

Stockage Virtuel VMFS

VMFS : la base du stockage virtuel.

Youssef Achehboune,

Mohammed Bakir

Introduction : Le Stockage Virtuel

Le Défi du Stockage Traditionnel

Le stockage physique direct pose des défis majeurs en virtualisation :

- Le stockage classique est difficile à faire évoluer.
- Difficulté à partager les fichiers VM (vmx, vmdk).
- Les données sont souvent séparées et mal partagées entre serveurs.

La Solution : Le Stockage Virtuel

vSphere introduit une couche d'abstraction essentielle :

- Abstraction des LUNs en Datastores.
- VMFS : Le système de fichiers clusterisé.
- Permet l'accès simultané et sécurisé des hôtes.

Sommaire

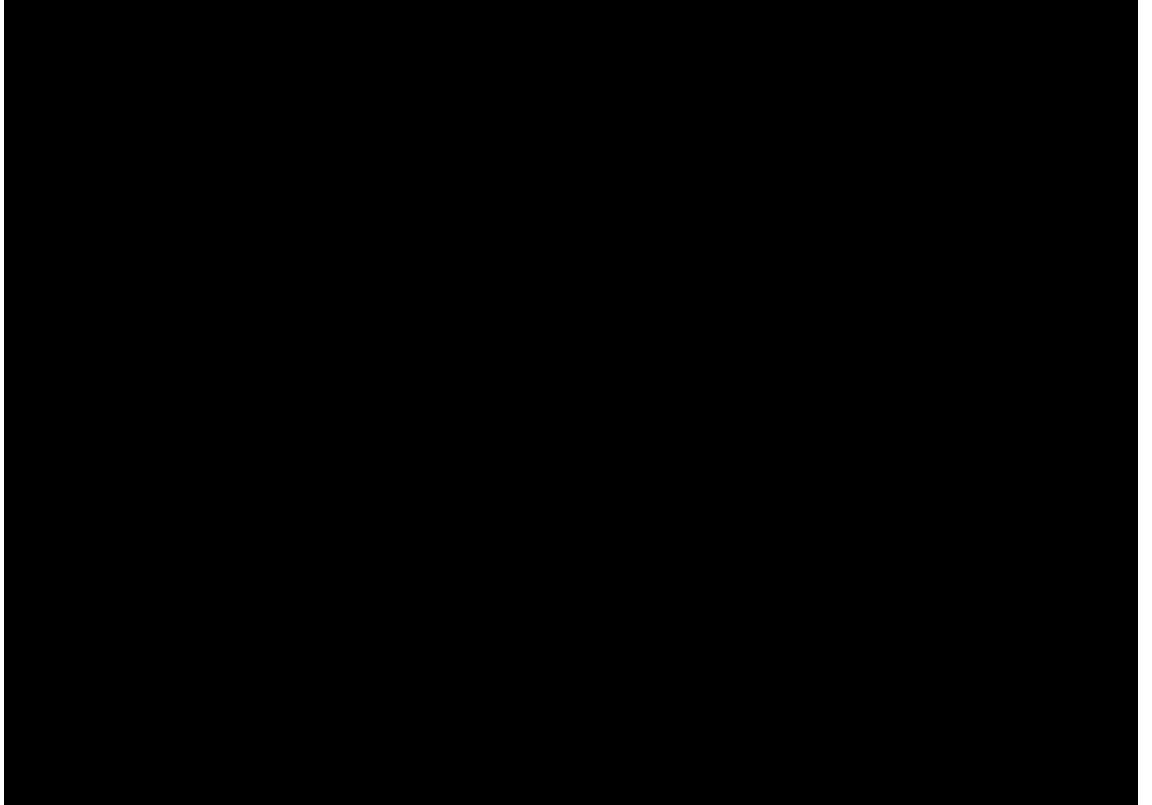
-  Qu'est-ce que VMFS ?
-  Architecture et Protocoles (SAN, LUN)
-  Provisioning de Disque (Thin vs Thick)
-  Snapshots et Évolutions (VMFS 5 vs 6)
-  Services Clés (vMotion, HA)
-  Limitations et Bonnes Pratiques

Qu'est-ce que VMFS ?

VMFS (Virtual Machine File System)

VMFS est un système de fichiers en cluster haute performance, conçu spécifiquement par VMware pour stocker les machines virtuelles.

- **Clustering** : Permet à plusieurs hôtes ESXi de lire et écrire sur le même volume partagé simultanément.
- **Verrouillage (ATS)** : Utilise Atomic Test & Set pour protéger l'intégrité des données sans bloquer tout le volume.
- **Optimisation I/O** : Gère les I/O spécifiques des VM (nombreux accès, fichiers volumineux).

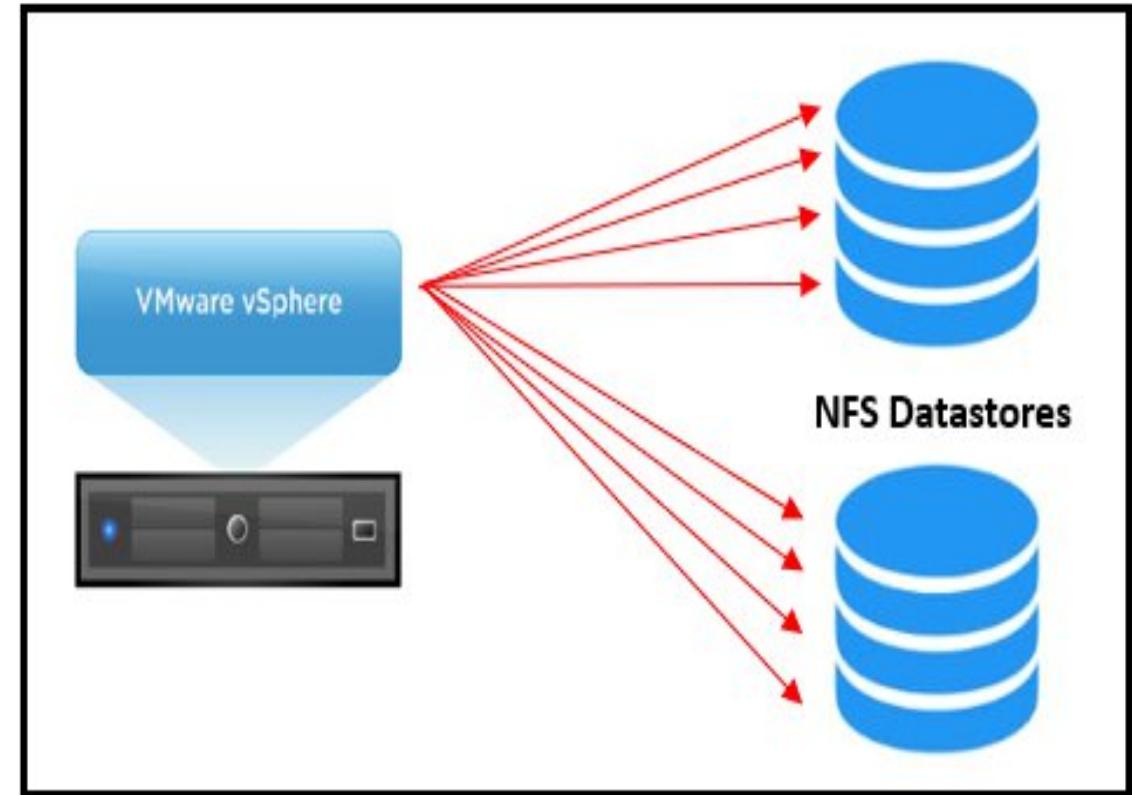


Le Datastore

Le Cœur de l'Abstraction

Le Datastore est l'abstraction vSphere pour le stockage. C'est un conteneur logique, similaire à un lecteur de disque, qui masque la complexité du matériel sous-jacent (FC, iSCSI, NFS). Il stocke les fichiers qui constituent une VM.

- Il est accessible par tous les hôtes du cluster.
- Il peut être étendu sur plusieurs LUNs (Extent).



Anatomie d'une Machine Virtuelle



.vmx

Fichier de **configuration**. Le "BIOS" et la définition matérielle de la VM (CPU, RAM, disques).



.vmdk

Le **disque dur virtuel** (Virtual Machine Disk). Contient l'OS invité et les données de la VM.



Autres Fichiers

.vmsn : Fichier d'état de snapshot.
.nvram : Fichier du BIOS/EFI.
.log : Fichiers journaux (logs) de la VM.

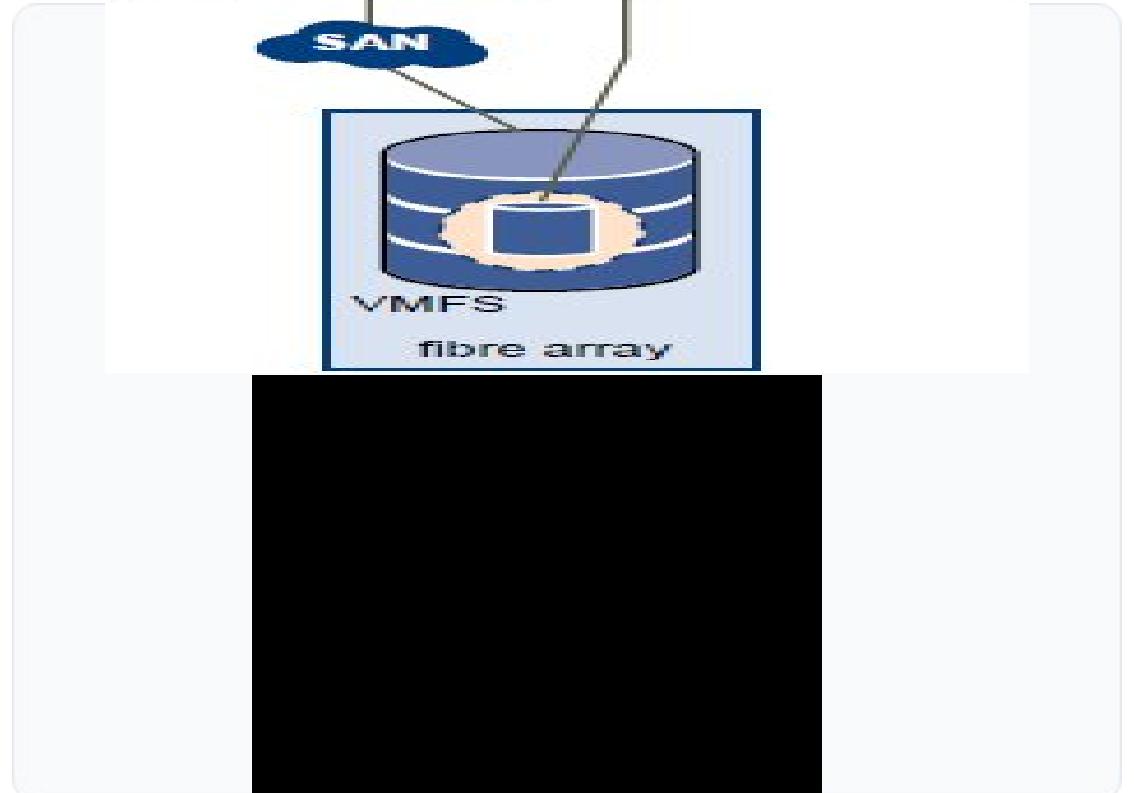
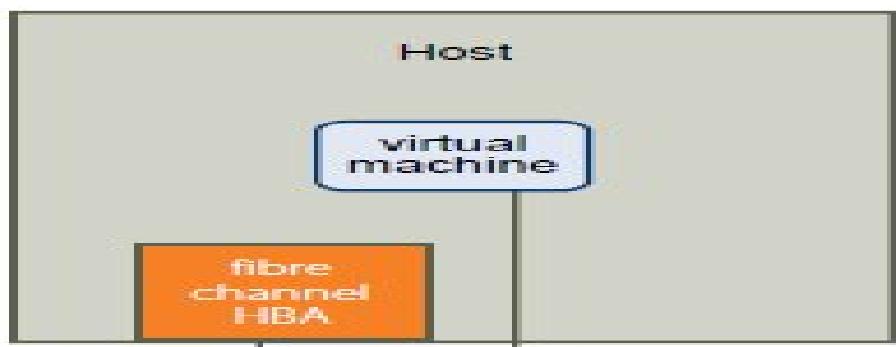
Définition SAN & LUN

SAN (Storage Area Network)

C'est le **réseau** dédié au stockage. C'est "l'autoroute" (composée de switchs et de câbles Fibre Channel ou Ethernet) qui relie les serveurs (ESXi) aux systèmes de stockage.

LUN (Logical Unit Number)

C'est une **unité logique** de stockage. Un "disque" virtuel provisionné depuis le système de stockage et présenté aux serveurs via le SAN. C'est sur le LUN que l'on crée le Datastore VMFS.

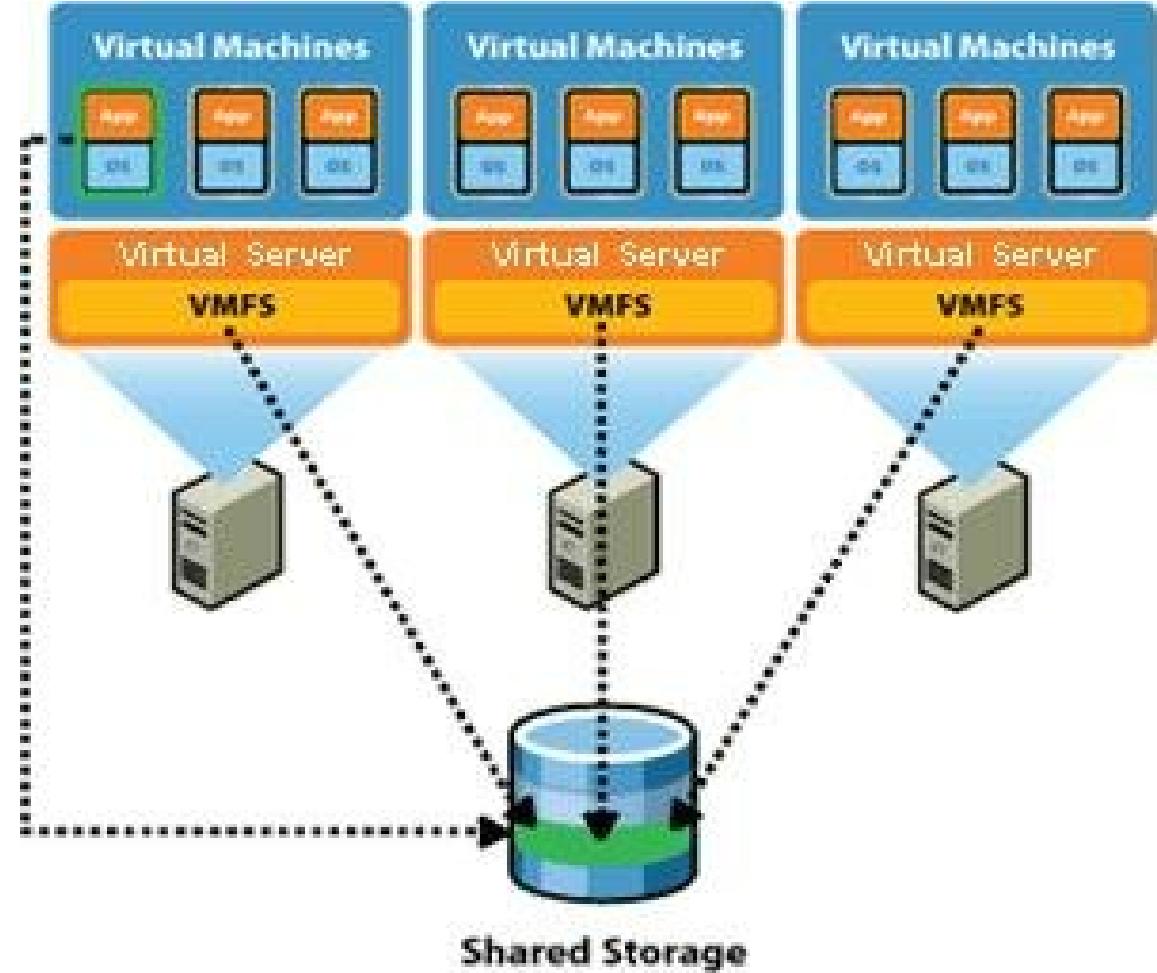


Protocole de Stockage

VMFS est un système de fichiers. Il a besoin d'un protocole de transport qui lui présente des blocs de données bruts (comme un disque dur physique).

Les principaux protocoles pour vSphere sont :

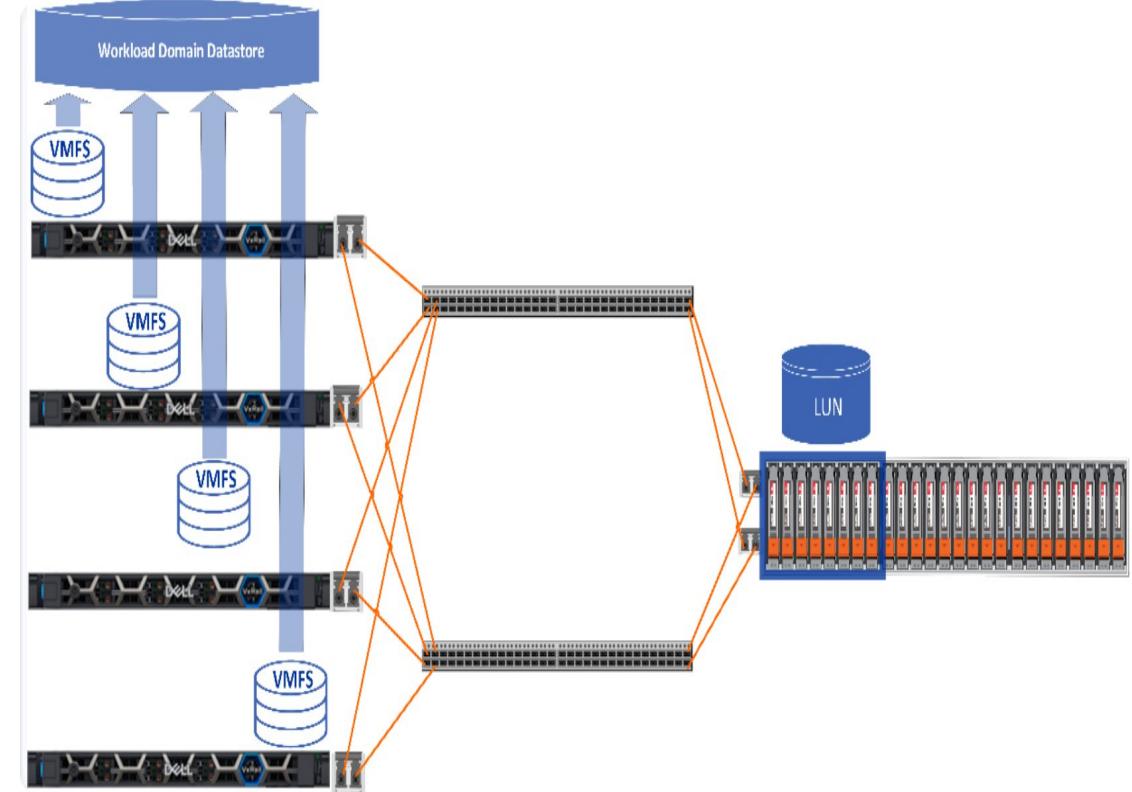
- Fibre Channel (FC)
- iSCSI (Internet SCSI)
- FCoE (Fibre Channel over Ethernet)



Fibre Channel (FC)

Protocole le plus utilisé pour les déploiements vSphere critiques. Très performant et fiable, il utilise un réseau dédié (le SAN).

- **HBA (Host Bus Adapter)** : Une carte dédiée dans l'hôte ESXi.
- **Switchs FC** : Switchs dédiés au trafic de stockage.
- **WWW (World Wide Web)** : L'adresse "MAC" d'un HBA, utilisée pour le Zoning.



FCoE et iSCSI (Le Stockage sur IP)

iSCSI (Internet SCSI)

Transporte les commandes "bloc" SCSI sur un réseau IP (Ethernet) standard.

- Utilise des NICs standards (Software) ou des HBA iSCSI (Hardware).
- **Avantage** : Coût réduit, utilise l'infrastructure Ethernet existante.

FCoE (Fibre Channel over Ethernet)

Transporte les trames Fibre Channel natives directement à l'intérieur de trames Ethernet.

- Nécessite un réseau Ethernet "Converged" (10Gb+) et des switchs compatibles (CNA).
- **Avantage** : Consolide le trafic LAN et SAN.

Provisioning de Disque

Qu'est-ce que le Provisioning ?

C'est la méthode d'allocation de l'espace disque pour un VMDK (disque virtuel) sur le Datastore VMFS.

La question : Si je crée un disque de 100 Go, est-ce que je réserve ces 100 Go **immédiatement**, ou **au fur et à mesure** ?

L'Approche : Thin vs Thick

VMFS permet deux approches principales :

- **Thin Provisioning (Fin)** : Le VMDK commence petit et "gonfle" avec les écritures.
- **Thick Provisioning (Épais)** : Les 100 Go sont réservés sur le Datastore dès la création.

Thin vs Thick

Thin Provisioning

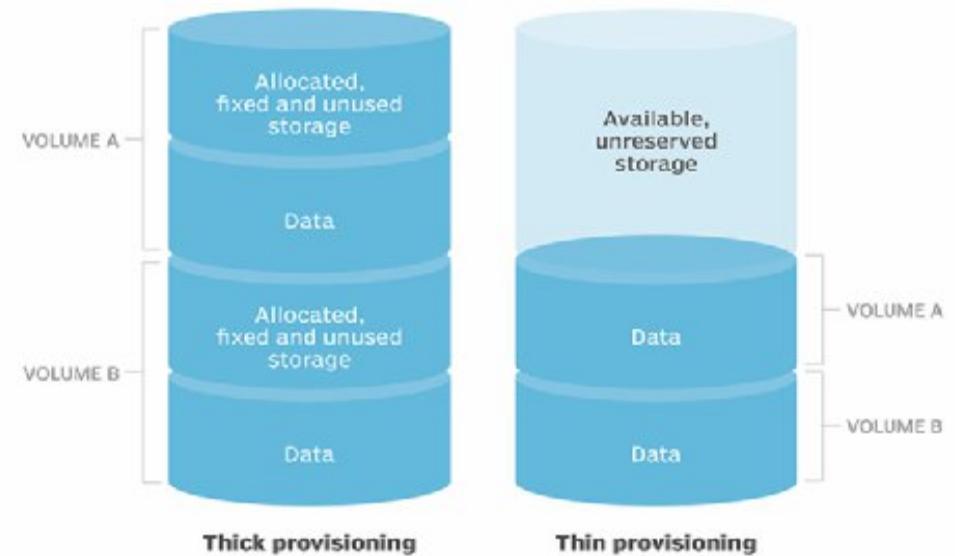
Avantage : Économie d'espace (sur-allocation). Idéal pour le Test/Dev.

Risque : "Over-provisioning". Si le Datastore est plein, les VM crashent.

Thick Provisioning

- **Lazy Zeroed (Défaut) :** Espace réservé. Rapide à créer.
- **Eager Zeroed :** Espace réservé ET mis à zéro. Lent à créer, mais performances maximales (recommandé pour BDD).

Thick vs. thin provisioning

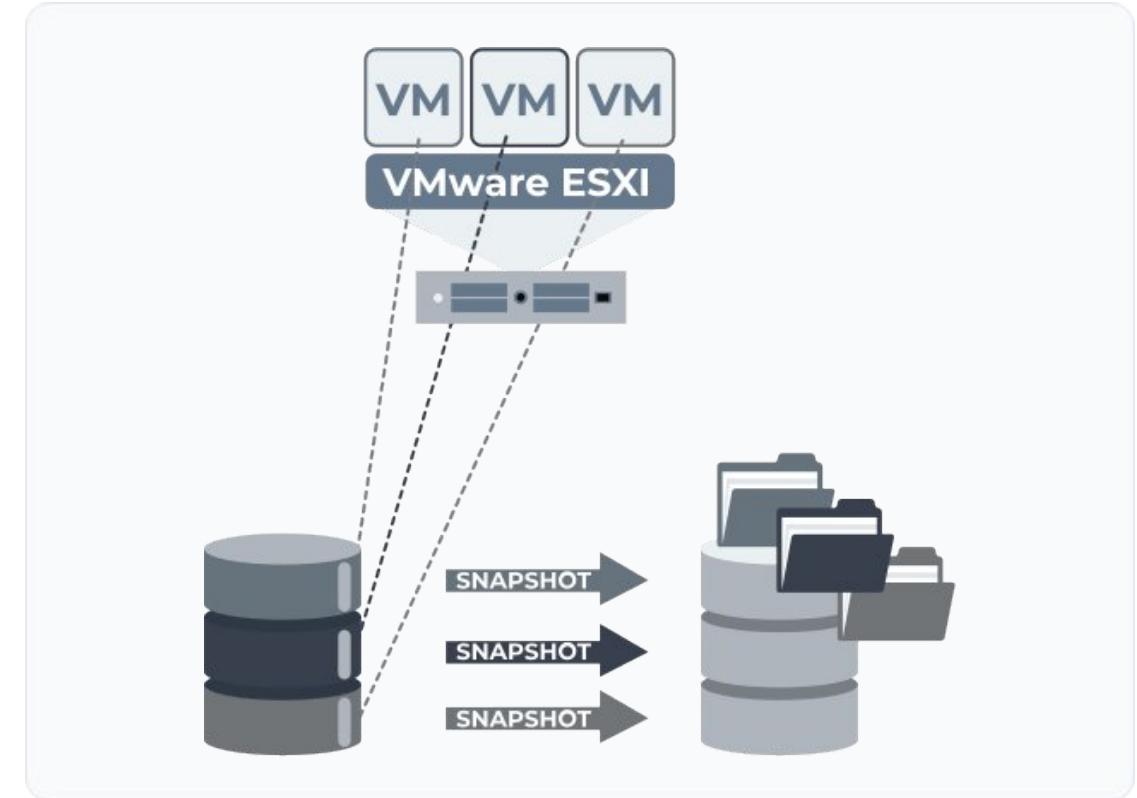


Snapshot

Un snapshot **n'est pas une sauvegarde**. C'est un moyen de "geler" l'état d'un VMDK à un instant T pour pouvoir y revenir.

Fonctionnement :

- Le VMDK original (disque de base) est mis en lecture
- ~~seule~~ Un nouveau fichier est créé : le **disque "Delta"**.
- Toutes les nouvelles écritures vont dans ce fichier Delta.



VMFS 5 vs VMFS 6

Caractéristique	VMFS 5 (vSphere 5.5 / 6.0)	VMFS 6 (vSphere 6.5+)
Récupération d'Espace (UNMAP)	Manuelle (commande CLI)	Automatique (en arrière-plan)
Taille de Bloc	1 Mo (unifié)	512b (petits blocs, plus efficace)
Disques 4K Natif (4Kn)	Non supporté	Supporté
Verrouillage (Locking)	ATS (conseillé) ou SCSI-2	ATS Uniquement (plus performant)
Mise à niveau	En ligne (depuis VMFS 3)	Hors ligne requise (migration)

Services Clés (Rendus possibles par VMFS)

vMotion / Storage vMotion

Migration à chaud d'une VM (RAM) ou de son stockage (Disque) sans interruption de service.

vSphere High Availability (HA)

Redémarrage automatique des VM sur un autre hôte en cas de panne matérielle du premier.

Distributed Resource Scheduler (DRS)

Équilibrage de charge automatique des VM (via vMotion) pour optimiser les performances CPU/RAM du cluster.

vMotion

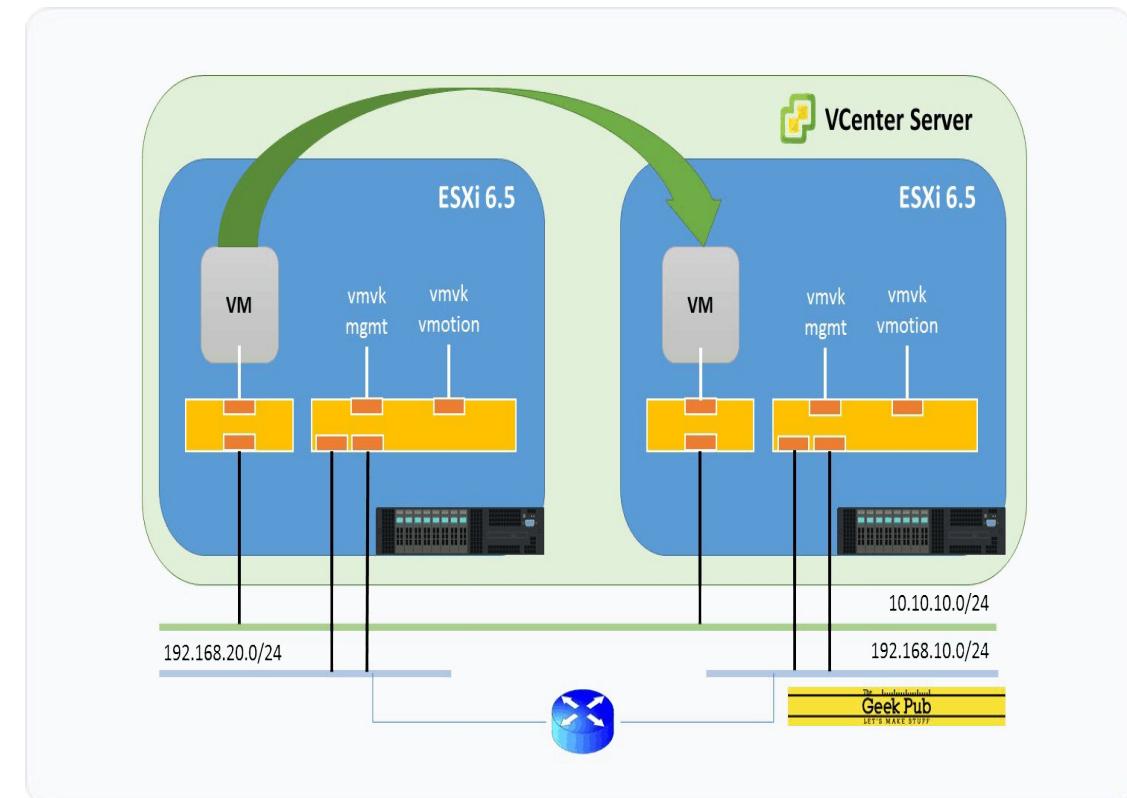
Migration à chaud

vMotion déplace une VM allumée d'un hôte ESXi A vers un hôte B.

Étapes clés :

- La **mémoire (RAM)** est copiée de A vers B via le réseau vMotion.
- L'état final (CPU/registres) est transféré.
- Le contrôle des fichiers VMDK (sur le Datastore VMFS partagé) est transféré à l'hôte B.

L'utilisateur ne voit qu'une ou deux pertes de ping.



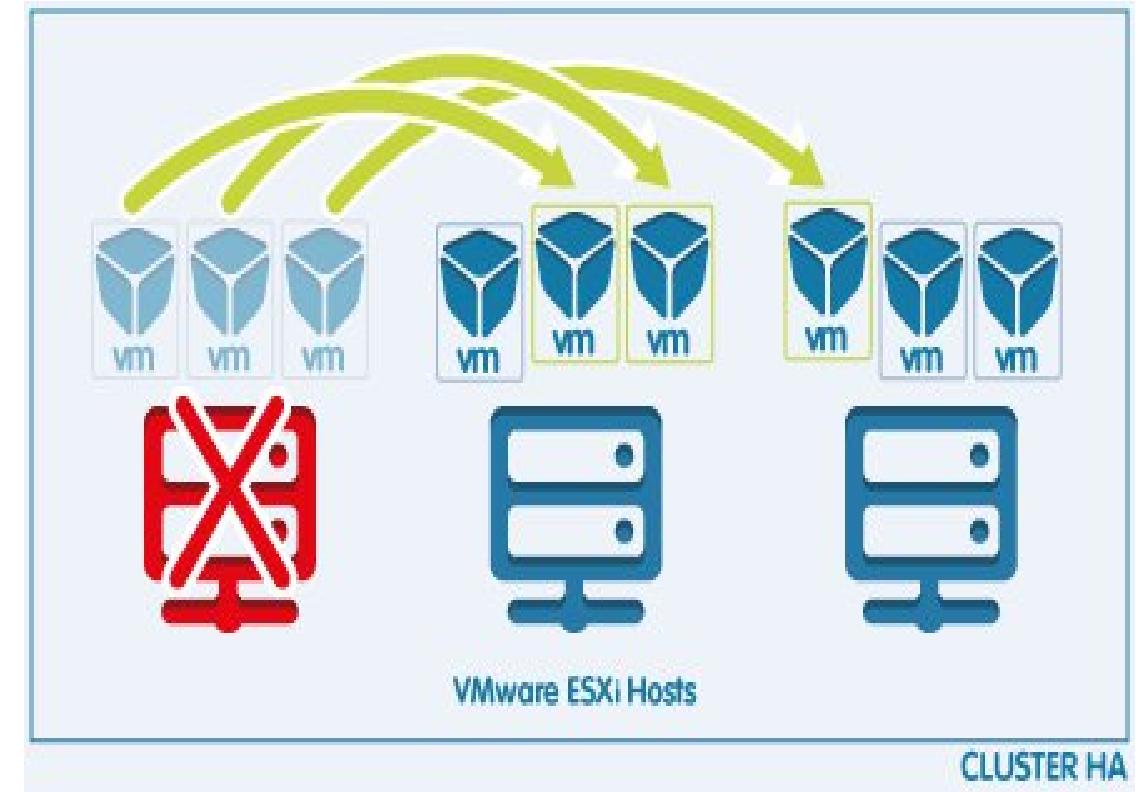
Service Clé : vSphere HA (Haute Disponibilité)

Reprise sur Panne Automatique

vSphere HA protège contre les pannes matérielles des hôtes ESXi.

Scénario de panne :

- L'hôte ESXi A tombe en panne (matériel).
- HA le détecte.
- HA redémarre automatiquement les VM de A sur d'autres hôtes (B, C) du cluster.
- **Rendu possible car** les hôtes B et C ont déjà accès aux fichiers VM sur le Datastore VMFS partagé.



Partie 6 : Limitations et Risques



SPOF (Single Point of Failure)

Si la baie de stockage (ou le Datastore) tombe en panne, *toutes* les VM s'arrêtent. La redondance du SAN est critique.



Latence & "Voisin Bruyant"

Une VM très gourmande en I/O peut impacter les performances de *toutes* les autres VM sur le même Datastore.



Saturation (Thin)

Le risque majeur du Thin Provisioning : si le Datastore est plein (100%), les VM qui tentent d'écrire crashent.

Partie 6 : Bonnes Pratiques

-  **Utiliser VMFS 6 :** Toujours privilégier VMFS 6 pour les nouvelles installations (UNMAP auto).
-  **Multipathing (MPIO) :** Configurer au moins 2 chemins (HBA, switchs, ports) pour chaque LUN.
-  **Monitoring :** Surveiller activement la latence (ms) et l'utilisation de l'espace.
-  **Gestion des Snapshots :** Ne *jamais* laisser un snapshot ouvert plus de 48-72h.
-  **Ségrégation des I/O :** Séparer les VM à haute I/O (BDD) des VM standards.

Synthèse et Conclusion

VMFS : Le Pilier

VMFS est le système de fichiers en cluster qui a rendu la virtualisation d'entreprise possible.

- C'est le **facilitateur** des services clés (vMotion, HA, DRS).
- Il repose sur un stockage "Bloc" partagé (SAN).
- Sa gestion (perf, capacité) est critique.

Le Futur (SDS/vSAN)

Le modèle SAN/VMFS traditionnel, bien que robuste, est complexe et coûteux.

- La tendance est au **Software-Defined (SDS)**.
- Des solutions comme vSAN (Hyperconvergence) ou vVols (Virtual Volumes) gagnent en popularité.

Questions ?

Merci de votre attention.