Les types de systèmes de fichiers pour stockage des machines virtuelles

1. RDM (Raw Device Mapping)

Concernant le RDM, veuillez consulter le chapitre Les machines virtuelles où ce sujet est traité.

2. VMFS (Virtual Machine File System)

Le système de fichiers VMFS est le système de fichiers principal (journalisé) de VMware. Les principaux objectifs étaient de fournir un système de fichiers capable de supporter de gros fichiers tout en bénéficiant en parallèle de performances maximisées. On dit de VMFS que c'est un système de fichiers clusterisé. Plusieurs ESX peuvent en effet lire/écrire en parallèle sur le même espace disque VMFS. Cependant, une VM ne doit être accédée que par un seul ESX à un instant donné. C'est pourquoi VMware a inclus dans VMFS la capacité de « disk locking », autrement dit une technique de verrouillage.

Afin de gérer les accès multiples, VMware utilise un mécanisme de réservation SCSI. Ce mécanisme peut être utilisé lorsque VM démarre, ou bien lorsque qu'un Snapshot a lieu ou encore lors d'un Storage vMotion (de manière non exhaustive).

Cette fonctionnalité permet concrètement que tous les volumes soient vus par tous les hôtes au même moment. Cette technologie autorise donc l'activation de vMotion, HA et DRS. Le VMkernel ne bloque pas le système de fichiers ce qui a pour effet de permettre les accès concurrents.

a. VMFS v5

Le VMFS v5 permet :

- La gestion des accès concurrentiels comme nous l'avons vu auparavant.
- 256 volumes par système.
- 2048 VM par volume VMFS.
- 64 hôtes connectés par volume.
- De gérer des volumes de 64 To (enfin !).
- D'avoir une taille de bloc unifiée de 1 Mo (avant il fallait choisir en 1, 2, 4 ou 8 MB suivant la taille du volume).
- L'extension.

Clairement, la fonctionnalité la plus attendue était la possibilité de créer des LUN dont la taille est supérieure à 2 TB (en VMFS v3). C'est désormais chose faite !

- La possibilité d'étendre un volume VMFS sur plusieurs disques ou LUNs.

  Il est possible de créer des volumes VMFS étendus sur plusieurs LUNs. Le bénéfice est le suivant : il devient possible d'augmenter facilement un volume sans perturber la production. Cependant, il faut faire attention à cette technique. Si l'un des LUN est constitué d'un simple RAID0, et qu'un disque subit une défaillance, le volume entier ne fonctionnera plus. Il faut donc vérifier que l'ensemble des LUNs possèdent une tolérance de panne.

  Un serveur ESXi (en version 6.5) gère jusqu'à 512 LUNs et 256 montages NFS. Généralement on en trouve bien moins par hôte physique. Être limité par le nombre de volumes maximum est un cas assez rare en production.

- Des volumes de 64 To maximum

  Chaque LUN vu par un serveur ESX peut avoir une taille maximum de 64 To. Du fait que VMware a redéveloppé

VMFS v5 en 64 bits (passage de MBR (*Master Boot Record*) en GPT (*GUID Partition Table*), il devient possible désormais de virtualiser n'importe quel serveur ayant un très grand nombre de données. Auparavant, en VMFS v3, il n'était possible que de créer des LUN de 2 To, avec 32 « extents » afin d'atteindre 64 TB.

L'upgrade de VMFS v3 vers v5 n'est pas sans effet de bord. Par exemple on conserve les anciennes tailles de bloc. Il est recommandé de créer directement un LUN en v5 afin de bénéficier de toutes les fonctionnalités.

On parle de 32 « extents » au total ce qui représente en fait une base (le master extent contenant les informations de partition) et 31 extensions.

L'utilisation d'un MBR est maintenue avec la limite de 2 To tant que la limite n'est pas dépassée. Au-delà, le basculement en GPT s'effectue de manière automatique.

L'API VAAI ne sera pas entièrement exploitée (voir les documentations VMware pour plus de détails).

b. VMFS v6

Le VMFS v6 reprend les capacités du VMFS v5 auquel, il ajoute les améliorations suivantes (https://storagehub.vmware.com/#!/vsphere-core-storage/vsphere-6-5-storage/vmfs-6/1) :

- Le support des périphériques de stockages avec la norme 4 ko et l'émulation 512e. Le VMFS 6 s'aligne sur des blocs de 4 ko. La taille des métadonnées se base sur des multiples de 4 k.
- Les fichiers de ressources système tels que les pointeurs de blocs, les sous-blocs et les fichiers de description ont chacun leur propre fichier de gestion créé lors de la création du datastore VMFS6. Ces fichiers grossissent de manière dynamique.
- L'utilisation du SFB 1 MB (*Small File Block*) et LFB 512MB (*Large File Block*). Le Thin Provisioning utilise le SFB. Le thick provisioning, lui utilise le LFB au maximum, puis complète avec des blocs SFB.
- L'utilisation du Space Efficient spare (SEspare), par défaut, pour la gestion des snapshots. Le SEspare intègre la récupération d'espace (unmap) qui permet de libérer l'espace suite à la suppression d'un snapshot.
- Des améliorations sur la parallélisation des opérations de (re)signature et de découverte des périphériques de stockage. Un plus grand nombre de périphériques (512) et de chemins d'accès (2000).

Contrairement au VMFS v3 où il était possible de faire la migration vers le VMFS v5. Il n'est pas possible de faire de migration du VMFS v5 vers le VMFS v6.

Création d'un datastore VMFS

iSCSI

Dans le cadre de l'utilisation de l'iSCSI dans une infrastructure VMware, VMware a défini trois types d'initiateurs :

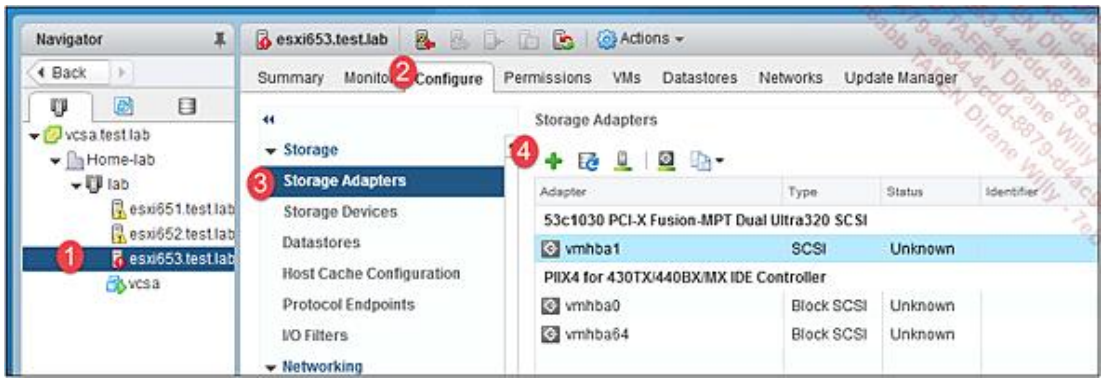
- Initiateur Software, il est inclus dans le code et s'intègre dans le VMKernel, et s'appuie sur la carte réseau présente dans l'ESXi. Il est représenté en tant que vmhba dans le type iSCSI software adapter. Dans le cas de l'utilisation de l'initiateur logique, il est important de prendre en compte les points suivants (https://kb.vmware.com/kb/2038869 et https://kb.vmware.com/kb/2045040) :
  - Utilisation du port binding ou liaison de port :
    - Dans le cas où les ports de la baie résident dans le même domaine de broadcast et le même sous réseau IP que le VMkernel port.
    - Les VMkernel ports sont dans le même sous-réseau.

- Actuellement, la liaison de port ne supporte pas le routage réseau.

- Cas où l'on n'utilise pas la liaison de port :
  - Dans le cas où les ports de la baie ne résident ni dans le même domaine de broadcast ni dans le même sous-réseau IP.
  - Tous les ports VMkernel utilisés dans le cadre de la connexion iSCSI sont présents dans des vSwitch/vDS, des sous réseau et des domaines de broadcast différents.
  - Dans le cas d'agrégation de lien réseau (LACP/Etherchannel - 802.3ad).

VMware fournit un document concernant la gestion du multipathing avec l'adaptateur logiciel iSCSI (http://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/techpaper/vmware-multipathing-configuration-software-iscsi-port-binding-white-paper.pdf).

Création d'une carte réseau iSCSI de type software :



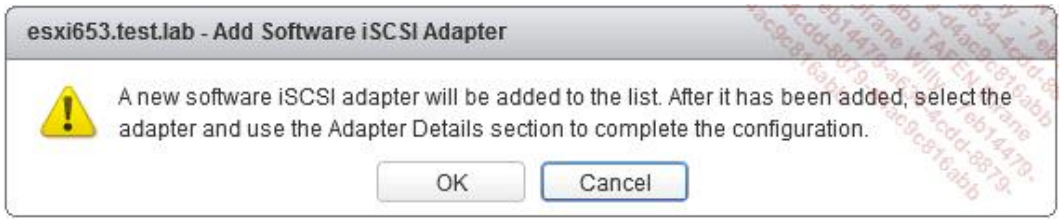
1- Sélectionner un hyperviseur.

2- Sur le plan central du Web Client, choisir l'onglet **Configure**.

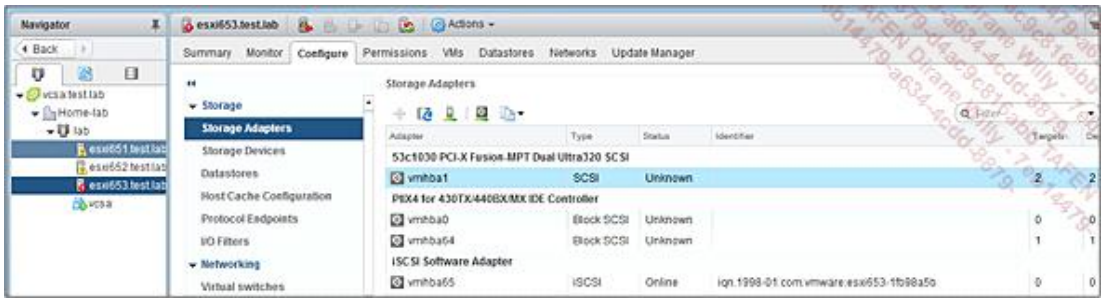
3- Choisir le menu **Storage Adapters**.

4- Ajouter un storage adapter : l'initiateur logiciel.

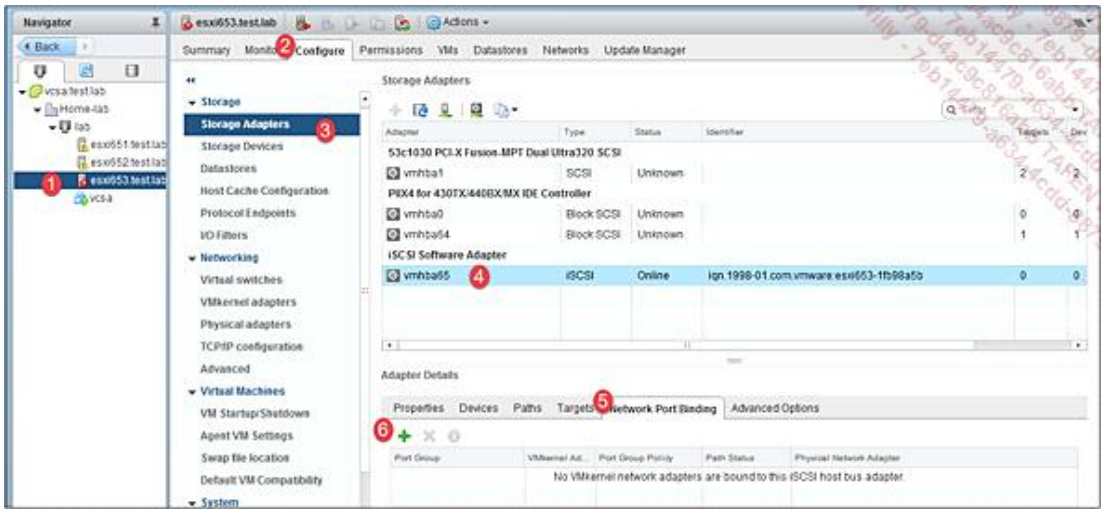
L'initiateur logiciel est disponible (on ne peut en créer qu'un). Les adaptateurs FCoE et FC ne sont pas disponibles car l'hyperviseur ne dispose pas de cartes compatibles. De plus dans le cadre de l'utilisation de carte CNA FCoE, il est nécessaire de créer un initiateur logiciel spécifique.



Un avertissement somme toute assez bénin apparaît. Après avoir cliqué sur **OK** , il sera possible de configurer l'adaptateur de stockage nouvellement ajouté.

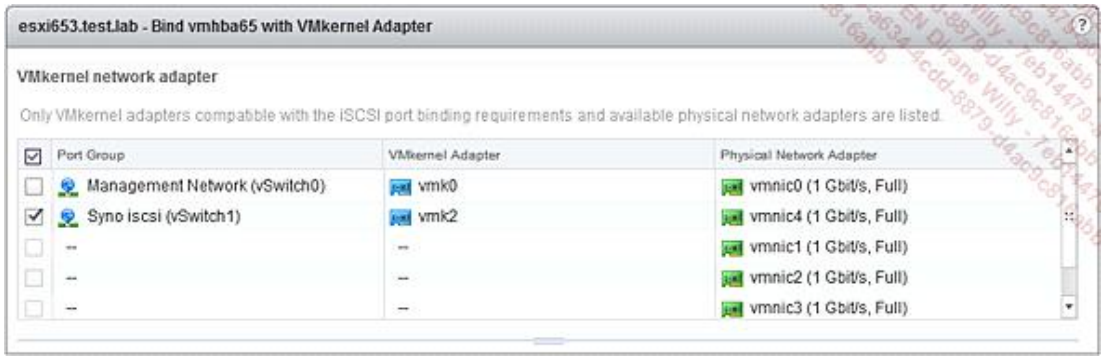


On peut voir l'adaptateur vmhba65 avec l'identifiant iqn.1998-01.com.vmware:esxi653-1b98a5b.

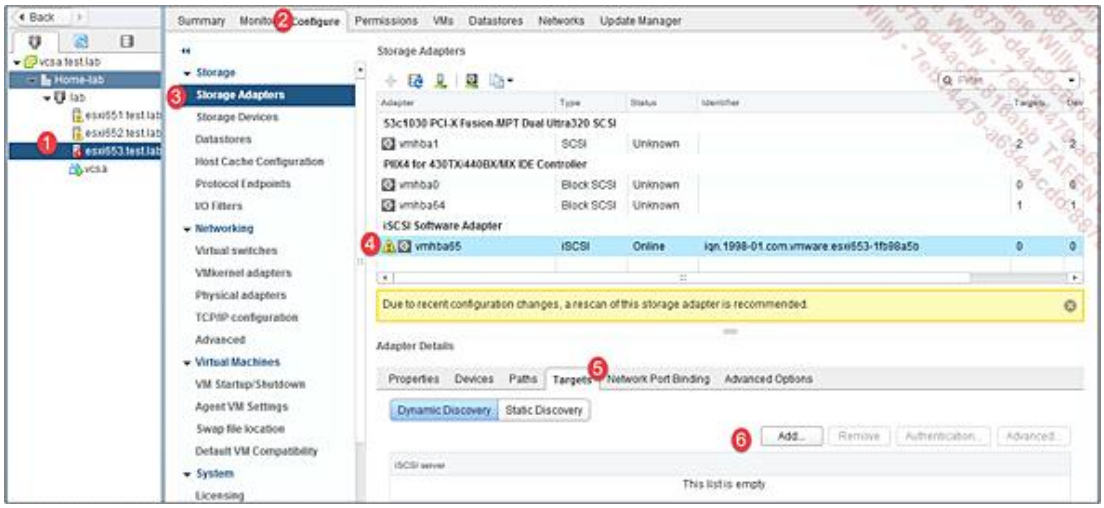


Il sera possible d'associer l'initiateur iSCSI logiciel à plusieurs VMkernel ports.

Ce n'est pas obligatoire mais préférable pour choisir les chemins physique et logique vers le stockage. (Les points numérotés représentent une indication du cheminement dans l'interface graphique pour pouvoir configurer la liaison de ports - network port binding).

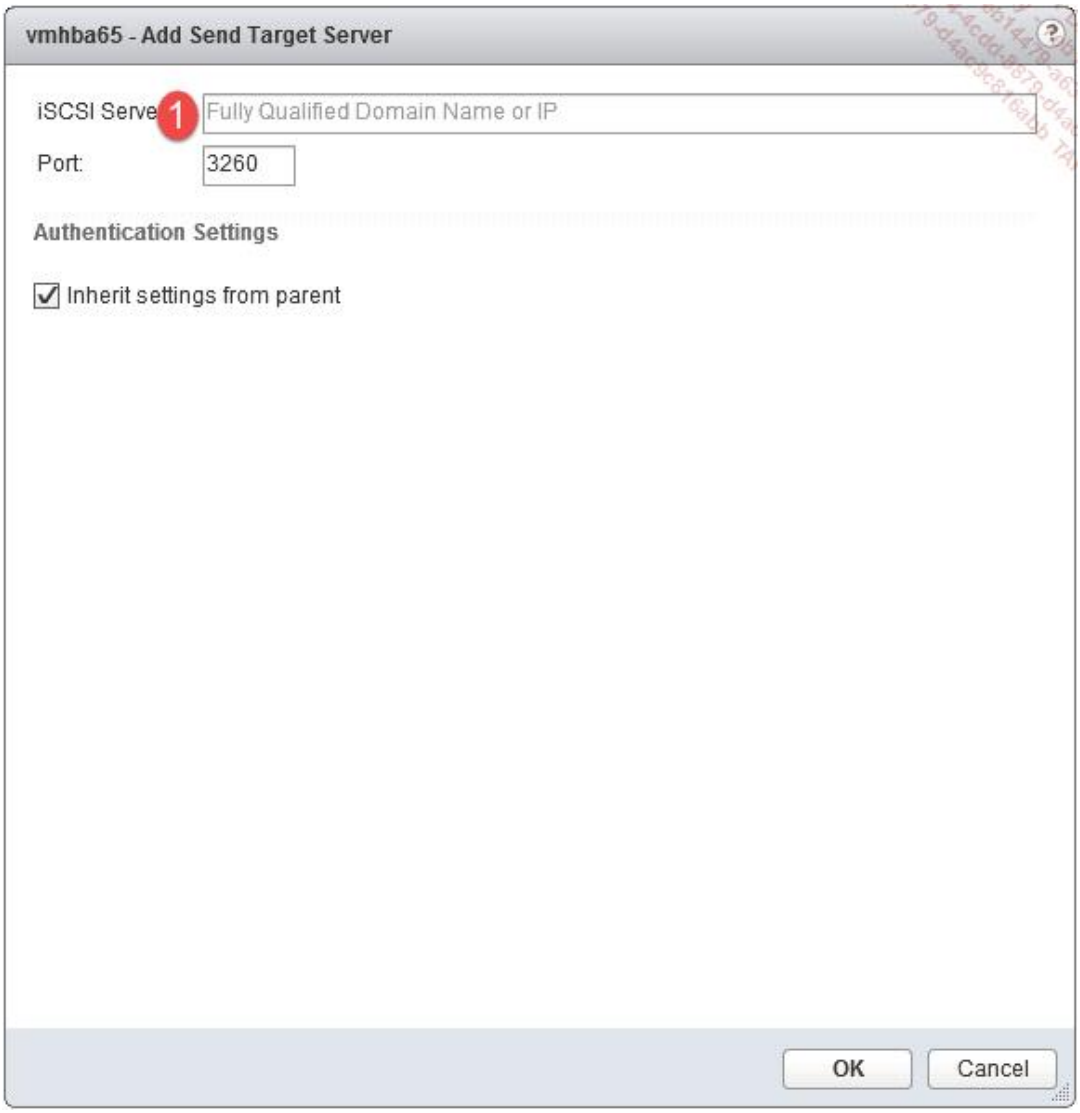


Ici, choisissez le Syno iscsi, le VMkernel port vmk2.

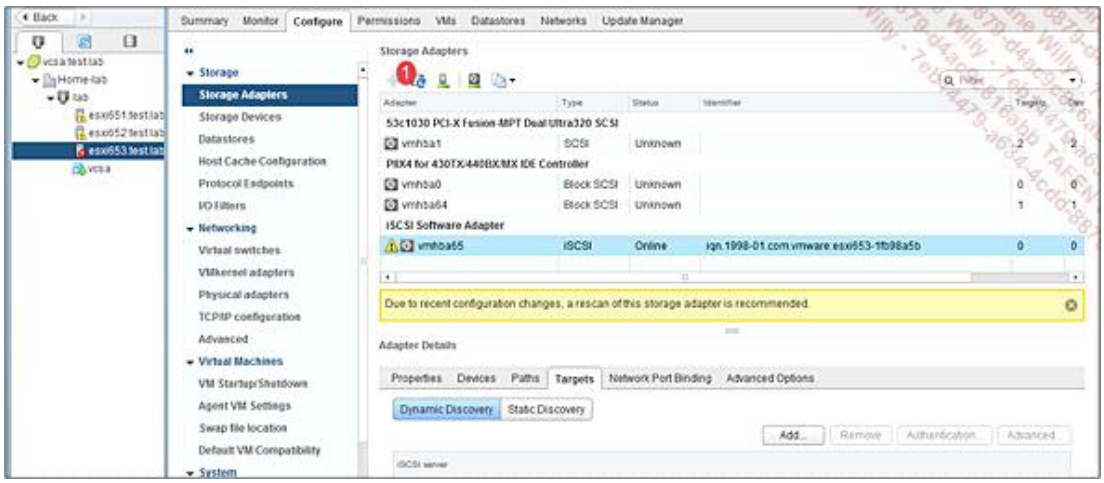


Pour configurer les adaptateurs de stockage (3), sélectionnez la carte iSCSI Software Adapter vmhba65 (4). En sélectionnant l'onglet **Target** (5), il est possible d'ajouter les baies de stockage en cliquant sur **Add** (6).

Après un rescan obligatoire afin de prendre en compte la configuration effectuée, il est possible d'ajouter une cible (une baie de stockage).

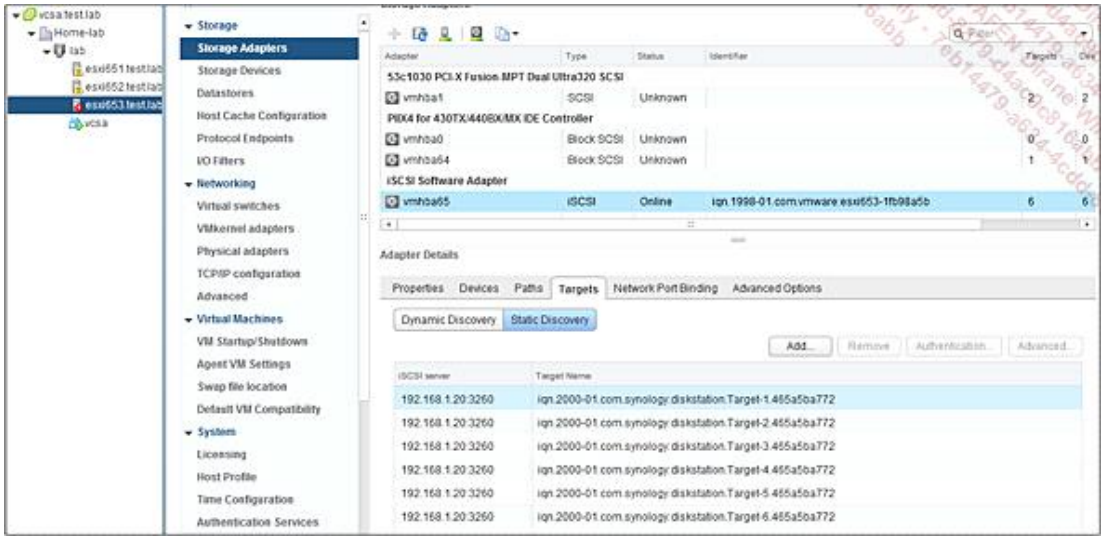


Inscrivez le nom ou l'adresse IP de la baie qui présente les LUNs.



Un rescan est nécessaire.

Le rescan porte sur les nouveaux volumes ainsi que les nouveaux serveurs de stockage ; effectivement, nous venons d'ajouter une cible.



- Initiateur dépendant du matériel (*hardware dependant*), ils sont dépendants de l'hôte tant pour la configuration de l'initiateur que pour sa gestion et accès réseau.
- Initiateur indépendant du matériel (*hardware indépendant*), bien que sa configuration puisse se faire via le client web vSphere, il est possible de faire la configuration directement à partir des paramètres de la carte.

Pour des besoins de sécurité, il est possible d'implémenter le Challenge Handshake Authentication Protocol ou CHAP. Le CHAP est un protocole d'authentification point à point. Il est défini par la RFC 1994.

Le CHAP fonctionne en trois étapes, mais à intervalle régulier réauthentifie le partenaire. Cette authentification au lieu lors de l'établissement de la première connexion, tout comme après chaque rétablissement du lien réseau.

Le protocole CHAP nécessite la configuration d'un mot de passe prépartagé sur le client et le serveur (pre-shared secret).

Les trois étapes sont les suivantes :

- Le serveur génère un nombre aléatoire (avec un compteur qui s'incrémente à chaque envoi, ce qui permet

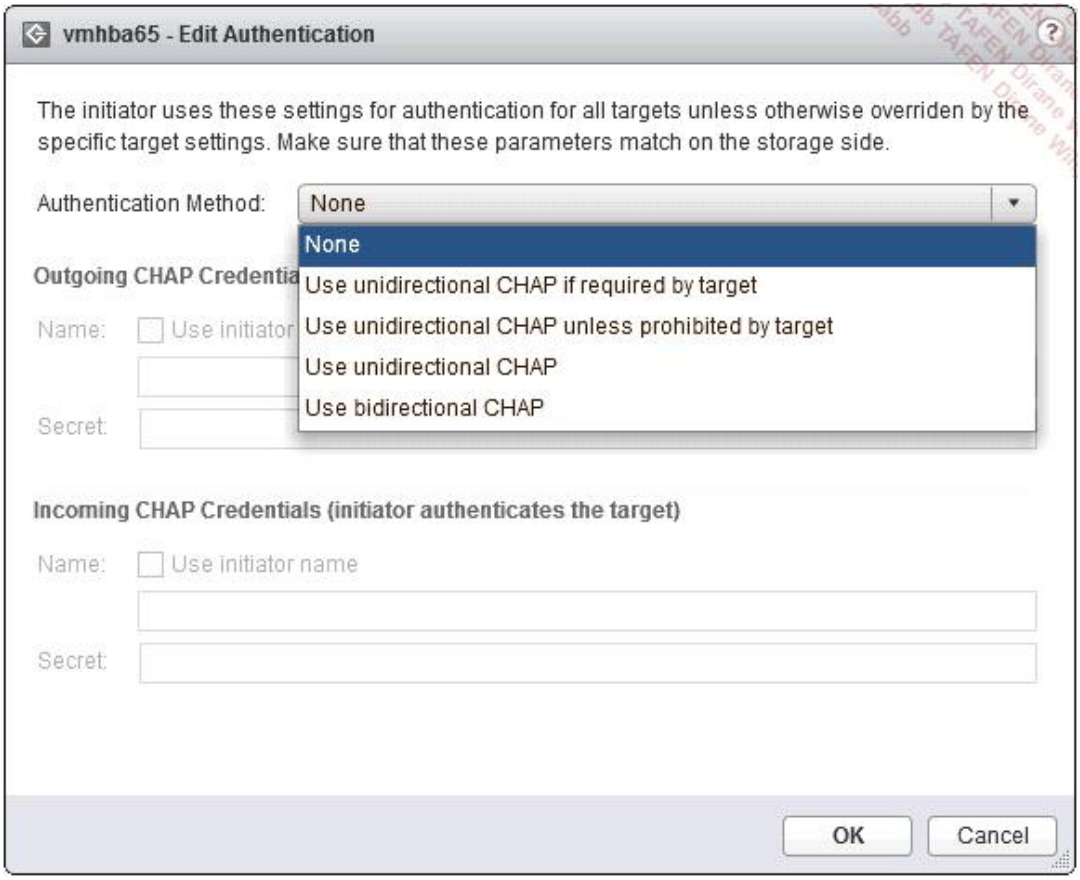
d'identifier la session) et envoie ce nombre au client.

- Le client répond avec le résultat de l'application d'une fonction de hash sur le mot de passe partagé que nous lui avons configuré ainsi que le nombre aléatoire.
- Le serveur compare la valeur envoyée par le partenaire à son propre résultat. Si les valeurs correspondent, l'authentification est validée et par extension la connexion active. Dans le cas contraire, la connexion est rompue.

Lors de la configuration CHAP, il faut faire attention au choix que l'on fait : unidirectionnel ou bidirectionnel. Dans le cas de la configuration unidirectionnelle, la cible authentifie l'initiateur. Dans le cas bidirectionnel, la cible et l'initiateur s'authentifient l'un par rapport à l'autre.

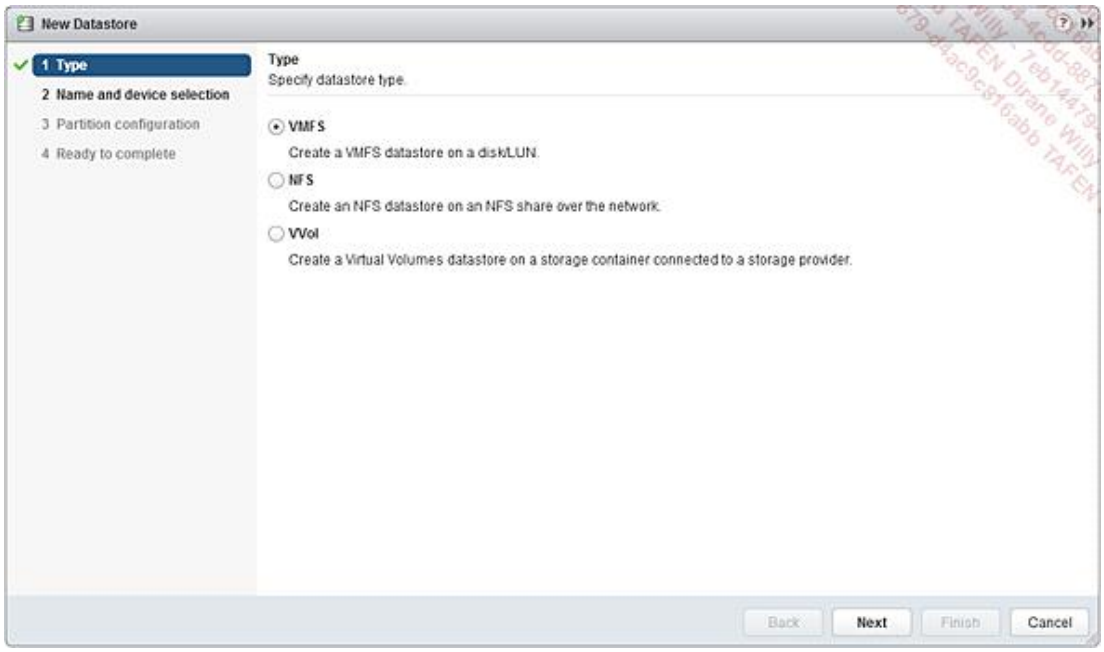
Les choix disponibles sont les suivantes :

- Unidirectionnel avec authentification CHAP : si c'est demandé par la cible signifie que par défaut le CHAP n'est pas utilisé, mais que si cela est demandé par la cible, le CHAP sera utilisé.
- Unidirectionnel avec authentification CHAP : si ce n'est pas interdit par la cible signifie que par défaut le CHAP est utilisé, mais que si la cible n'est pas apte à l'utiliser, le CHAP sera ignoré.

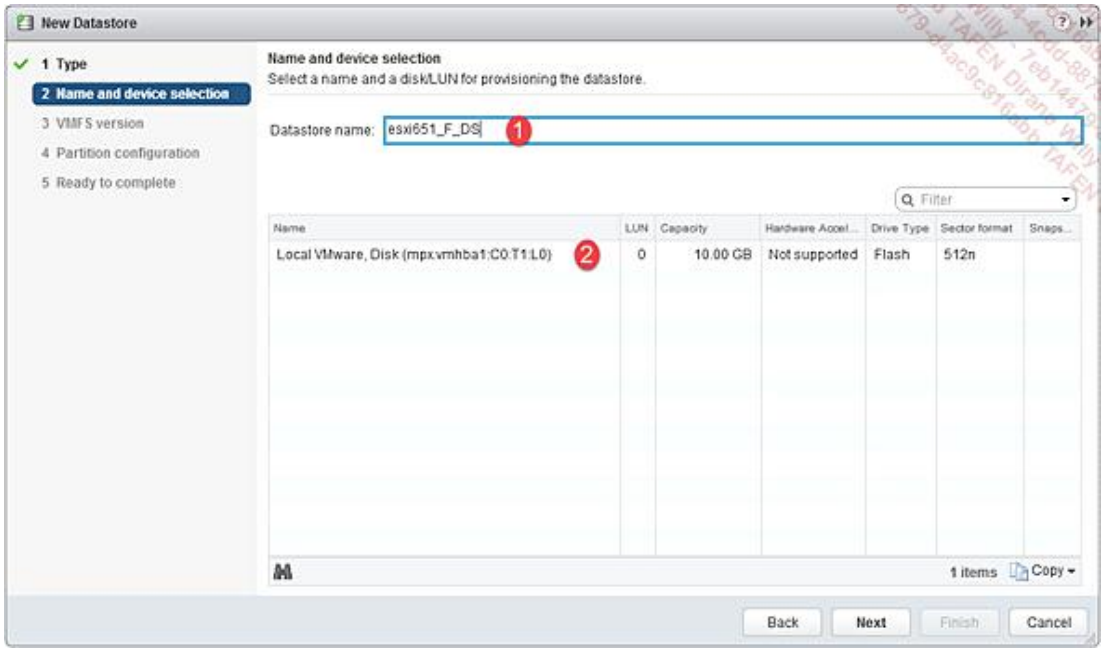


Pour créer un datastore, rendez-vous dans l'onglet configuration de l'ESXi (3), puis dans la catégorie **Datastore** (4). Lançez ensuite l'aide à la configuration (5). Il est conseillé de respecter la recommandation d'un datastore par LUN afin d'avoir une seule région de métadonnées VMFS par LUN.

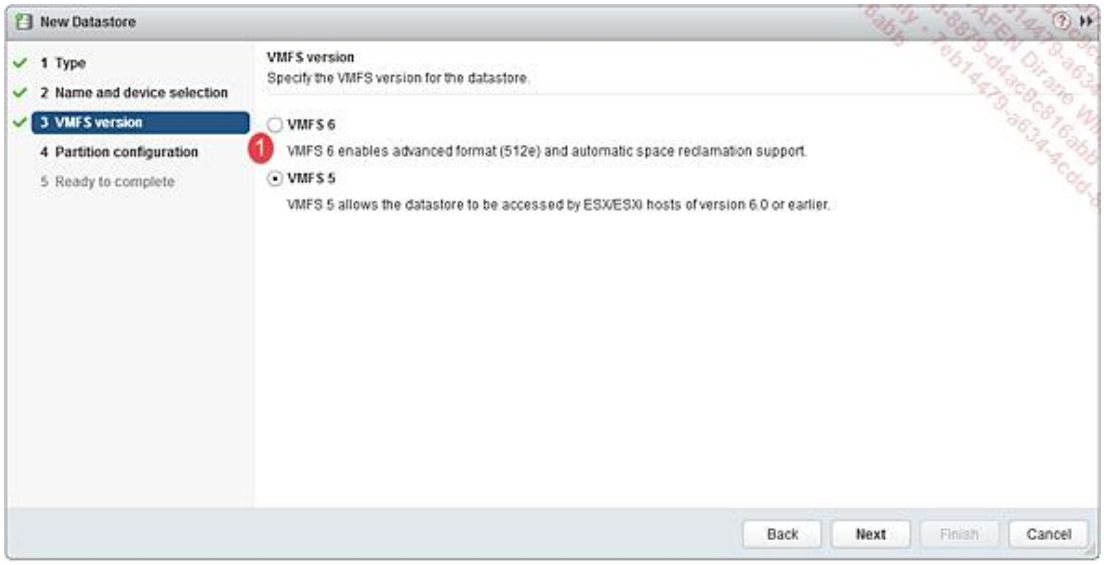
Créer plusieurs datastores dans une LUN est possible et supporté mais fortement déconseillé par VMware car cela peut poser des problèmes de performances et de cohérence (notamment en termes de thin-provisionning par la baie de stockage).



Choisissez le type (VMFS).

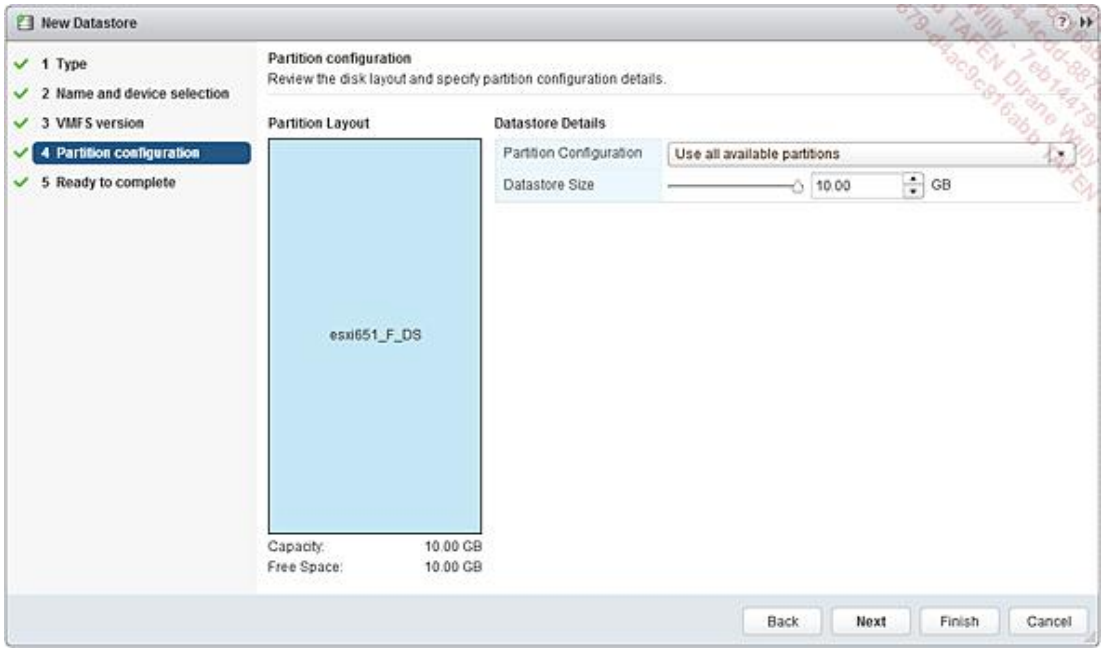


Indiquez le nom du datastore (en fait de la partition) et sur quel volume on le crée (ici le disque local).

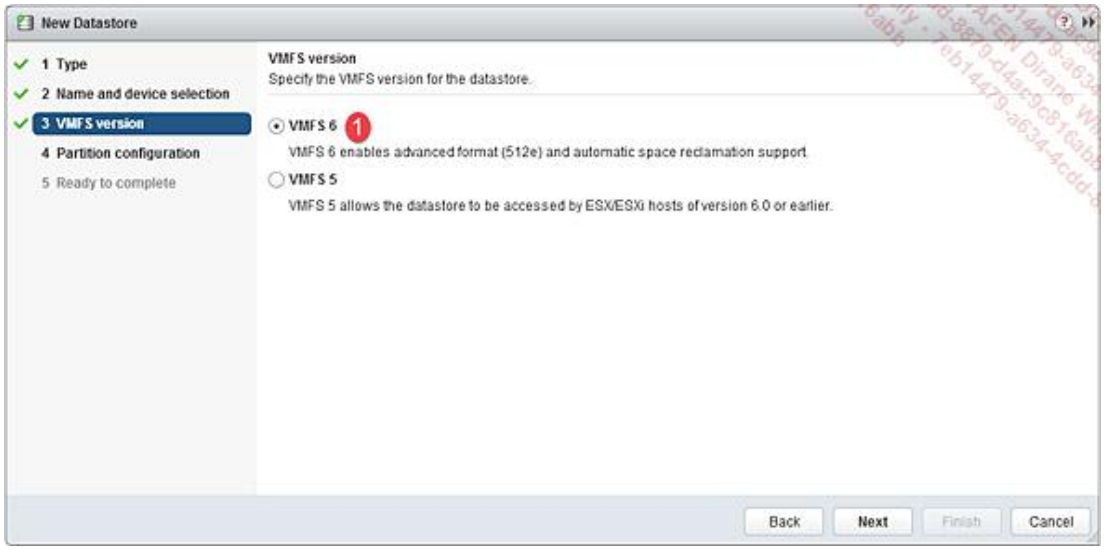


Ici vous verrez la différence entre la configuration de VMFS 5 et la configuration VMFS 6.

Choisissez la version (ici VMFS5 pour assurer la compatibilité avec les ESXi 6.0 et antérieurs).

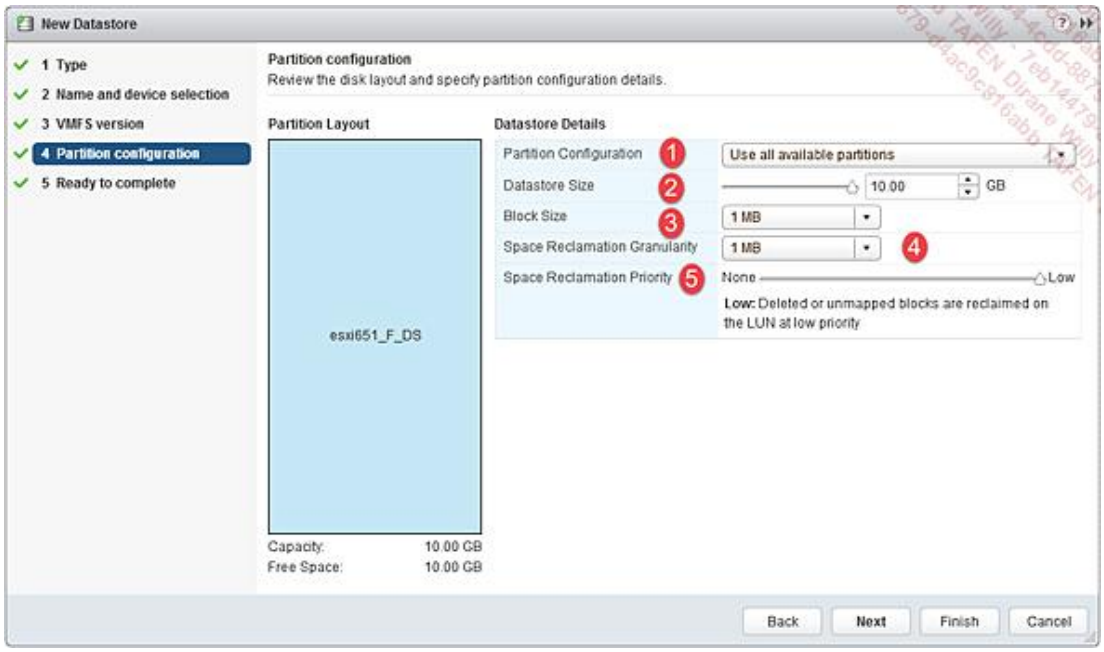


Choisissez la taille de la partition et c'est terminé.



Au choix de la version de VMFS on peut choisir VMFS 6 pour des accès à partir d'hyperviseurs ESXi en version 6.5 minimum.

Il est possible de spécifier la taille du datastore (2), la taille des blocks (3), la finesse de la récupération d'espace - unmap (4) ainsi que sa priorité (5).



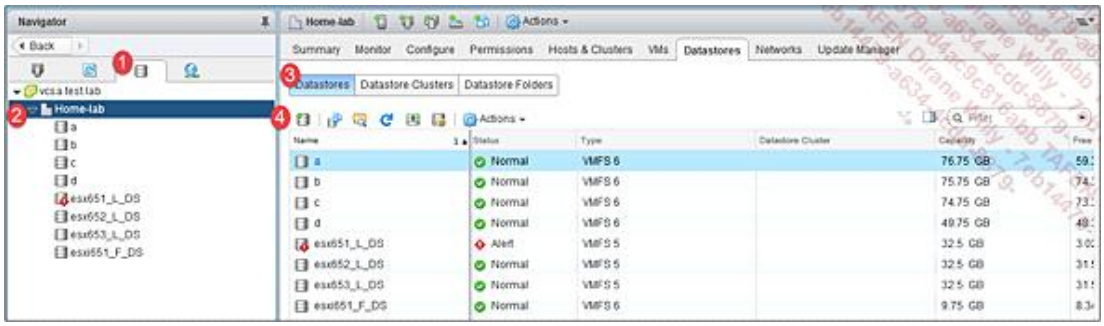
FC

L'ajout de stockage FC diffère de l'ajout de stockage iSCSI/FCoE sur les points suivants :

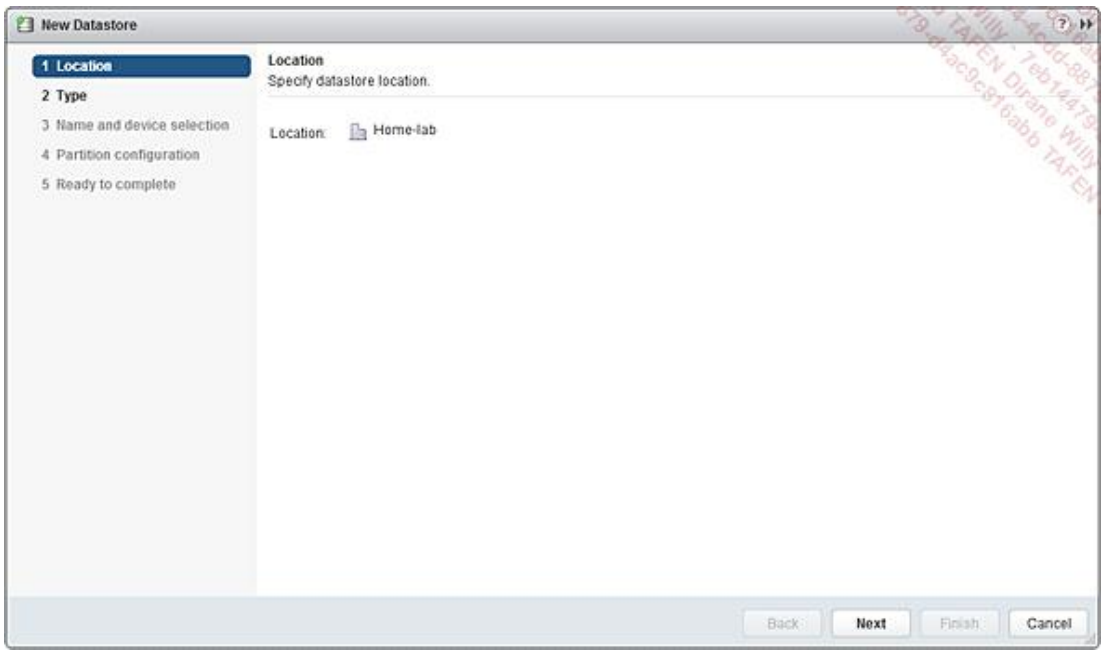
- Il n'est pas nécessaire d'ajouter un initiateur logiciel car il y a obligatoirement un HBA matériel nommé vmhba#.
- Le plus gros de la configuration est effectué au niveau de la fabrique FC (zoning et LUN masking).
- En général un rescan après démarrage de l'ESXi suffit à découvrir les LUNs accessibles.

Création de datastore NFS

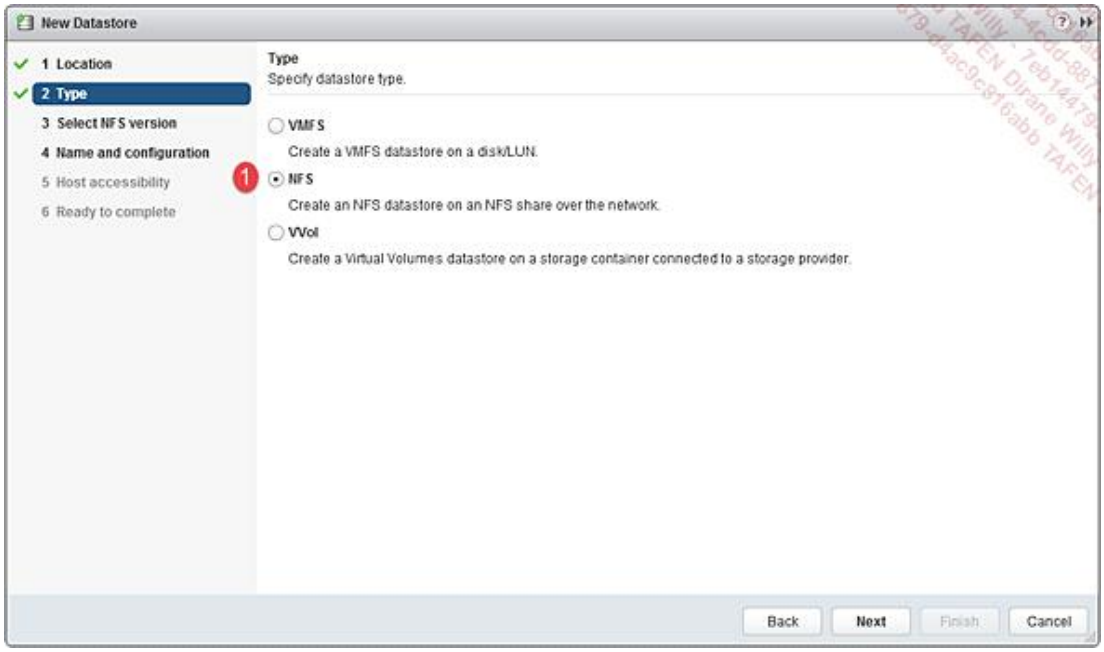
Lors de la configuration d'un datastore NFS, il ne faut pas oublier d'autoriser les ESXi à avoir un accès simultané en écriture au partage.



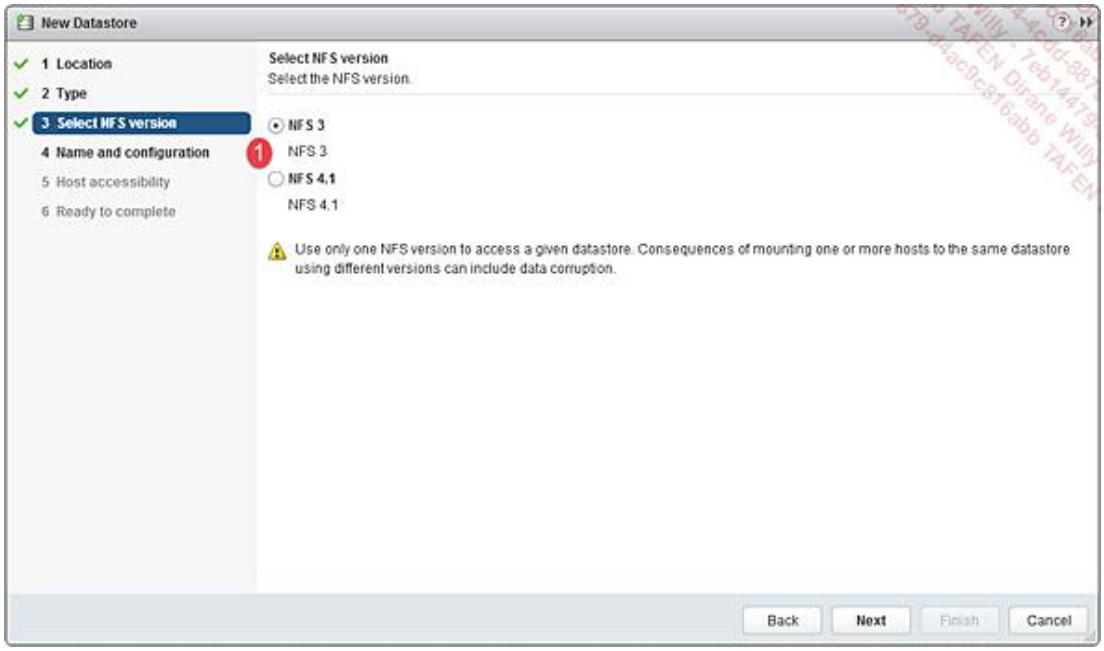
L'ajout de datastore NFS se fait à partir du même menu que pour un datastore VMFS.



Choisissez le datacenter (ici, il n'y en a qu'un).



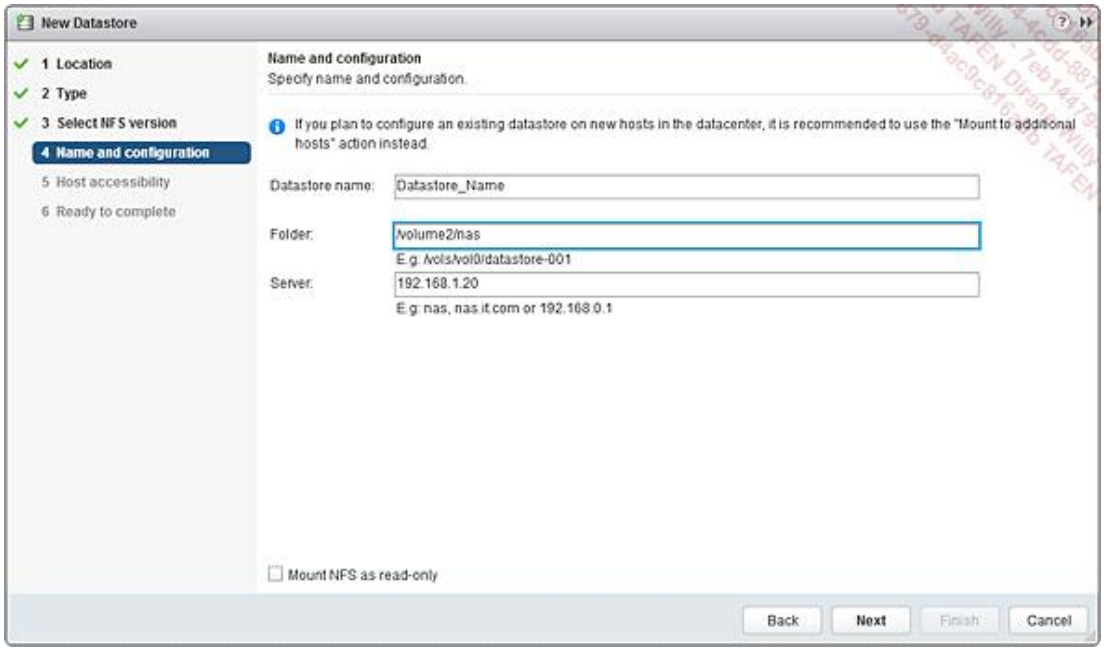
La différence se situe dans le choix du type de datastore : choisissez NFS.



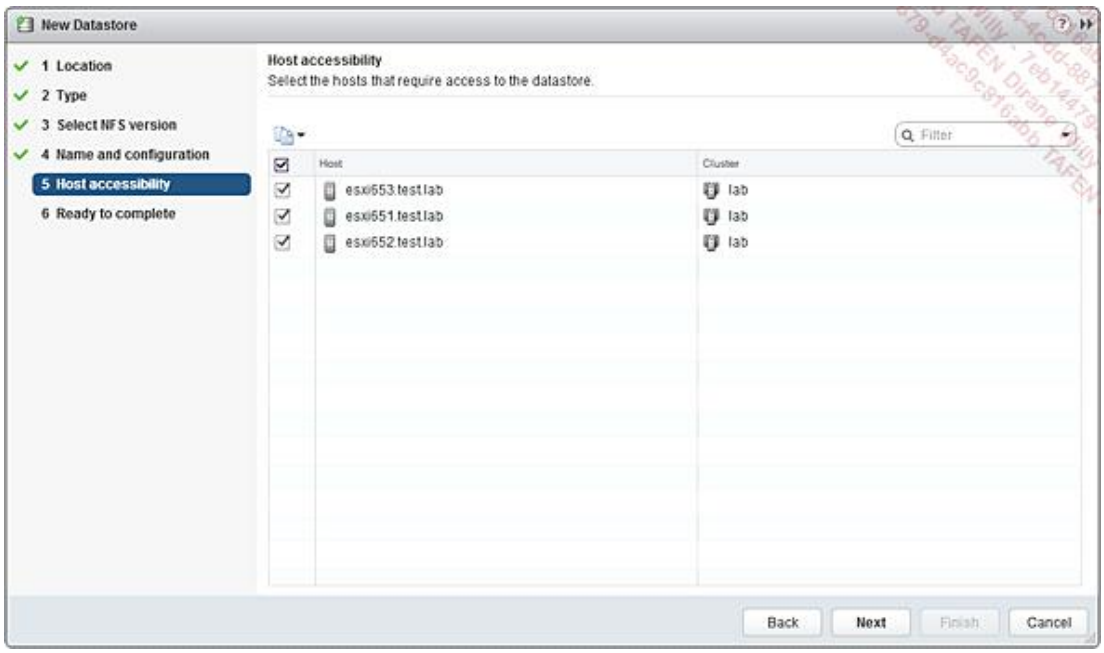
NFS3 ou NFS4.1 : veillez à bien choisir la version selon le serveur. Pour un serveur et pour tous ses volumes, on utilise une seule version de NFS. Les différences entre les versions 3 et 4.1 entraînent des limitations concernant certaines fonctions. Le tableau ci-dessous montre les différences entre le NFS3 et le NFS4.1 :

	NFS3	NFS4.1
Mécanisme de sécurité	AUTH_SYS	AUTH_SYS et Kerberos (krb5 and krb5i)
Algorithme de chiffrement avec Kerberos	N/A	AES256-CTS-HMEC-SHA1-96 et AES128-CTS-HMAC-SHA1-96
Multipathing	Non supporté	Supporté via les trunks
Mécanisme de verrouillage	Verrouillage du côté client	Verrouillage du côté serveur
IPv6	Supporté	Supporté pour AUTH_SYS et Kerberos

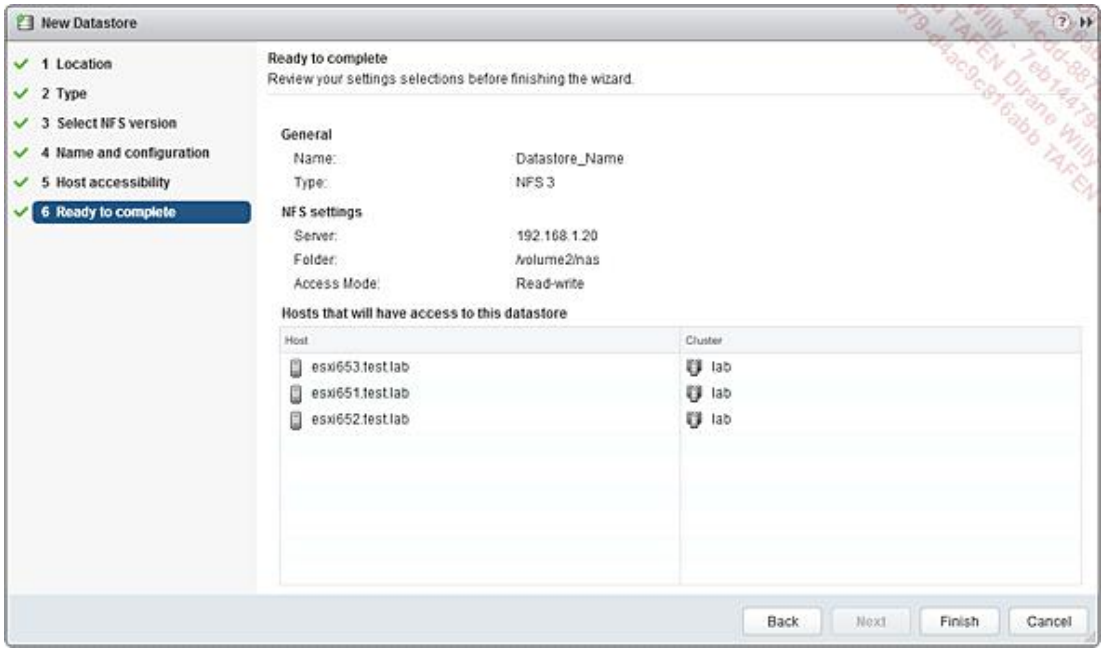
Storage DRS	Oui	Non
Storage I/O Control	Oui	Non
Site Recovery Manager	Oui	Non



Les informations telles que le serveur, le dossier partagé (on dit exporté pour du NFS) ainsi que le nom du datastore (à indiquer par l'administrateur) sont demandées.



Choisissez les serveurs qui auront accès au datastore. Il est à noter que le Web Client permet d'éviter la configuration individuelle pour chaque hyperviseur.



Un récapitulatif est présenté avant la création du datastore (attention, la création du datastore n'est pas du tout liée à la création du volume, nous sommes en mode fichier et il s'agit seulement de donner l'accès aux hyperviseurs).

On voit donc le datastore créé en NFS3. Dans l'exemple, le datastore a été nommé : Datastore_name.

En NFS 4.1, il faut configurer le compte kerberos qui aura accès au partage. Pour cela, il faut pour chaque hyperviseur aller dans les services d'authentification et après la partie lockdown nous avons la possibilité de renseigner le compte qui sera utilisé dans le cadre de l'authentification NFS Kerberos.

Lors de la configuration d'un datastore NFS en v4.1, nous devons sélectionner le bon protocole, puis après avoir nommé le datastore NFS nous devons activer l'authentification via Kerberos.

N'oubliez pas d'ajouter vos hyperviseurs à un domaine Active Directory et de bien configurer la synchronisation de temps (serveur NTP).

c. Virtual Flash File System (VFFS)

Le VFFS est dérivé du VMFS et est spécifique aux périphériques de stockage Flash, qui sont reconnus comme tels par l'ESXi (https://kb.vmware.com/kb/2013188). Il est possible d'agréger le stockage flash de plusieurs ESXi afin de fournir un stockage VFFS utilisable par l'ensemble des ESXi du cluster. Cela permet de faire les actions suivantes :

- vSphere Flash Read Cache (vFRC).
- Host swap cache.

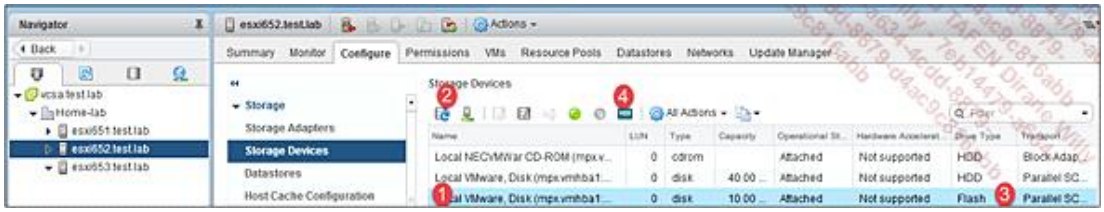
- I/O Cache Filtering.

Dans le cas où un périphérique Flash est détecté comme étant un périphérique de stockage magnétique il est possible de corriger cela de la manière suivante :

Sélectionnez le disque (vous voyez que le type de périphérique est HDD (3)). Cliquez sur le F pour changer le marquage du périphérique de HDD en Flash.

Attention au guide de compatibilité matériel : si les SSD ne sont pas directement reconnus, vérifiez le hardware compatibility guide : http://www.vmware.com/resources/compatibility/search.php

Nous avons un message d'avertissement.



Faites une mise à jour de la vue. Voyez que le type du périphérique de stockage est devenu Flash.

Pour revenir à l'état initial, sélectionnez le périphérique (1) et cliquez sur HDD (4)

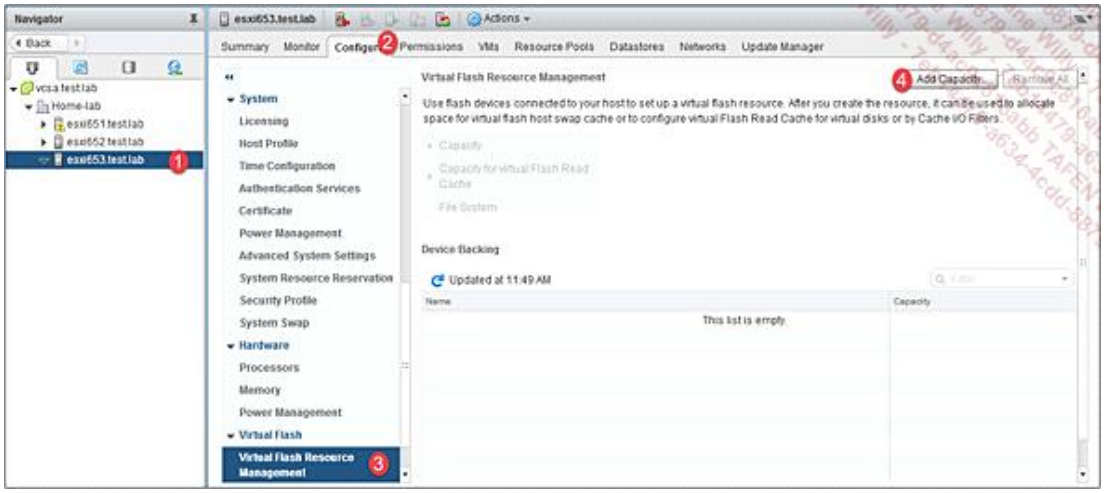
Attention, il ne s'agit pas de datastores permettant de stocker tous les fichiers d'une machine virtuelle. Ici le stockage n'est utilisé que pour créer et utiliser un cache de lecture.

vSphere Flash Read Cache (vFRC)

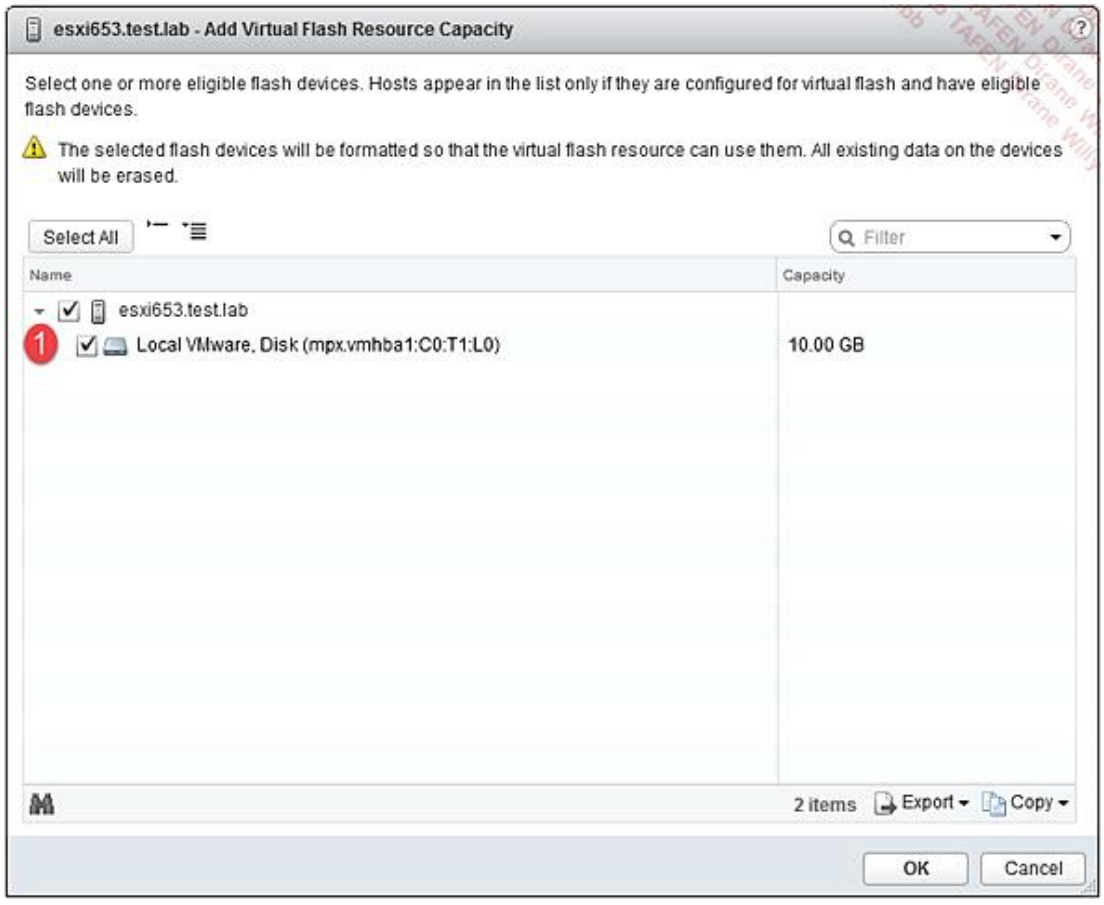
Le vFRC permet d'augmenter la performance de certaines VM (dans le cadre des opérations de lecture) grâce à l'utilisation des périphériques flash. Il prend en charge la mise en cache des données d'écriture immédiate ou de lecture, la lecture se fait à partir du cache si l'information est présente, tandis que les écritures sont envoyées au stockage. Le Flash Read Cache ne prend pas en charge les RDM en mode de compatibilité physique. Le flash Cache

Read est transparent pour vMotion et DRS sous réserve que le serveur hôte cible ait la capacité (présence de stockage flash, et la configuration adéquate). Attention à HA, qui ne redémarra pas une VM avec Flash Cache Read s'il n'y a pas suffisamment d'espace Virtual Flash disponible sur l'hyperviseur concerné. Il faudra la supprimer manuellement. L'utilisation du Flash Read Cache exclut l'I/O Caching Filtering.

En partant du principe que le stockage Flash est bien détecté sur le serveur ESXi, dans l'onglet **Configure** (2), allez jusqu'à **Virtual Flash Resource Management** (3).

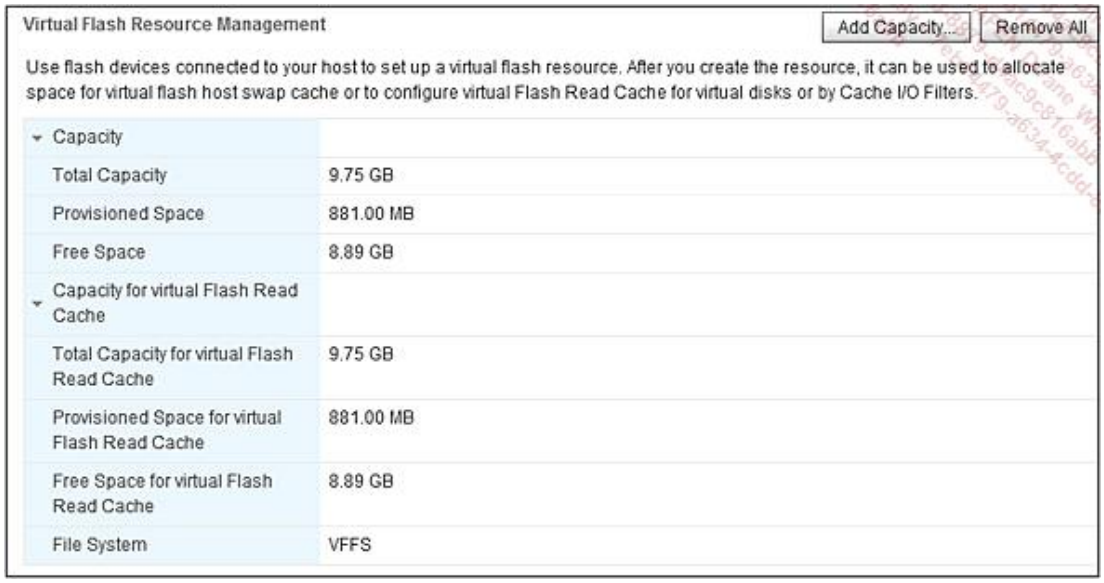


Ajoutez de la capacité pouvant être utilisée comme cache de lecture au niveau de l'hyperviseur (4).

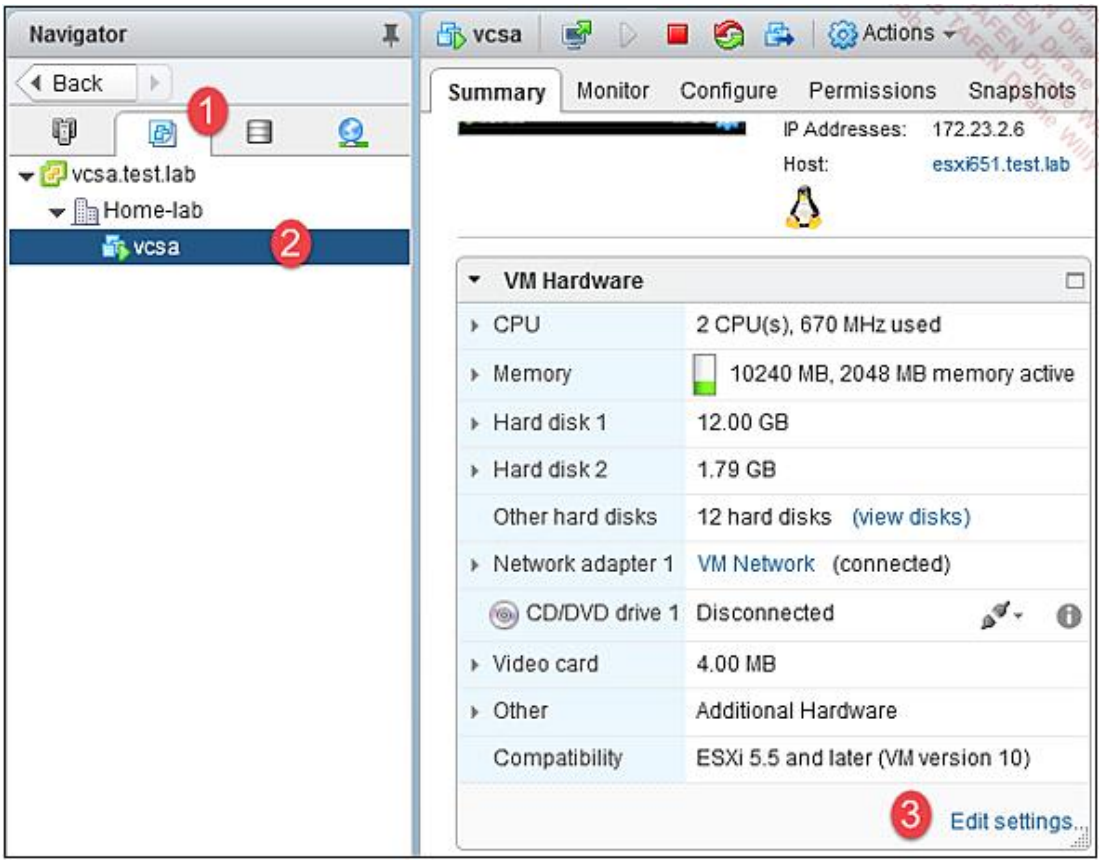


Choisissez le volume concerné (il sera formaté en VFFS).

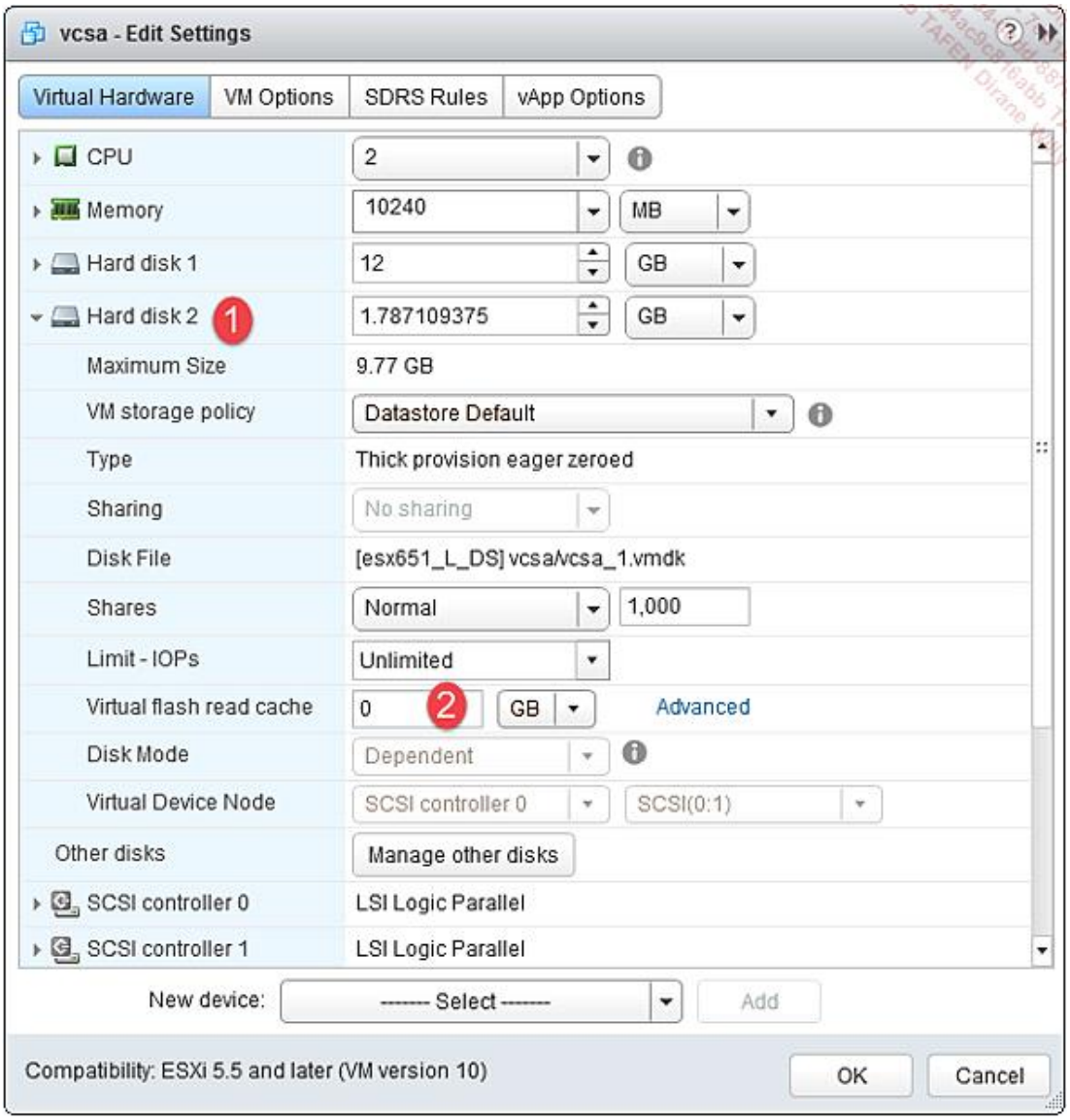
Le menu de Virtual flash présente la capacité disponible au niveau du serveur ESXi.



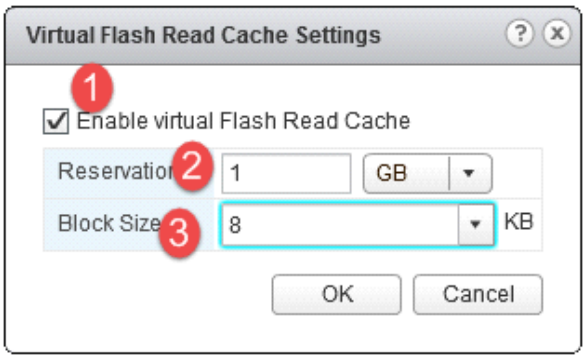
Vous pourrez ensuite ajouter du cache à différentes machines virtuelles.



Sélectionnez la machine virtuelle et éditez les réglages afin de lui ajouter une quantité d'espace pouvant servir de cache d'I/O.



Configurez la quantité d'espace réservé à cette VM à des fins de cache d'I/O.



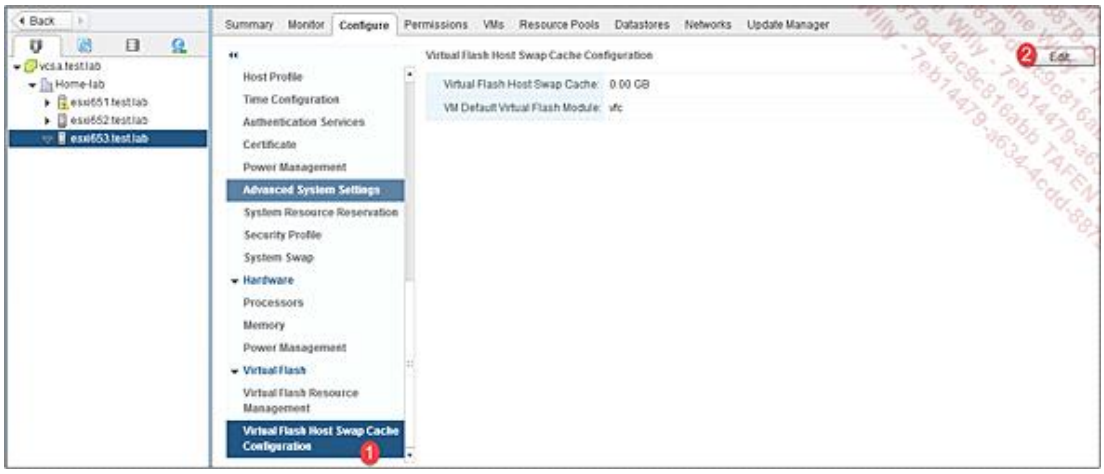
Selon le type de données gérées par la machine virtuelle, il est intéressant de changer la taille des blocs dans le volume VFFS. Cela nécessite une connaissance du fonctionnement des applications dans le système invité.

Host swap cache

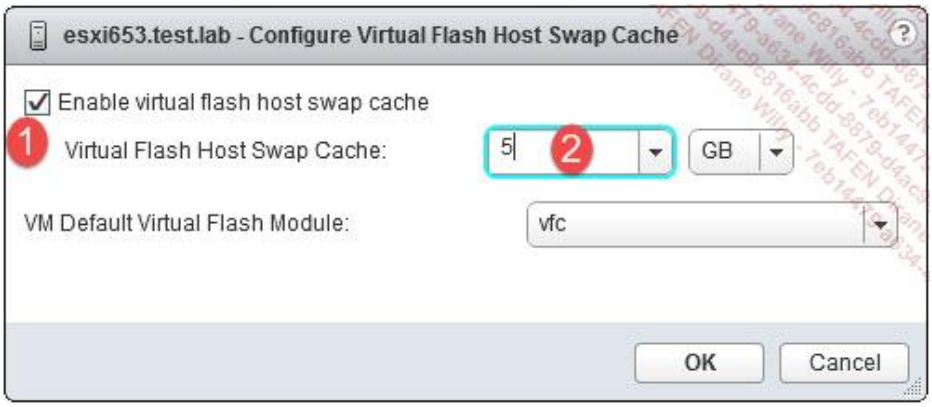
Le Host swap cache est une réservation d'espace dédié à l'ESX en cas de surcharge mémoire après que les

diverses techniques d'économie et récupération de la mémoire aient échouées ou ne soient pas suffisantes. Le stockage flash étant bien plus performant que les disques durs traditionnels, l'effet du swapping peut être moins impactant que prévu.

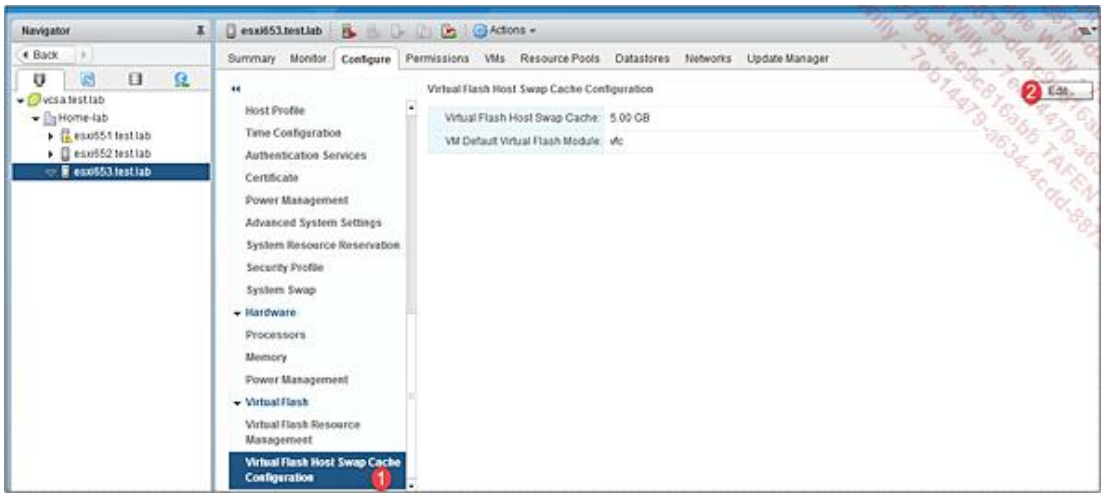
Le Host Swap Cache se configure par hôte ESXi de la manière suivante :



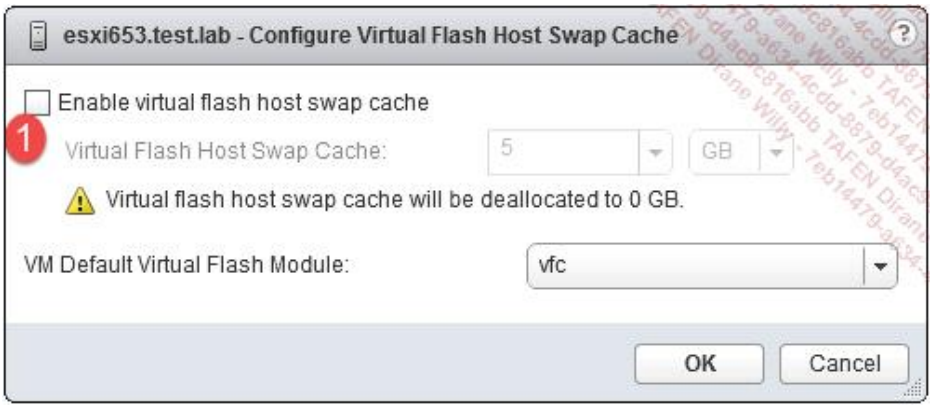
Sélectionnez le serveur hôte et dans l'onglet **Configure** , choisissez **Edit** (2) dans la partie **Virtual Flash Host Swap Cache Configuration** (1).



Activez le cache et configurez la capacité.

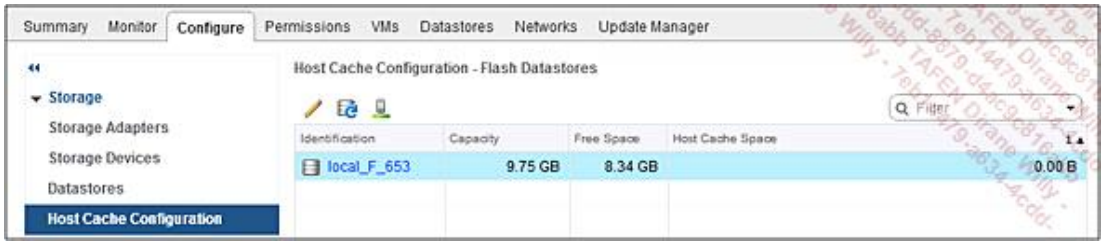


Si vous désactivez le cache, celui-ci sera vidé et non utilisable par le serveur ESXi.

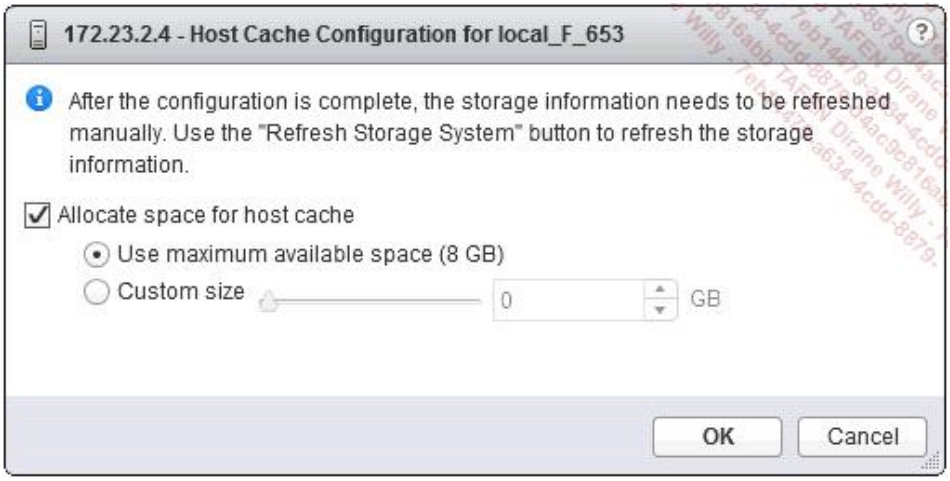


Host Cache

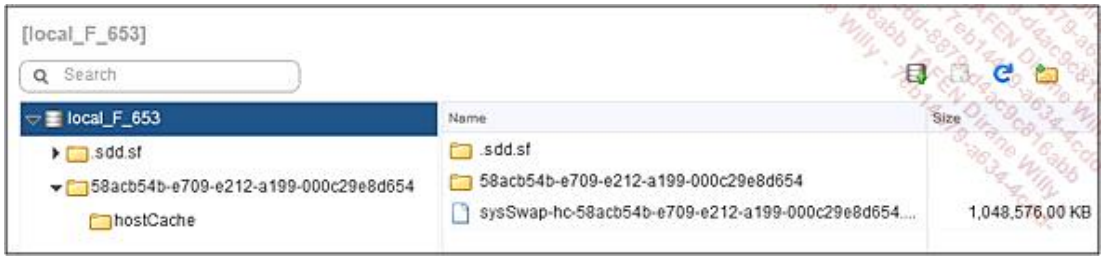
Cette option nécessite, outre un SSD, que ce dernier soit formaté en VMFS (voir le point concernant la création d'un datastore VMFS). Ce SSD formaté au format VMFS est visible pour le serveur qui le contient dans Host cache configuration.

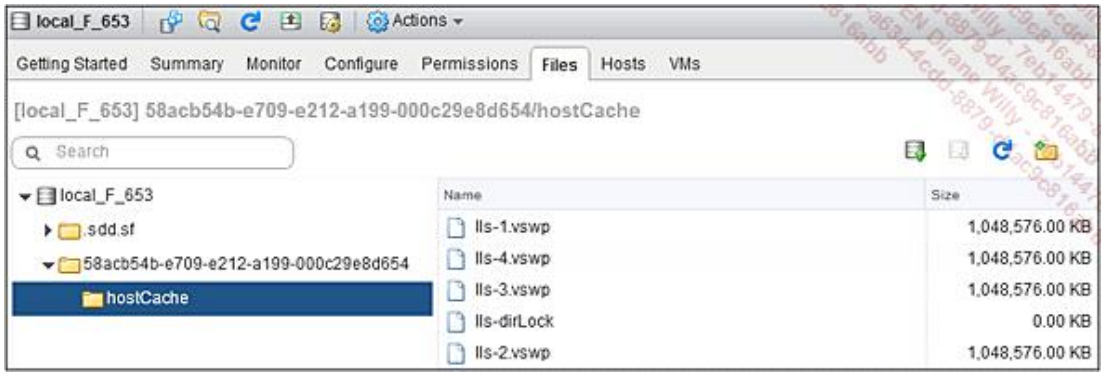


Lorsque vous éditez les propriétés de ce datastore, vous pouvez définir la quantité d'espace pouvant être allouée pour créer le cache de l'hyperviseur.



On peut voir dans le datastore configuré que le serveur ESXi utilise le swap configuré sur le datastore alloué. Si vous regardez de plus près, vous voyez à la racine un fichier nommé sysSwap-hc-GUID. Il existe un répertoire au nom du GUID.





I/O Cache Filtering

I/O Cache Filtering rentre dans le VAIO afin d'être utilisé en tant que tampon entre les machines virtuelles et l'ESXi. Il s'appuie sur les politiques de stockage et VASA. Dans ce cas le volume VFFS est utilisé pour donner un cache afin de stocker les données de machines virtuelles avant qu'elles soient envoyées à la baie de stockage. Le volume Flash ne peut pas être utilisé dans le cadre du vSAN et du VMFS.

3. Virtual Volume

Le vVol est la technique de VMware permettant l'utilisation de manière agnostique du stockage, permettant de passer les fichiers constituant les machines virtuelles à des objets connus et manipulables par la baie. Du point de vue de l'hyperviseur, il ne s'agit que d'un nouveau type de datastore, permettant l'accès aux machines virtuelles. La nouveauté avec vSphere 6.5 et le VASA 3 est que le vVol supporte maintenant la réplication de baie à baie.

a. Le Storage Container (SC)

Le vVol s'appuie sur un container de stockage (*Storage Container*). Là où historiquement l'équipe stockage fournissait des LUNs à dimension figée à formater avant utilisation, le Storage Container, lui, est dynamique et ne nécessite aucun formatage, il permet aussi de faire de l'isolation entre client ou service en n'allouant qu'une partie du volume de stockage.

b. Le Protocole Endpoint (PE)

Pour communiquer avec l'extérieur, le Storage Container s'appuie sur le Protocole Endpoint. Il agit en tant que proxy pour le cheminement des I/O entre la baie et l'hyperviseur. Il s'agit de ce que l'on nomme le Data Path. Lorsque le PE est utilisé par un protocole de type SCSI, le PE est vu comme une LUN. Attention, aucune donnée n'est écrite dedans ! Lorsque le PE utilise le NFS, il est vu comme étant un point de montage (un partage avec une adresse IP).

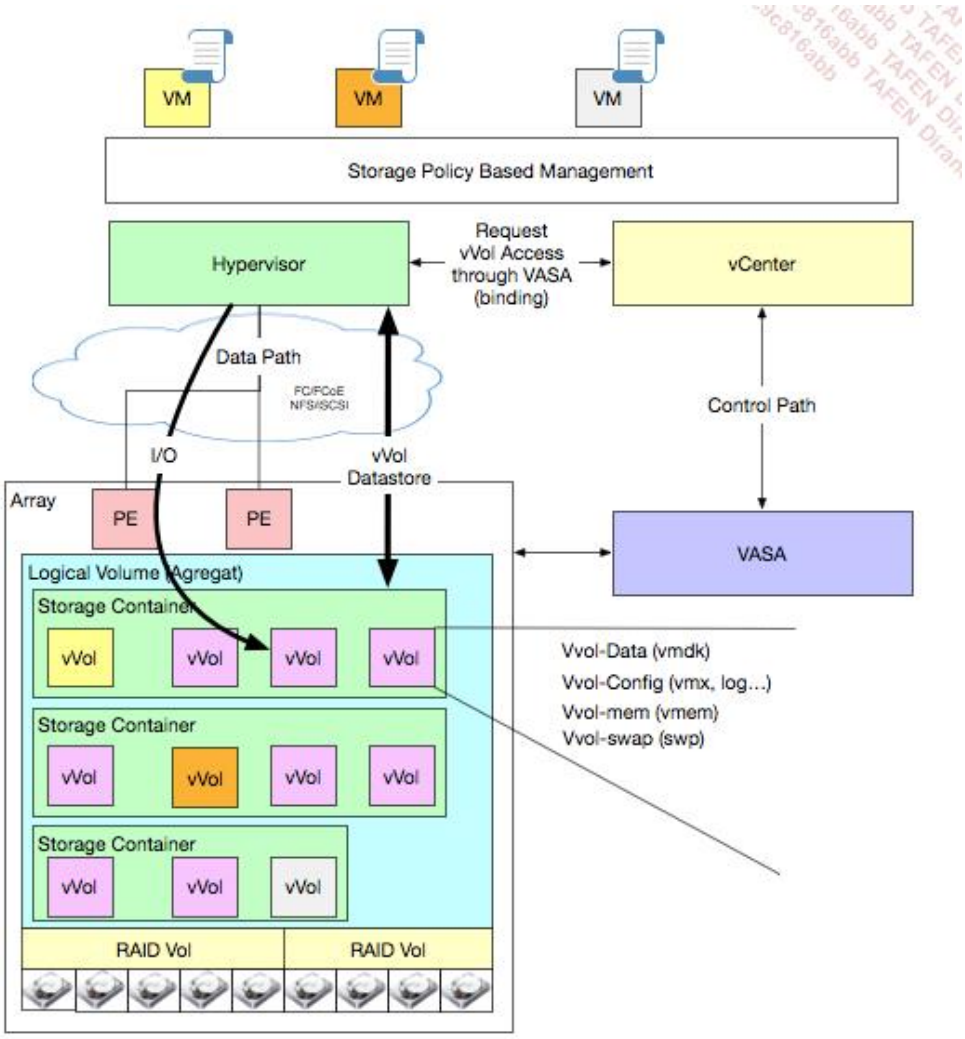
c. Le virtual Volume

Le vVol est en lui-même un objet qui peut être de plusieurs types :

- Config : ce type correspond aux fichiers de configuration (vmx), metadata, fichier de description des disques, fichiers de log. Lorsque l'hyperviseur utilise le protocole SCSI pour se connecter au stockage, les objets de type Config sont formatés (VMFS). Dans le cas de l'utilisation du protocole NFS, ce type d'objet est vu comme un répertoire. Ce type d'objet est provisionné en mode thin.
- Data : ce type correspond aux fichiers VMDK. Ce type d'objet est directement présenté aux machines virtuelles comme des disques SCSI. Ils peuvent être provisionnés en mode Thin comme en mode Thick.
- Swap : ce type d'objet se crée au démarrage de la machine virtuelle. Il est provisionné en mode Thick (par défaut), sa taille se base sur la taille de la réservation de la mémoire au niveau de la machine virtuelle.
- Mem : ce type d'objet représente la mémoire (vmem) ainsi que les snapshots (avec le snapshot de la mémoire). Le provisionnement est de type thick.
- Other : ce type correspond à des fonctionnalités spécifiques (par exemple CBRC - *Content-Based Read Cache* qui est une fonction de cache pour Horizon View...).

d. VASA et SPBM

Précédemment dans ce chapitre nous avons parlé des composants VASA et SPBM. Pour rappel VASA ou *VMware API for Storage Awareness* qui permet de faire prendre conscience au vCenter des capacités de la baie de stockage (type de disque, déduplication, compression...). SPBM ou *Storage Policy Based Management* sont l'équivalent VMware des GPO Windows appliqué au stockage. On peut ainsi appliquer les politiques de stockage (SPBM) au niveau de la machine virtuelle en se basant sur les capacités de la baie de stockage remontées par VASA.



À titre d'exemple, Dell fournit un document expliquant la mise en place des vVOLs avec des baies EMC VMAX étape par étape (https://www.emc.com/collateral/white-papers/h14576-vmware-virtual-volumes-emc-vmax3-vmax-all-flash.pdf).

4. vSAN - l'hyperconvergence selon VMware

a. Introduction

Aujourd'hui, le modèle des infrastructures distribuées est largement utilisé quand il s'agit de déployer des ressources techniques (calcul, réseau, stockage...).

Quand vous y pensez, la virtualisation de serveurs a pourtant remis en cause une partie de ce modèle bien établi et largement déployé. Par exemple mettre en place un serveur physique par usage (serveur d'annuaire, serveur mail, serveur web...) et potentiellement par client. Malgré cet ancrage profond au sein des architectures (et des esprits), nous ne déployons plus nos infrastructures techniques comme quinze ans auparavant.

Cependant, il est intéressant de constater que quand bien même les serveurs physiques à usage dédié sont désormais rares, nous continuons à utiliser des serveurs dédiés à un nombre restreint d'applications mais d'une autre manière. La consolidation des ressources est passée par là. Au lieu d'utiliser un serveur physique exécutant un OS client type (Windows, Linux, BSD...), nous faisons appel à un hyperviseur.

Bien entendu, la seule idée de faire converger des ressources en un même point pose la question de la disponibilité, de la sécurité des données, au fond de savoir si vous souhaitez mettre « tous vos œufs dans le même panier ». Quel que soit le modèle utilisé, la question se pose et est adressée. L'exemple du serveur physique non virtualisé obligeait en son temps les administrateurs à déployer plusieurs serveurs (d'annuaire par exemple) pour assurer une haute disponibilité. Les hyperviseurs peuvent assurer une haute disponibilité du serveur virtualisé, allant de la simple disponibilité jusqu'au basculement quasi immédiat selon le besoin. Que le serveur soit virtualisé ou non, ces modèles de déploiement n'auraient pas été considérés sérieusement par les professionnels s'ils n'avaient pas offert cette fonctionnalité essentielle.

En complément de la disponibilité, ce trait « distribué » se retrouve dans nos façons d'utiliser ces hyperviseurs par le biais de clusters. Là aussi, nous distribuons la charge parmi les membres d'un même groupe d'hyperviseurs.

L'hyperconvergence constitue une autre évolution du modèle distribué vers un modèle où nous consolidons d'autres ressources utilisant auparavant des équipements dédiés. De la perspective des hyperviseurs, nous offrons plus de responsabilités à ces derniers, dans des domaines où la gestion et l'opération étaient encore externalisées. Dès lors, un ensemble d'hyperviseurs est ainsi capable d'opérer un réseau ou un stockage **entièrement** virtualisé ! Ainsi, les ressources de virtualisation, de stockage, de réseau peuvent être portées par une seule et unique architecture !

Pour faire un parallèle avec les infrastructures dont nous avons parlé précédemment, une architecture convergée doit vous évoquer (à nouveau) la concentration des ressources et donc la potentielle faiblesse de ce modèle. Et pourtant, à nouveau, cela ne veut pas dire pour autant qu'il s'agisse de solutions sans mécanisme de haute disponibilité, nous y reviendrons plus tard dans ce chapitre.

Une architecture convergée implique une utilisation beaucoup plus efficace des ressources, de l'espace dans lequel sont installés les serveurs et une simplification de l'administration de l'architecture. Concernant ce dernier point, ce type d'architecture permet aux administrateurs de la ferme d'hyperviseurs d'installer, de configurer, de maintenir et surveiller l'ensemble de leur parc (y compris le réseau ou le stockage virtualisé) via une interface unifiée.

L'hyperconvergence est la révolution qui est en train de changer le monde du stockage. À l'instar de ses concurrents (Nutanix, Simplivity...), VMware propose une solution d'hyperconvergence, via la technologie nommée vSAN.

b. VMware vSAN

vSAN permet de faire porter le stockage non plus par une ou plusieurs baie(s) de stockage traditionnelles et ses disques mais par les hyperviseurs eux-mêmes et leur stockage local.

L'hyperviseur devient donc en charge du stockage de données (machines virtuelles, fichiers annexes, images ISO...). La disponibilité et l'intégrité des données sont assurées contre les défaillances de disques, d'hôte, de réseau, ainsi que l'optimisation des I/O selon les accès disques. Une fois la solution vSAN configurée, il apparaît aux hôtes membres du cluster une banque de données (*datastore*) partagée et accessible à tous. Le vMotion est rendu possible, ainsi que le Storage vMotion si vous souhaitez migrer les données de ou vers un datastore vSAN.

Pour être exhaustif, voici un récapitulatif des fonctionnalités clefs et capacités de vSAN 6.5 citées par VMware :

- **Intégré avec vSphere** : vSAN est intégré dans le noyau vSphere, ce qui permet d'optimiser le chemin I/O pour offrir les performances les plus hautes, avec une emprunte processeur et mémoire minimale.
- **Optimisé par flash** : vSAN accélère les lectures / écritures et en général les opérations I/O en utilisant un cache reposant sur de la mémoire flash. En parallèle, vSAN est capable de gérer les clusters all-flash, c'est-à-dire des clusters dotés uniquement de stockage flash. Une nouveauté de vSAN 6.5 est la prise en charge des architectures all-flash dans la licence de base.
- **Augmentation du pool de ressources vSAN non-disruptive** : augmentez les performances et la capacité de votre datastore (ajout de disque, ajout d'hôte) vSAN à chaud, sans arrêt de production.
- **Déduplication et compression** : dans les configurations all-flash, la déduplication et la compression optimisent la capacité du stockage, avec des empreintes processeur et mémoire minimale. Cette fonctionnalité est disponible à partir de la licence « Advanced ».
- **Erasure Coding** : la fonctionnalité « Erasure Coding » permet d'aller jusqu'à doubler la capacité de stockage, en gardant les données de résilience intactes. Ce mode est capable de supporter la simple (RAID-5) ou double parité (RAID-6) avec au minimum 4 (3 bandes de données + 1 parité) ou 6 (4 bandes de données + 2 parités) hôtes. Si la configuration le permet, le mode « Erasure Coding » peut être utilisé en alternative au mode vSAN par défaut « miroir », type RAID-1. Même si cela n'est pas requis, il est recommandé (https://blogs.vmware.com/virtualblocks/2016/02/12/the-use-of-erasure-coding-in-virtual-san-6-2/) d'utiliser des architectures all-flash, principalement à cause de la surcharge apportée par un tel mécanisme (calcul et stockage de parité, reconstruction de donnée…). Prêtez également attention à la licence que vous achetez si vous comptez utiliser cette fonctionnalité, celle-ci est disponible à partir de la version « Advanced ».
- **Qualité de service (QoS)** : vous pouvez limiter, contrôler et surveiller les IOPS au niveau d'une machine virtuelle, limitant l'effet de machines trop gourmandes en IOPS. Attention cependant, cette fonctionnalité n'est disponible que sur la licence « Enterprise » de vSAN.
- **Service de santé vSAN** : le service de santé vSAN surveille la compatibilité matérielle, les performances, la capacité du stockage et présente des diagnostics types fournis par VMware en cas de problème.
- **Target iSCSI** : vSAN 6.5 permet l'exposition de l'espace de stockage en tant que target iSCSI. Tout cela administré dans le même espace de gestion, par le biais de vCenter.
- **2-Node « Direct Connect »** : vous pouvez connecter directement deux hôtes vSAN à l'aide d'un câble croisé, évitant de passer par un commutateur.
- **Modèle de déploiement pour les filiales** : VMware propose une formule de déploiement pour les filiales, nommée ROBO (*Remote Office / Branch Office*). En résumé, vous pouvez porter une architecture locale vSAN à partir de deux nœuds localisés dans la filiale et d'un nœud témoin hébergé sur l'infrastructure centrale, sous conditions d'une liaison de qualité entre la filiale et l'architecture centrale. Des licences spécifiques pour ces usages sont proposées par VMware.
- **API vSAN / PowerCLI** : vSAN propose des API vSAN et des commandes PowerCLI (PowerShell) pour vous permettre d'automatiser vos déploiements, installation, configuration et maintenance de vos clusters. Ainsi, sans aucune formation, un opérateur peut administrer un cluster vSAN à l'aide de scripts écrits par vos soins.
- **Politiques de stockage appliquées aux machines virtuelles** : vSAN permet d'utiliser des politiques de stockages appliquées directement aux machines virtuelles pour automatiser et équilibrer les opérations utilisant des

ressources de stockage. Ainsi, grâce aux politiques, vous pouvez décider qu'une machine virtuelle doit utiliser tel mode de redondance et exploiter un stockage type tier flash ou mécanique. C'est d'ailleurs à l'aide de ces politiques que vous pouvez également limiter les IOPS de lecture ou écriture d'une machine virtuelle.

- **Tolérance à la panne et disponibilité avancée** : vSAN exploite les mécanismes RAID et de protection de cache pour assurer que la donnée stockée ne sera pas perdue, quand bien même un disque, un hôte ou le réseau venait à dysfonctionner. vSAN supporte les autres fonctionnalités de disponibilité vSphere comme Fault Tolerance et HA. Vous pouvez également utiliser vSphere Replication pour vSAN, vous permettant d'offrir une protection équivalente à un RPO (*Recovery Point Objective* - maximum de donnée(s) que vous êtes prêt à perdre après une interruption de service) de 5 minutes.
- **Cluster étendu vSAN** : vous pouvez étendre un cluster vSAN à deux sites géographiquement éloignés, sous conditions de la bande passante disponible et la latence de la connexion. Cet usage vous permet de minimiser le potentiel arrêt de production et perte de donnée en cas de défaillance d'un site entier. Vous devez utiliser une licence « Enterprise » pour utiliser cette fonctionnalité.

Après avoir décrit les principales fonctionnalités de vSAN, un peu d'historique sur l'évolution de cette technologie depuis son annonce il y a quelques années.

vSAN a commencé son parcours en étant annoncé lors du premier jour du VMworld 2013, en août de la même année. Nous avons parlé d'hyperconvergence dans la partie précédente, sachez que, de façon très intéressante, l'annonce initiale de vSAN coïncide d'ailleurs avec une autre, celle de NSX, la solution de virtualisation de réseau et sécurité de VMware (SDN - *Software Defined Networking*).

Pour revenir à l'annonce concernant vSAN, VMware a également mis à disposition en septembre 2013 une bêta publique, se basant sur la nouvelle version de l'hyperviseur qui venait tout juste d'être de sortir, vSphere 5.5. Il faudra attendre avril 2014 pour voir la publication officielle de la première version de vSAN (vSAN 1.0), intégrée dans vSphere 5.5 Update 1 (https://www.vmware.com/support/vsphere5/doc/vsphere-esxi-55u1-release-notes.html).

« Intégrée » ai-je dit ? Oui ! Vous le verrez quelques pages plus loin, vSAN est directement intégré à la build des hyperviseurs. Si vous avez une version vSphere supérieure ou égale à la 5.5 U1, vous n'avez donc rien à installer de plus ! Vous l'avez également deviné, le jour où vous mettez à jour votre build d'hyperviseur, vous mettez à jour **automatiquement** vSAN. On a vu plus compliqué en termes de processus d'installation et de mise à jour.

vSAN est aujourd'hui adopté par 7 000 entreprises (http://www.yellow-bricks.com/2017/01/27/__trashed/), en à peine trois ans à l'écriture de ces lignes. La quête continuelle de l'optimisation du TCO va continuer de pousser les entreprises à s'intéresser de plus près aux technologies d'hyperconvergence, dont vSAN.

Les versions de vSAN se sont succédées avec celles des releases vSphere, un tableau récapitulatif des versions et nouveautés vous est présenté ci-dessous :

Version	Version vSphere minimale	Principales nouveautés
vSAN beta	5.5	
1.0	5.5 U1	
6.0	6.0	Configurations All-Flash Domaines de panne Domaines de chiffrement / hachage matériel Nouveau format de disque
6.1	6.0 U1	ROBO vSphere Replication Cluster étendu Fault Tolerance Service de santé w/ vCenter Plug-in vSAN vROM/vROPS
6.2	6.0 U2	Déduplication et compression RAID-5/6 (Erasure Coding)

		Qualité de service - IOPS Service de performance Service de santé amélioré
6.5	6.5	APIs vSAN / PowerCLI améliorées Nouveau modèle de licensing (all-flash...) iSCSI Target Containers Support des disques 512e

Voici le modèle de licensing qui s'applique aujourd'hui pour un cluster vSAN 6.5 :

Fonctionnalité vSAN	Standard	Advanced	Enterprise	ROBO Std.	ROBO Adv.
Management basé sur les politiques de stockage	☑	☑	☑	☑	☑
Caching flash en lecture & écriture	☑	☑	☑	☑	☑
RAID distribué	☑	☑	☑	☑	☑
Switch virtuel distribué	☑	☑	☑	☑	☑
Snapshots et clones vSAN	☑	☑	☑	☑	☑
Rack Awareness Availability	☑	☑	☑	☑	☑
vSphere Replication	☑	☑	☑	☑	☑
Somme de contrôle logicielle	☑	☑	☑	☑	☑
Architectures all-flash	☑	☑	☑	☑	☑
Service iSCSI Target	☑	☑	☑	☑	☑
Déduplication et compression		☑	☑		☑
RAID-5/6 Erasure Coding		☑	☑		☑
Cluster étendu (multisite)			☑		
QoS - Limites IOPS			☑		

Comme évoqué précédemment, l'un des grands changements dans le mode de licensing est la prise en charge des architectures all-flash dès la licence « Standard », quand il fallait auparavant une licence « Advanced » pour en bénéficier (en vSAN 6.2 notamment). Cette partie introductive étant terminée, intéressons-nous de plus près aux modes de fonctionnement de vSAN.

c. Mode de fonctionnement

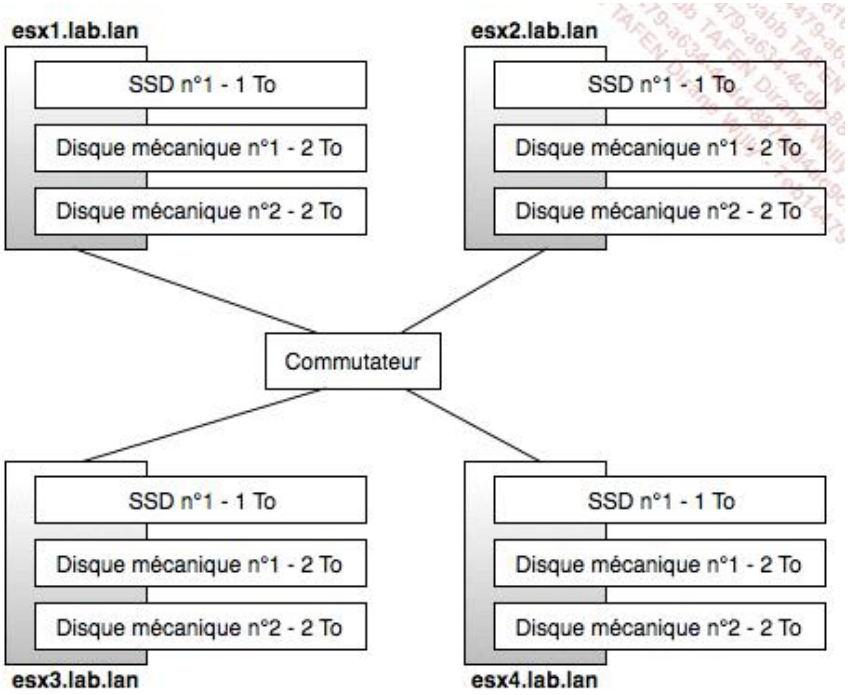
vSAN peut être déployé de plusieurs façons, selon le besoin et les ressources disponibles. Nous en retiendrons trois dans cette section, à savoir :

- un déploiement typique avec 4 nœuds

- un autre plus particulier avec (seulement) 2 nœuds
- un cluster étendu entre deux sites

Pour concentrer nos démonstrations sur les concepts de vSAN, nous considérons ici que les hôtes appartiennent au même sous-réseau et sont reliés par un commutateur Gigabit.

Le déploiement avec 4 nœuds est illustré avec la figure ci-dessous :



La redondance de données est assurée dans ce scénario à l'aide d'un mode mirroring, type RAID-1.

Plusieurs questions se posent ici. Pourquoi mélangeons-nous un SSD avec des disques mécaniques sur chaque hôte ? Pourquoi mettre un SSD de cette taille ? Quelle est la capacité utilisable ? Pour répondre à ces questions, il faut comprendre comment fonctionne vSAN.

vSAN différentie deux types de ressources (tier), celles destinées à la mise en cache (*caching*) et celles destinées à stocker de la donnée (*capacity*). vSAN exploite ces deux types de ressources simultanément pour faire fonctionner le stockage partagé.

Le but de la couche caching est de permettre d'accélérer les opérations de lecture / écriture quand celui de la couche capacity est de permettre le stockage de données. Un cluster mélangeant d'ailleurs deux types de disques (SSD + HDD) est appelé hybride. Vous l'avez deviné, dans le déploiement d'un cluster hybride, nous aurons intérêt à associer des SSD dont les performances sont élevées à la couche caching, quand nous utiliserons les disques mécaniques en couche capacity. Par ailleurs, il est intéressant de voir que la partie de mise en cache réserve 70 % de la capacité de cache disponible pour les accès lecture quand le reste est consacré aux écritures (http://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/products/vsan/virtual-san-6.2-design-and-sizing-guide.pdf) .

La question de la taille du SSD disposé sur chaque hôte suit une recommandation de VMware, qui consiste à disposer d'au minimum 10 % de la capacité totale pour la mise en cache, comparé à l'espace utilisé pour stocker des données, sans tenir compte de l'espace consommé par la redondance d'information.

Pour calculer ce besoin, vous devez évaluer l'utilisation potentielle de votre datastore par vos machines virtuelles.

Ici nous prévoyons d'utiliser dans notre *cluster* 75 machines virtuelles qui occuperont en moyenne 100 Go d'espace disque. La formule est donc simple :

$$75\ machines\ virtuelles * 100\ Go = 7{,}5\ To$$

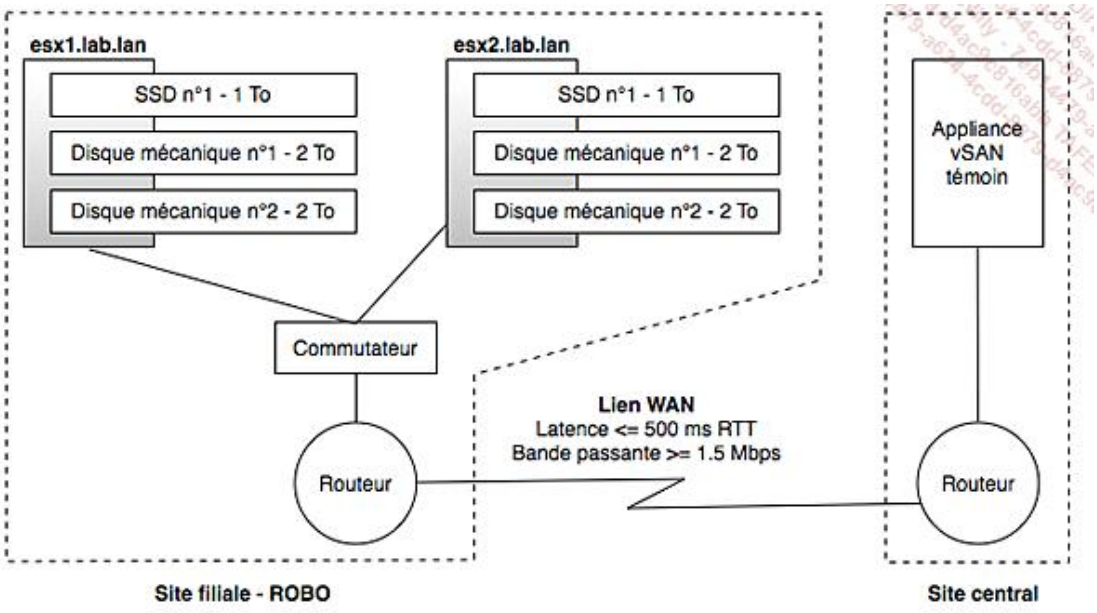
Si l'on calcule 10 % de cette capacité, nous obtenons 750 Go. N'importe quel SSD avec cette capacité (ou plus) aurait ici fait l'affaire. Les SSD 1 To étant plus courants, le choix s'est porté sur cette capacité.

Attention cependant, quand nous faisons ce calcul, nous partons du principe que l'ensemble des hôtes et disques fonctionneront de façon nominale. Mais il faut penser à ce qui peut se produire en cas de panne. Ainsi, en cas de défaillance d'un hôte, l'espace disponible sur la banque de données partagée en sera réduit. Dans le cas de notre configuration, celui-ci « tombera » au maximum à 12 To et non plus 16 To comme auparavant, hors protection.

D'autant plus que cet espace peut être utilisé par vSAN pour assurer la protection des données après une panne. Dans notre cas, vSAN sera incapable de protéger l'ensemble de vos données si vous dépassez les 6 To de données à protéger. C'est pourquoi il est toujours recommandé de prévoir au minimum, l'équivalent de la capacité d'un hôte en espace libre. D'ailleurs, le cluster déclenchera automatiquement un avertissement si l'utilisation de votre banque de données (données + protection) dépasse 80 % pour vous informer de l'arrivée d'un potentiel problème en cas de panne d'un hôte. L'avertissement deviendra un « danger » si le seuil de 90 % est dépassé.

Cette topologie supporte la panne d'un hôte ou d'un disque (FTT - *Failures To Tolerate* = 1).

Le déploiement avec 2 nœuds diffère en plusieurs points avec l'approche 3 nœuds, en voici un aperçu via un schéma :



Il est possible d'utiliser depuis vSAN 6.1 un type de déploiement nommé ROBO (*Remote Office Branch Office*). Un modèle, comme son nom l'indique, à destination des filiales qui souhaitent porter des clusters vSAN, sans avoir à disposition un data center. Le modèle ROBO présenté consiste en deux nœuds. Avant l'arrivée de ce modèle, vous deviez avoir au minimum trois nœuds pour former un cluster vSAN.

À cette étape de la lecture, vous pouvez vous demander comment un cluster peut fonctionner avec uniquement 2 nœuds. En théorie, un cluster a constamment besoin d'une majorité, appelée communément un quorum, pour fonctionner, déterminer le statut de ce dernier et de ses ressources associées. Cette majorité est constituée par les membres de ce même cluster qui votent. Cependant, une majorité ne peut pas être déterminée par deux

hôtes, le cluster se retrouve dans une situation de statu quo.

VMware vous propose d'utiliser un hôte ESXi pour jouer le rôle de témoin (*witness* en anglais). Son rôle est d'assurer le quorum en cas de panne d'un des deux hôtes mais également de stocker les objets témoins des données ainsi que les métadonnées du cluster vSAN.

L'utilisation d'un hôte physique pour ce rôle semble disproportionnée car celui-ci doit être dédié à cette tâche. En d'autres mots, vous ne pouvez pas exécuter de machines virtuelles sur ce dernier.

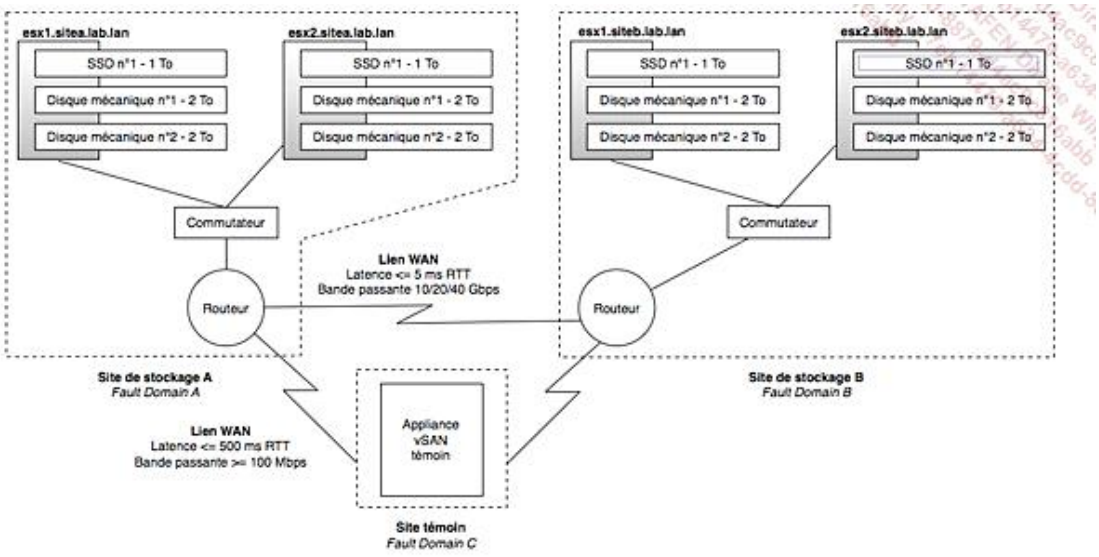
C'est là où intervient l'appliance vSAN témoin que vous voyez disposée sur le site central dans le schéma. L'idée est de virtualiser un hôte ESXi pour occuper ce rôle de témoin, ce qui permet l'utilisation plus souple de l'hôte support. Nous faisons donc de la virtualisation de serveur ESXi (*nested*) supportée par VMware pour un usage en production ! L'appliance témoin est fournie gratuitement par VMware, sous la forme d'un fichier OVA à déployer à l'aide de vCenter sur votre site central.

Si vous comptez déployer une telle architecture, vous devez vous assurer de la qualité de liaison entre les deux sites (connectivité L3 (unicast) entre le témoin et les hôtes :

- bande passante minimale de 1.5 Mbps,
- latence inférieure ou égale à 500 ms aller-retour - RTT).

Cette architecture supporte la panne d'un des deux hôtes ou d'un disque (FTT = 1).

Un cluster étendu peut se présenter sous la forme suivante :



Le cluster est éclaté en deux sites actifs-actifs, disposant de chaque côté des mêmes ressources. Il est intéressant de constater que dans ce scénario précis, en cas de panne d'un des sites, nous ne pourrons dégager une majorité et donc atteindre le *quorum* . Nous aurons donc besoin d'un troisième site, jouant le rôle de témoin, comme c'était le cas dans notre scénario ROBO, à la différence que nous aurons besoin ici de trois sites séparés !

Chacun de ces sites forme un fault domain, c'est-à-dire un domaine de panne. Un fault domain consiste en un groupement d'hôtes vSAN, réunis en fonction de leur proximité physique. Par exemple, on aura tendance à regrouper des serveurs situés dans une même armoire au sein d'un seul et même fault domain. L'objectif ? Indiquer à vSAN de protéger les objets entre les fault domains et de les rendre disponibles en cas de panne d'un ou de la totalité des éléments d'un fault domain. Si nous avons défini une tolérance à la panne (FTT) de 1, vSAN fera donc le nécessaire pour rendre la donnée disponible en dehors du fault domain d'origine.

Les domaines de pannes jouent un rôle particulier dans ce scénario puisqu'ils identifient le rôle de chaque site. Ainsi, le premier fault domain correspondra au site principal (A), le deuxième au site secondaire (B) et le troisième au site témoin (C). Dans notre cas, il s'agira en toute logique de répliquer les données du site principal vers le site secondaire. vSAN essaiera cependant en cas de panne locale (hôte, disque…) de favoriser une solution locale avant de faire appel au site secondaire. Dans cette configuration, il ne peut exister que trois fault domains au maximum.

VMware recommande une bande passante supérieure ou égale à 10 Gbps et une latence de maximum 5 ms (RTT) entre deux sites portant de la donnée. Les sites peuvent être connectés en L2 (multicast) ou L3 (unicast) via des routeurs. Concernant les caractéristiques de la liaison entre les sites de stockage et le site témoin, VMware recommande une bande passante d'au moins 100 Mbps et une latence (RTT) inférieure ou égale à 500 ms. La connectivité entre les sites stockage et témoin est assurée en L3 (unicast).

La tolérance à la panne est ici d'un hôte ou d'un disque à cause de la limitation sur le nombre de fault domains dans ce scénario (FTT = 1 maximum).

Nous avons vu trois implémentations vSAN type, abordons maintenant comment le stockage des objets vSAN fonctionne. Chaque donnée stockée dans vSAN est en réalité un objet. Prenons l'exemple d'une machine virtuelle stockée sur vSAN, ses objets pourront être de type suivant :

- **VM Home Namespace**

  Cette entité contient les fichiers de configuration et associés de la machine virtuelle (VMX, NVRAM, logs...). Elle est créée à la création d'une machine virtuelle sur une banque de données vSAN.

- **VMDK**

  Disques de stockage de la machine virtuelle.

- **VM Swap Object**

  Fichier d'échange de la machine virtuelle. Créé au lancement de la machine virtuelle

- **Snapshot Delta VMDK**

  Fichier créé en cas de snapshot, lié au(x) disque(s) de stockage de la machine virtuelle.

- **Memory object**

  Fichier créé en cas de snapshot sur une machine virtuelle lancée. Contient le memory dump de la machine au moment du snapshot.

Dans le cadre d'une redondance permettant la panne d'un élément de notre cluster (FTT = 1, RAID-1), un objet sera répliqué une fois. L'objet ainsi que son réplica occupent le rôle de « Component ». Les métadonnées pouvant permettre le quorum pour un objet en cas de panne sont appelées « Witnesses ». Ces dernières ne contiennent aucune donnée applicative.

Reprenons l'exemple d'un déploiement à 4 nœuds avec un objet vSAN. Cet objet pourra avoir la répartition suivante :

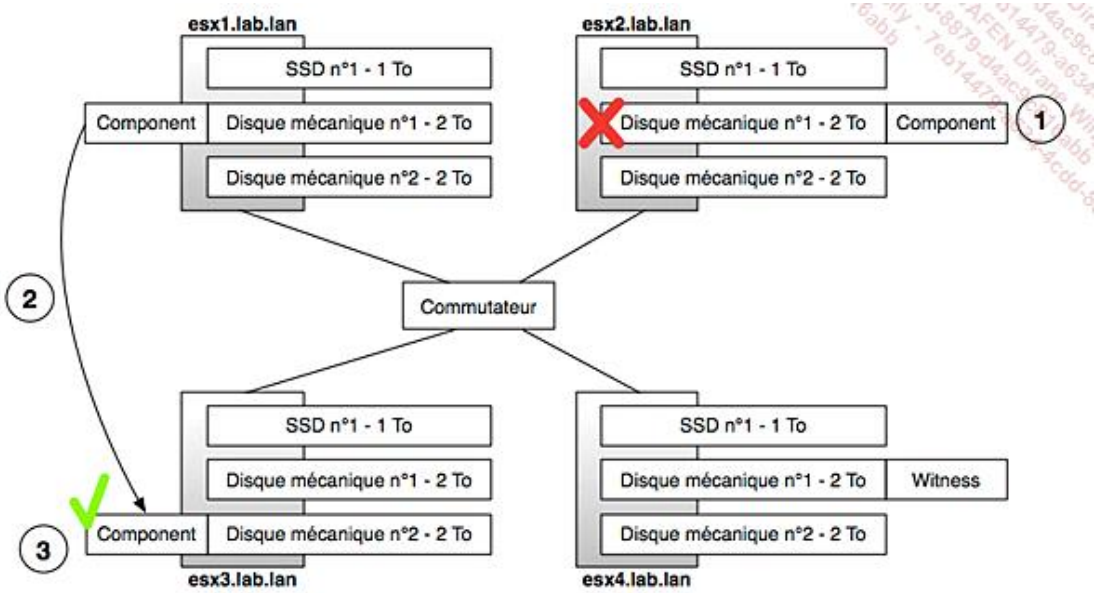
- esx1.lab.lan - *Component*
- esx2.lab.lan - *Component*
- esx3.lab.lan - *Witness*

Pour chaque objet, la répartition peut être différente. Ainsi, les objets qui composent une machine virtuelle pourraient être stockés sur des hôtes différents. La répartition ne se fait pas au hasard et vise à équilibrer la charge stockage sur l'ensemble des hôtes. Si malgré ces mesures, un hôte venait à dépasser une utilisation de 80 % de la couche capacity, vSAN peut déclencher un rééquilibrage (*rebalance*) automatique visant à restaurer cet

équilibre. Cette action peut également être déclenchée de façon manuelle par l'administrateur.

En cas de panne d'un hôte ou d'un disque, vSAN pourra réagir pour garantir la disponibilité de l'ensemble des objets stockés. Voyons comment vSAN réagit dans ces cas.

Analysons tout d'abord comment vSAN gère la panne d'un disque d'un hôte, grâce à ce schéma :



Avant la panne, nous avions un fichier disque d'une machine virtuelle (VMDK) stockée sur les disques mécaniques (tier *capacity*) des serveurs esx1.lab.lan et esx2.lab.lan. Voici ce qui se produit en cas de panne :

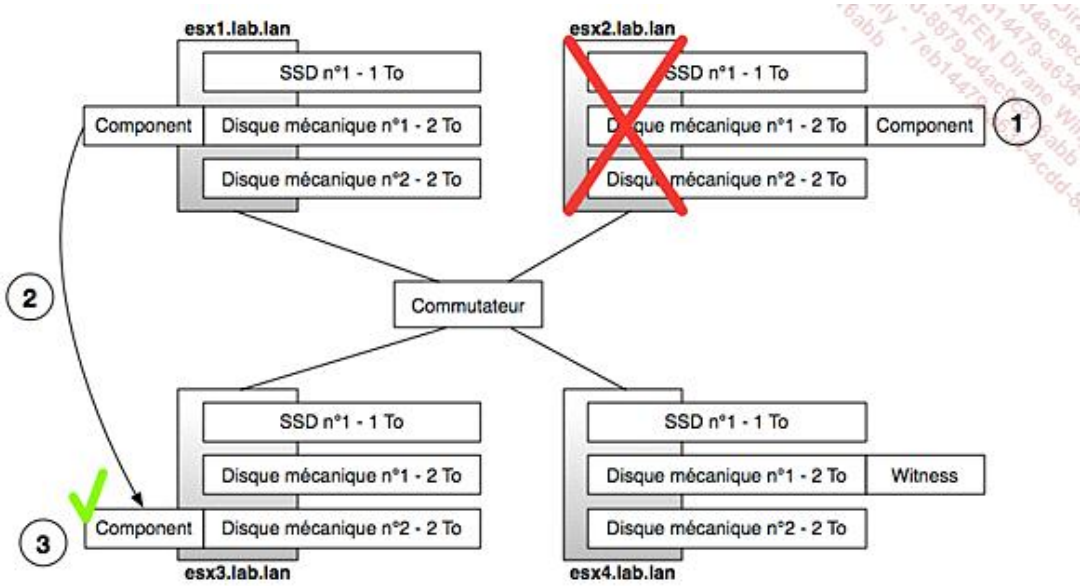
1- Le disque mécanique n°1 d'esx2.lab.lan tombe soudainement en panne.

2- L'objet composent côté esx2.lab.lan rentre en mode « degraded ». vSAN tente alors immédiatement de protéger l'objet en le copiant sur un autre hôte, ici sur esx3.lab.lan.

3- Une fois l'information dupliquée à nouveau, ce serveur devient *component* pour cet objet dont vSAN vient d'assurer la réplication.

Cette protection de l'objet, entreprise de façon automatique par vSAN, est effectuée via une opération de resynchronisation. Malgré la panne d'un disque, la machine virtuelle n'a « rien vu » et continue à s'exécuter normalement. À cause de la charge générée par cette opération, il n'est pas impossible cependant que les performances de cette dernière aient pu être altérées pendant la copie du réplica.

Nous avons parlé ici de la panne d'un disque ? Mais que se passerait-il si un hôte venait à dysfonctionner ? Utilisons un autre schéma pour l'expliquer :



Avant la panne d'un hôte, esx2.lab.lan portait là aussi un objet vSAN VMDK. Voici ce qu'il arrive en cas de panne d'un hôte :

1- L'hôte esx2.lab.lan ne répond plus. vSAN décide de déclencher un compteur (60 minutes par défaut) pour s'assurer que l'hôte ne reviendra pas entre temps avant de lancer une opération de protection de l'objet. L'objet qui était en *component* passe en statut « *absent* ».

2- Le compteur expiré, vSAN décide d'assurer la protection de la donnée sur un autre hôte, ici esx3.lab.lan.

3- Une fois l'information dupliquée à nouveau, ce serveur devient *component* pour cet objet dont vSAN vient d'assurer la réplication.

La même procédure aurait été appliquée si l'hôte esx2.lab.lan avait perdu sa connexion au réseau.

Vous avez peut-être remarqué que selon la panne (disque, hôte), la réaction de vSAN n'est pas similaire.

- « *Degraded* » signifie pour vSAN une erreur irrécupérable, comme une panne disque, ce qui explique la réaction immédiate de vSAN.
- « *Absent* » considère que l'erreur peut être récupérable (l'hôte peut tout à fait revenir avec les données intactes en cas de redémarrage, de perte réseau...). En toute logique, c'est le compteur de 60 minutes qui s'applique ici. Sachez que ce dernier est modifiable sur chaque hôte vSphere, via l'option avancée VSAN.ClomRepairDelay même s'il est recommandé de laisser la valeur par défaut.

Voici un tableau qui résume les pannes et la classification opérée par vSAN :

Raison	Réaction
Panne de disque capacity (flash ou mécanique)	*Degraded*
Panne de disque *caching*	*Degraded*
Panne d'un contrôleur de stockage	*Degraded*
Débranchement d'un disque *capacity* (flash ou mécanique)	*Absent*
Débranchement d'un disque *caching*	*Absent*
Panne d'un hôte	*Absent*
Perte de la connectivité réseau	*Absent*

Panne d'un adaptateur réseau	*Absent*

Indépendamment de la décision prise par vSAN, sachez qu'il est souhaitable d'avoir au minimum dans votre architecture vSAN quatre hôtes pour permettre la protection active des données en cas de panne via une resynchronisation.

Si vous ne déployez qu'une infrastructure à 2 (ROBO) ou 3 nœuds, ces derniers seront incapables de lancer une telle opération, ils se contenteront d'assurer la disponibilité des ressources avec le réplica existant. Comme nous l'avons vu précédemment, vous ne pouvez supporter dans ces modes qu'une tolérance à la panne de 1 au maximum (FTT = 1).

La resynchronisation ne vise pas seulement à protéger l'objet mais aussi à évacuer ou rééquilibrer les données portées par les membres du cluster. Voici la liste exhaustive des événements déclenchant une telle opération dans votre cluster vSAN :

- Suite à une modification de la politique de stockage des machines virtuelles, par exemple, si vous modifiez votre politique pour protéger les objets stockés dans votre *datastore* avec un FTT = 2 quand la politique mentionnait avant FTT = 1, une opération de réplication de l'objet est lancée pour assurer la conformité des objets avec cette nouvelle politique.
- Le redémarrage d'un hôte après une panne
- La survenue d'une panne sur un hôte. En cas de panne considérée comme irrécupérable (*degraded*) ou récupérable (*absent*), vSAN cherchera à protéger à nouveau les objets.
- L'évacuation des données d'un hôte avant la mise en maintenance de ce dernier, avec l'option « Full data migration ».
- L'utilisation à plus de 80 % de la couche *capacity* sur un hôte. Dans ce cas, comme évoqué précédemment, vSAN déclenchera automatiquement une opération de rééquilibrage.

Associons les tolérances à la panne avec les modes de réplication dont nous avons parlé auparavant à l'aide du tableau suivant :

	RAID-1 (*Mirroring*)		**RAID 5/6 (*Erasure-Coding*)**	
FTT (Nombre de pannes à tolérer)	**Nombre minimal d'hôtes**	**Taux de réplication des données**	**Nombre minimal d'hôtes**	**Taux de réplication des données**
FTT = 0	3	x1	N/A	N/A
FTT = 1	3	x2	4	x1.33
FTT = 2	5	x3	6	x1.5
FTT = 3	7	x4	N/A	N/A

Les taux de réplication de données s'entendent ici sans prendre en compte la compression et la déduplication, possible dans une architecture *all-flash* .

On remarque que le mode de stockage le plus efficace est comme prévu celui utilisant la méthode *Erasure Coding* , avec un taux de réplication assez bas pour assurer une protection des données, comparé au mode classique RAID-1. Il nécessite pour rappel au minimum 4 hôtes dans votre cluster vSAN, un argument de plus en direction d'une architecture dotée au minimum de 4 nœuds !

Pour être encore plus précis, regardons de plus près quelle sera votre capacité utilisable en fonction des disques et de la méthode employée, à l'aide des tableaux suivants :

RAID-1 (*Mirroring*) - *tier capacity*						
FTT (nombre de pannes à tolérer)	**Nombre d'hôtes**	**Taille des disques capacity** (capacitée non-formatée en To)	**Nombre de disques capacity**	**Taux de réplication des données**	**Capacité nominale du datastore** (hors-protection) en To	**Capacité utilisable du datastore en To**
FTT = 0	3	2	2	1	12	12
FTT = 1	3	2	2	2	12	6
FTT = 2	5	2	2	3	20	6,66
FTT = 3	7	2	2	4	28	7

RAID-5/6 (*Erasure-Coding*) - *tier capacity*						
FTT (nombre de pannes à tolérer)	**Nombre d'hôtes**	**Taille des disques capacity** (capacitée non-formatée en To)	**Nombre de disques capacity**	**Taux de réplication des données**	**Capacité nominale du datastore** (hors protection) en To	**Capacité utilisable du datastore en To**
FTT = 1	4	2	2	1,33	16	12,03
FTT = 2	6	2	2	1,5	24	16

On remarque là aussi que le modèle RAID-1 est très coûteux en termes d'espace disque et de réplication, d'autant plus lorsqu'on souhaite avoir un FTT supérieur à 1. En revanche, l'approche *Erasure Coding* nous permet d'avoir beaucoup plus d'espace disponible, tout en continuant à protéger nos données. Même si cela peut impacter les performances de notre datastore virtuel, le gain substantiel de place peut jouer un rôle déterminant dans votre choix entre un déploiement RAID-1 et celui-ci.

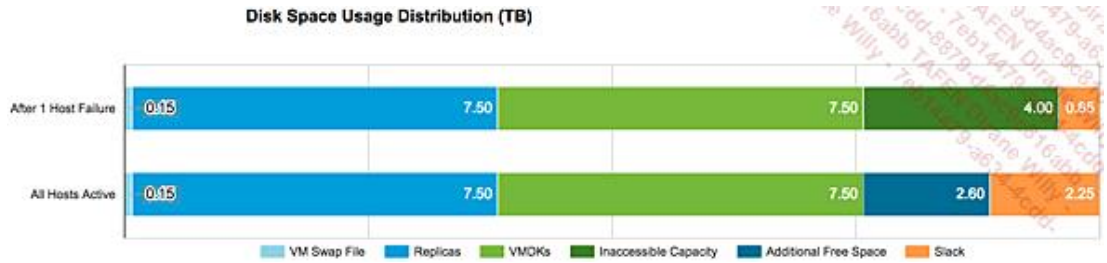
L'équivalent FTT = 1 dans un RAID-5/6 pourrait être comparé à la protection du RAID-5, quand le FTT = 2 correspond à celle offerte dans un mode RAID-6. Avant de parler de la mise en place d'un cluster vSAN, découvrons comment vous pouvez planifier un déploiement en amont.

d. Planification

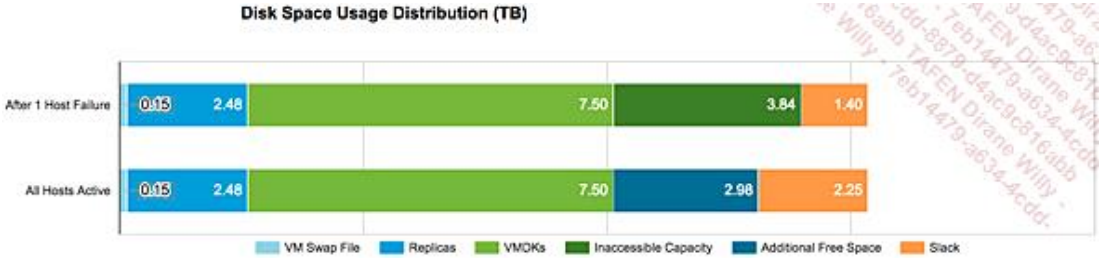
Il peut être compliqué de planifier un déploiement vSAN, tant sur les ressources matérielles (serveur, processeur, contrôleur de stockage, réseau) que sur les disques à employer.

Pour vous aider dans ces décisions, VMware vous propose un site qui vous permet de simuler le déploiement d'un cluster vSAN et prévoir les ressources nécessaires en fonction du nombre de machines virtuelles à déployer, de l'espace disque occupé par ces dernières, des IOPS et d'autres caractéristiques.

Deux exemples ici de simulation de stockage en RAID-1 puis RAID-5/6 :



Simulation avec scénario FTT = 1, RAID-1, sur la base de 75 machines virtuelles (1vCPU, 1 Go de vMem), occupant chacune 100 Go d'espace disque



Simulation avec scénario FTT = 1, RAID-5/6, sur la base de 75 machines virtuelles (1vCPU, 1 Go de vMem), occupant chacune 100 Go d'espace disque

Comme vous le voyez, le site calcule également pour vous la capacité disponible après la panne d'un hôte pour vous permettre d'estimer la marge disponible si jamais un tel événement devait se produire dans votre cluster.

À la fin de la simulation concernant le stockage, l'outil vous propose même une liste de composants (BOM - *Bill of Materials*) pour constituer votre cluster vSAN.

Default Server Configuration of Selected Performance Profile		
Components	Qty	Details
CPU	2	Intel® Xeon® Ivy Bridge 10C E5-2650V3 2.6G
MEM	8	16GB RDIMM, 2133MT/s, Dual Rank, x4 Data Width
Controller	1	PERC H730 RAID Controller, 1GB NV Cache
NIC	2	Intel X520 DP 10Gb DA/SFP
Flash Cache	1	200GB SSD-SATA Mix Use 6G
Persistent Disk	4	960GB SSD-SATA Read-Intensive 6G

Il vous permet aussi de calculer des aspects plus financiers (TCO, CAPEX, OPEX) de cette solution si nécessaire. Le site est accessible au public à l'aide du lien suivant : https://virtualsansizing.vmware.com.

Concernant le matériel et les serveurs à utiliser dans le cadre d'un déploiement vSAN, vous devez utiliser du matériel validé par VMware, par le biais de la HCL (*Hardware Compatibility List*). Là aussi, VMware vous propose un site pour vous guider dans l'achat de vos serveurs (Dell, Cisco, Supermicro...) selon vos besoins. Vous trouverez ci-après le lien du site : http://vsanreadynode.vmware.com

Enfin, pour vous lancer sans avoir à disposer de ressources importantes de votre côté, n'hésitez pas à essayer vSAN à l'aide de ressources virtualisées par VMware, les laboratoires à distances appelés HoL (*Hands-on-Labs*) ici : http://www.vmware.com/go/try-vsan-hol

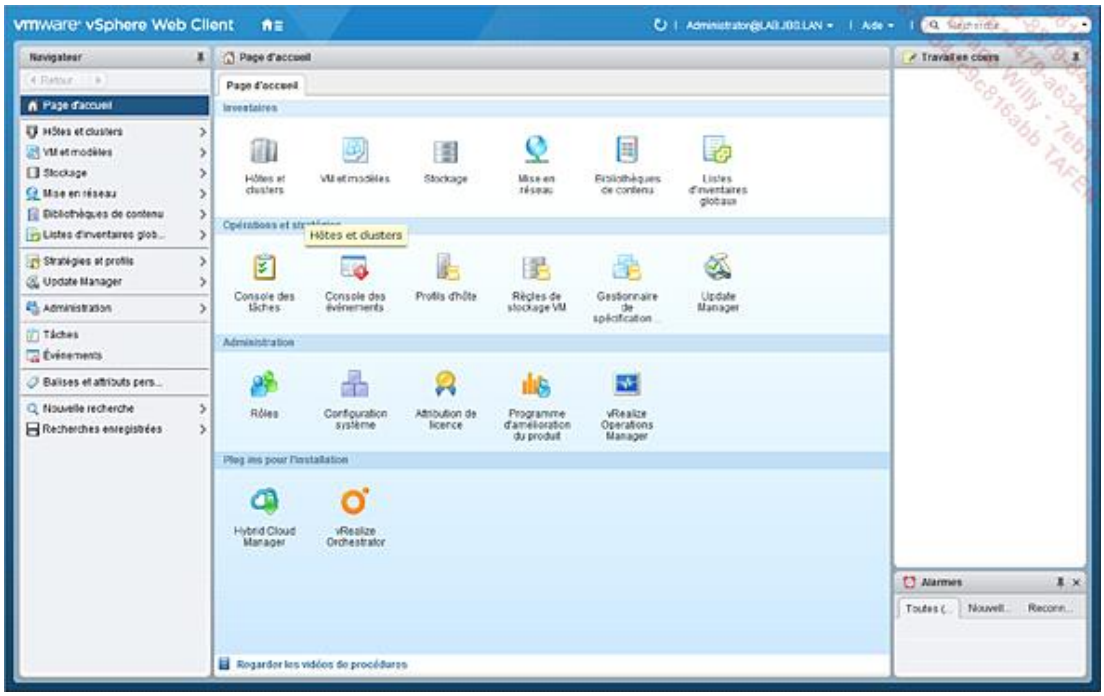
Assez parlé de théorie, passons à l'implémentation de vSAN avec une base vSphere 6.5 !

e. Mise en œuvre avec vSphere 6.5

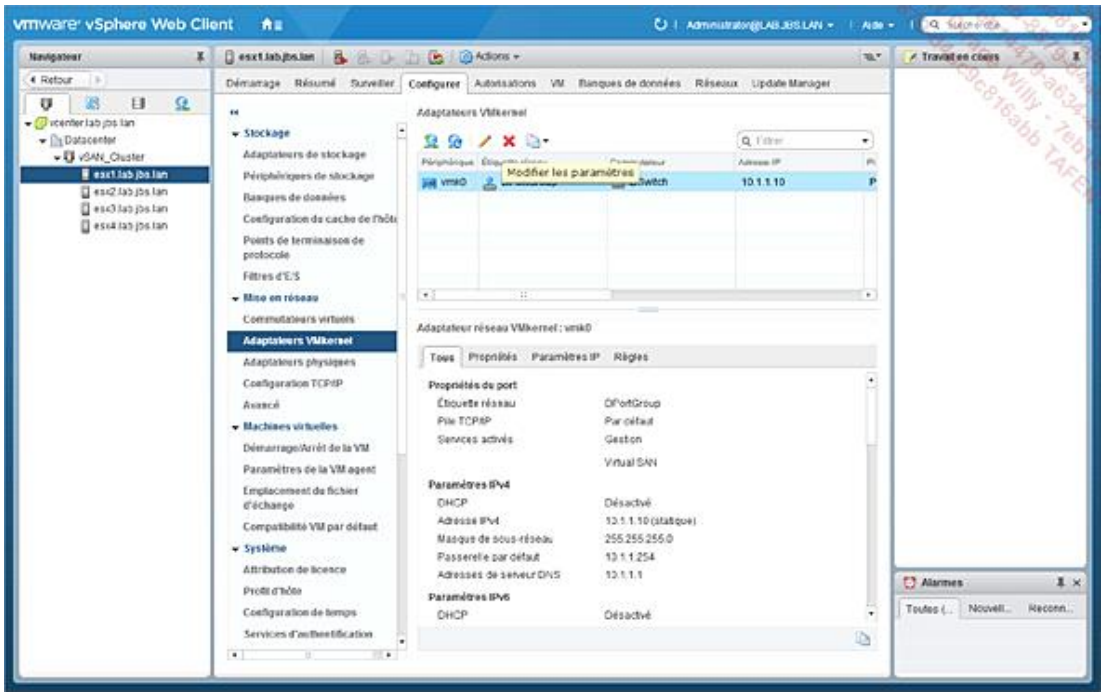
La mise en œuvre d'un cluster vSAN est assez rapide, pour peu que vous disposiez déjà d'un serveur vCenter et d'un cluster existant. L'activation du cluster vSAN se fera d'ailleurs au niveau du cluster et non pas sur les hôtes individuellement.

Nous allons ici implémenter à l'aide du vSphere Web Client, un cluster vSAN all-flash à quatre hôtes, dotés chacun de 40 Go de cache (*caching*) et 80 Go de stockage (*capacity*).

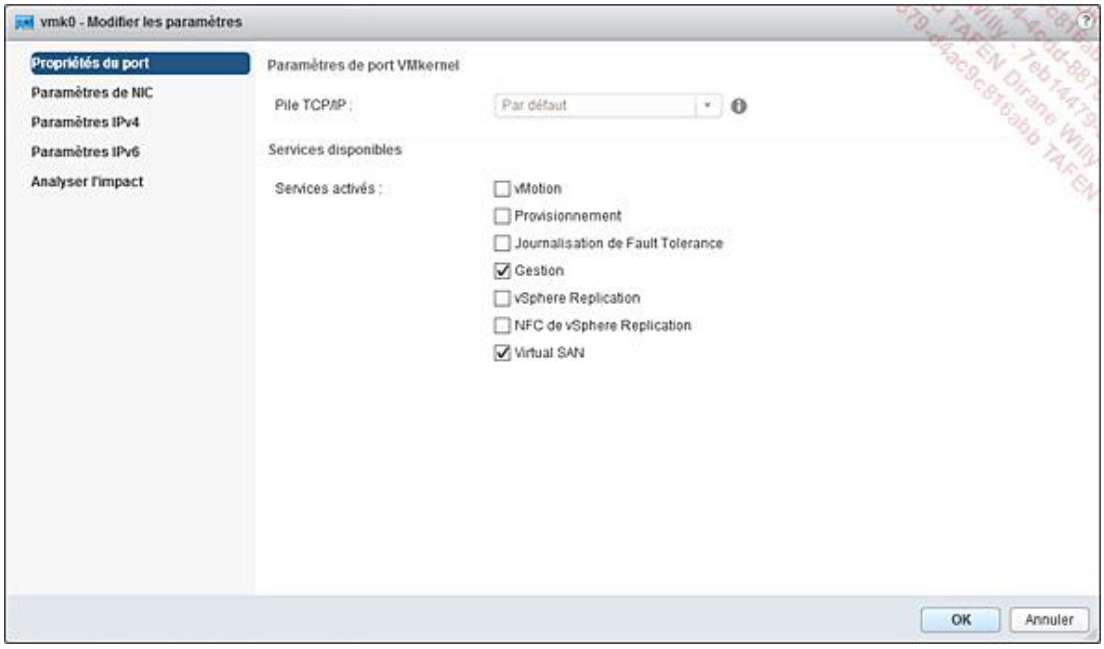
Avant de lancer la création du cluster vSAN, nous allons indiquer, comme c'est le cas avec d'autres services (FT, vMotion...) quelle interface utiliser pour le trafic vSAN par le biais d'interfaces VMkernel. Rendez-vous donc pour commencer dans la partie « Hôtes et clusters ».



Puis, vous pouvez effectuer la configuration de l'interface VMkernel sur l'hôte « esx1.lab.jbs.lan », à l'aide de l'onglet **Configurer et Modifier les paramètres** une fois l'interface VMkernel sélectionnée.



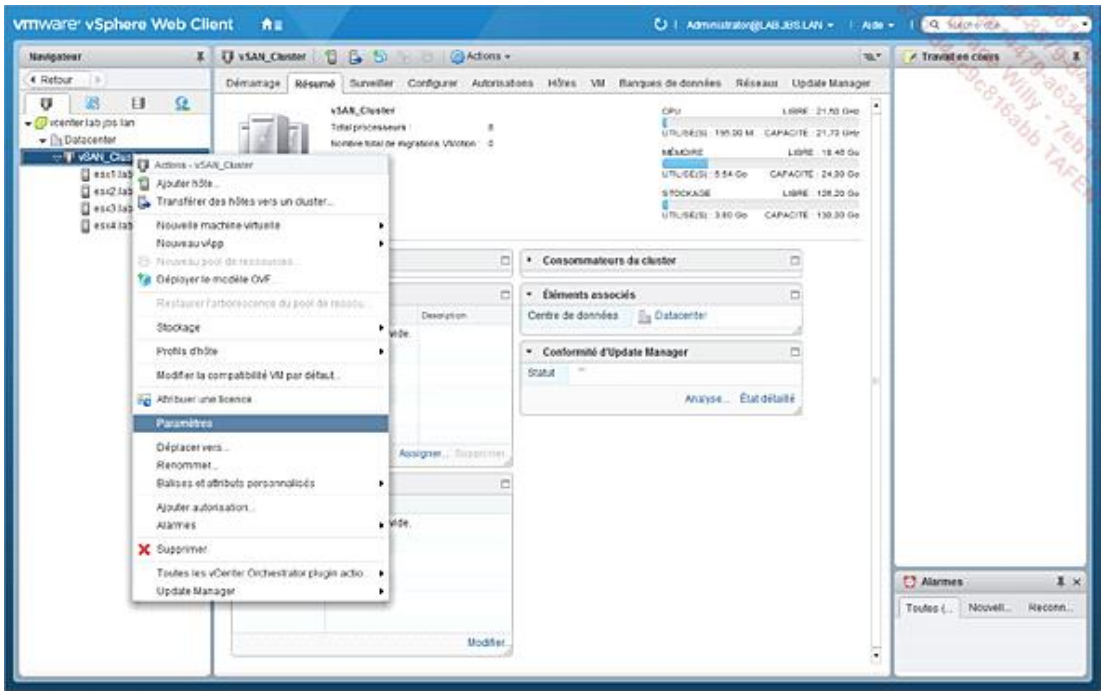
Puis cochez la case **Virtual SAN** pour activer le trafic sur cette interface VMkernel.



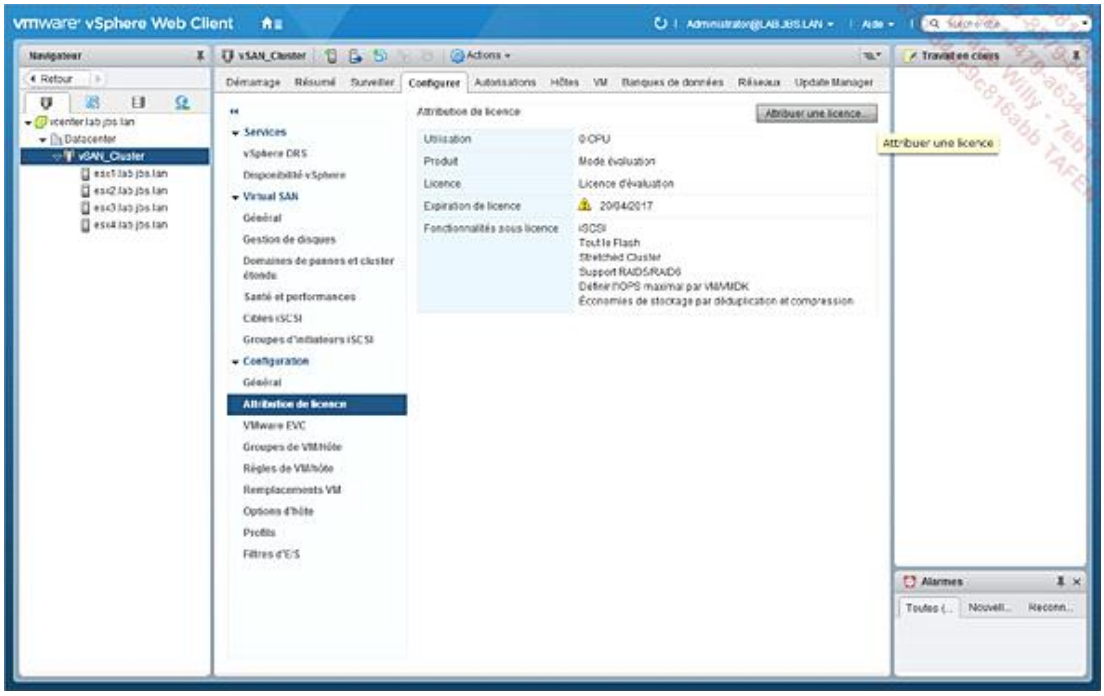
Pour la simplicité de la démonstration, notez que nous utilisons ici l'interface VMkernel par défaut, vmk0. Même si cette configuration est possible techniquement, il est recommandé cependant de dédier une (ou plusieurs) interface(s) physique(s) (ainsi que son interface VMkernel) pour un tel trafic.

Cette manipulation est à répéter sur tous les hôtes membres du cluster qui participeront à vSAN.

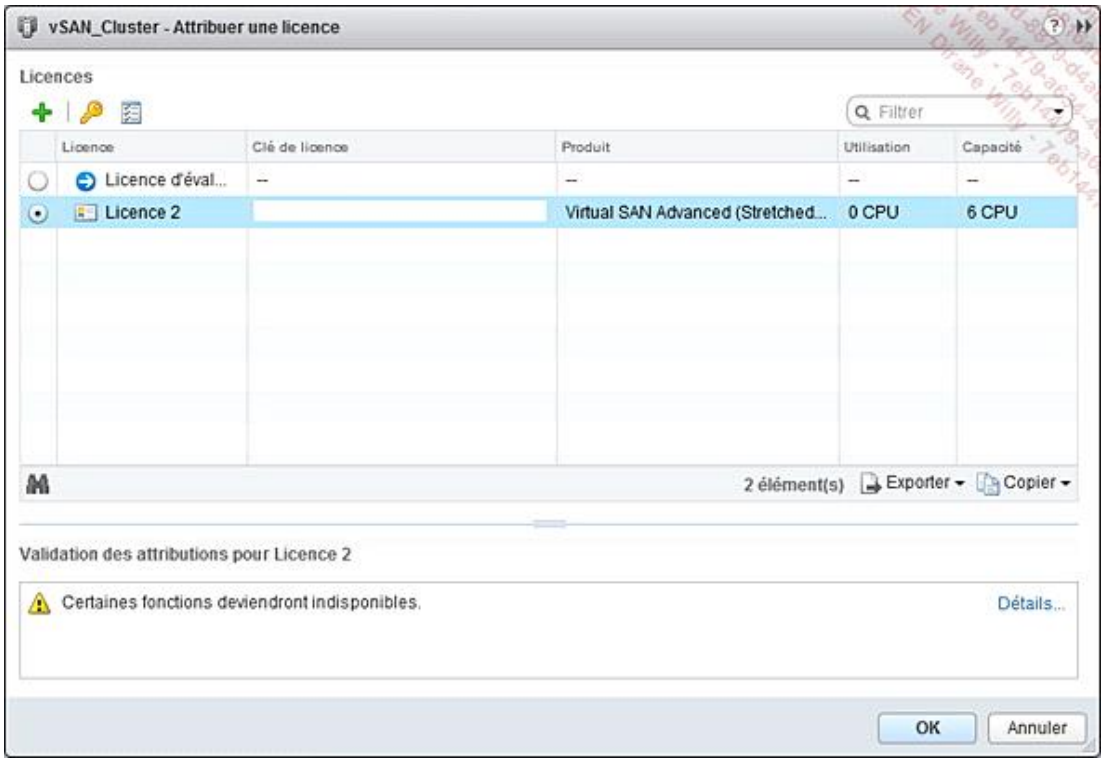
Rendez-vous dans la partie **Paramètres** du cluster ensuite pour attribuer votre licence vSAN à votre cluster, avant de lancer sa création.



Cliquez sur le bouton **Attribuer une licence** .

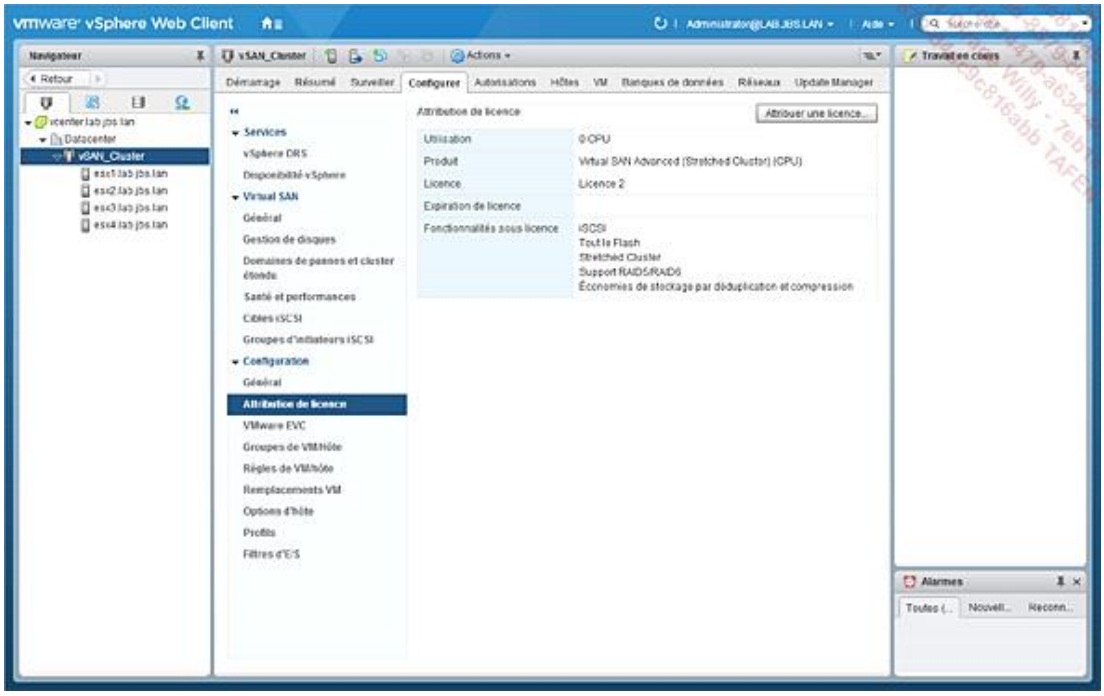


Puis sélectionnez la licence à attribuer.

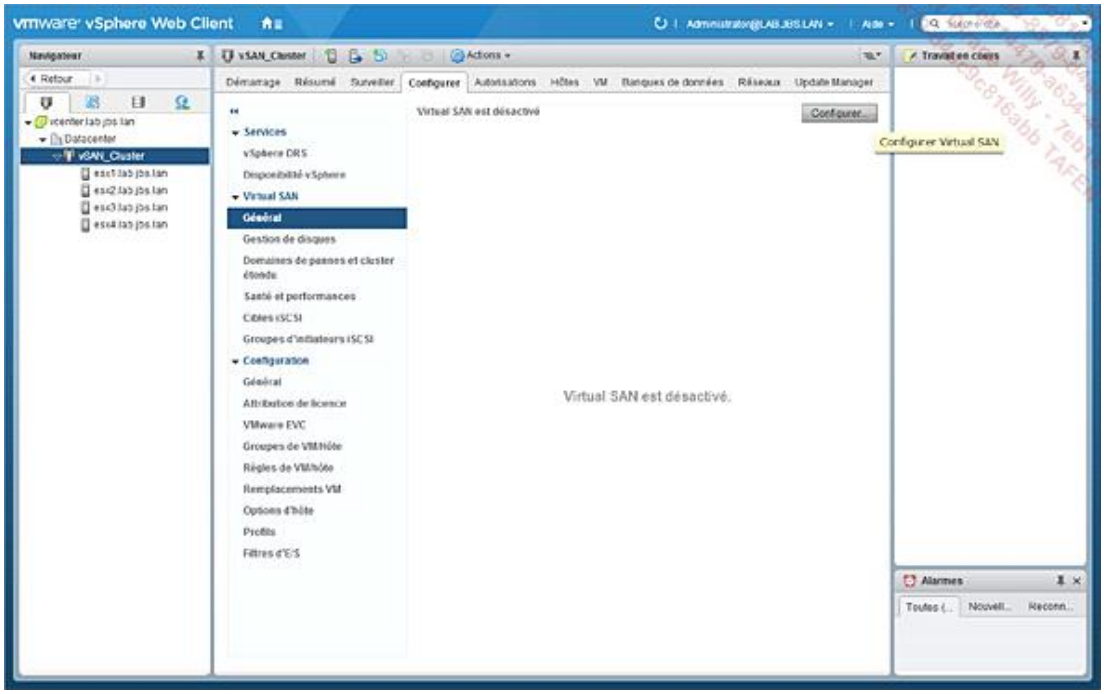


Notez que vous devrez au préalable importer votre licence à l'aide de la section **Licences** de vCenter.

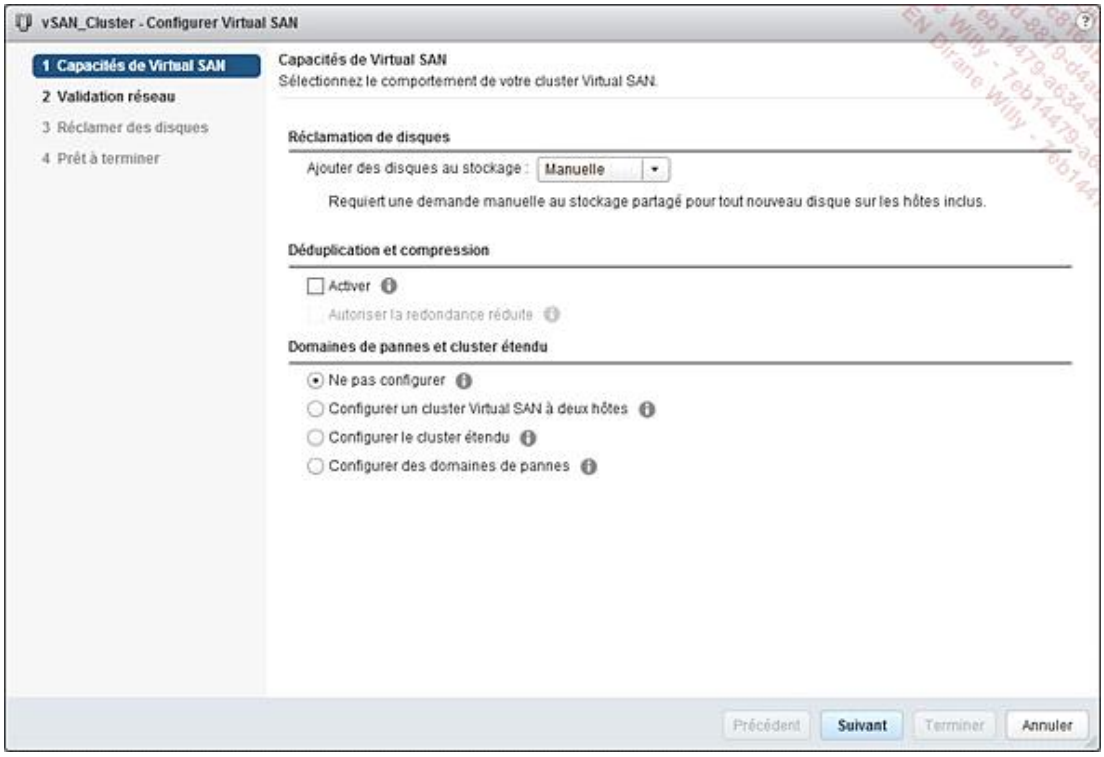
Vous pouvez ensuite voir la licence attribuée à votre cluster.



Une fois cette partie terminée, rendez-vous dans la partie **Général** de la section **Virtual SAN** pour lancer la configuration de votre cluster vSAN à l'aide du bouton **Configurer** .



Dans cet écran, vous pouvez déjà définir les premières caractéristiques de votre cluster vSAN.



La première section consiste à vous demander si vous souhaitez identifier manuellement ou automatiquement les disques qui feront partie de votre cluster vSAN. Il est préférable de choisir l'option **Manuelle** si vous souhaitez maîtriser précisément votre attribution.

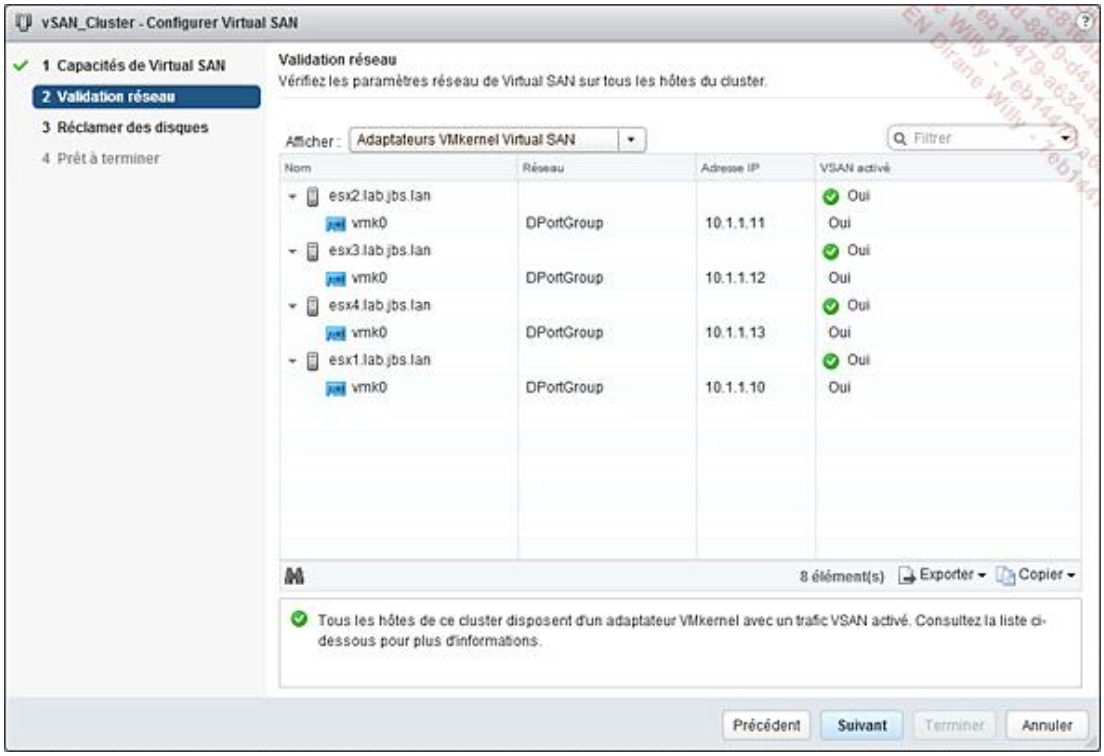
La deuxième section vous permet d'activer la déduplication et la compression, nous le laissons désactivé pour le moment mais sachez que vous pourrez l'activer après à l'aide de la page de configuration dédiée.

La dernière section concerne les scénarios suivants, déjà évoqués plus hauts :

- Configurer un cluster Virtual SAN à deux hôtes - ROBO. Vous devrez désigner un witness host juste après dans ce cas de figure.
- Configurer le cluster étendu - relatif au *stretched cluster*
- Configurer des domaines de pannes - relatifs aux *faults domains*

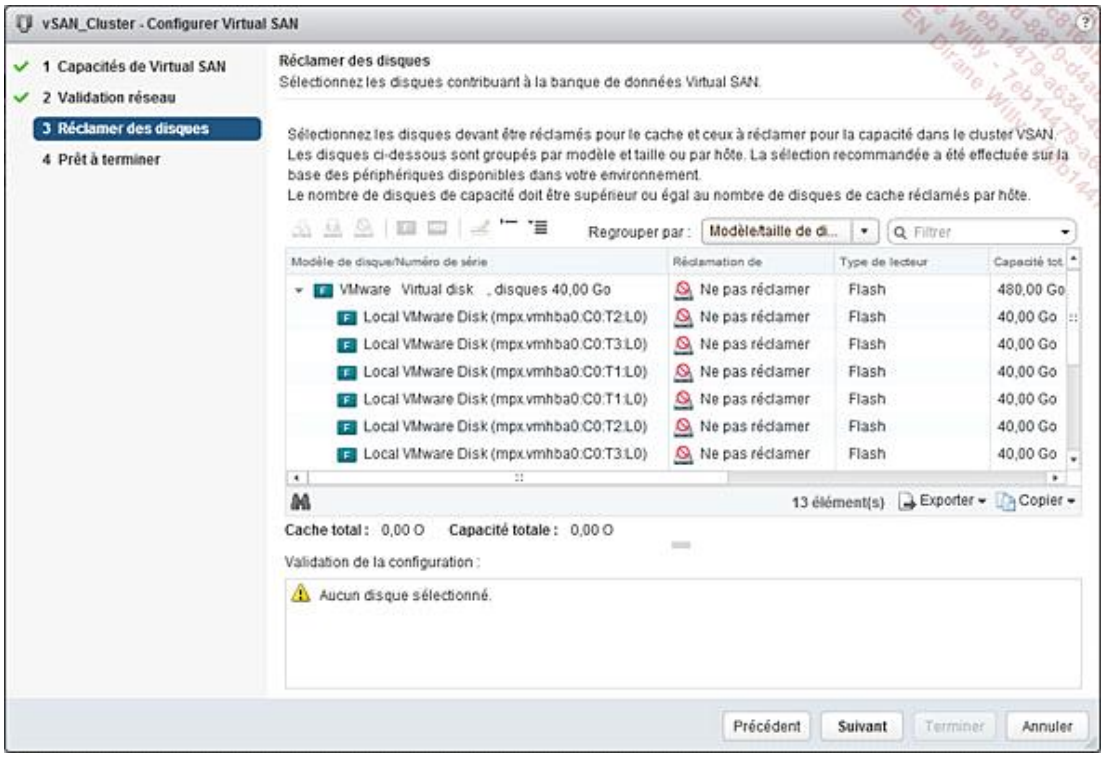
Cliquez sur **Suivant** pour poursuivre.

L'assistant vérifie à cette étape que l'ensemble des hôtes du cluster disposent d'une interface VMkernel sur laquelle le trafic vSAN est activé.

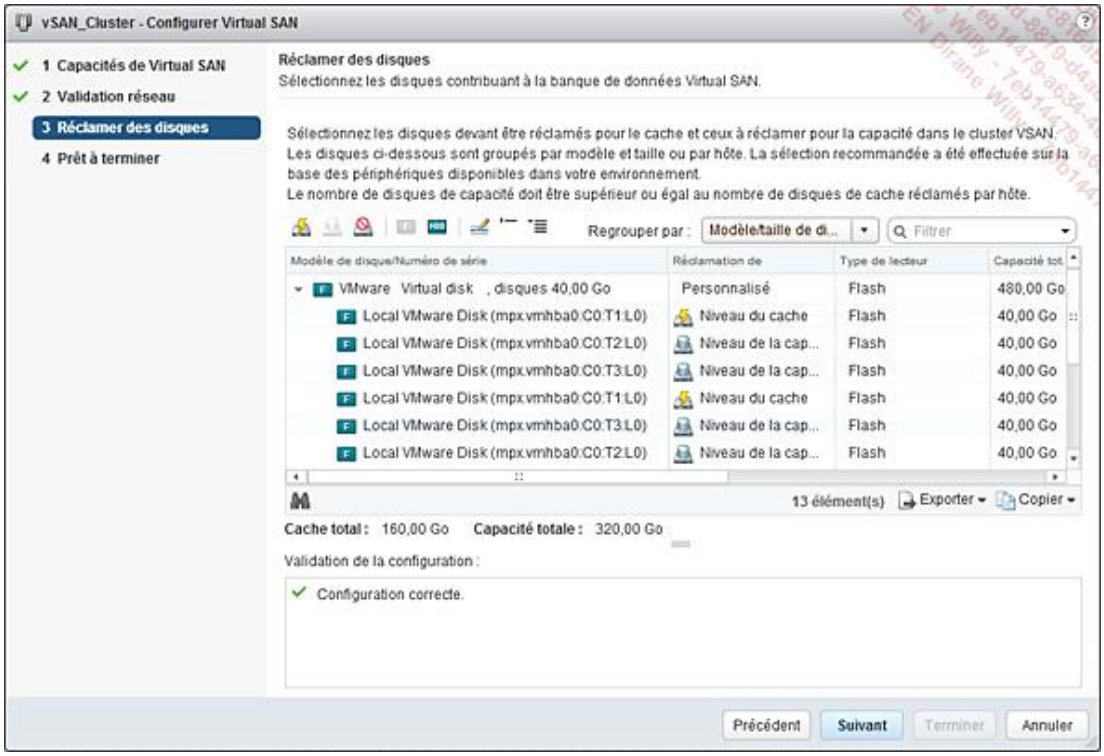


Cliquez sur le bouton **Suivant** pour poursuivre.

Dans cet écran, vous devez définir pour chaque disque le rôle qu'il occupera (*caching* ou *capacity*).

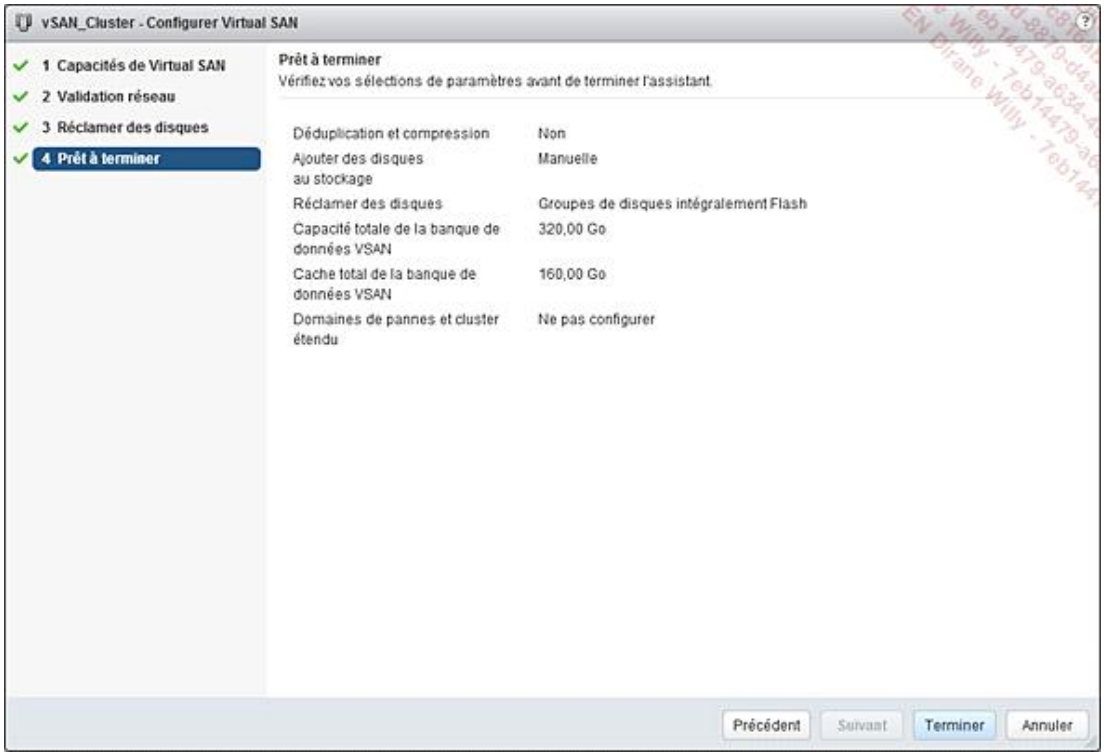


Une fois ce travail terminé, vous devriez obtenir un écran tel que celui illustré par la capture ci-dessous :

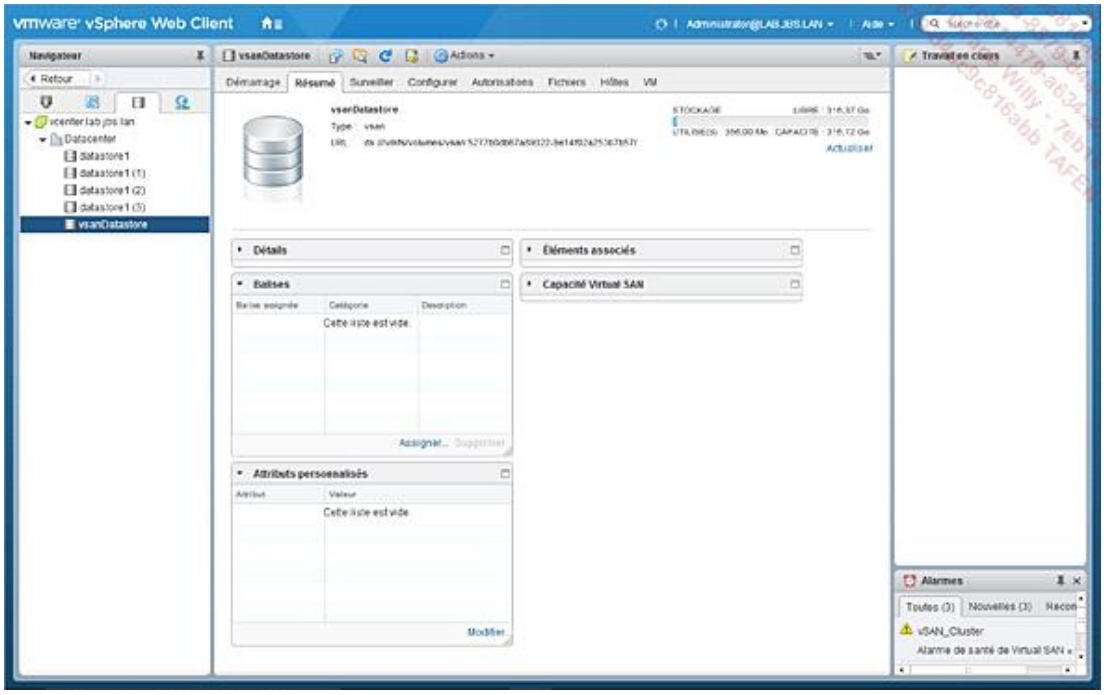


Comme vous pouvez le constater, le ratio de répartition est ici de un disque *caching* et de deux disques *capacity* pour chaque hôte. Une fois cette attribution terminée, cliquez sur **Suivant** pour poursuivre.

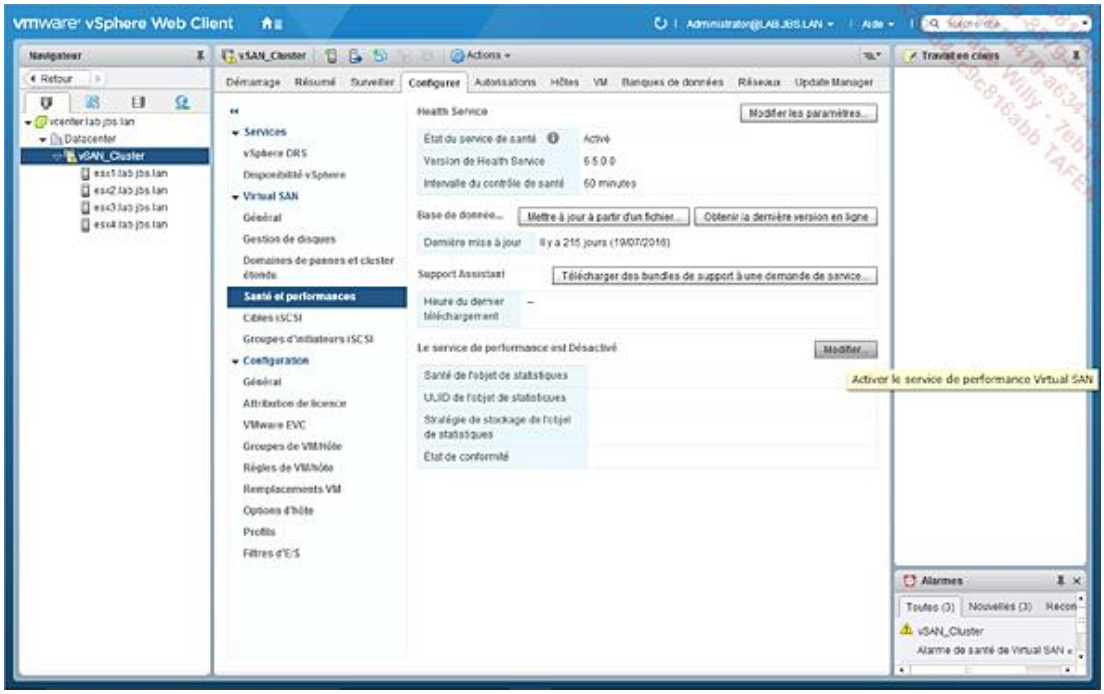
Dans ce dernier écran, vous obtenez un récapitulatif des paramètres que vous avez mentionnés, cliquez sur le bouton **Terminer** pour lancer la création de votre cluster vSAN.



Une fois l'assistant terminé et après avoir attendu un petit moment (selon les performances de votre système et de vos disques) le temps de la création du *datastore* partagé, vous pouvez voir apparaître ce datastore, nommé « vsanDatastore ».

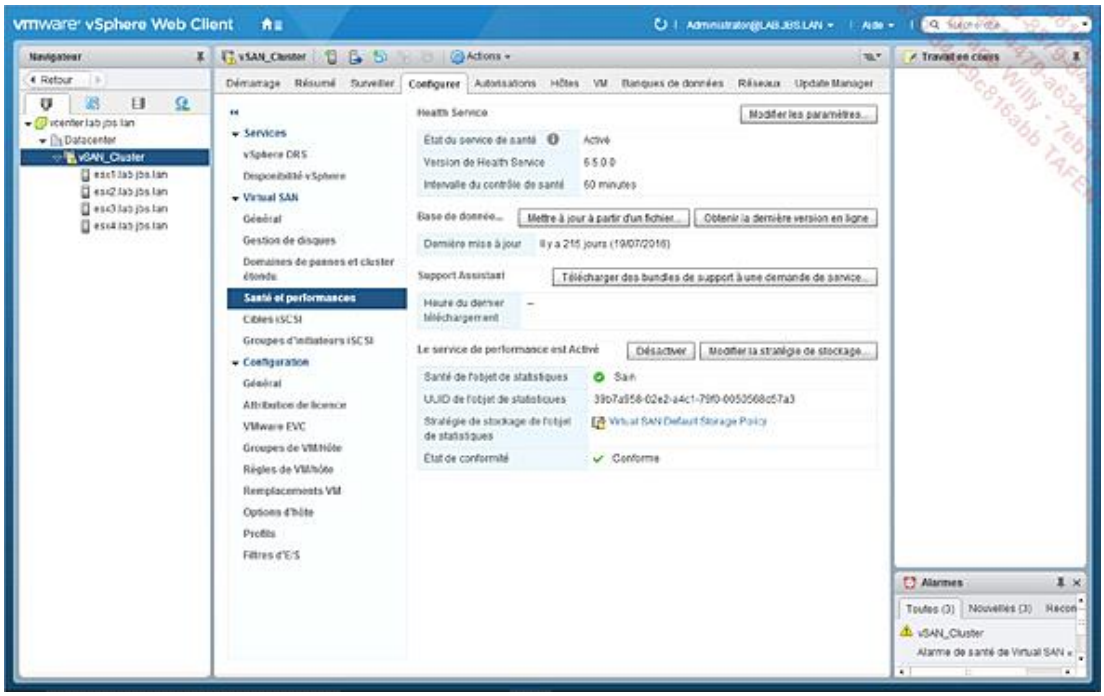


Concernant les premières étapes à effectuer d'emblée, il est intéressant d'activer le service de performances (qui vous fournit des informations sur les usages de votre datastore, IOPS, throughput...) qui ne l'est pas par défaut. Rendez-vous dans la section **Santé et performances** pour l'activer. Cliquez sur le bouton **Modifier** .



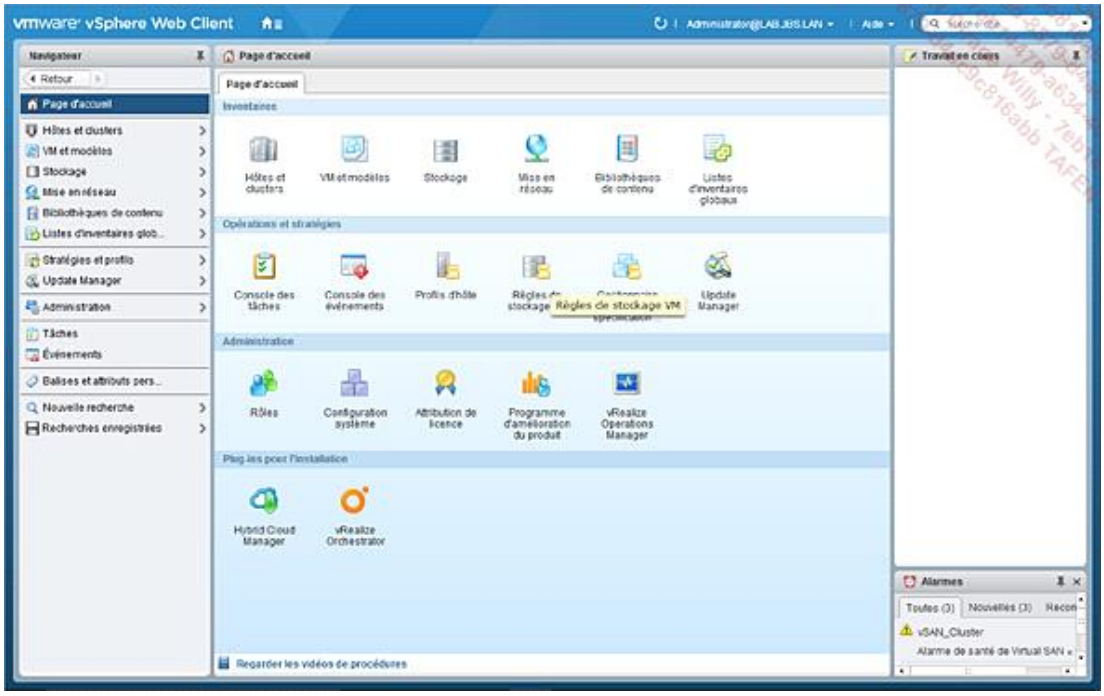
Puis, cochez la case **Activer le service de performance Virtual SAN** pour activer le service. Validez ensuite à l'aide du bouton **OK** .

Une fois validé, vous devriez obtenir un écran tel quel ci-dessous pour vous indiquer que l'objet vSAN dédié aux statistiques est activé, sain et conforme à la politique de stockage sélectionnée en vigueur dans votre cluster vSAN.

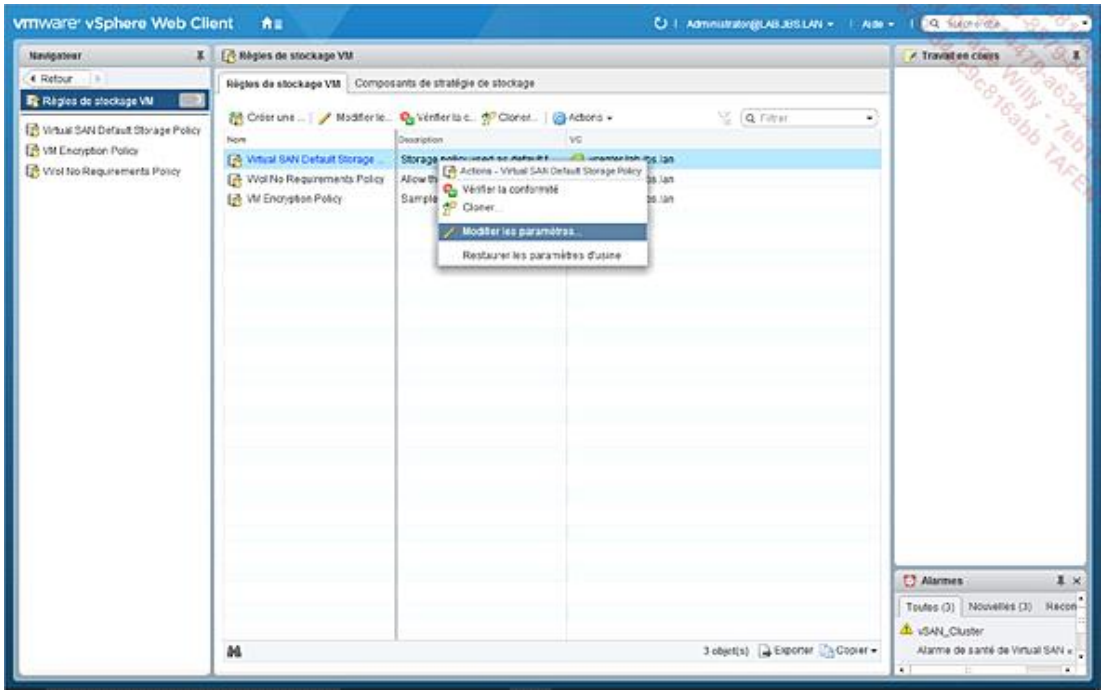


Depuis le début de la configuration, vous pouvez vous demander à quel endroit pouvez-vous spécifier la redondance souhaitée (FTT) ainsi que son mode (RAID-1 ou RAID-5/6). Souvenez-vous que cette partie n'est pas configurée ni opérée dans la configuration globale de vSAN mais dans les politiques de stockage ! Comme énoncé dans l'introduction, l'avantage principal est de pouvoir offrir aux machines virtuelles différentes tolérances à la panne et modes de réplications au sein d'un même cluster vSAN.

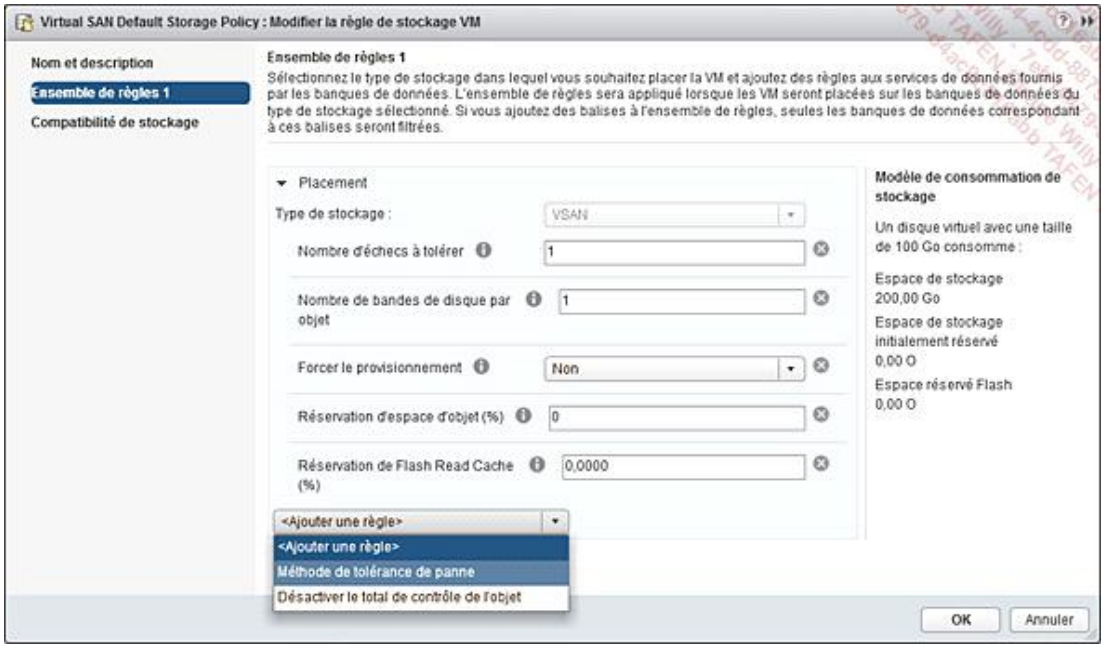
Rendez-vous donc dans la partie **Règles de stockage VM** pour éditer la politique de stockage en vigueur.



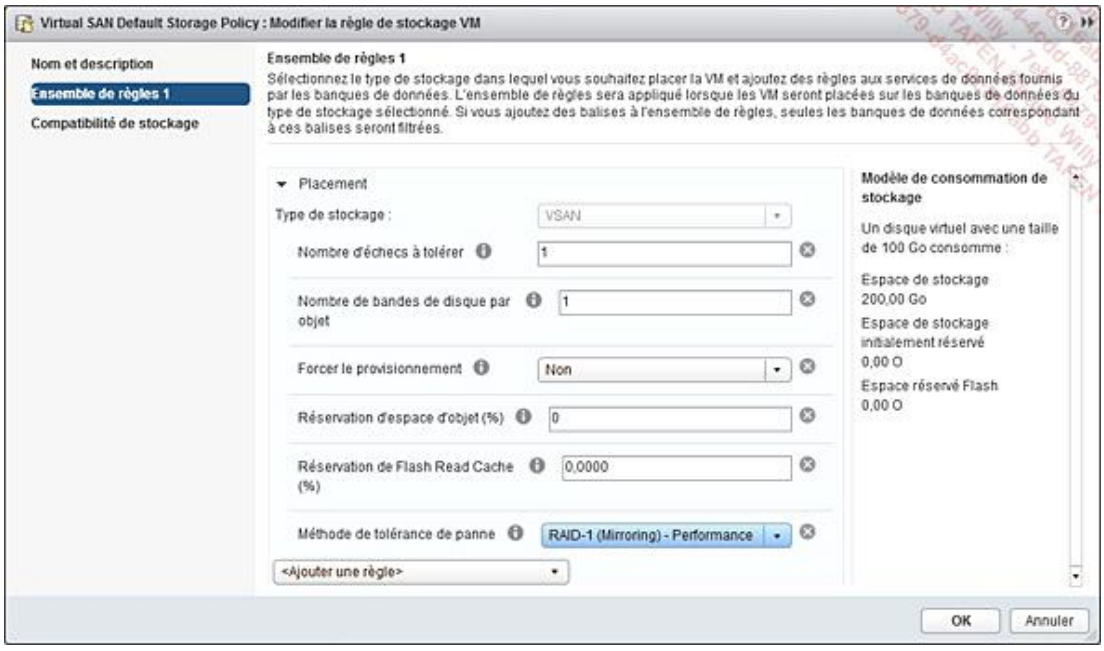
Virtual SAN Default Storage Policy représente la politique par défaut appliquée dans un cluster vSAN. Pour l'éditer, effectuez un clic droit sur celle-ci puis sélectionnez l'option **Modifier les paramètres** .



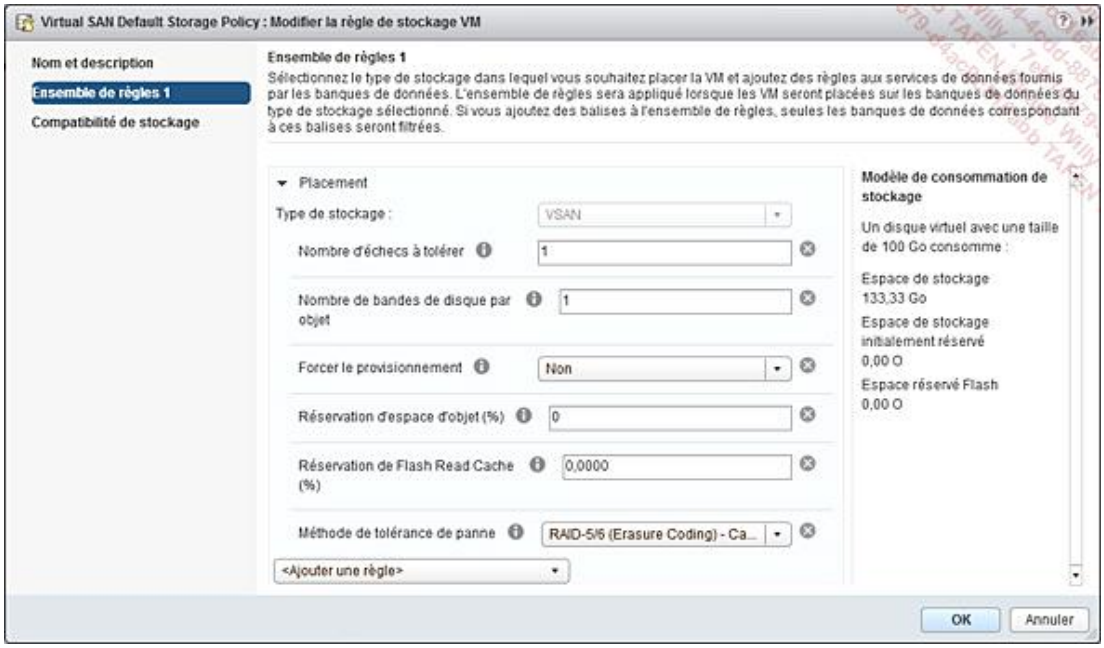
La partie représentant la FTT ici est la ligne **Nombre d'échecs à tolérer** , par défaut il est à 1 comme attendu. Pour configurer le mode de redondance, ajoutez une règle **Méthode de tolérance de panne** .



Puis sélectionnez le mode de redondance voulu. Sachez que le mode RAID-1 - *Performance* s'applique par défaut, il n'est pas très utile de le mentionner explicitement dans la politique, sauf pour l'identifier aisément.



Vous pouvez changer la méthode de tolérance de panne vers du RAID 5/6 en sélectionnant la valeur **Capacity** .

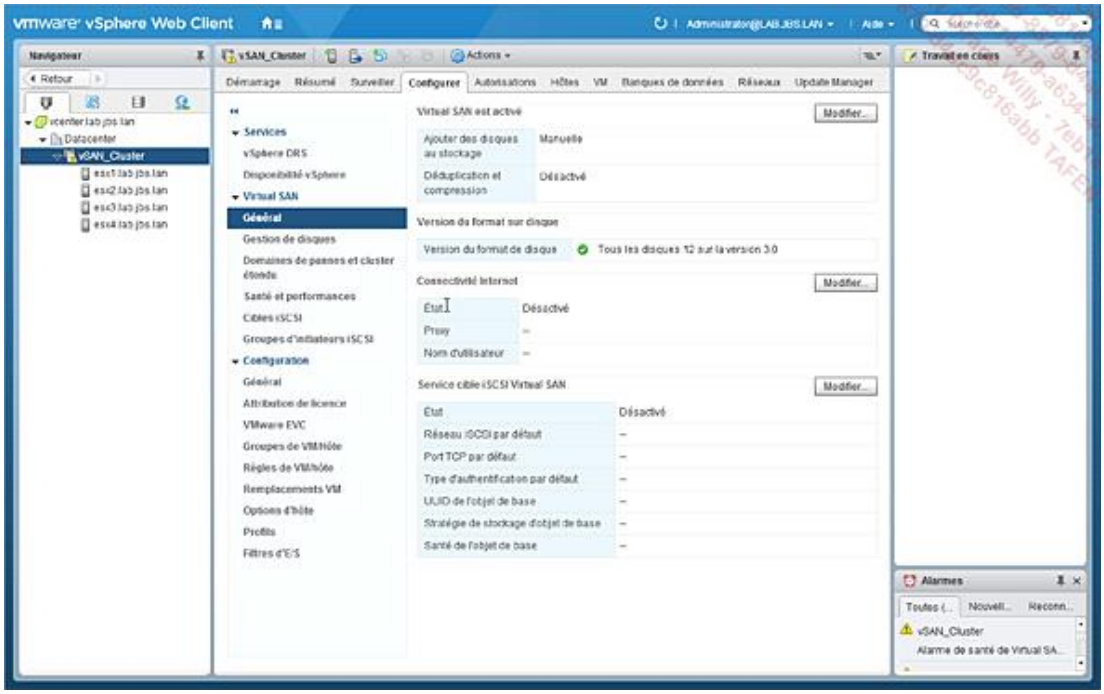


Notez que l'espace de stockage consommé (avec protection) d'un disque virtuel de 100 Go passe de 200 Go avec le mode *performance* à 133,33 Go avec le mode *capacity* .

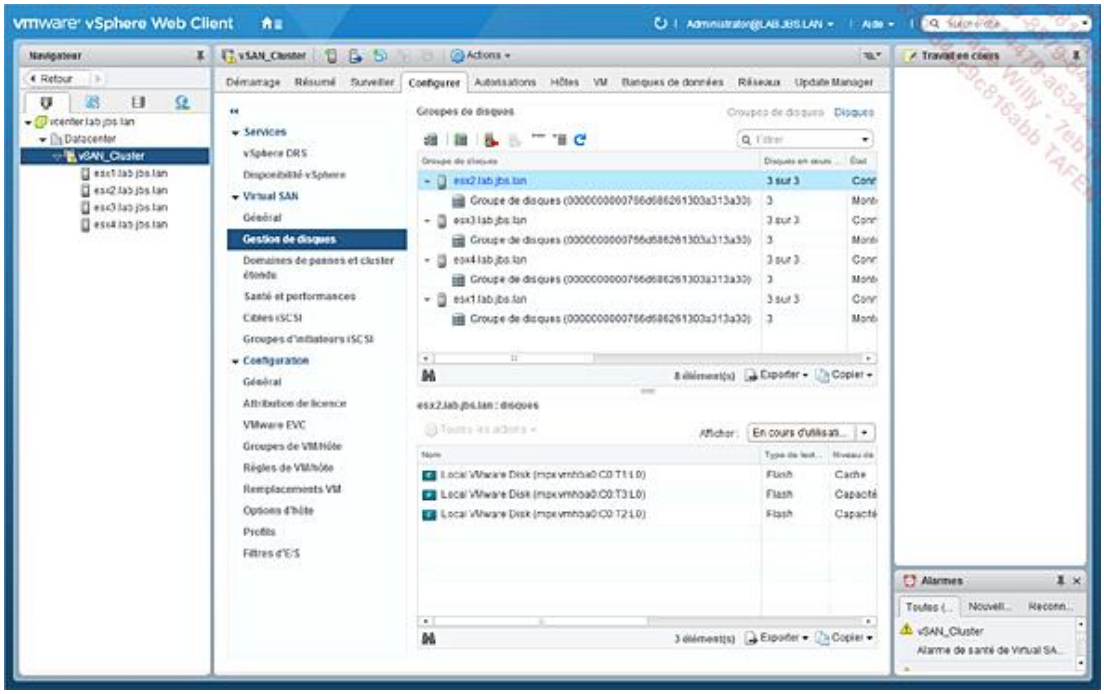
Vous pouvez valider la modification de la politique de stockage en cliquant sur le bouton OK. Prêtez attention à ce que vous faites puisque cette politique s'appliquera à l'ensemble du cluster et pourra déclencher des resynchronisations. Si vous souhaitez répondre à un besoin particulier, autant créer une nouvelle politique de stockage et l'assigner ensuite à votre machine virtuelle.

Balayons rapidement les sections de configuration de vSAN.

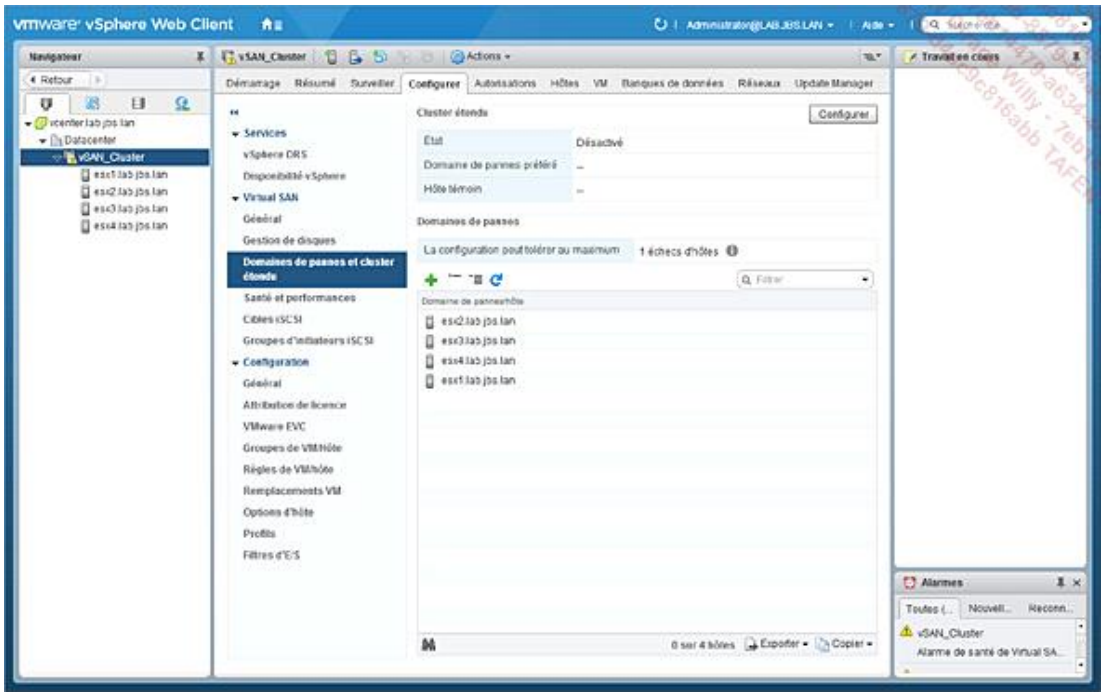
La section **Général** vous permet de définir le statut global de vSAN ainsi que les formats de disques, la connectivité Internet (notamment pour les mises à jour de listes de compatibilités matérielles, les HCL) et le service iSCSI qui est une nouveauté de vSAN 6.5.



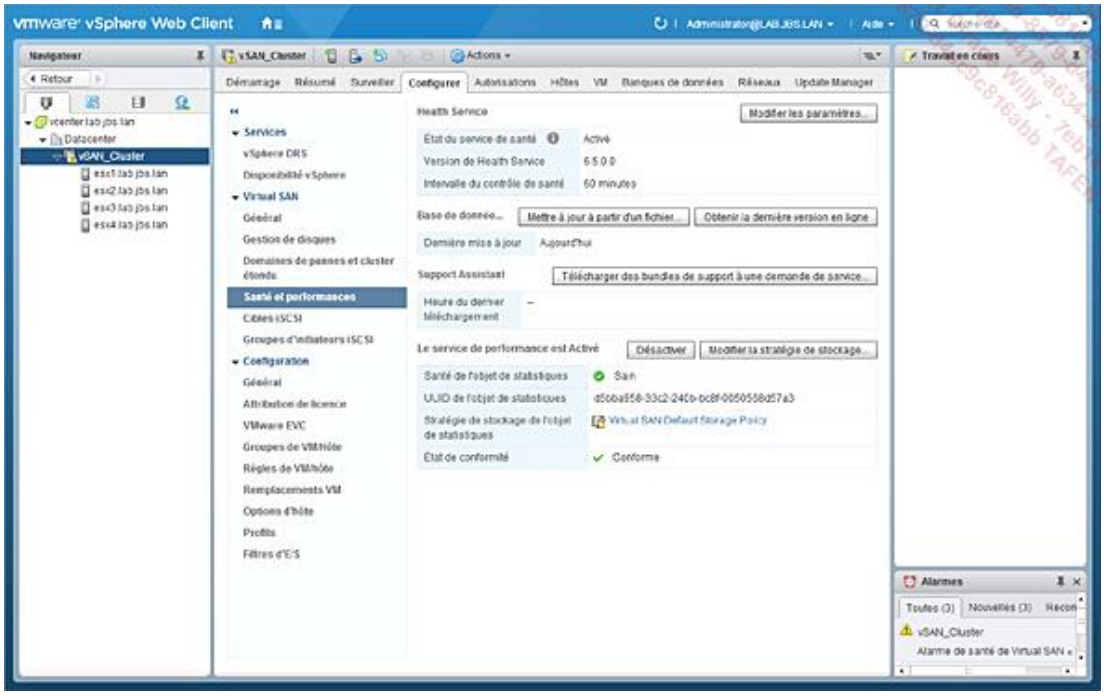
La partie **Gestion des disques** vous permet d'identifier les groupes de disques (*caching* + *capacity*) définis sur chaque hôte. Cette section vous permet également de vérifier le statut opérationnel de chaque disque.



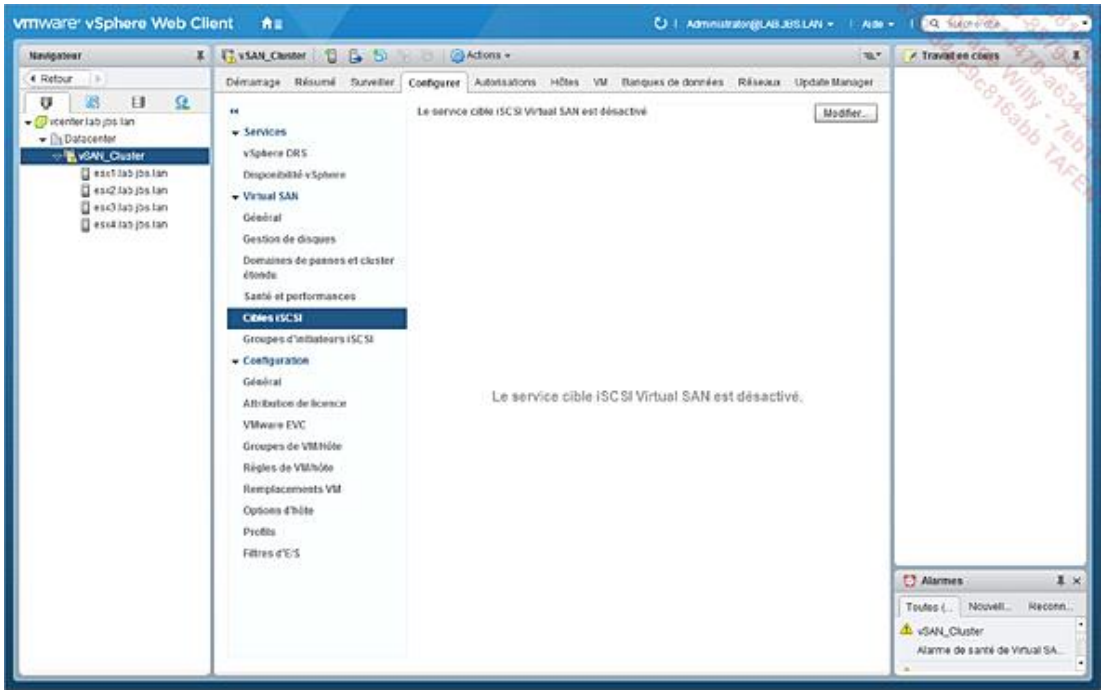
La section **Domaines de pannes et cluster étendu** affiche le statut de vos *fault domains* ainsi que le témoin nécessaire dans ces scénarios. Il affiche également votre tolérance à la panne de la configuration courante.



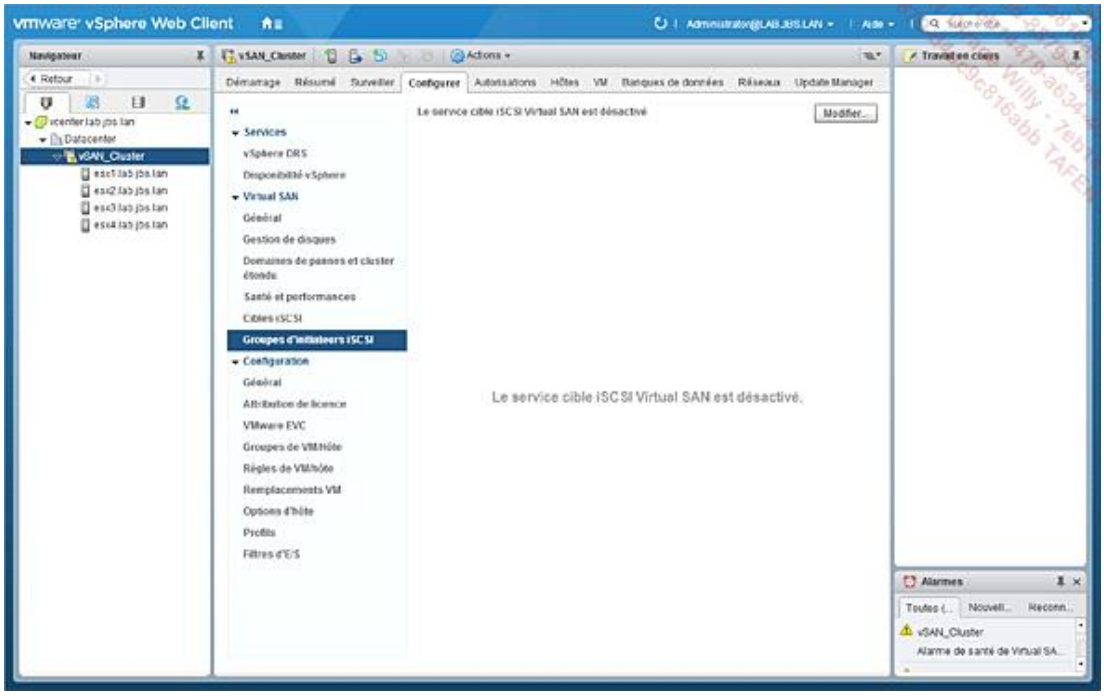
La section **Santé et performances** vous affiche le statut du service de santé et de performances vSAN. Il permet également de mettre à jour la base de données HCL ainsi que d'exporter les logs à la demande du support VMware directement sur leurs serveurs.



Les parties iSCSI vous permettent de créer vos cibles et initiateurs iSCSI pour exposer une partie de votre *datastore* vSAN partagé, par le biais de LUNs.



La partie **Groupes d'initiateurs iSCSI** vous permet de gérer plus facilement.



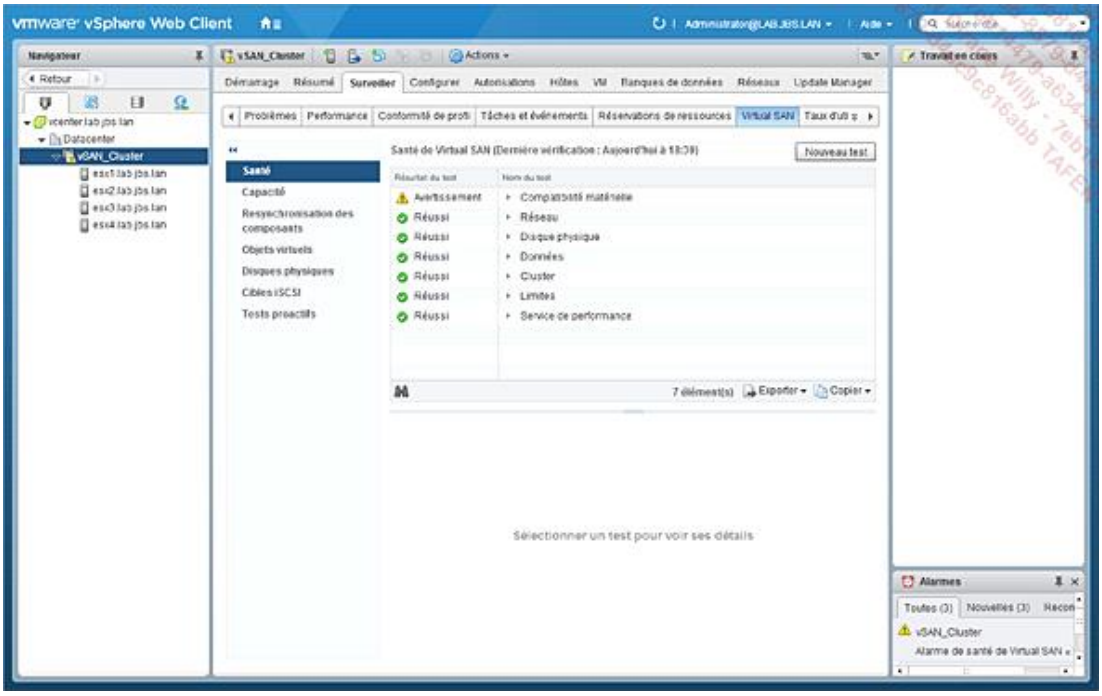
Terminons ce chapitre par une partie de supervision opérationnelle de votre cluster vSAN ainsi que les erreurs de production les plus courantes.

5. Supervision et résolutions de problèmes

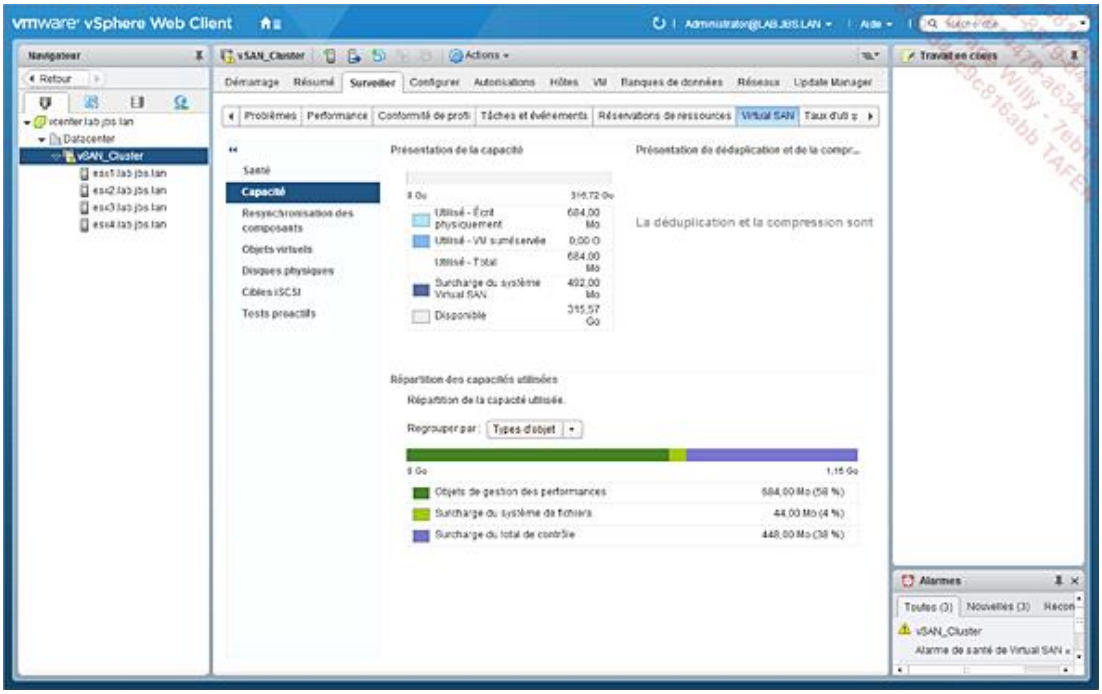
Une fois votre infrastructure mise en place, il est important de la surveiller pour vous permettre de vous assurer que celle-ci fonctionne toujours de façon nominale et que les besoins n'excèdent pas les ressources installées.

vSphere dispose de services pour vous assurer de la condition opérationnelle de vSAN. Faisons un tour d'horizon des différents écrans de supervisions que vous offre le vSphere Web Client. Rendez-vous dans l'onglet **Surveiller** au niveau de votre cluster pour faire apparaître ces options.

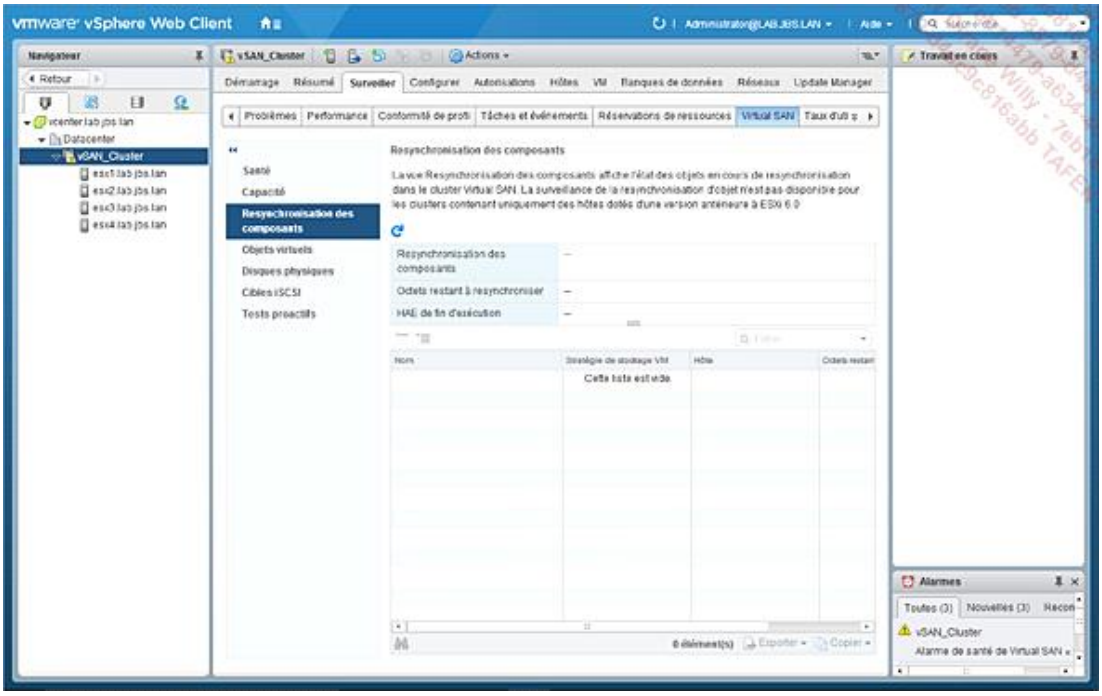
L'écran **Santé** vous permet de voir en un coup d'œil si l'ensemble de votre cluster vSAN est opérationnel, tant sur le statut réseau, des disques ou des données.



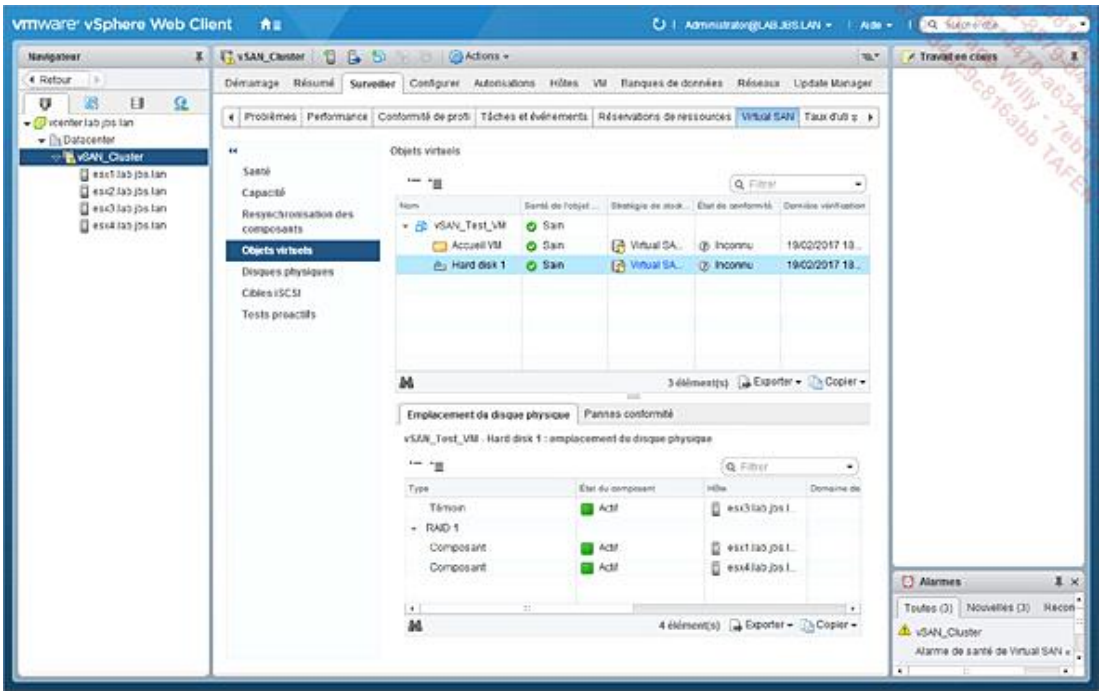
La rubrique **Capacité** vous permet de voir l'utilisation de votre espace de stockage, autant par les objets vSAN que les réplicas de ces derniers.



La rubrique **Resynchronisation des composants** permet à l'administrateur de visualiser les opérations en cours de resynchronisation et leur durée estimée de complétion.

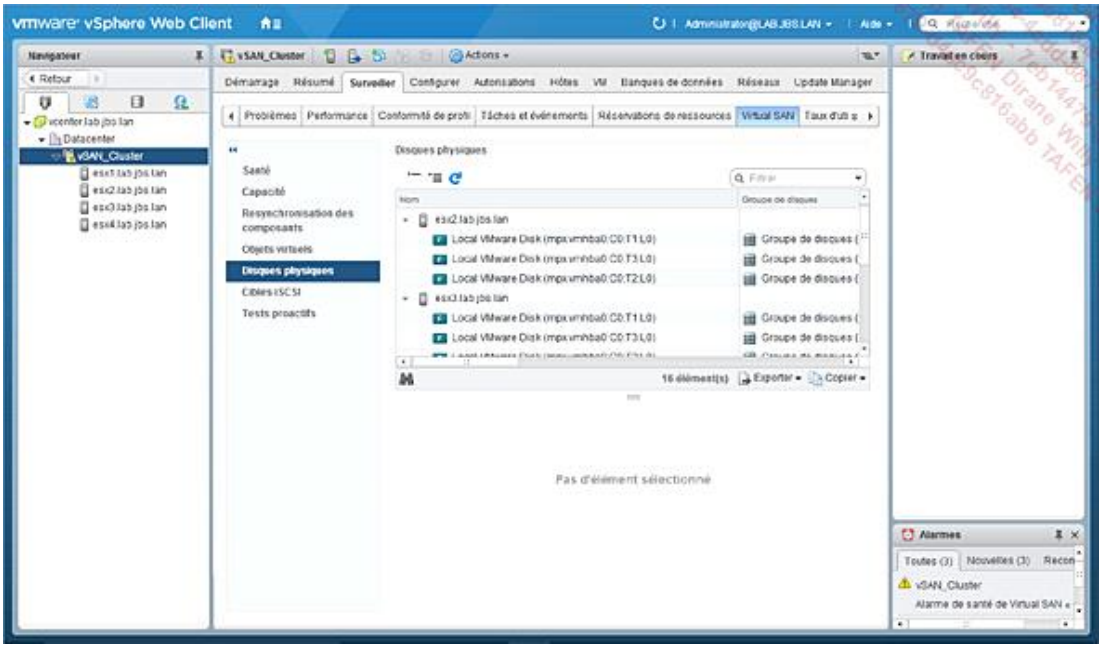


La rubrique **Objets virtuels** permet d'identifier le rôle de chaque hôte dans le stockage de vos machines virtuelles.

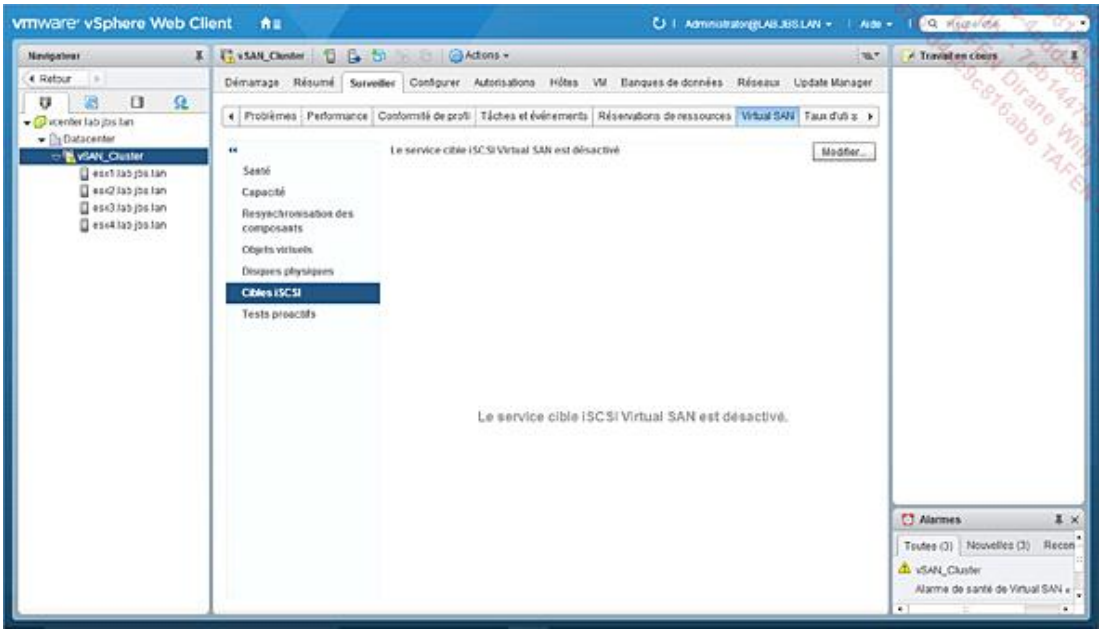


Nous constatons dans cet écran que le disque virtuel 1 (VMDK) de la machine virtuelle « vSAN_Test_VM » est porté par esx1.lab.jbs.lan et esx4.lab.jbs.lan qui stockent la donnée, quand esx3.lab.jbs.lan joue le rôle de témoin. Notez que le « *VM Homespace* », ici **Accueil VM** peut être porté par les mêmes hôtes ou d'autres, sans que cela pose problème.

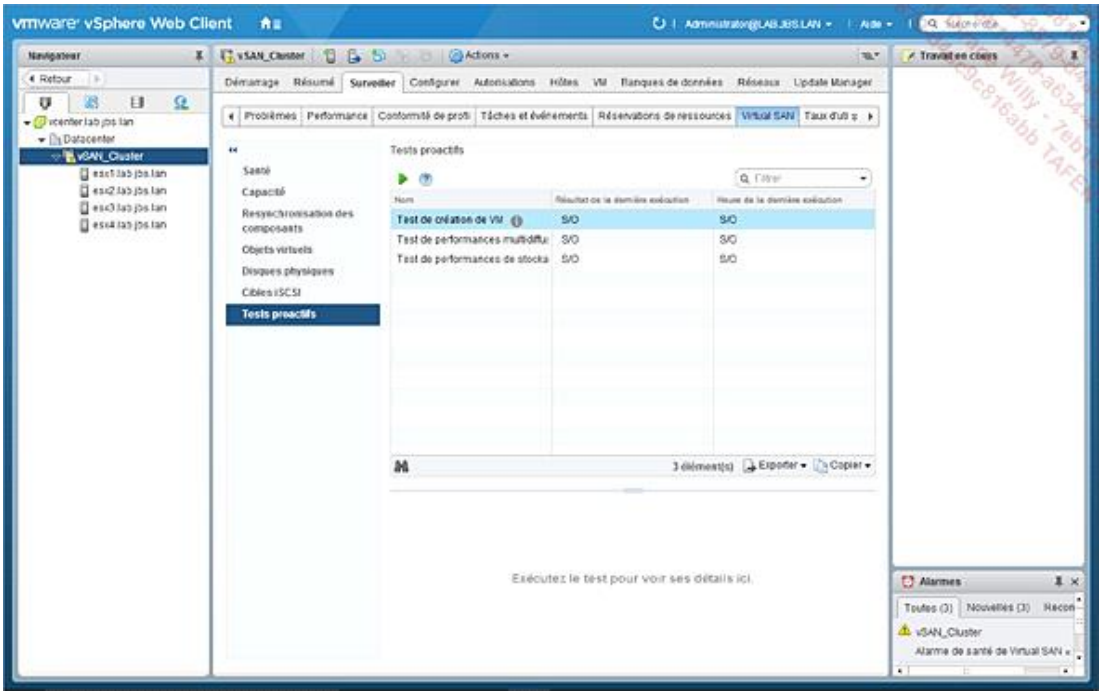
La section **Disques physiques** vous permet de visualiser les disques de chaque hôte et le statut de ceux-ci.



Les cibles iSCSI peuvent être surveillées à l'aide de la section dédiée.

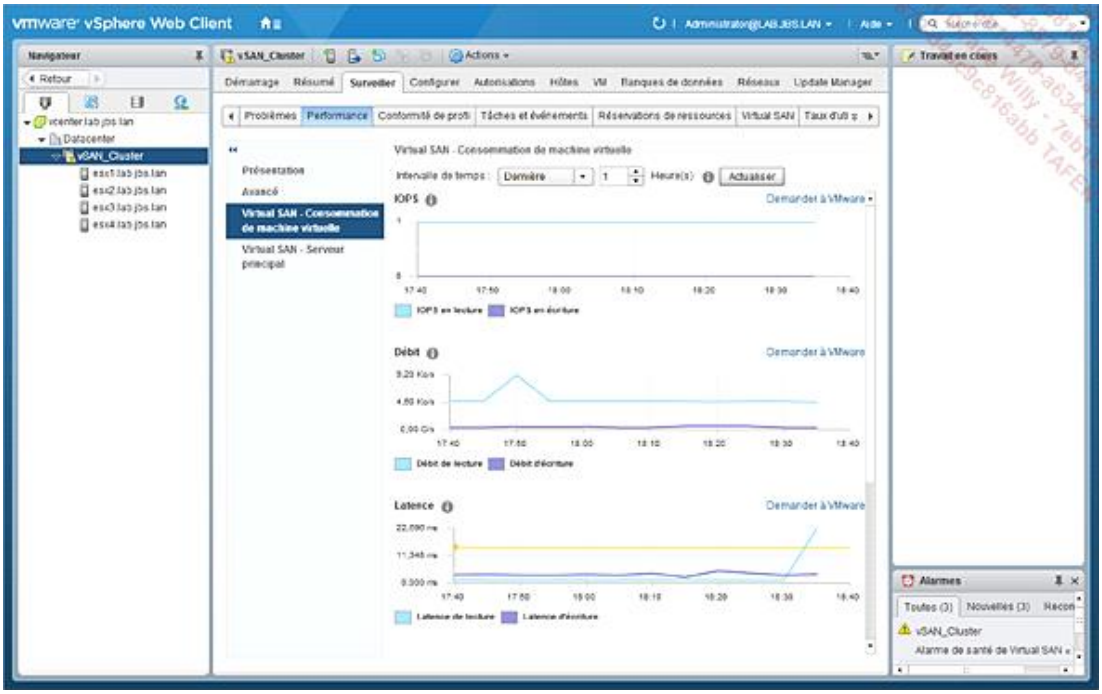


Vous pouvez mener des tests de création de machine virtuelle et de performances (parmi lesquels des stress tests) sur votre datastore virtuel pour vous assurer que ses performances sont alignées à vos besoins.

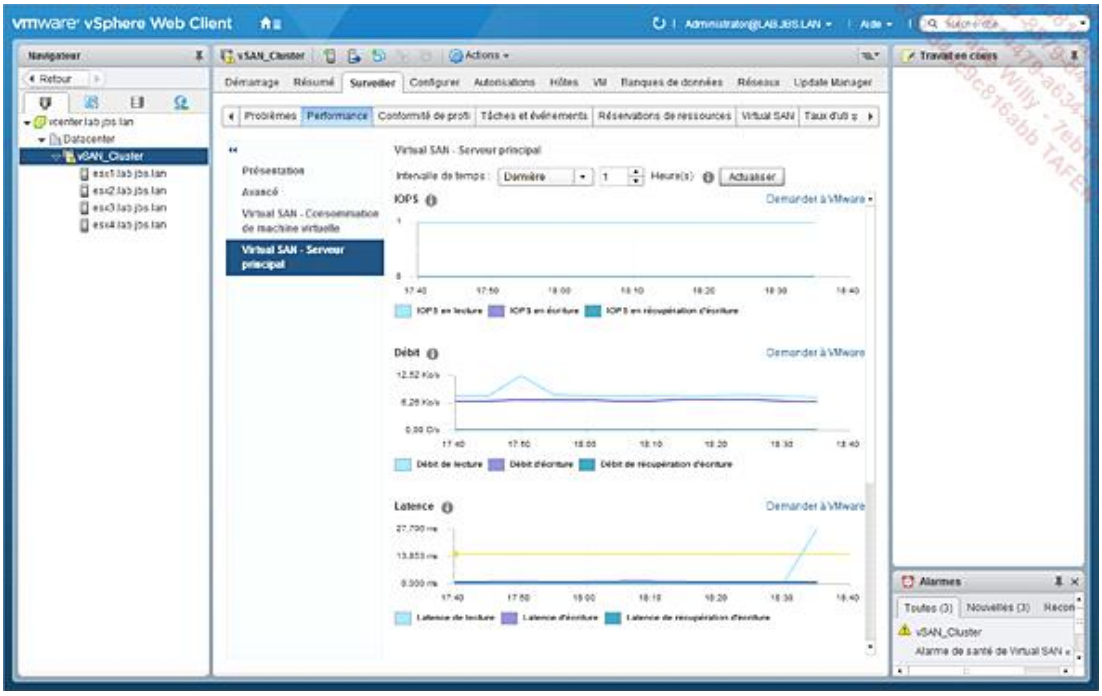


Enfin, la section **Performance** de l'onglet **Surveiller** vous permet de déterminer l'utilisation de votre datastore au niveau du cluster, d'un hôte ou d'une machine virtuelle.

La première rubrique **Consommation de machine virtuelle** graphe l'utilisation du datastore dans les machines virtuelles, tant sur les IOPS, le débit, la latence, les encombrements, et les I/O en attente.

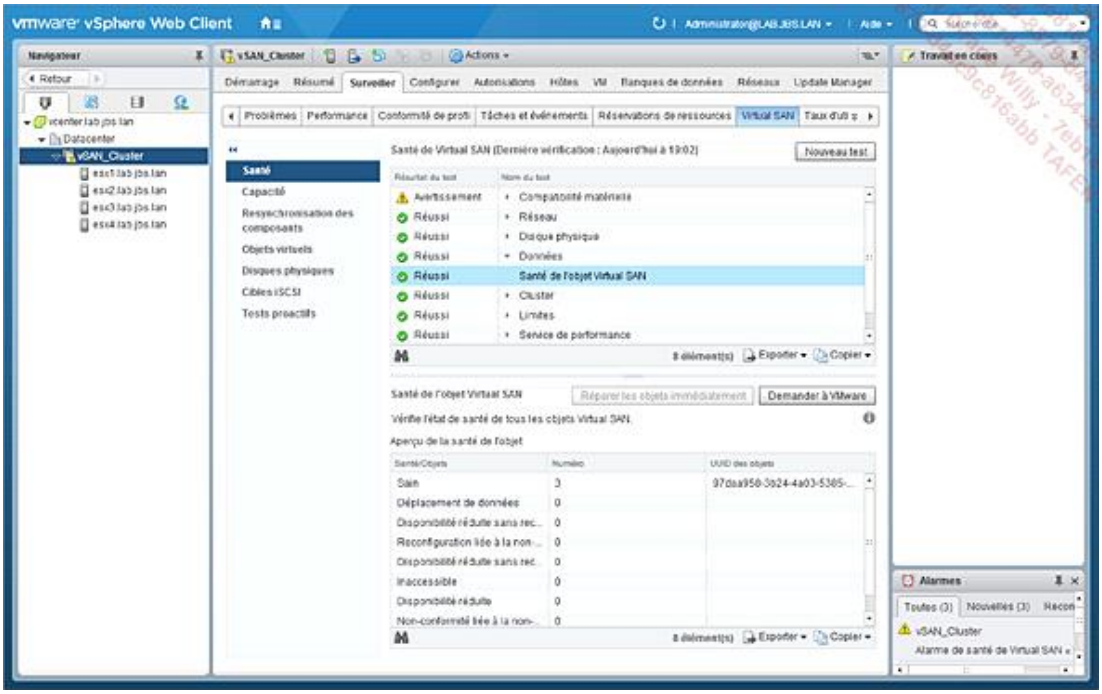


La partie nommée **Serveur principal** est plus explicite en anglais (nommée *backend*). Elle analyse la charge disque générée par le fonctionnement de vSAN, hors consommation donc des machines virtuelles utilisant le *datastore* .



Pour terminer cette partie, intéressons-nous aux erreurs générées par vSAN.

Les erreurs peuvent être présentes à tous les niveaux, vous l'avez vu dans la section **Santé** de l'interface de supervision de vSAN. Cependant, une section doit attirer votre attention, une dont le sens et les statuts peuvent être difficiles à interpréter correctement, celle relative au statut des objets stockés dans votre *datastore* .



Le tableau ci-dessous vous propose un résumé des différents statuts et comment y répondre en tant qu'administrateur :

Statut	Signification	Action à prendre
Sain	L'objet est dans une condition parfaite, aligné avec la politique de stockage et n'est pas sujet à quelconque opération de la	Aucune

	perspective de vSAN.	
Déplacement de données	L'objet est en bonne santé et conforme à la politique. Il est en train d'être déplacé, suite à un rééquilibrage ou à une évacuation des données.	Aucune
Disponibilité réduite sans reconstruction	L'objet a subi un échec mais vSAN a été en mesure de le tolérer. L'objet est toujours accessible et fonctionnel. Cependant, vSAN ne travaille pas à la protection de l'objet, à cause de la limitation de ressources ou d'un échec de protection.	Vérifier l'utilisation des ressources et leur statut. Vous devez résoudre le problème **au plus vite** pour éviter une indisponibilité (en cas de panne supplémentaire).
Disponibilité réduite sans reconstruction - minuteur du délai	L'objet a subi un échec mais vSAN a été en mesure de le tolérer. L'objet est toujours accessible et fonctionnel. Cependant, vSAN ne travaille pas à la protection de l'objet, à cause du minuteur (60 minutes par défaut) pour tenter de protéger les objets.	Vérifier la panne survenue et restaurer le statut de la ressource (hôte, réseau...) **avant** que le minuteur n'expire (pour éviter une reconstruction non nécessaire).
Disponibilité réduite	L'objet a subi un échec mais vSAN a été en mesure de la tolérer. L'objet est toujours accessible et fonctionnel. vSAN travaille activement à la protection de l'objet pour ramener l'objet à la conformité.	Restaurer la ressource manquante au plus vite pour éviter la reconstruction de l'ensemble des objets du *datastore* .
Reconfiguration liée à la non-disponibilité	vSAN travaille activement à la reconstruction des objets suite à une modification de l'administrateur de la politique de stockage. Cet événement n'est pas lié à une indisponibilité de ressources.	Aucune
Non-conformité liée à la non-disponibilité	Ce statut s'applique dans les autres statuts ne s'appliquent pas. Un objet dans un tel état n'est pas conforme à la politique de stockage mais fournit la FTT attendue.	Aucune. Selon VMware, il n'y a pas de cas documenté dans lequel cet état serait applicable.
Inaccessible	L'objet a subi plus de pannes qu'il était censé tolérer. Il est inaccessible.	Trouver l'origine et résoudre le problème au plus vite pour minimiser le temps d'indisponibilité (*downtime*) de la ressource.

Quand vous avez des questions relatives au fonctionnement de vSAN, n'hésitez pas à consulter les ressources VMware telles que la documentation ainsi que les KB (*Knowledge Base*).

Pour avoir plus d'information concernant ce chapitre, n'hésitez pas à lire les livres suivants :

- Essential Virtual SAN (VSAN) : Administrator's Guide to VMware Virtual SAN, Second Edition de Cormac Hogan et Duncan Epping.
- Storage Implementation in vSphere 5.0 Technology Deep Dive de Mostafa Khalil.
- Le guide de gestion du stockage, ainsi que le guide de gestion du VSAN de VMware disponible dans la documentation officielle (https://pubs.vmware.com/vsphere-65/index.jsp).