

FICHE DE RÉVISION — LES CLUSTERS

1. Les deux scénarios : PRA et PCA

Plan	Situation	Exemples
PRA / PRI	Plan de Reprise d'Activité / Informatique. J'ai une infra qui fonctionne, un incident grave intervient, l'activité est interrompue , je dois pouvoir reprendre l'activité.	La sauvegarde entre dans le PRA
PCA / PCI	Plan de Continuité d'Activité / Informatique. J'ai un incident, mais ça n'interrompt pas mon activité , ça continue de fonctionner.	RAID 5 , réplication de DNS , de contrôleur de domaine , de DHCP

Les CLUSTERS entrent dans la CONTINUITÉ d'activité (PCA).

2. Cluster d'hyperviseurs et failover

Un hyperviseur qui contient toutes nos VM. Un autre hyperviseur ailleurs, vide, qui en cas de problème va **accueillir les VM ayant migré automatiquement** sur ce second hyperviseur. On a donc une **continuité de service**.

2.1 FailOver — basculement de charge

On en parle ici pour les VM, mais cela fonctionne aussi pour d'autres ressources :

- FailOver de **base de données** (SQL, etc.)
- FailOver de **serveur de messagerie** (Microsoft Exchange, etc.)

L'idée est d'assurer la HIGH AVAILABILITY (HA), ou Fault Tolerance (FT) — en français : la HAUTE DISPONIBILITÉ.

On ne peut pas faire des clusters de tout et n'importe quoi. La rapidité de basculement est tributaire du matériel utilisé. En pratique, en tant que TSSR, il y a très peu de chance qu'on y soit confronté : c'est plutôt du ressort des admins.

2.2 LoadBalancer — équilibrage de charge

C'est de l'**équilibrage de charge réseau**. Nous avons **deux serveurs de production en même temps**, et les utilisateurs sont répartis automatiquement de manière à équilibrer la charge réseau.

2.3 Les nœuds

Lorsqu'on monte des clusters, on paramètre nécessairement des NŒUDS. Jusqu'à 32 nœuds maximum.

3. NAS, SAN et LUN

Lorsqu'on monte des clusters Hyper-V, on doit monter une zone de ressources partagées qu'on nomme **SAN**.

Élément	Définition
NAS	1 boîtier dans lequel il y a 1, 2 ou plus de disques durs, connecté au réseau et accessible sur le réseau. Quand on achète un NAS, il a une limite physique : il est acheté pour une volumétrie.
SAN	C'est comme un switch réseau, mais pour le stockage . Un SAN est vide , et on achète des baies de disques. Du matériel qui coûte très cher (10 k€ facilement), mais l'avantage est qu'il est granulaire : il a de l'extension possible.
LUN	Il s'agit de la partition du SAN . Au même titre qu'on fait une partition sur un disque dur, sur un SAN on appelle ça une LUN. Quand le serveur est connecté à la LUN, c'est comme si le serveur avait un disque dur d'une taille définie en physique : on le voit dans la Gestion des disques, mais il n'est pas physiquement présent dans le serveur.

3.1 Répartition des ressources dans un cluster d'hyperviseurs

Où	Quoi
Physiquement utilisé	Processeur, RAM et carte réseau
Sur le réseau	Le stockage, dans la LUN

4. Le HeartBeat (battement de cœur)

L'analogie est exacte et volontaire : comme un cœur qui envoie un signal électrique en continu, **les nœuds se pinguent en permanence entre eux**.

Signal	Interprétation
Le signal arrive	tout va bien
Le signal s'arrête	le nœud est considéré comme mort → basculement automatique

Le HeartBeat passe sur un RÉSEAU DÉDIÉ, séparé du trafic de production — justement pour ne pas confondre « le nœud est mort » et « le réseau de prod est saturé ».

5. Les 4 réseaux d'un cluster

Il faut également un réseau dédié pour le stockage : comme les hyperviseurs se connectent aux données, il faut une connexion **solide et rapide**. En pratique, pour assurer la performance de ce réseau, c'est de la **fibres optique**.

Le SAN n'est pas sur le réseau de production : il a son propre réseau isolé, dédié uniquement aux échanges entre le SAN et les hyperviseurs. Les utilisateurs n'y ont aucun accès, il n'y transite que du stockage.

Réseau	Qui parle à qui	Rôle
Production	utilisateurs ↔ hyperviseurs	le trafic métier
HeartBeat	hyperviseur 1 ↔ hyperviseur 2	surveillance mutuelle
SAN	hyperviseurs ↔ baie de stockage	les données des VM
Cluster	communication inter-nœuds	coordination : topologie, état, ressources

Chacun a son rôle, aucun ne se mélange avec les autres. C'est ça qui garantit la stabilité et les performances du cluster.

5.1 Le teaming — et pourquoi ça coûte cher

Comme une carte réseau peut tomber en panne, **on double tout**. C'est du **teaming** (ou agrégation de cartes) : chaque carte a son doublon, et si l'une tombe, l'autre prend le relais automatiquement.

4 réseaux × 2 cartes	= 8 cartes réseau par hyperviseur
8 cartes × 2 hyperviseurs	= 16 cartes réseau pour les nœuds

Et ce, sans compter le SAN ni les équipements réseau intermédiaires. Ce qui explique pourquoi ce genre d'infra représente un budget conséquent et reste du ressort des admins.

6. Maquette CLUSTERPA.LAN — monter le SAN

Nom du domaine : **CLUSTERPA.LAN**. On transforme **SRVPA04** en SAN.

6.1 iSCSI

iSCSI = Internet Small Computer System Interface. Un contrôleur de disque SCSI À TRAVERS LE RÉSEAU.

C'est un protocole de connexion de disque dur, mais dans un contexte de connexion réseau : entre le SRVPAxx et le SAN, **c'est comme s'il y avait un câble réseau**.

Gestionnaire de serveur > Services de fichiers et de stockage > iSCSI

6.2 Créer la LUN

Dans notre labo, vu qu'on n'a pas de vrai SAN, la nomenclature diffère : on va faire des Virtual Hard Disk (VHDX). Dans un vrai SAN, on aurait « Créer une nouvelle LUN ».

Pour créer la LUN, on utilise donc la fonction **Créer un nouveau disque dur virtuel**, et on crée un **VHDX**. La **cible iSCSI** contient les LUN.

6.3 Côté clients

Sur nos « clients » — les serveurs clients de notre LUN — on va **initialiser iSCSI**.

Pour que les deux SRVPA puissent accéder en lecture / écriture sur la LUN, il faudra installer la fonctionnalité CLUSTERING DE BASCULEMENT.

Une fois l'outil installé, on ouvre le **Gestionnaire de cluster de basculement**.

7. Le QUORUM

S'il y a MOINS de 50 % des votants, le cluster NE BASCULE PAS. Les votants, ce sont les NŒUDS.

7.1 Situation de départ — le problème

SRVPA02 (1 vote) + SRVPA03 (1 vote) = 2 votants
Quorum requis = 2

Astéroïde sur SRVPA02 → SRVPA03 tout seul → **1/2 = 50 %** → **pas de quorum** → le cluster se coupe → **pas de basculement**.

1 votant n'est pas SUPÉRIEUR à 50 % des votants. Il faut donc tricher : on va utiliser un truquage (la LUN) pour avoir un votant de plus.

7.2 La solution — le Disk Witness (disque témoin)

On ajoute un **disque témoin** (*Disk Witness*) sur un stockage partagé (LUN iSCSI ou FC), qui obtient **1 vote supplémentaire** dans la configuration du quorum.

```
SRVPA02 (1) + SRVPA03 (1) + Disk Witness (1) = 3 votants  
Quorum requis = 2
```

Astéroïde sur SRVPA02 → SRVPA03 + Disk Witness → **2/3 > 50 %** → **quorum atteint** → **basculement OK**.

2 votants sur 3, c'est plus de 50 % : le cluster bascule. C'est TOUT l'intérêt du disque témoin.

7.3 Mise en œuvre

Une fois qu'on ajoute un nouveau disque dans notre SAN (le SRVPA04), on dispose **déjà de la cible** — donc pas besoin de créer une nouvelle cible (**Target-01**) — et le nouveau disque **remonte automatiquement** dans les gestionnaires de disques de SRVPA02 et SRVPA03.

Il faut ensuite aller dans le **Gestionnaire du cluster** et **ajouter le disque** que l'on vient de créer. Pour optimiser l'espace, on change le disque témoin du quorum :

```
Témoin de Quorum  
> Témoin de Disque  
> Sélectionner le disque voulu
```

8. Partager la LUN et créer des VM en haute disponibilité

Une fois le Disk Witness du quorum changé, on ajoute le fait de **partager notre LUN** (ce qui n'était pas possible jusque-là).

Ce qui change : ce n'est PAS un lecteur qui est partagé, c'est un « dossier ». Il faudra aller dans C:\ClusterStorage\Volume1.

Pour créer des machines en haute disponibilité, il faut passer par la CRÉATION DE VM DANS LE GESTIONNAIRE DE CLUSTER.

9. Pièges et points à retenir pour l'ECF

PRA = l'activité est INTERROMPUE, on la reprend (la sauvegarde en fait partie). PCA = l'activité CONTINUE malgré l'incident (RAID 5, réplication, clusters).

- **FailOver** = basculement de charge. **LoadBalancer** = équilibrage de charge. Ce n'est pas la même chose.
- L'objectif d'un cluster, c'est la **haute disponibilité** (HA / Fault Tolerance).
- **32 nœuds maximum** dans un cluster.
- **NAS** = un boîtier avec une volumétrie fixe. **SAN** = un « switch de stockage », vide, extensible, cher. **LUN** = une partition du SAN.
- Une LUN est vue par le serveur comme un **disque physique**, alors qu'elle n'est pas dans le serveur.
- **iSCSI** = Internet Small Computer System Interface : du SCSI **à travers le réseau**.
- En labo, la LUN est simulée par un **VHDX** créé comme un disque dur virtuel, exposé via une **cible iSCSI**.
- Le **HeartBeat** doit être sur un **réseau dédié**, sinon on confond « nœud mort » et « réseau saturé ».
- **4 réseaux distincts** : production, HeartBeat, SAN, cluster. Doubles par teaming → **8 cartes par hyperviseur**.
- Pour l'accès concurrent en lecture / écriture sur la LUN, il faut la fonctionnalité **Clustering de basculement**.
- **Quorum** : il faut **plus de 50 %** des votants. Avec 2 nœuds seulement, la perte d'un nœud = 50 % = **pas de basculement**.
- Le **Disk Witness** apporte le **3e vote** qui débloque la situation : 2/3 > 50 %.

- Le stockage partagé du cluster n'est **pas une lettre de lecteur** mais un dossier :
`C:\ClusterStorage\Volume1`.
- Une VM en haute disponibilité **doit être créée depuis le Gestionnaire de cluster**, pas depuis Hyper-V directement.