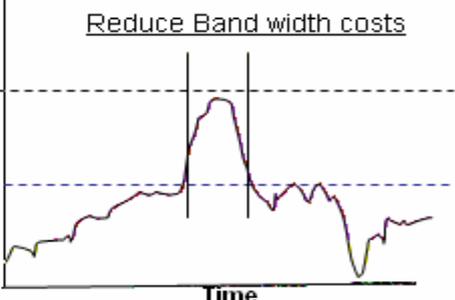
Reduced Cache utilization



Pull from Secondary Site

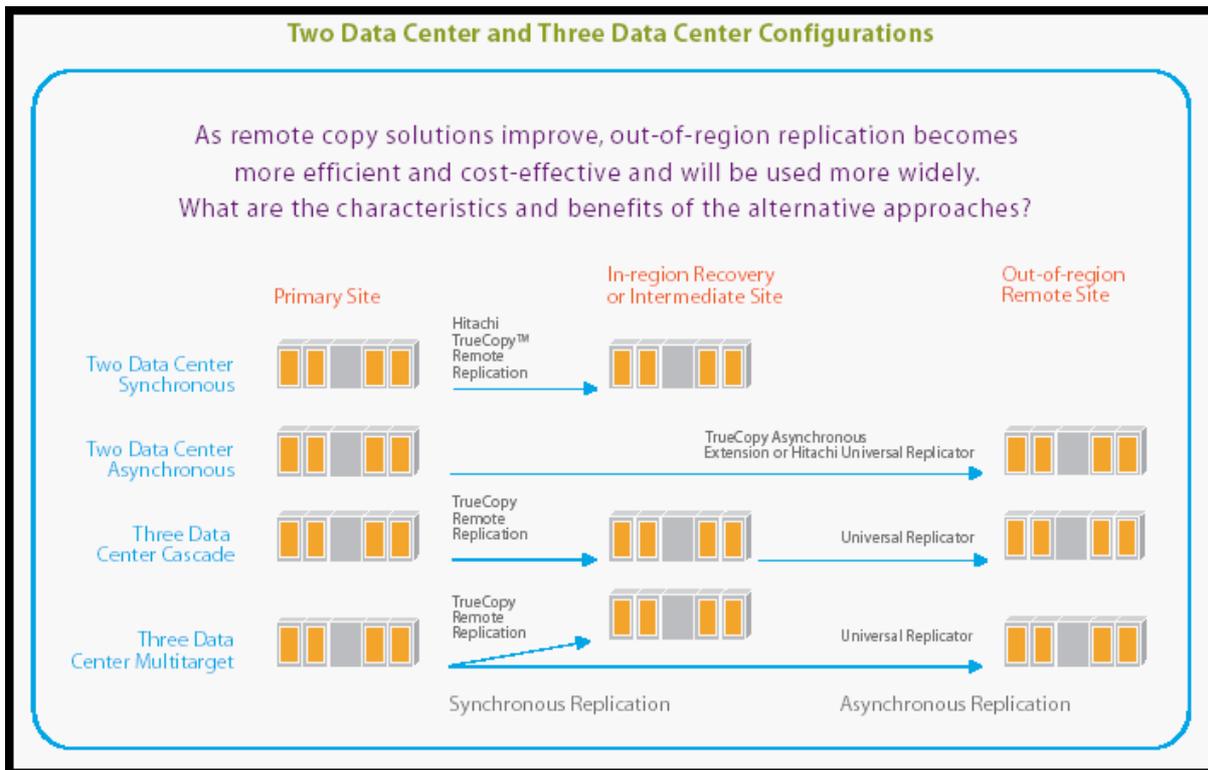


Reduce Band width costs



Time

10.6.5 Continuité de service avec HUR (Virtualisation du stockage)



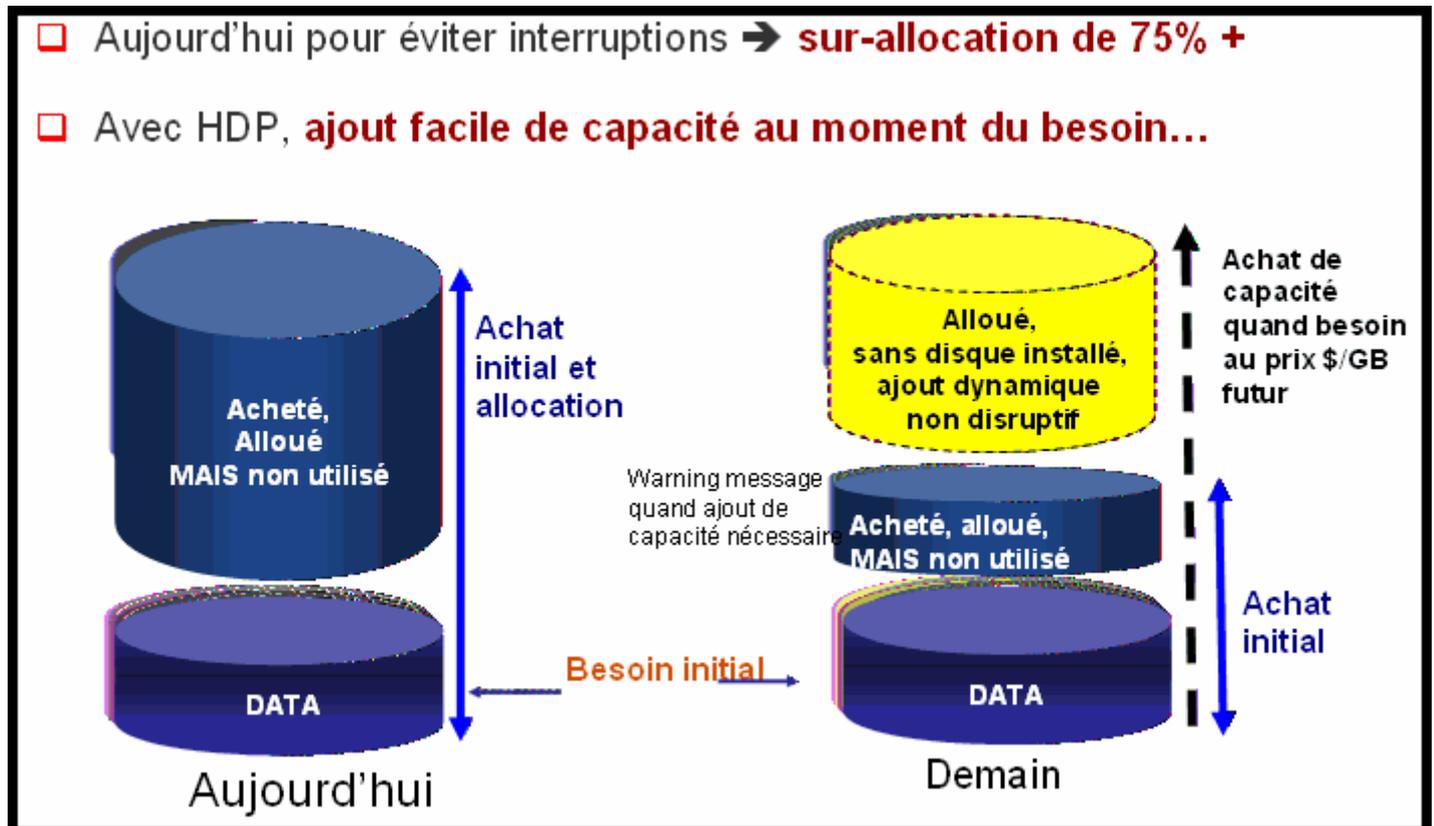
Continuité de services

- ▲ Sans perte de données, Sans multiplication de copies intermédiaires
- ▲ Indépendance / vitesse liens télécoms
- ▲ Processus simplifiés sur 2 à 3 sites

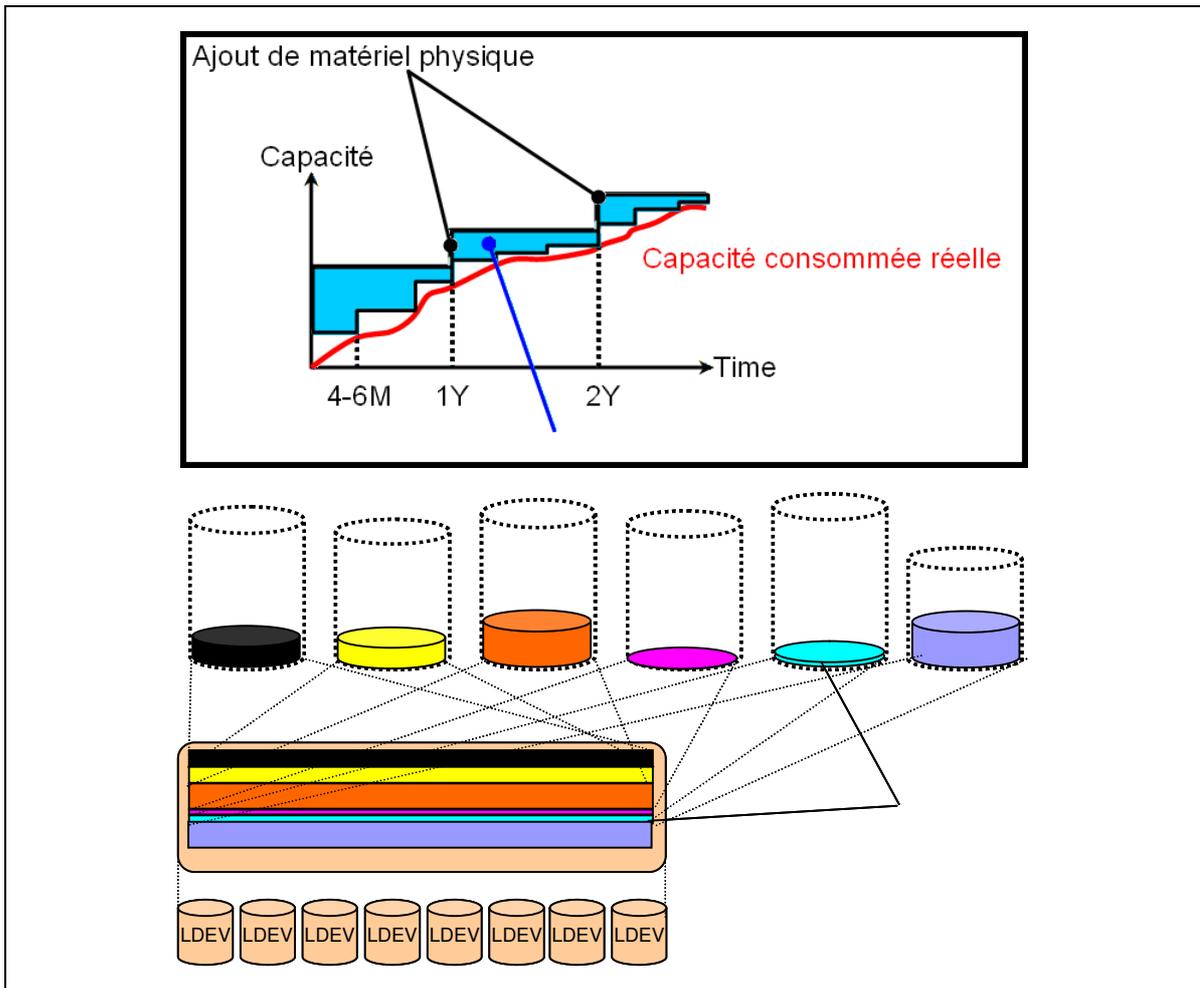
10.7 Technologie de « Thin provisioning »

En illustration de cette technologie, on prendra l'exemple d'HDS qui a été le premier à introduire cette technologie sur ses baies de stockage virtualisée USP.

10.7.1 Hitachi Dynamic Provisioning (HDP) sur USP V

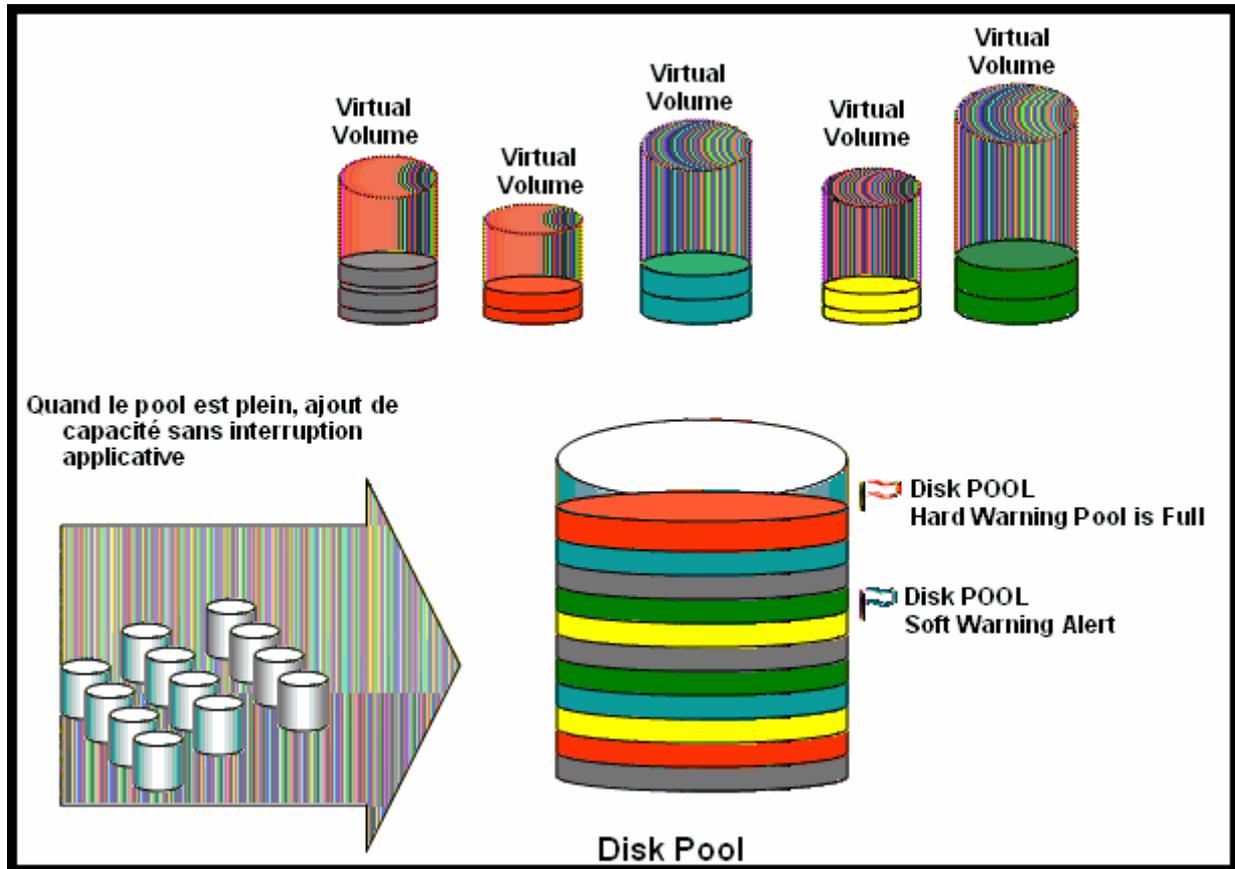


10.7.2 Hitachi Dynamic Provisioning (HDP) / Gestion d'approvisionnements de volumétrie



- HDP permet la gestion d'approvisionnements de volumétrie en flux tendu (« Thin provisioning »), permettant de lisser les besoins d'évolution de la volumétrie.
- Grâce à HDP, les serveurs (aux données en couleurs) disposent d'une volumétrie (virtuelle) qui anticipe leurs besoins, tandis que leurs données (réelles) sont stockées dans un Storage Pool mutualisé.

10.7.3 Hitachi Dynamic Provisioning (HDP) / – Seuils d’alerte



10.8 La Sauvegarde virtualisée / VTL / Technique de déduplication

VTL → Virtual Tape Library

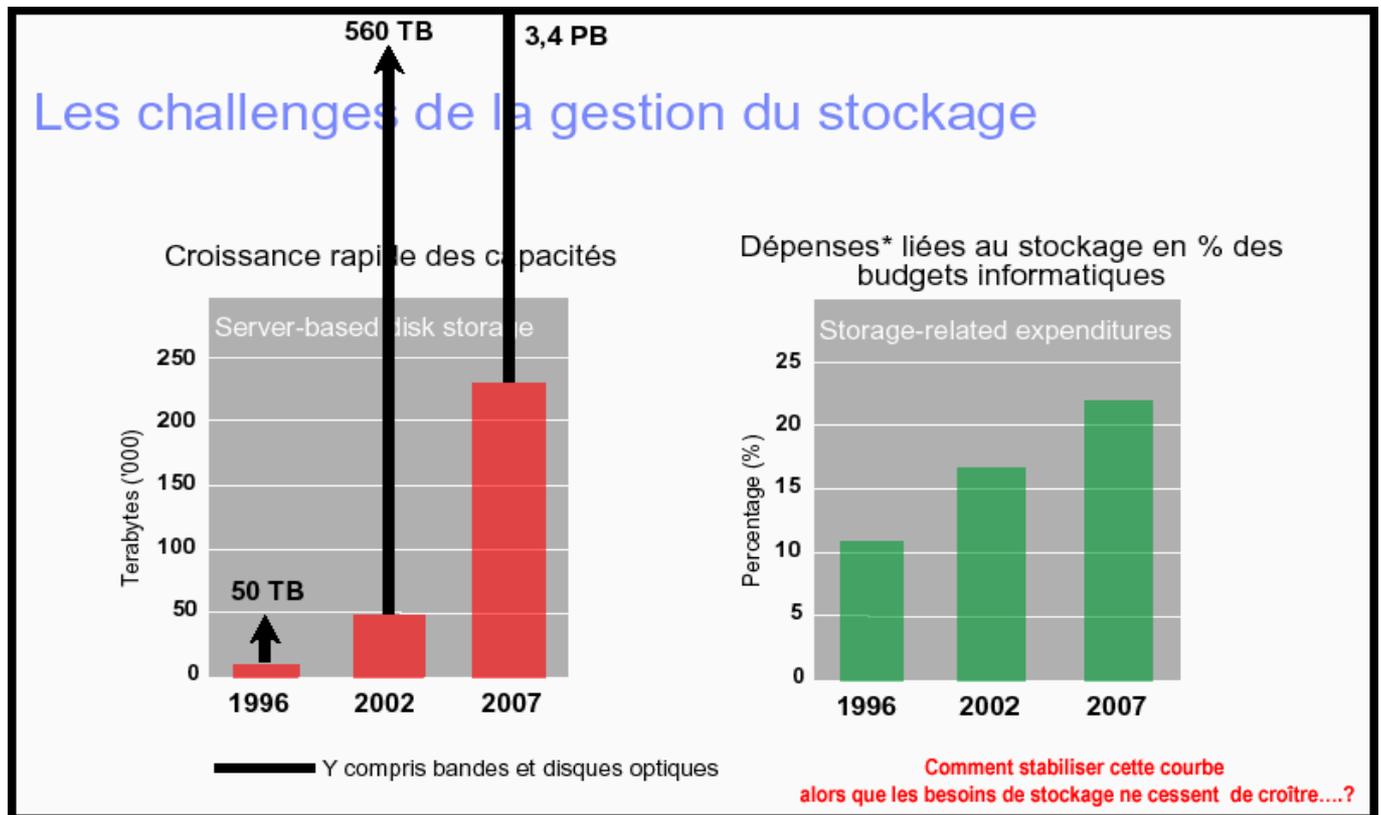
10.8.1 Data Center / La Sauvegarde

L'infrastructure de Stockage et sa Sauvegarde du Data Center sont le garant de la disponibilité et de la conservation des données, patrimoine de votre organisation.

C'est un des facteurs très important de la pérennité de votre Entreprise.

→ D'ici 2011, 60% des sauvegardes utiliseront des Appliances Disques / VTL (environ 30% à ce jour), et toutes ces VTL auront une option de Déduplication (Objectif : Réduire la capacité de stockage sur disques / « Green IT »).

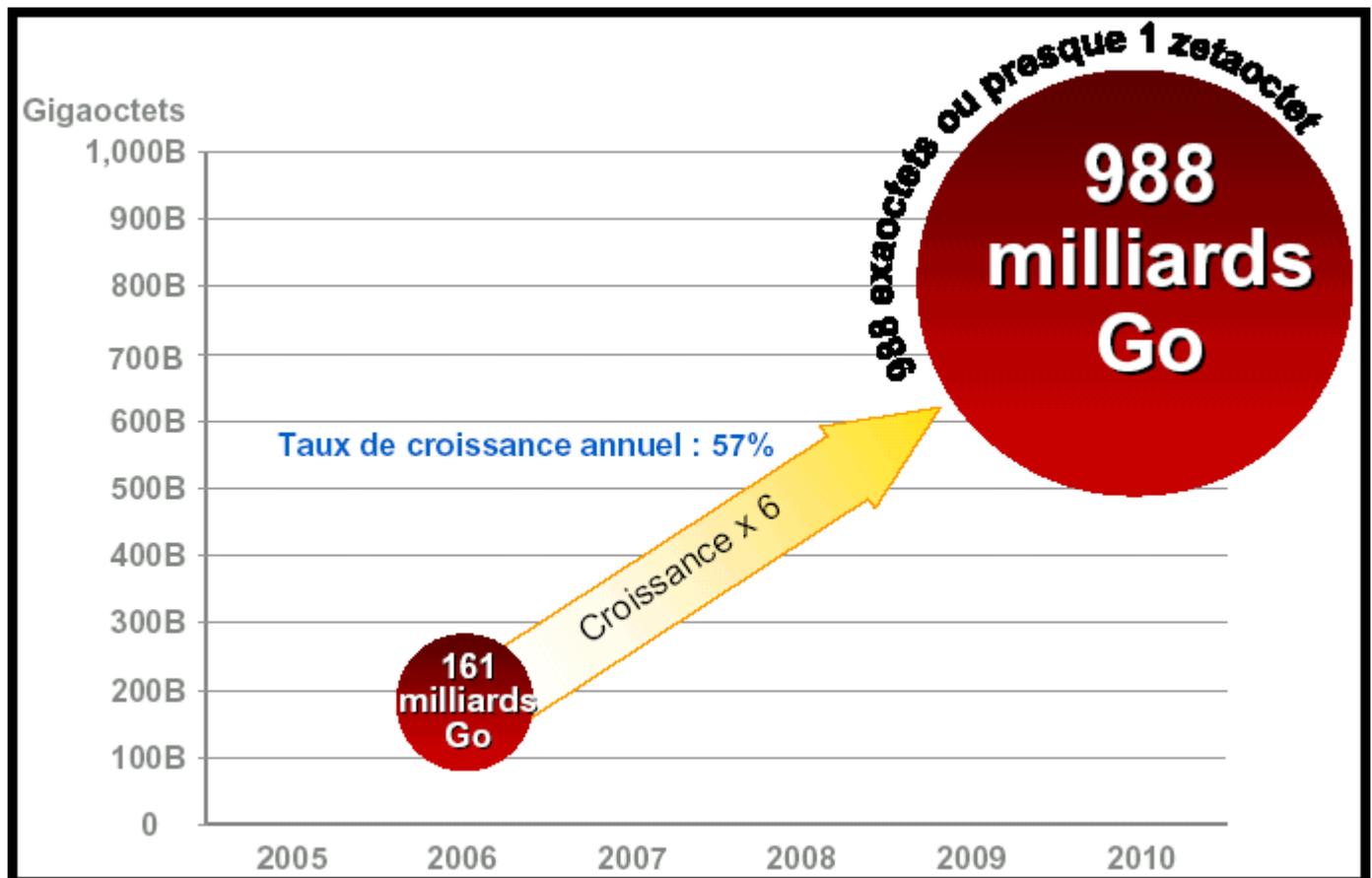
10.8.2 Les challenges de la gestion du Stockage et de sa Sauvegarde



Source : International Technology Group

(*) : Matériel, logiciel, SAN, personnel, opérations de sauvegardes, restaurations, sécurité...

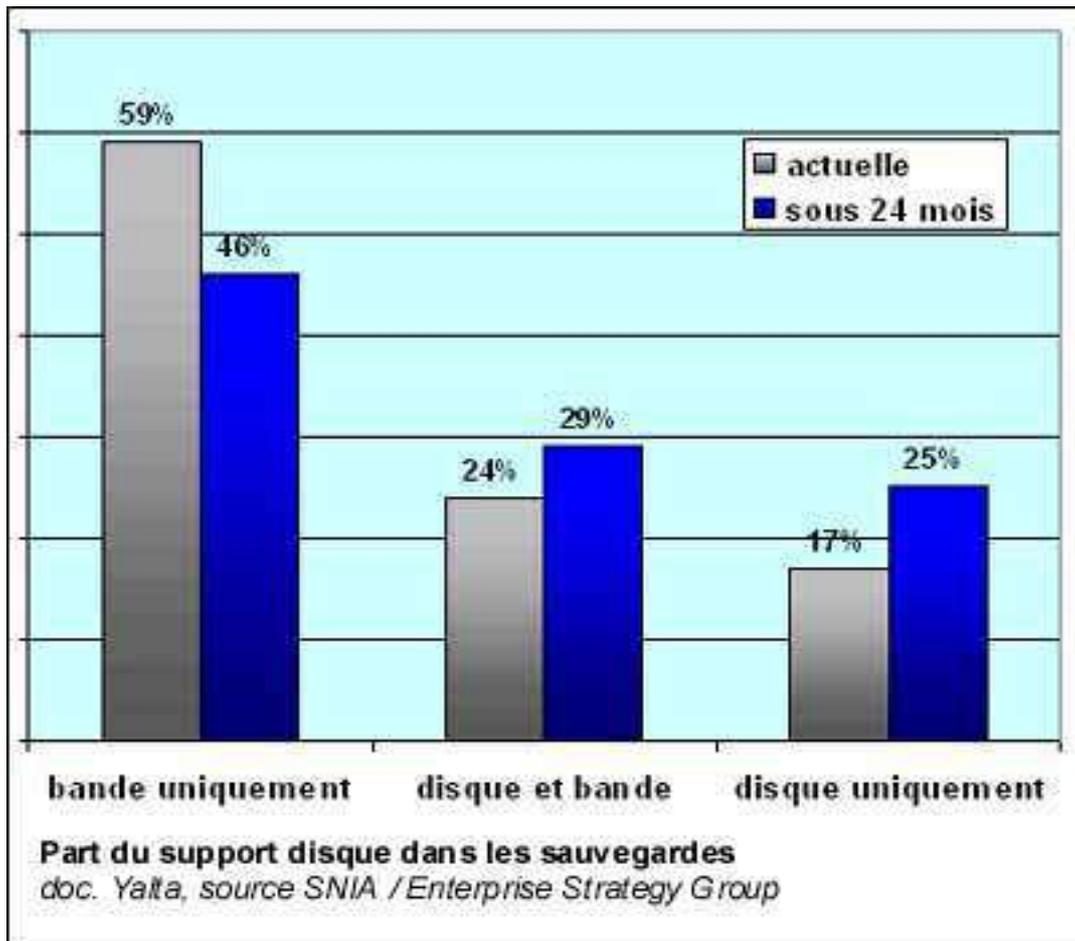
→ **Taux de croissance annuel : 56 à 57%**



10.8.3 Les acteurs majeurs (Constructeurs et OEM) VTL

Fournisseurs	Références	Descriptif
Data Domain datadomain.com	DD460 Restorer	Solution rack 5 U. Capacité physique de 6 To. Capacité virtuelle de 24 To à 210 To. Fonctions de compression et de déduplication intégrées.
EMC emcfrance.com	EDL 4200	Solution de sauvegarde et restauration sur base Clariion CX3. Intégration de Symantec NetBackup 5 et NetWorker 7.3 d'EMC pour la virtualisation de bandes.
Falconstor falconstor.com	iPStor	Solution logicielle pour plate-forme de virtualisation. Support de 32 sauvegardes simultanées sur SAN FC ou iSCSI. Gestion de 16 bibliothèques virtuelles. Support de 2 048 cartouches virtuelles.
Fujitsu Siemens fujitsu-siemens.fr	CentricStor VTC	Plate-forme VTL pour moyennes entreprises, deux lecteurs de bandes intégrés. 16 lecteurs virtuels extensible à 32. Cache système de 830 Go. Support des médias IBM 3490.
HDS hds.com	VT for Open Systems	Solution modulaire sur base de technologie Diligent. Logiciel de sauvegarde sur disque VTF Open. Logiciel de déduplication ProTec Tier optionnel.
HP hp.com	VLS 6000	Solution pour PME sur serveur ProLiant à partir de la technologie Sepaton. Capacité de 2,5 To, 2 ports Fibre Channel. Logiciel de gestion Command View.
IBM ibm.fr	Virtualization Engine TS7510	Solution permettant l'emploi de deux serveurs de virtualisation. 56 To de mémoire cache dont 46 To utilisables. Supporte 64 bibliothèques et 512 lecteurs virtuels.
Network Appliance www-fr.netapp.com	NearStore VTL 300	4 versions possibles de disques SATA (250 Go, 320 Go, 500 Go). Capacité maximum utile de 128 To sans compression. RAID 4 Netapp. De 2 à 8 interfaces host Fibre Channel 4 Gigabits. Support maximum de 10 000 bandes virtuelles.
Overland Storage overlandstorage.com	REO 1500	Solution de sauvegarde et de restauration sur disque. Logiciel REO Protection OS. Fonction pour la création de cartouches virtuelles de capacité variable. Interface web, module d'alerte par e-mail.
Quantum quantum.com/fr	DX 13500	Solution pour PME et sites déportés. Protection de 250 Go à 10 To de données brutes. Rétention sur disque jusqu'à 540 To. Débit de données de 80 Go/heure. Connectivité FC ou IP. Fonctions de réplication multisites via le WAN.
Sun StorageTek sun.com/storagetek	VTL 1140	Serveur modulaire à architecture RAID qui émule les principaux formats de bandes. Contrôleurs redondants. Débit de 400 Mo/s. Support de 2048 cartouches virtuelles. Capacité de 5,4 à 42 To utiles hors compression.

10.8.4 Le Marché du Disque Dur et de la Bande dans les Sauvegardes



10.8.5 Les chiffres du marché VTL :

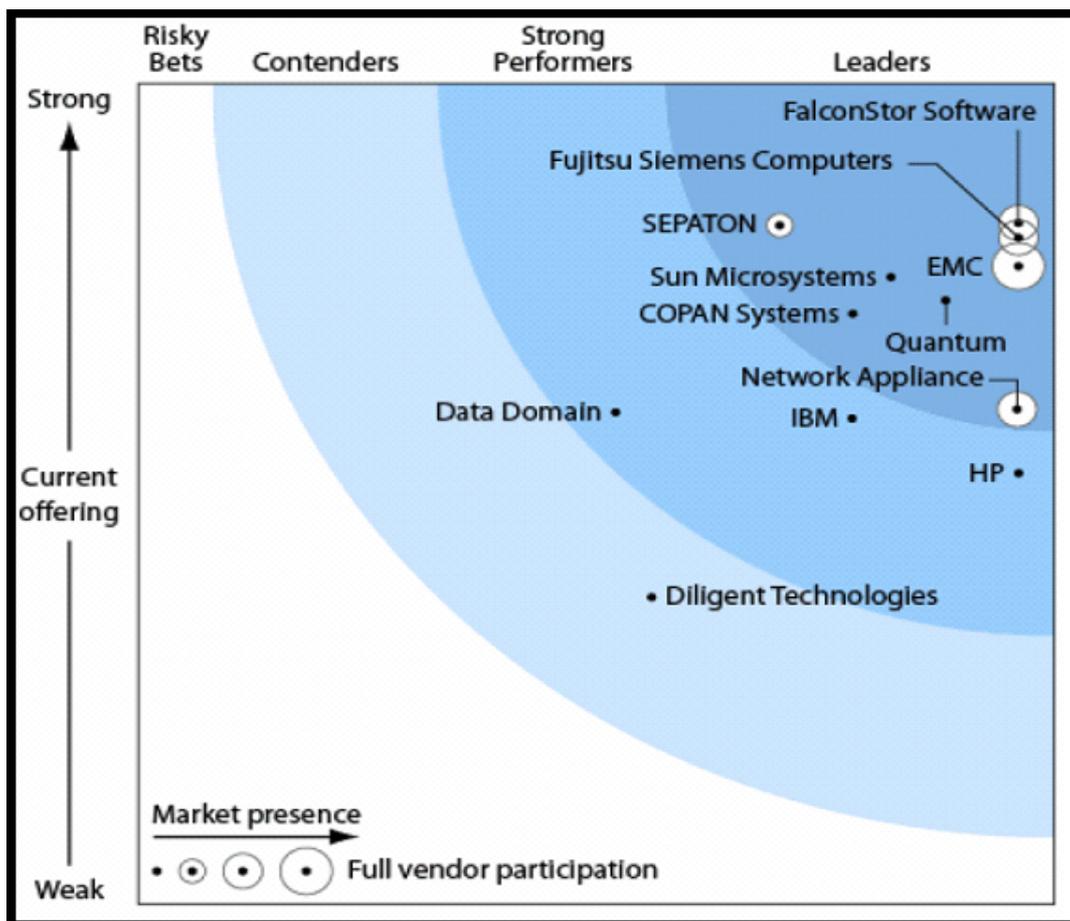
Le marché des librairies de bandes virtuelles (VTL) devrait doubler d'ici 2011 pour atteindre \$1,5 milliard

➔ **Le marché des systèmes VTL n'est pas uniquement celui du remplacement des librairies de bandes physiques.**

Les systèmes VTL seront également utilisés :

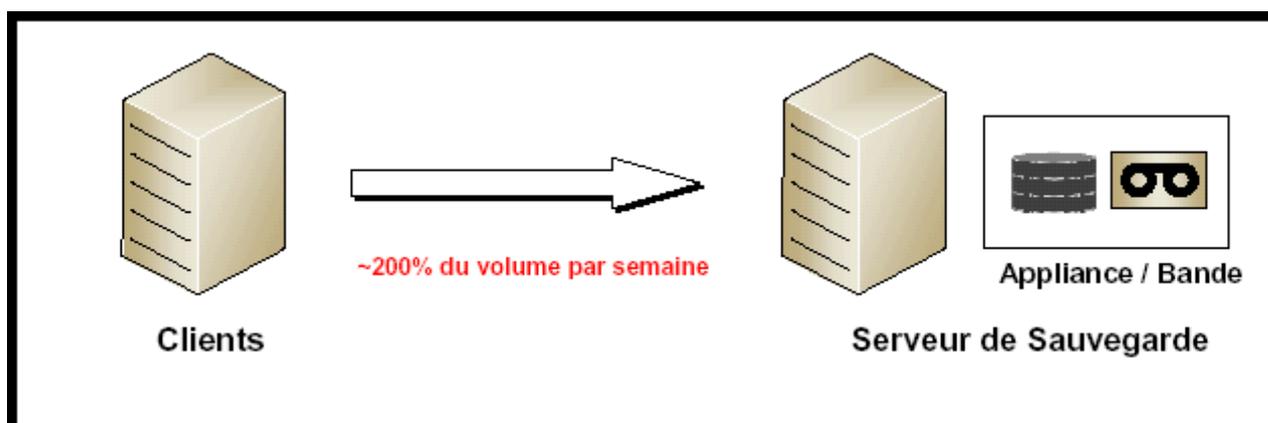
- pour fiabiliser la sauvegarde,
- pour diminuer (ou contenir) la fenêtre de sauvegarde,
- pour améliorer la fiabilité et les temps de restauration (accès aléatoires),
- pour optimiser l'usage de cartouches de plus en plus capacitives,
- pour diminuer l'exposition aux évolutions technologiques,
- pour prendre en compte les architectures de stockage multi-niveaux des clients...

10.8.6 Le Marché du VTL et de la Déduplication : « Carré magique »



10.8.7 Sauvegarde / Data Center

Les sauvegardes conduisent à transférer environ 200% de la capacité visible par semaine.



Constat :

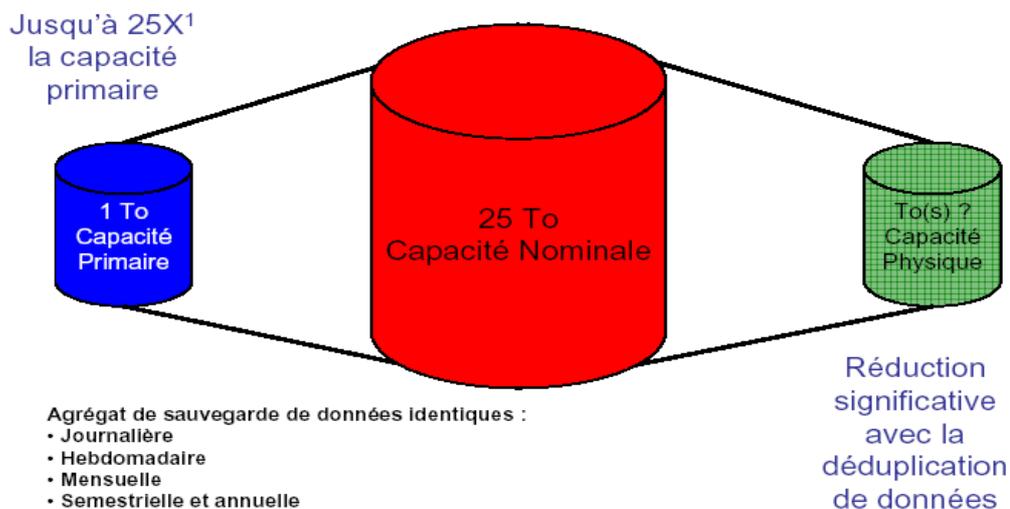
En sauvegarde réseau, en général, «chaque semaine, 70 à 95% du volume total de données sauvegardées est identique».

De plus, la **sauvegarde des messageries** génère beaucoup de **données identiques** car celles-ci ont beaucoup de messages pratiquement identiques, phénomène lié aux **multidestinatoires**, phénomène encore plus important pour des courriels comportant des **fichiers joints** que l'on retrouve sauvegardés en multiples exemplaires identiques...

En résumé et première approche :

➔ **Les données à sauvegarder sont « pratiquement » identiques à celles qui ont déjà été sauvegardées précédemment :**

- Données du type grand-père, père, fils, etc.,
- Sauvegardes « Totale » le week-end et « Incrémentales » la semaine,
- **Sur un an, chaque Téraoctet de disque primaire génère environ 25 TOctets sur bande,**
- Beaucoup de données similaires et dupliquées.



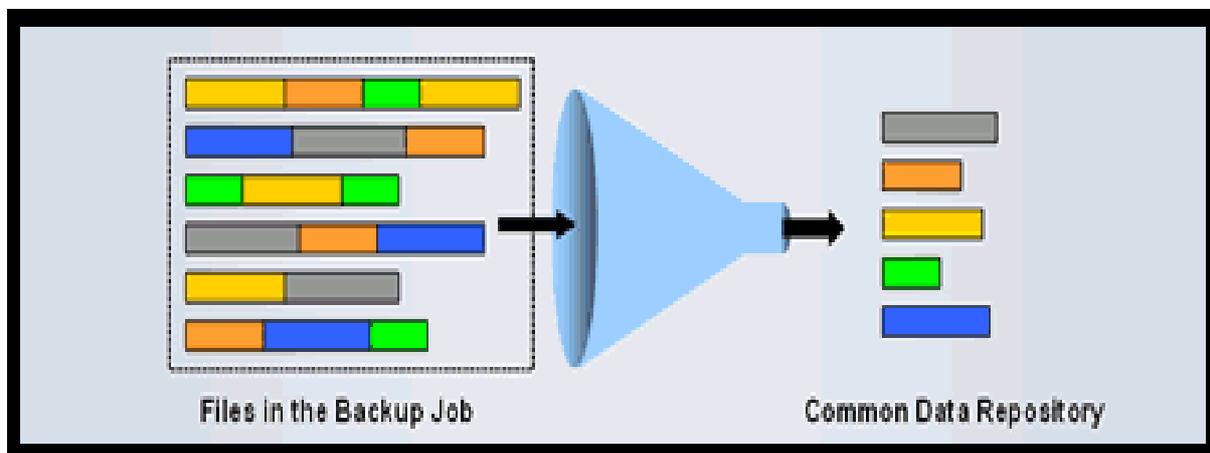
¹ Source: Strategic Research, *Strategic Profile: The New Metrics of Disk-based Data Protection*

(1) ➔ Sur un an, chaque Téraoctet de disque primaire génère environ 25 TOctets sur bande

D'où le développement de la technologie « Single-Instance-Store » ou de **Déduplication** principalement, dans un premier temps, en sauvegarde et en VTL :

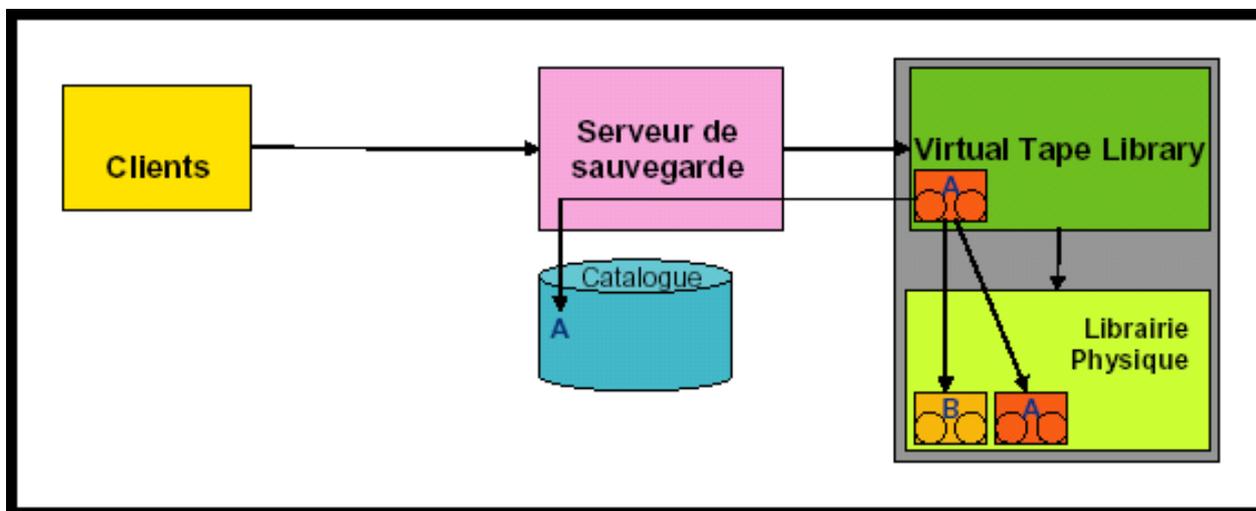
- Éliminer (ou éviter de générer en sauvegarde) la redondance des blocs identiques (mapping et segmentation des données).

La technologie de Déduplication / Vision simplifiée



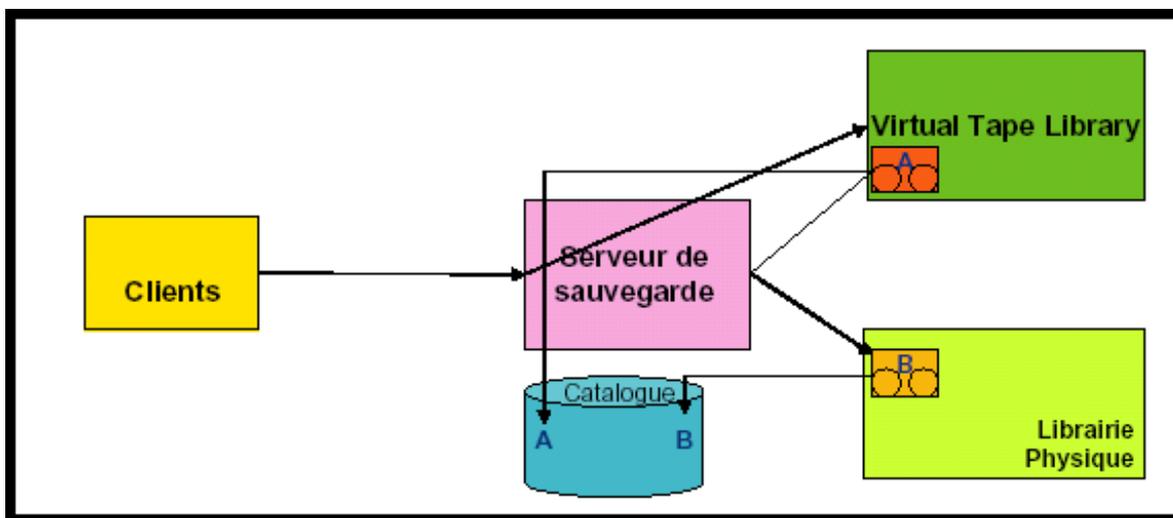
10.8.8 Architecture VTL (indépendance ou non / aux bibliothèques)

10.8.8.1 Architecture avec dépendance du type réel de la bibliothèque gérée



Architecture « type » gestion du stockage HSM (Hierarchical Storage Management)

10.8.8.2 Architecture avec indépendance du type réel de la bibliothèque « gérée »



→ Il y a indépendance totale entre la visibilité des « applicatifs de sauvegarde » avec les types des bibliothèques / bandes émulées via le SAN et indépendance totale avec les bibliothèques réelles connectées au VTL

10.8.9 La technologie de Déduplication (SIS, SIR, CAS)

SIS (Single-Instance Storage)

SIR (Single Instant Repository)

CAS (Content-Addressed Storage)

10.8.9.1 Introduction à la technologie de Déduplication (SIS, SIR, CAS)

C'est une technologie qui permet de **limiter la taille des données à transmettre** sur une liaison à distance ou **à écrire sur disques** par la **suppression des blocs / patterns / profils redondants** de données.

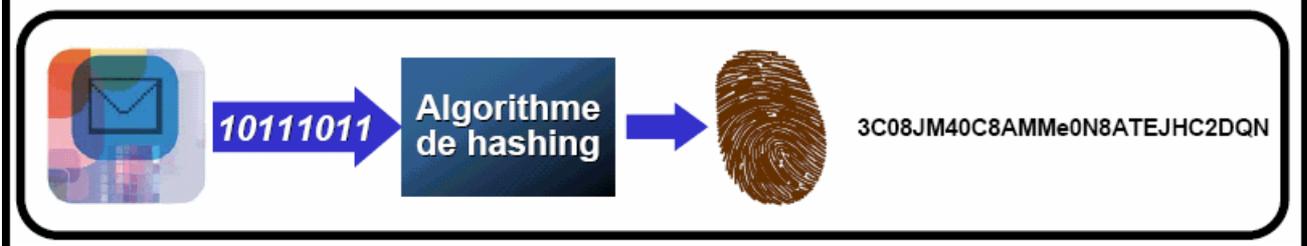
Elle vise à diminuer la capacité de stockage / sauvegarde sur disques surtout en sauvegarde en continu, en ILM et sur VTL...

Le concept de la déduplication est lié à la gestion d'un cache de données géré par le contenu et non un cache de données géré par les adresses comme le sont depuis très longtemps les caches d'Entrées / Sorties (RAID et disques) et les caches de code instruction CPU (Unité Centrale / Processeur).

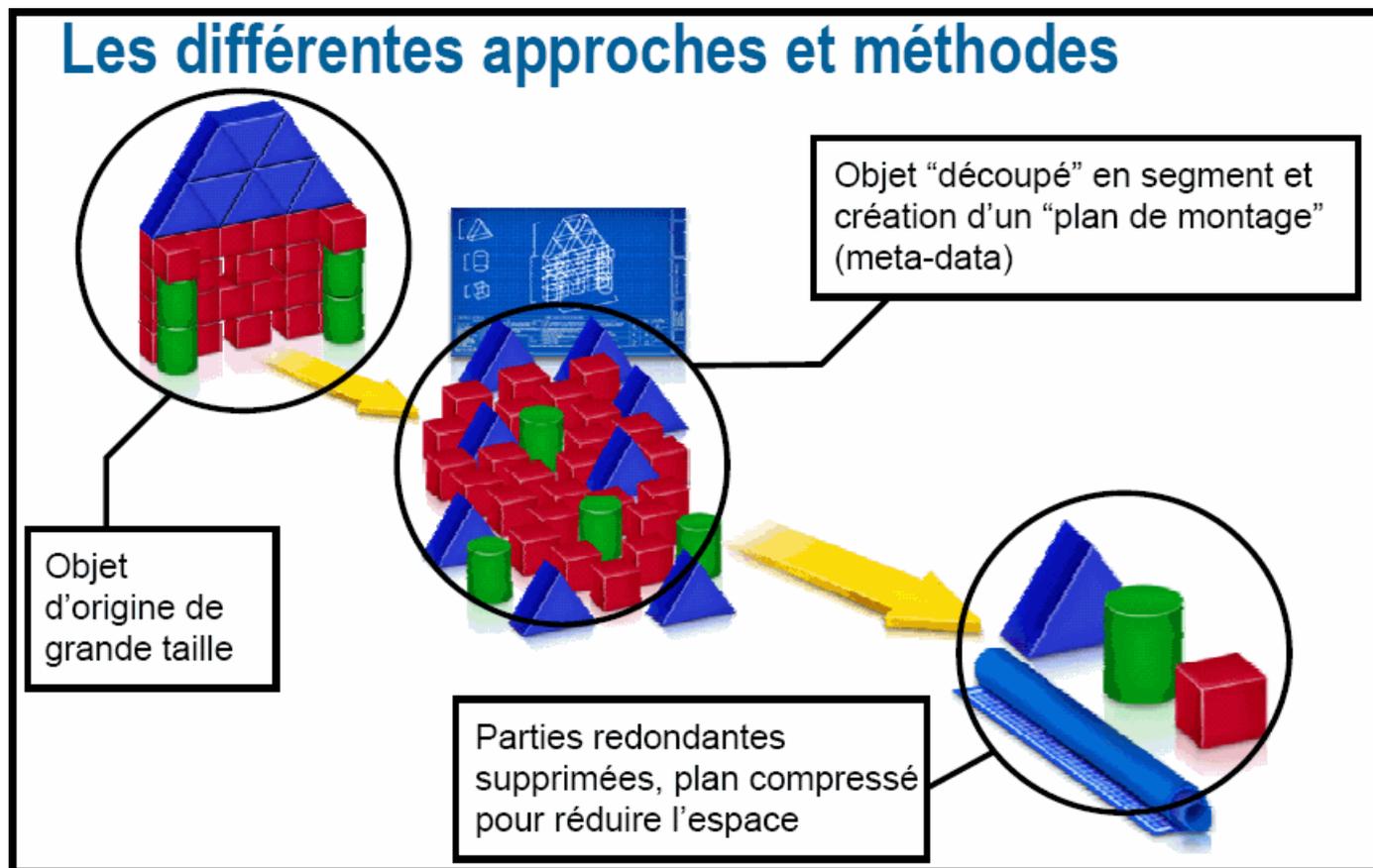
Déduplication, dé-doublonnage, factorisation, segmentation, **suppression de la redondance de profils** de données identiques... La technologie, qui consiste à factoriser des séquences de bits afin d'économiser drastiquement la consommation d'espace disque, emporte l'adhésion des acteurs du stockage et de la sauvegarde.

10.8.9.2 La technologie de Déduplication / Concept général

Dé-doublonner les données entre elles grâce au calcul de leur signature établie par un algorithme de *hashing* et par la comparaison de ces signatures pour garantir l'unicité du jeu de données à stocker



10.8.9.3 La technologie de Déduplication / Illustration (Phase 1)



10.8.9.4 La technologie de Déduplication / Illustration (Phase 2)

Les différentes approches et méthodes

Objet "découpé" en 4 segments, comme en première phase.

Plusieurs objets peuvent être ainsi assemblés avec les mêmes 4 pièces de base tant que le plan d'assemblage est conservé.

3 des 4 pièces sont déjà présentes (stockées) sur le support, seul l'ajout des nouvelles est nécessaire.

10.8.9.5 La technologie de Déduplication / « Hashing »

Name	Output size L (bits)	Performance (cycles/byte) Intel Xeon: C / assembly*	Collision chance 50% (or greater) when these many chunks (or more) are generated **	Chance of one collision in a 40 PB archive*** (using 4KB / chunk)	Year of the standard
MD5	128	9.4 / 3.7	$2^{64} \approx 10^{20}$	$0.5 * 10^{-20}$	1992
SHA-1	160	25 / 8.3	$2^{80} \approx 10^{24}$	$0.5 * 10^{-28}$	1995
SHA-256	256	39 / 20.6	$2^{128} \approx 10^{40}$	$0.5 * 10^{-60}$	2002
SHA-512	512	135 / 40.2	$2^{256} \approx 10^{80}$	$0.5 * 10^{-140}$	2002
Whirlpool	512	112 / 36.5	$2^{256} \approx 10^{80}$	$0.5 * 10^{-140}$	2003

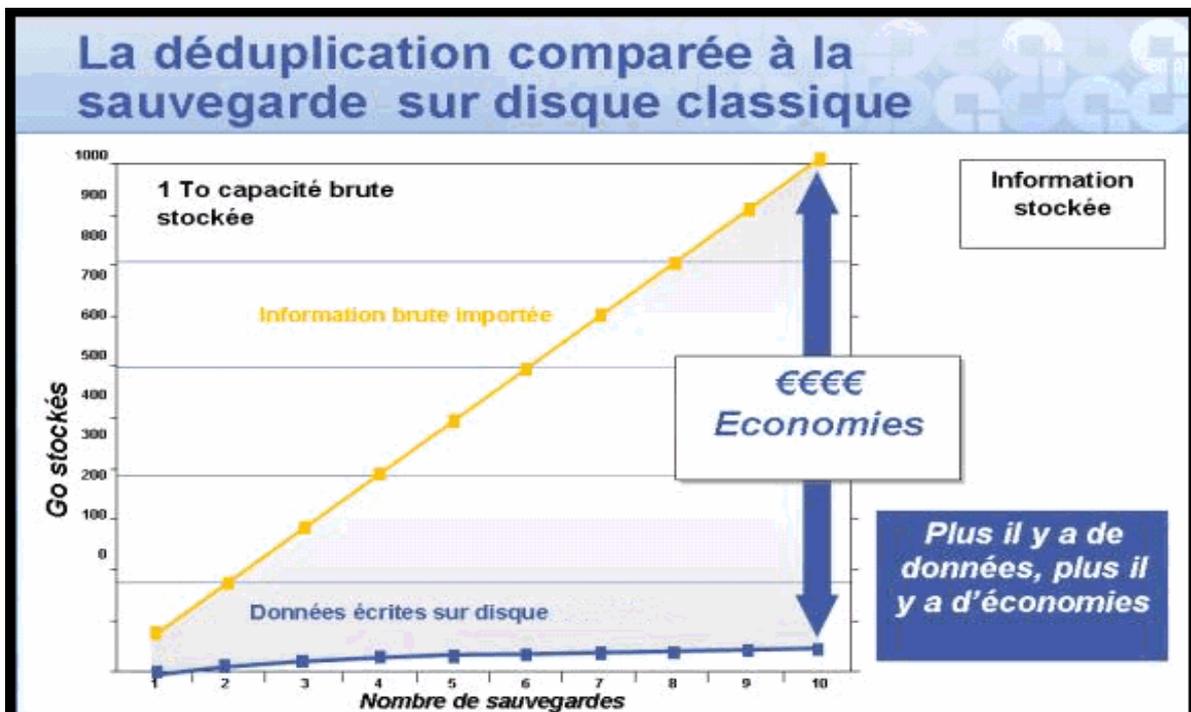
Ratio / Gain de la déduplication : 10 à 50 (REX en moyenne 20 à 25)

Remarques :

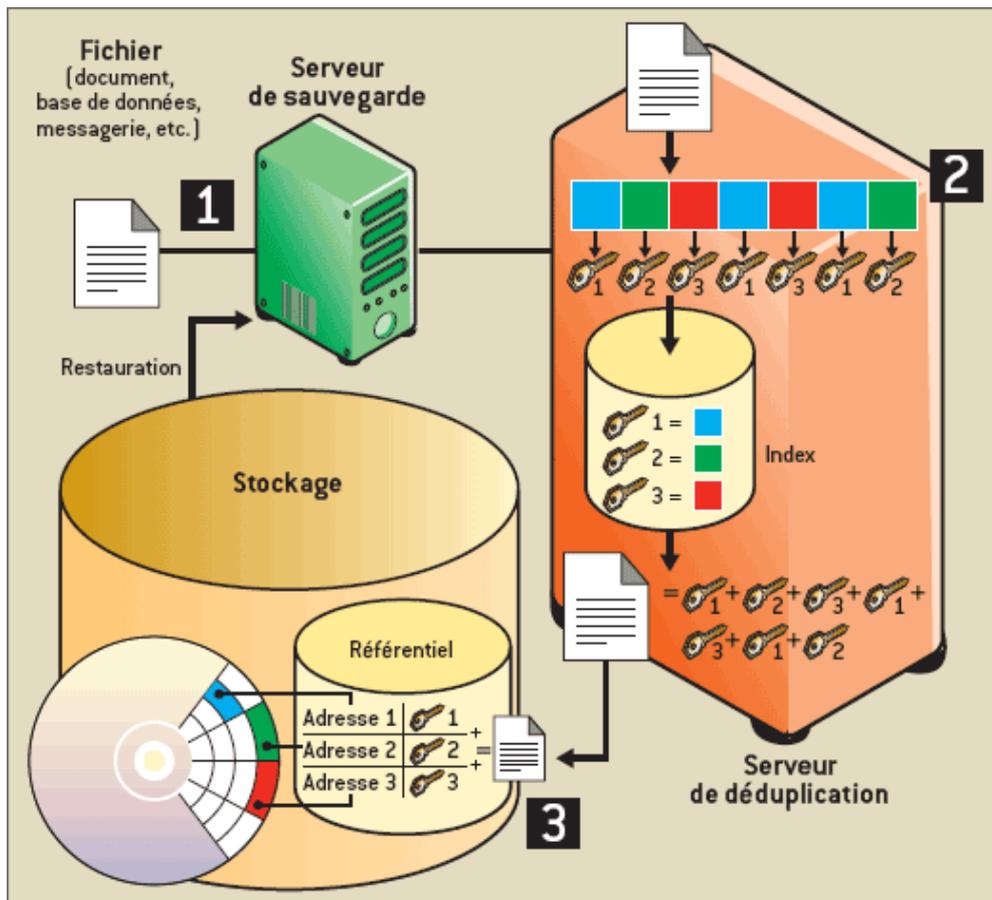
SHA-1 (Secure Hash Algorithm) est une fonction de **hachage cryptographique** conçue par la National Security Agency des États-Unis (NSA),

*** The probability of one hard-drive bit-error is about 10^{-14}

10.8.9.6 Déduplication / Sauvegarde sur disques

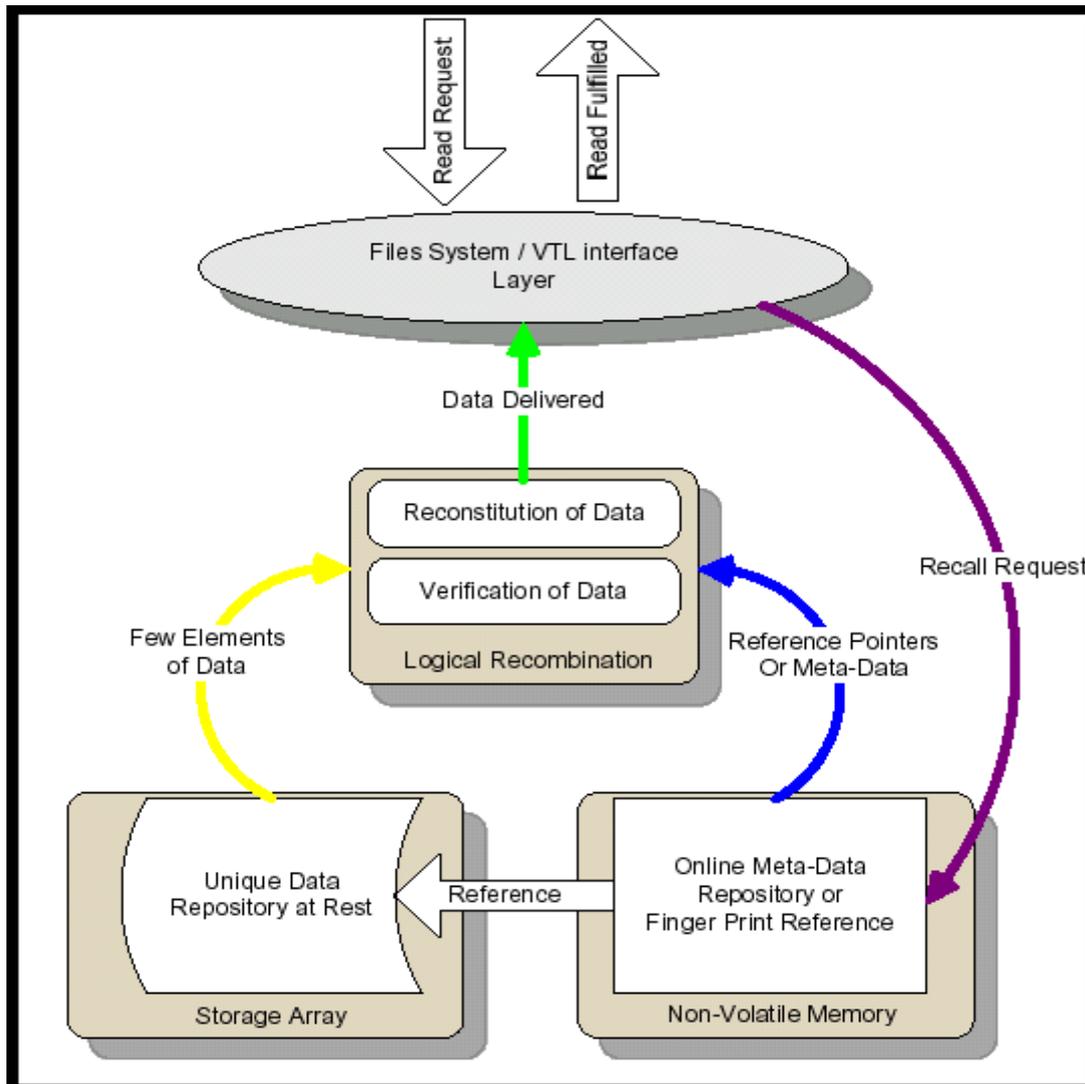


10.8.10 La technologie de Déduplication « Temps réel »

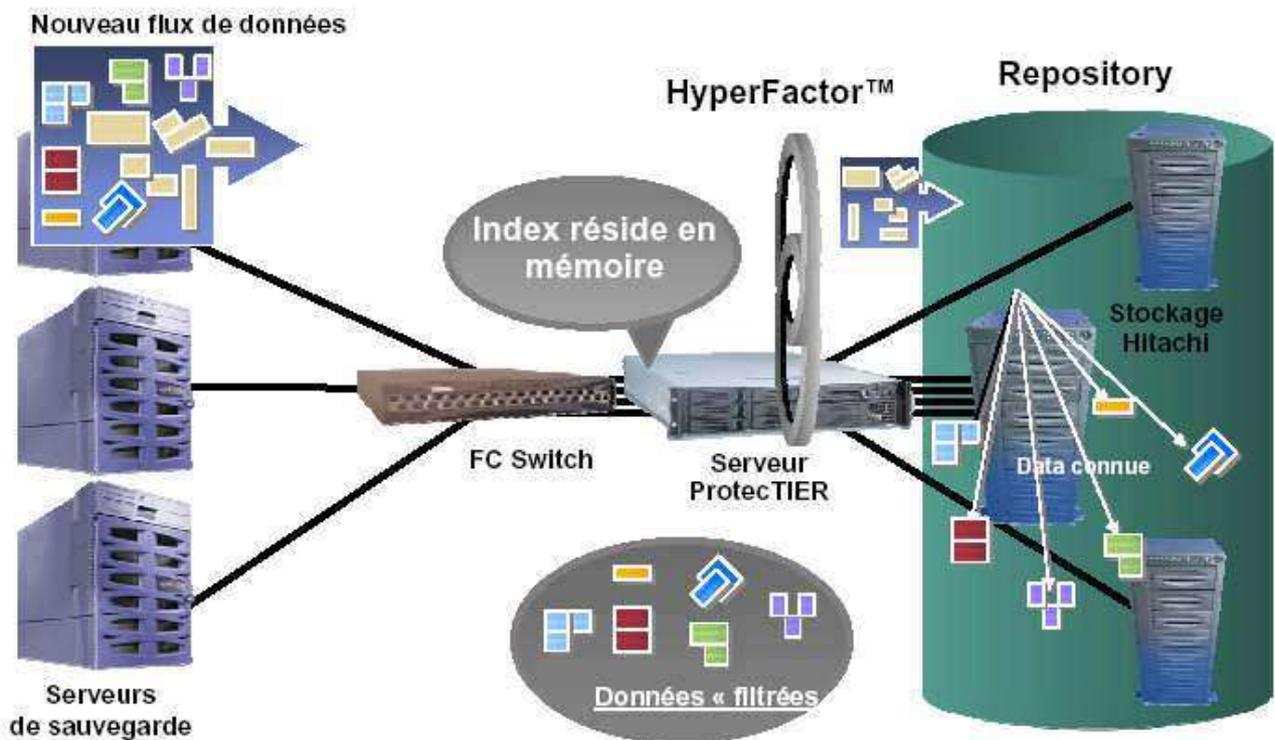


La technologie de déduplication est la gestion d'un cache du type « **Mémoire Associative sur le contenu** », cache géré par le contenu et non comme une antémémoire (cache instructions) d'UC gérée par les adresses d'instructions ou comme un cache d'E/S géré par les adresses des données secteurs.

10.8.11 Technologie de Déduplication / Reconstitution des données

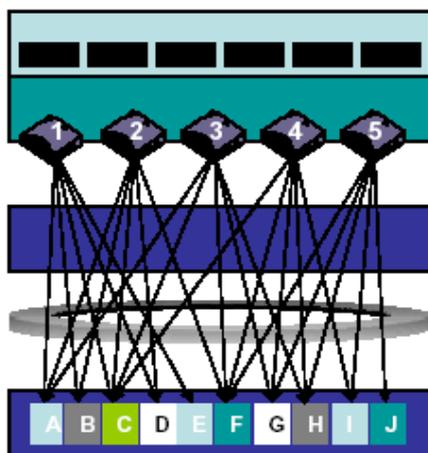


10.8.12 **Exemple d'implémentation : Bloc diagramme (Diligent Technologies / IBM)**



L'illustration ci-dessus montre le mapping des segments par rapport aux données redondantes initiales

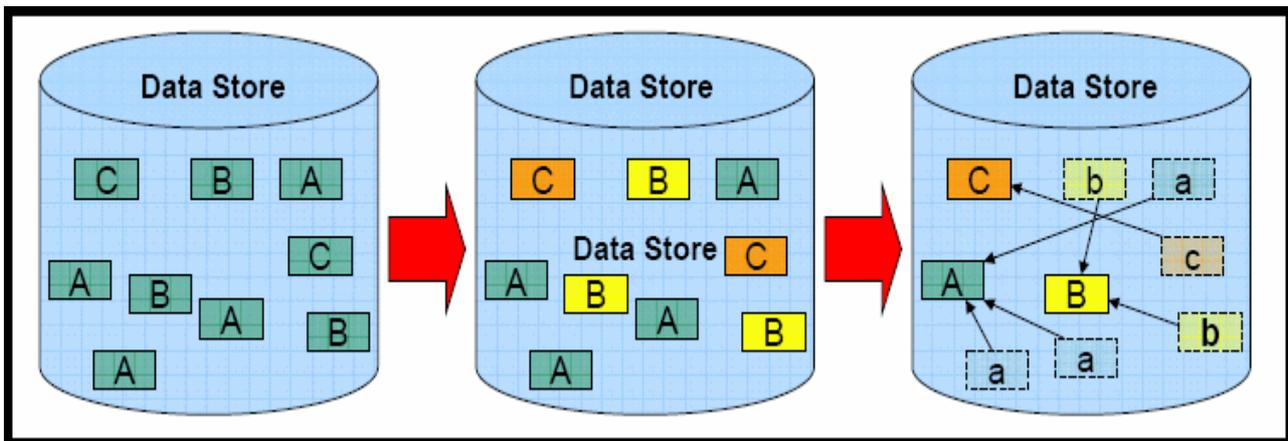
Bloc diagramme « suite » (Diligent Technologies / IBM)



- Les données sont enregistrées dans ProtecTIER via une bande virtuelle
- Les données traverses HyperFactor et seules les nouvelles données sont enregistrées
- Les données déjà présentes sont uniquement référencées
- En cas de réécriture les références concernées sont supprimées
- L'espace disque est remis à disposition quand le nombre de référence est descendue à zéro

Technologie : Ne sauvegarder sur disques qu'une seule fois les patterns de données identiques (segments), les données originales sont reconstituées par les pointeurs qui permettent de retrouver les patterns identiques qui n'ont été stockés sur disques qu'une seule fois (cache de contenu).

10.8.13 La technologie de Déduplication / Résumé



Une signature est calculée pour tous les blocs des données (qui sont segmentées)

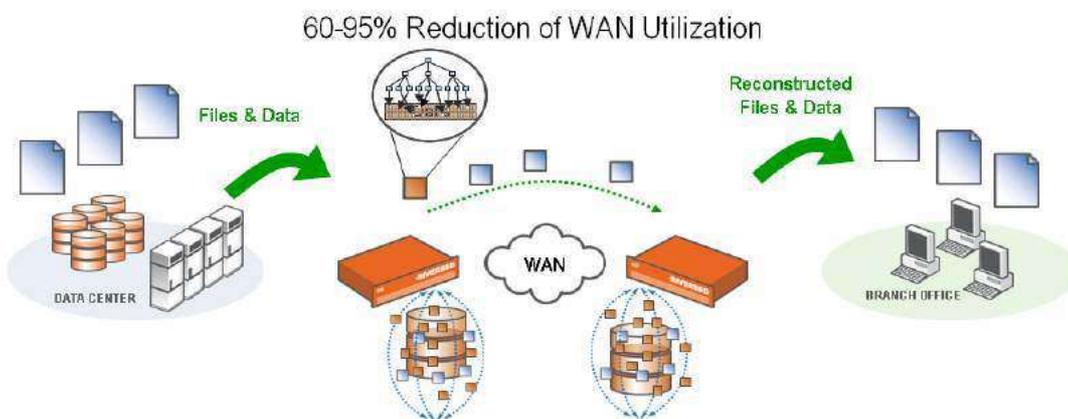
La valeur des signatures calculées est utilisée pour indexer la table des signatures déjà mémorisées afin d'identifier les blocs de données identiques et de déterminer éventuellement les collisions sur les signatures

Après avoir traité le problème des collisions éventuelles, les strings de données sont remplacés par des pointeurs afin de conserver sur disques une seule occurrence des blocs des données trouvés identiques

10.8.14 La Technologie de Déduplication en TELCO's

A propos des liens WAN, RIVERBED annonce une "Diminution de l'utilisation de la bande passante de l'ordre de 80%" par l'utilisation de ses Appliances « Steelhead », Appliances qui utilisent également la technologie de déduplication.

Riverbed Optimization System



10.8.15 Les Acteurs majeurs avec « De-Dupe » (Constructeurs / Éditeurs et OEM)

Constructeur	Produit	Commentaire	Tech / Hashing pour indexation de Déduplication	Tech / Content Aware (File Name, Data Format, Structure, Version...)	Tech / Hyper Factor (Déduplication temps réel)
Avamar Technologies http://www.avamar.com/ → http://software.emc.com/	Avamar a été rachetée par EMC	Avamar Technologies has been renamed and is now offered by EMC		X (Voir EMC)	
COPAN Systems http://www.copansystems.com/	Revolution 300 VTL Series	Single Instance Repository for MAID (SIR-M)	FalconStor Software (OEM)		
Data Domain http://www.datadomain.com/	DD 690 1,4 TB/H	« Online Reference Data »	In-band / Hash en temps réel « Inline Deduplication »		
EMC http://france.emc.com/	DL 1500, DL 3000 et DL 4000	Avec Déduplication « Avamar Data Store » / Déduplication globale à la source		X (Tech Avamar)	
EMC http://france.emc.com/	EMC Disk Library 6000 Series	New version of its de-duplication software from Avamar		X (Tech Avamar)	
FalconStor http://www.falconstor.com/	IPStor	FalconStor VTL Enterprise	Post-processing		
FSC / Fujitsu Siemens Computers http://www.fujitsu-siemens.fr/	CentricStor (VTC)	Architecture Grid avec option Déduplication 1,6 TB/H			In-band Over 450 MB/sec (Tech Diligent)

Constructeur	Produit	Commentaire	Tech / Hashing pour indexation de Déduplication	Tech / Content Aware (File Name, Data Format, Structure, Version...)	Tech / Hyper Factor (Déduplication temps réel)
HDS http://www.hds.com/	Virtual Tape Library Appliance Model 1000E et 1000L	Option Déduplication "Diligent ProtecTIER" 1,6 TB/H			In-band Over 450 MB/sec (Tech Diligent)
HP http://welcome.hp.com/cont/fr/fr/	VLS 6000	« de-duplication available early 2008 »		X (Tech Sepaton)	
IBM http://www.ibm.com/fr/fr/	TSM	Sauvegarde incrémentale et mode VTL	Déduplication en tâche de fond. → Vers fin 2008		Ne pas oublier le rachat de Diligent par IBM !
NetApp http://www.netapp.com/fr/	FAS6000 (NAS) NearStore VTL	Déduplication « Data ONTAP »	Déduplication en tâche de fond		
Overland http://www.overlandstorage.com/	REO 9500D	Deduplicating Virtual Tape Library appliance			In-band 200 MB/sec (Tech Diligent)
Quantum http://www.quantum.com/	DXi7500 Disk-based	VTL avec option de déduplication : "Up to 8 TB per hour" hors déduplication	Déduplication en tâche de fond		
SEPATON www.sepaton.com	S2100®-ES2 VTL et DeltaStor Déduplication	VTL avec Déduplication « Content Aware » Architecture Grid		In-band (Traitement temps réel) 34,5 TB/Heure	
SUN (& STK) http://www.sun.com/	Sun StorageTek Virtual Tape Library Plus ("De-Dup appliance soon")	X4500 VTL with Diligent's ProtecTIER			In-band Over 450 MB/sec (Tech Diligent)
SYMANTEC / VERITAS	Veritas NetBackup PureDisk			Déduplication globale des données à la source	

11 Concept de la virtualisation du poste de travail

11.1 Le « Bureau Virtuel »

Vous pouvez gérer facilement vos données (e-mails, contacts, rendez-vous, documents, tâches, ...) dans votre bureau virtuel, accessible dans un simple navigateur Web à partir de tout ordinateur connecté à Internet ou d'un client léger au sein de votre entreprise.

Par exemple, les Technologies Terminal Server de Microsoft et « Presentation Server » de CITRIX permettent d'implémenter les bureaux virtuels.

11.2 Le « Streaming Applicatif »

Le streaming applicatif permet de déployer sur des parcs informatiques des applications sans les installer. A l'exécution de chaque application, une bulle applicative virtuelle est montée sur le système d'exploitation. L'application exécutée dans cette bulle est isolée des autres applications et fonctionne de manière indépendante.

Les technologies principales utilisées sont « Presentation Server » de CITRIX et « SoftGrid » de Microsoft.

11.3 La Virtualisation du Poste de Travail

11.3.1 Introduction

Le recours à une infrastructure de virtualisation du poste de travail permet de mettre en œuvre une infrastructure « client léger » qui optimise l'utilisation, l'administration, le coût total de propriété et la souplesse.

Des environnements de bureau complets peuvent être exécutés dans des machines virtuelles sur des serveurs de centre de données et les utilisateurs finaux peuvent y accéder depuis tout PC ou client léger connecté au réseau d'entreprise.

Cette solution permet au service informatique de contrôler de façon centralisée les ressources informatiques des postes de travail et leurs données, de consolider des machines virtuelles et d'optimiser l'utilisation des ressources dans le centre de données.

Les utilisateurs ont ainsi la possibilité d'accéder à leur environnement de travail complet depuis n'importe quel site et n'importe quel client.

11.3.2 Présentation

Objectif : Dématérialiser le poste de travail avec uniformisation du matériel (universalité du modèle virtuel).

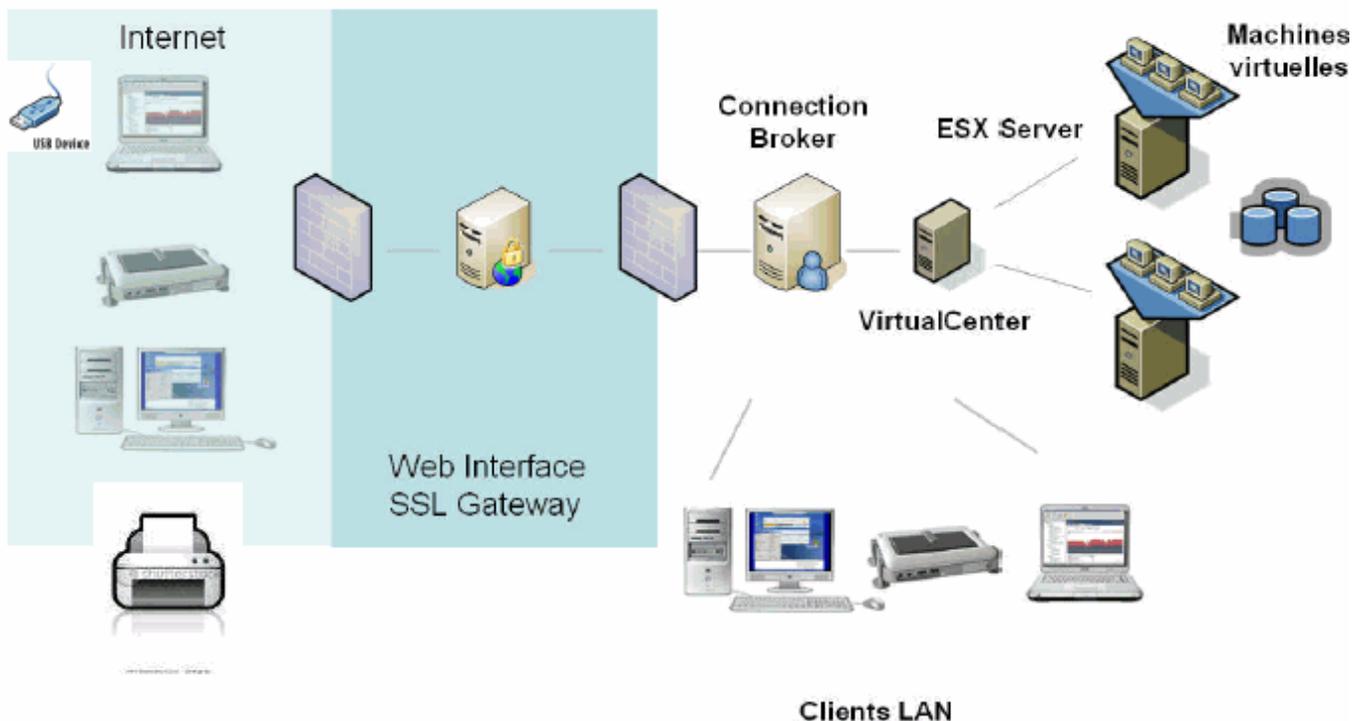
Les enjeux principaux des postes de travail sont :

- Déploiement
- Mobilité
- « Remote Office » / « Branch Office »
- Sécurité
- Périphériques et multimédia
- Green

11.3.3 Concept VDI (« voix, données, images » Câblage et connectique normalisée) :

- Connexion « un à un » entre le poste de travail et la machine virtuelle en session de bureau à distance,
- Les connexions s'effectuent à travers le réseau local d'entreprise,
- La session de bureau à distance est établie en connexion RDP (« **Remote Desktop Protocol** » est un protocole qui permet à un utilisateur de se connecter sur un ordinateur faisant tourner Microsoft Terminal Services. Des clients existent pour la quasi-totalité des versions de Windows, et pour d'autres systèmes d'exploitation, comme Linux.) depuis un poste de travail ou un terminal léger.

11.3.4 Architecture VDI

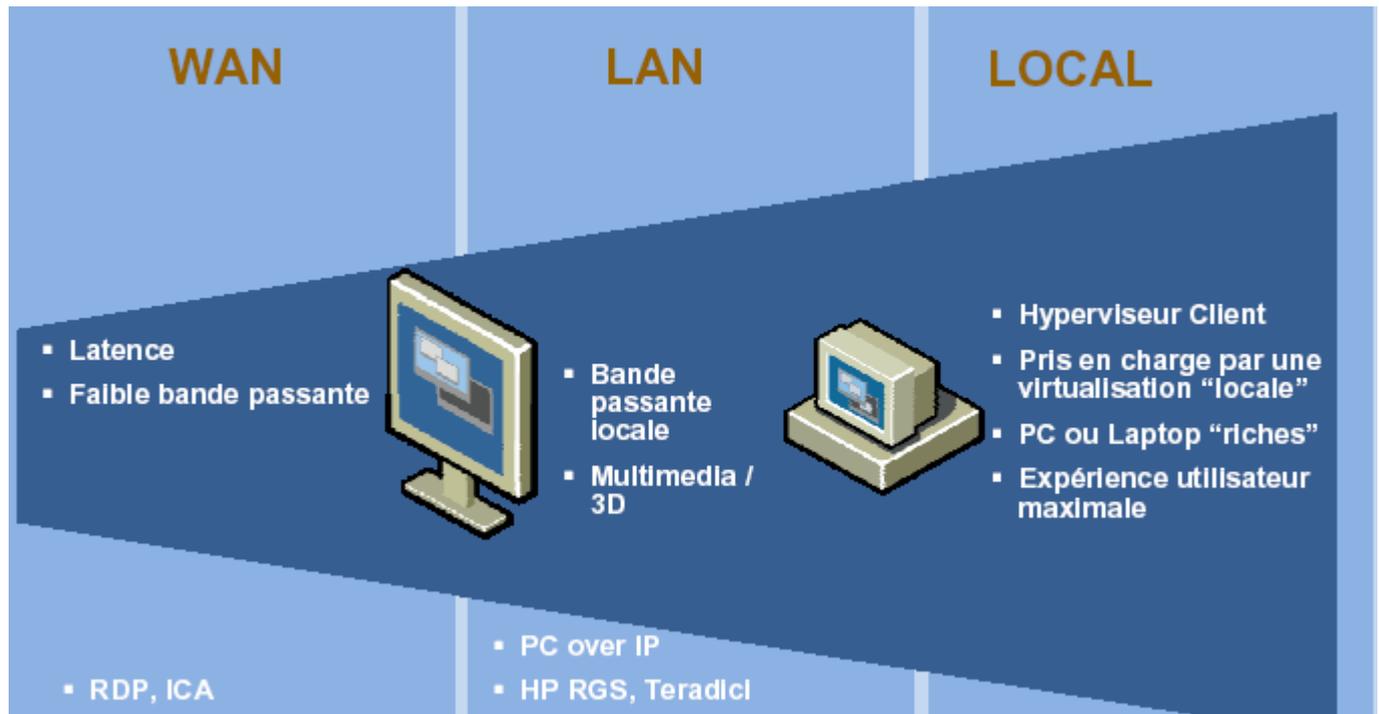


Remarque :

L'Appliance « Connection Broker » : Le "Connection Broker" gère tous les types de postes clients. Il peut ainsi allouer à un utilisateur un PC virtuel, une session applicative avec déport d'affichage ou un PC Lame. Le « Connection Broker » est le « service » qui est chargé de fournir à un utilisateur une connexion vers une machine virtuelle cliente. Il peut aussi, selon les éditeurs, gérer et provisionner ces mêmes machines virtuelles.

Exemple avec VMware : L'utilisateur se connecte au « Connection Broker » qui lui alloue une connexion vers une machine virtuelle hébergée sur la ferme ESX à laquelle est associé le « Connection Broker ». Exemples de « Connection Brokers » : VMware VDM 2, Leostream, Provision Networks).

11.3.5 Expérience Utilisateur / Architecture WAN, LAN et Local



11.3.6 Virtualisation du Poste de Travail : Principales technologies

Technologie	Description	Acteurs
Client léger Windows	Applications exécutées sur un serveur, déport de l'interface utilisateur sur le PC.	Microsoft, Citrix (et fabricants de clients légers).
Virtualisation d'applications	Applications exécutées sur le PC mais isolées entre elles et du système. Distribution de l'application en mode streaming par un serveur.	Microsoft, Citrix, Altiris.
Virtualisation du PC	Virtualisation de l'environnement système de chaque PC, sur un serveur. Déport de l'interface utilisateur sur le PC physique.	VMware
Streaming de système d'exploitation	Mise à disposition d'un système d'exploitation, auprès des PC, par un serveur.	Ardence (racheté par Citrix).
PC lame	Centralisation physique de chaque PC sur un serveur lame. Déport de l'interface utilisateur sur un client léger.	ClearCube et HP

12 Concept de la virtualisation applicative

12.1 Généralités de la virtualisation applicative

Le système d'exploitation du poste de travail / La virtualisation du poste de travail :

- La solution XEN au boot,
- La virtualisation des applications,
- La virtualisation du PDA et du téléphone : une bonne idée pour les applications.

Le principe de la virtualisation applicative consiste à générer un exécutable à partir du programme d'installation (par exemple Word). Celui-ci se lance alors physiquement en local, à partir d'une clé USB par exemple, (et non pas depuis un autre point du réseau) sans que l'on installe quoi que soit.



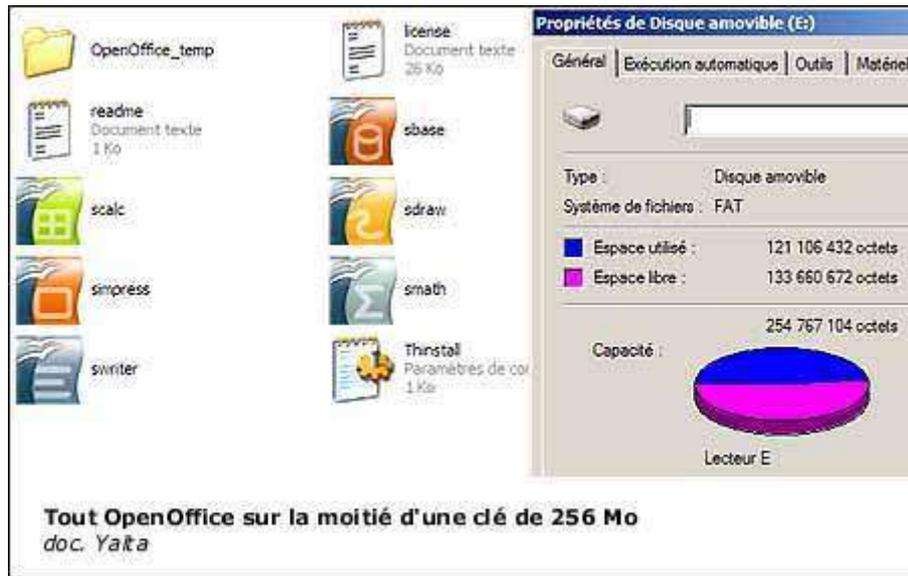
SoftGrid, Altiris, ThinStall, CITRIX proposent différentes variantes de ce principe.

12.2 Principe de la virtualisation applicative

La totalité des programmes et fichiers nécessaires se trouvent sur un espace de stockage quelconque (clé USB, mais aussi CD, email, unité disque locale ou réseau).

Le logiciel de virtualisation se positionne entre le système et l'application et fournit à cette dernière la totalité des ressources d'environnement qui lui sont nécessaires pour fonctionner, sans qu'elle n'accède

jamais au système réel (donc sans impacter la stabilité du poste hôte et sans altérer ses registres systèmes).



Par exemple « OpenOffice » virtualisé avec ThinStall n'occupe que la moitié d'une clé USB 256 MOctets et fonctionne immédiatement sur n'importe quelle machine qui possède un port USB. L'application, bien qu'un peu ralentie, reste parfaitement utilisable.

12.3 Intérêt de la virtualisation applicative

Cette technologie présente 2 intérêts principaux :

- L'installation « virtuelle » d'une nouvelle application sur un poste est considérablement raccourcie et simplifiée,
- Important : rien n'est modifié sur la machine et il ne restera aucune trace après le travail.

RAPPEL

Ce document est un support de cours et un complément d'information pour le stagiaire, il sera à compléter par les explications et commentaires au cours des 2 jours de ce module.

Remarque :

Beaucoup d'illustrations sont issues de « la toile », et une partie de la prose de certains paragraphes est également issue de la même origine. Ces documents sources sont donc du domaine public. Ces pages Web peuvent être retrouvées par le moteur de recherche favori du lecteur en utilisant certains mots clés associés aux pages recherchées. Pour éviter de donner des adresses de liens obsolètes, aucune bibliographie ne sera intégrée à ce support de cours.

D'autre part, en faisant ses propres recherches, le stagiaire peut donc trouver d'autres illustrations appropriées à ce support de cours.

→ Bonne lecture et bonne navigation !

FIN DU DOCUMENT