

Storage DRS

Le SDRS fournit la même action au niveau des datastores (gestion des espaces et gestion des I/O) que le DRS au niveau de la répartition de charges CPU et mémoire. Cependant, contrairement au DRS qui s'exécute toute les 5 minutes, le SDRS s'effectue par défaut toutes les huit heures. Il existe deux modes de fonctionnement pour le Storage DRS :

- Manuel : le SDRS fait des suggestions qui doivent être validées (approuvées) par l'administrateur VMware, afin qu'elles soient effectuées.
- Automatique : le SDRS prend en charge la suggestion et la validation avant d'effectuer les modifications suggérées.

Quelques caractéristiques de SDRS :

- Il prend en compte une agrégation logique des datastores - cluster de datastores : l'ensemble des datastores contenu dans un cluster est vu comme un objet logique unique. La répartition de charge s'occupe du placement initial, ainsi que des futures migrations (svMotion) en fonction des besoins en espace et en charge (I/O).
- Le placement initial définit l'emplacement initial de la machine virtuelle dans le cluster de datastore, id= id est sur quel datastore physique se trouvera la machine virtuelle. Cela se fait en fonction de l'espace disponible et de la latence d'I/O (si SIOC est activé) sur le datastore physique.
- L'équilibrage de charge de DRS et du Storage DRS fonctionne de la même manière. Là où le DRS équilibre les machines virtuelles entre les différents ESXi du cluster en se basant sur l'usage de la mémoire et du CPU, le Storage DRS lui, équilibre les machines virtuelles entre les datastores, en se basant sur la capacité en espace et/ou en IOPS.
- Les règles d'affinité/anti-affinité pour les disques durs (VMDK) des machines virtuelles. Par défaut les disques durs des machines virtuelles sont tous stockés ensemble dans le répertoire de la machine virtuelle sur le même datastore.
- La mise en maintenance des datastores : cela permet comme un pour cluster ESXi de pouvoir mettre de manière unitaire un datastore physique en maintenance. SDRS s'occupera de faire le Storage vMotion (svMotion) des machines virtuelles présentes sur le datastore physique.

Il est intéressant de noter que la remontée d'informations (statistiques de latence) se fait via SIOC. Ce dernier (SIOC) est utilisable aussi sans le Storage DRS. En se basant sur les statistiques de latence (SIOC) et de disponibilité d'espace, le sDRS peut utiliser le svMotion (<https://blogs.vmware.com/vsphere/2011/12/storage-drs-and-storage-array-features.html>) afin de répartir la charge en I/O et le besoin en espace entre les datastores. Par défaut, le Storage DRS ne fait de l'équilibrage de charge que sur l'espace disque (vSphere 5.5), sous vSphere 6.5 SIOC est activé par défaut.

1. Le cluster de datastores

Un cluster de datastores peut être considéré comme un dossier. L'activation de storage DRS n'est pas obligatoire. Cependant, activer storage DRS permet d'équilibrer l'utilisation des disques et la latence des IOPS (*Input Output Operations Per Second*).

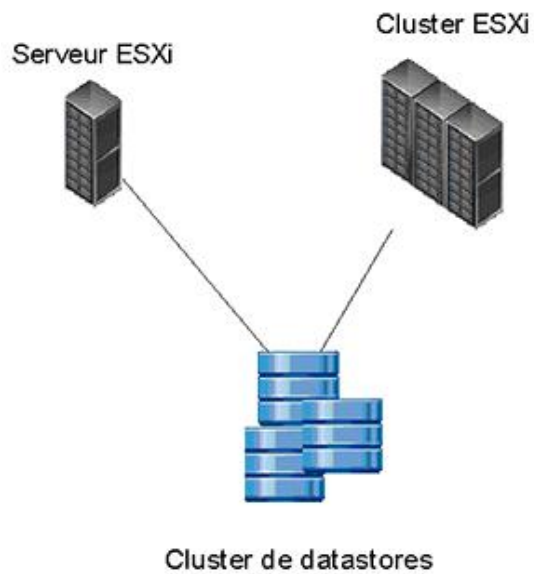
Un cluster de datastores peut contenir des datastores de plusieurs serveurs de stockage différents, cependant il convient de vérifier les performances car les LUN les plus lents ont un impact défavorable sur les performances générales. Le cluster peut être composé de datastores VMFS ou NFS, et ne supporte que des serveurs hôtes en version 5.0 minimum. Les serveurs ESX ou ESXi de versions précédentes ne peuvent être connectés à un cluster. Le mélange de systèmes de fichiers VMFS-3 et VMFS-5 au sein d'un même cluster est possible. Par contre un datastore en VMFS6 ne pourra être qu'au sein d'un cluster de datastore en VMFS6.

Dans le cas d'une activation de storage DRS sur un cluster, des règles plus précises sont à respecter :

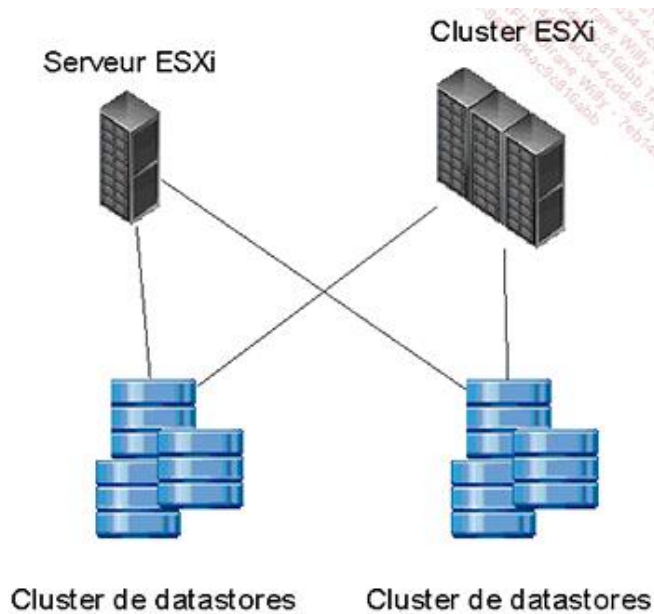
Storage DRS ne migre pas les données de machines virtuelles d'un datastore VMFS vers un datastore NFS, donc les clusters doivent être constitués d'un seul type de datastores (VMFS ou NFS).

Il ne faut pas mélanger de datastores répliqués et des datastores non répliqués car la réplication peut engendrer des latences supplémentaires et cela fausserait la configuration du cluster car les performances doivent être identiques. Storage DRS peut fonctionner en même temps que HA et DRS (cluster de serveurs hôtes).

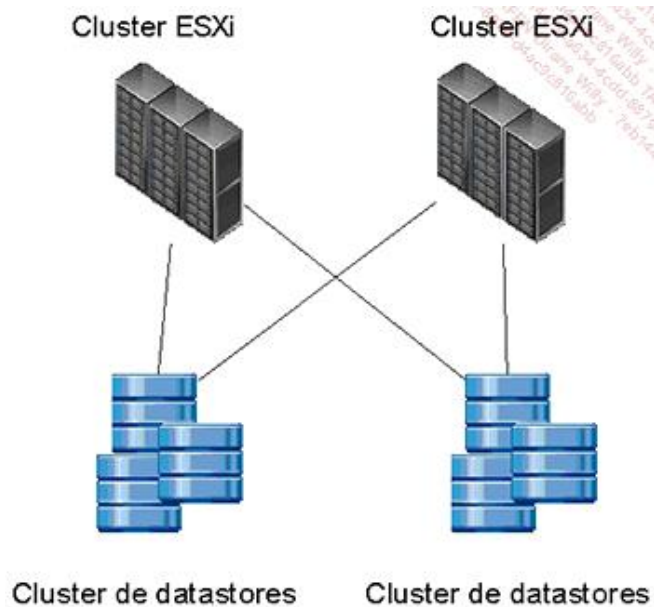
Un serveur et un cluster de serveurs peuvent être connectés à un cluster :



Un serveur hôte et un cluster de serveurs hôtes peuvent utiliser les mêmes clusters de datastores :

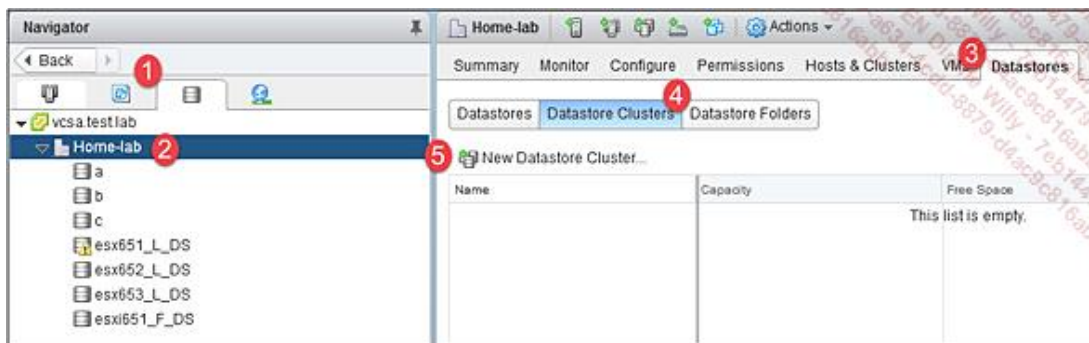


Plusieurs clusters de serveurs hôtes peuvent être connectés aux mêmes clusters de datastores :

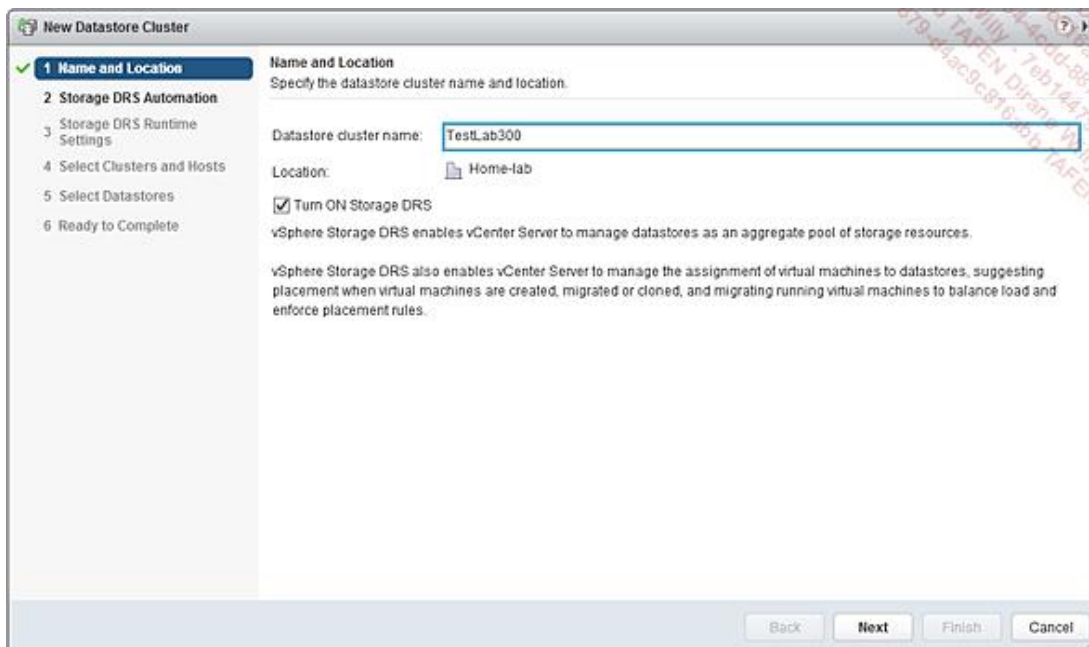


La création d'un cluster de datastores est assez simple via l'interface graphique, mais prend un peu de temps.

Lorsque vous êtes sur l'onglet des datastores de l'inventaire (1), sélectionnez le Datacenter (2), puis cliquez sur l'onglet **Datastores** (3). Vous pouvez maintenant créer notre datastore. Etant donné que vous voulez créer un cluster de datastores, vous devez vous mettre dans le cluster de datastores (4), pour enfin lancer l'interface de configuration du cluster de datastores (5).

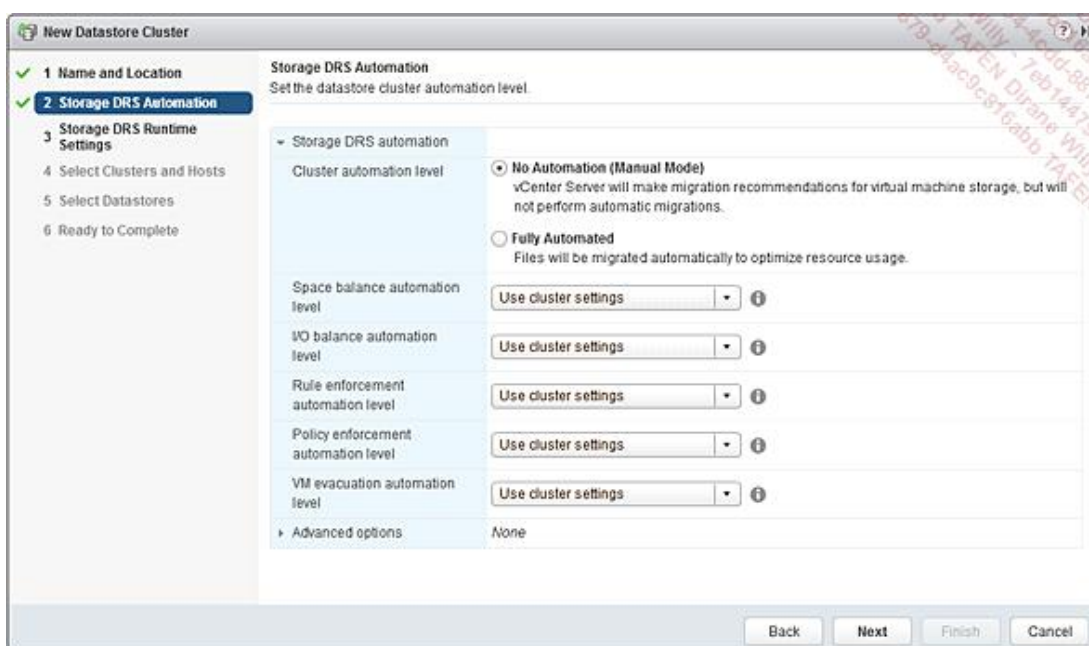


Renseignez tout d'abord le nom du cluster de datastores, puis validez.



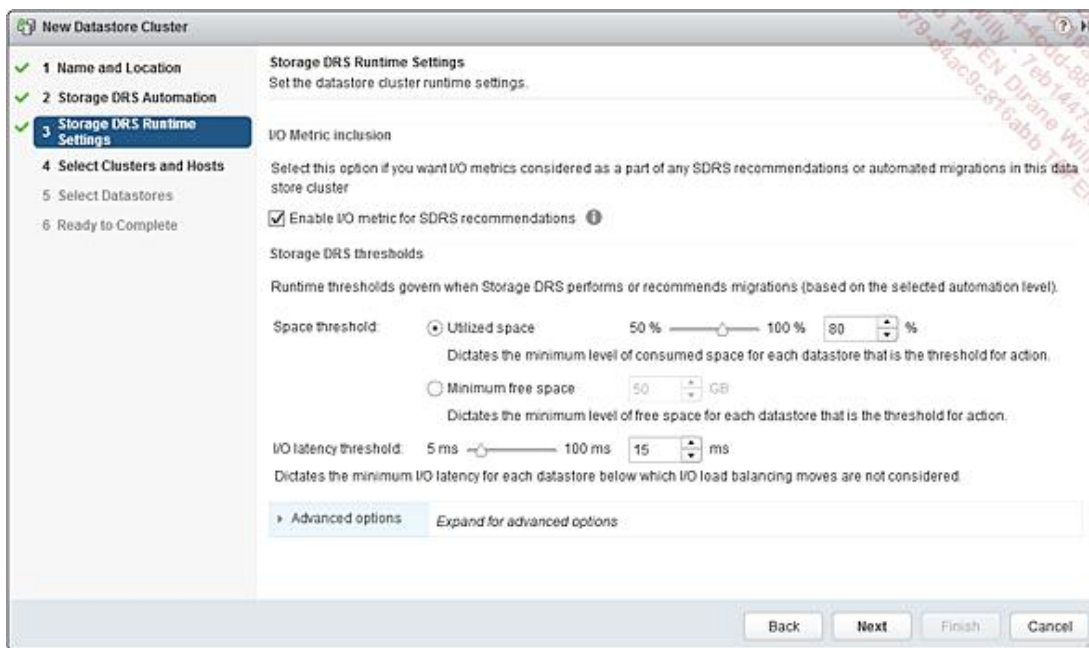
Si la configuration par défaut convient (**No Automation**) cliquons sur **Suivant**. sinon il est possible de configurer chacun des paramètres de manière individuelle.

- **Space Balance automation level** : niveau d'automatisation de l'équilibrage de charge en espace disque entre les différents membres du cluster de datastore.
- **I/O Balance automation level** : niveau d'automatisation de l'équilibrage de charge en I/O entre les différents membres du cluster de datastore.
- **Rule enforcement automation level** : niveau d'automatisation dans la correction de violation des règles d'affinités.
- **Policy enforcement automation level** : niveau d'automatisation dans la correction de violation des règles de placement des VMs dans le cluster de datastore.
- **VM evacuation automation level** : niveau d'automatisation de l'évacuation des machines virtuelles lors de la mise en maintenance d'un datastore.

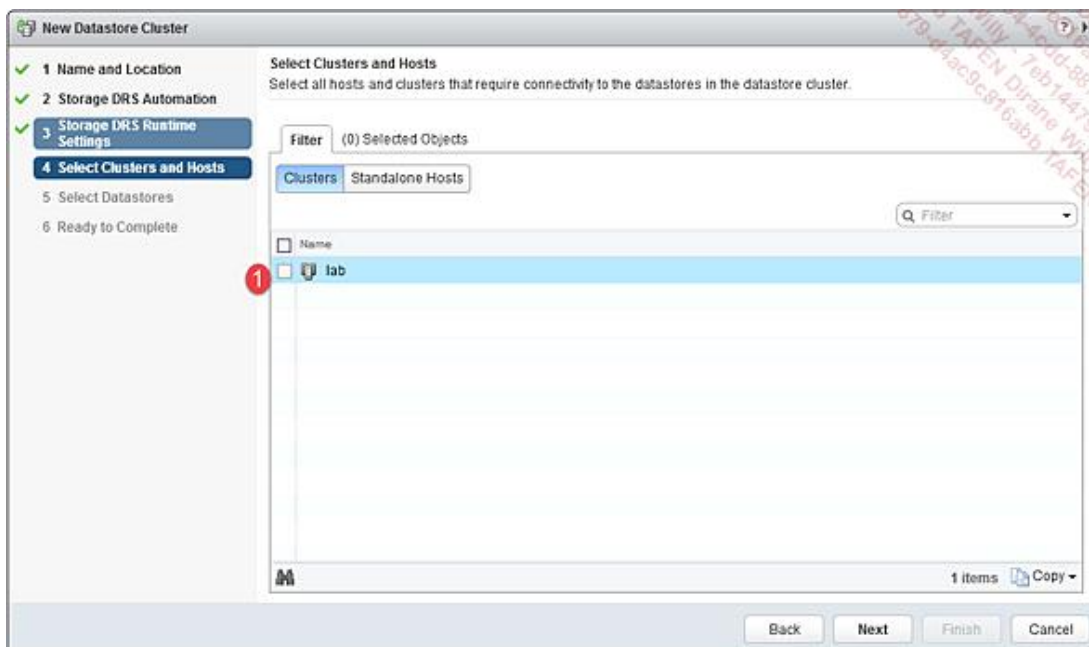


À l'étape suivante, configurez les seuils de détection et d'activation du Storage DRS. Par défaut, SIOC est actif en version 6.5 (I/O metric for SDRS recommendation), dans le cadre de la remontée de performance des datastores (latence).

Attention, le seuil de latence I/O au niveau Storage DRS ne doit pas être supérieur au seuil de la latence I/O de SIOC.



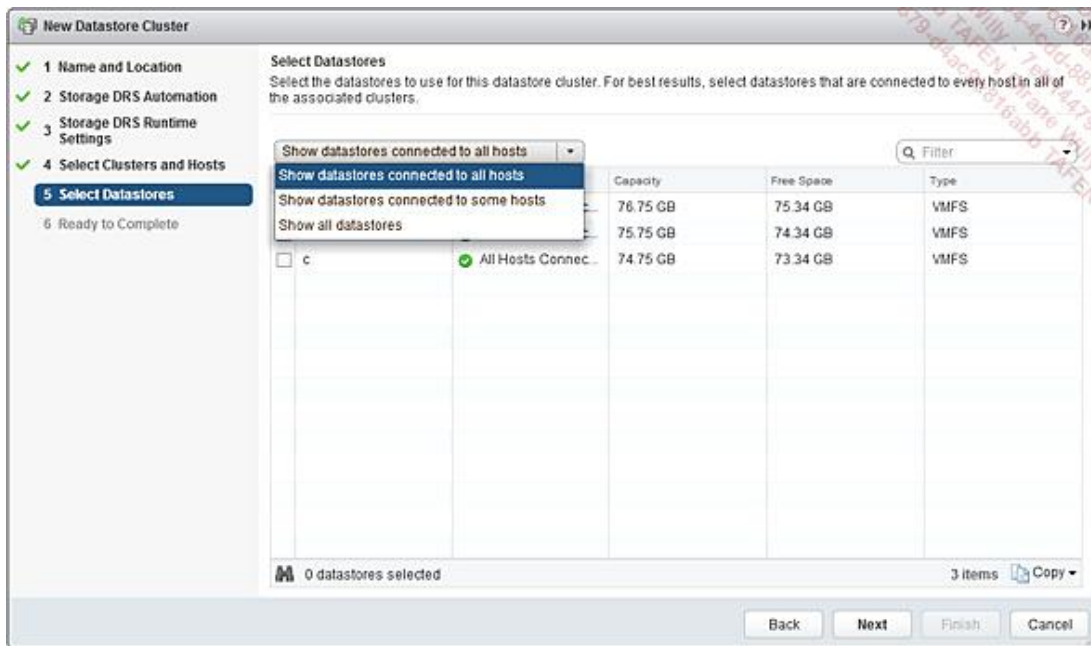
Sélectionnez le cluster d'hôtes pour lequel ce cluster de datastore est disponible.



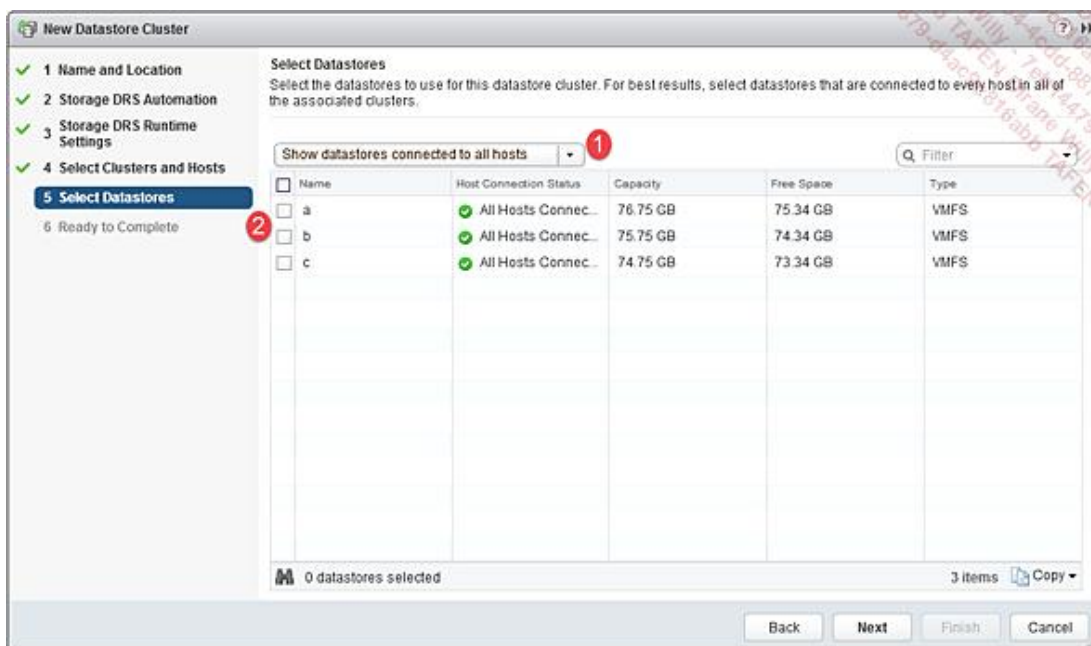
Puis filtrez l'accessibilité des datastores par les ESXi, il y a trois possibilités (<http://frankdennean.nl/2011/10/07/partially-connected-datastore-clusters/>) :

- Montrer tous les datastores, montre tant les datastores locaux (« disques » physiques du serveur), que les datastores distants.

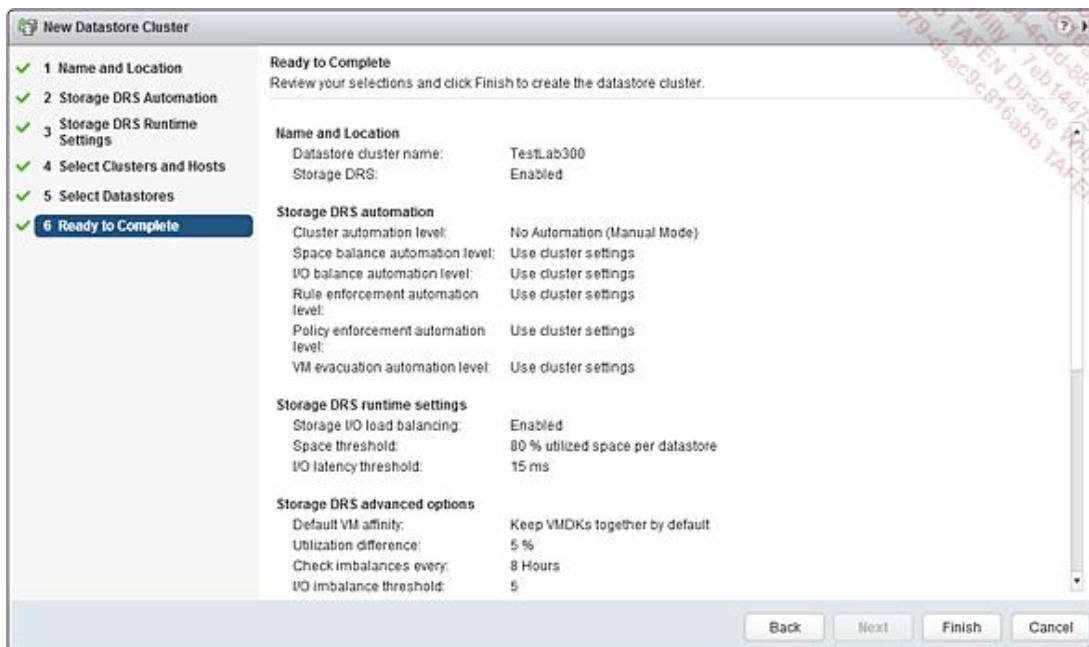
- Montrer tous les datastores connectés à certains ESXi, un datastore est dit partiellement connecté lorsqu'il n'est pas vu par l'ensemble des ESXi du cluster.
- Montrer tous les datastores connectés à tous les ESXi, il s'agit d'un datastore qui est vu par l'ensemble des ESXi du cluster.



En fonction du critère de sélection, sélectionnez les datastores qui nous intéressent.



Et validez le résumé des opérations à venir.



En ligne de commande via PowerCLI :

```
New-datastorecluster -name TestLab300 -location Home-lab
Set-datastorecluster -datastorecluster TestLab300 -
IOlatencyThresholdmillisecond 15 -IOLoadBalanceEnabled 1 -
SdrsAutomationLevel Manual SpaceUtilizationThresholdPercent 80
```

Le paramètre `IOLoadBalanceEnabled` (SIOC) accepte les valeurs suivantes :

- 0 ou `disable`
- 1 ou `fullyautomated`
- 2 ou `manual`



2. Les règles d'affinités

Il est possible de créer des règles d'affinité et anti-affinité pour rassembler ou séparer différents disques virtuels sur un datastore :

- Affinité intra-VM : tous les fichiers disques d'une machine virtuelle sont conservés sur le même datastore. C'est la règle par défaut.

- Anti-affinité intra-VM : les différents disques d'une même VM sont stockés dans des datastores différents. Ce type de règles s'applique sur tous les disques d'une VM ou certains d'entre eux.
- Anti-affinité inter-VM : les disques de plusieurs machines virtuelles sont stockés sur des datastores différents. Ce type de règle est similaire à l'anti affinité (entre machines virtuelles) configurable au niveau du cluster DRS.

Pour vider un datastore, on peut utiliser le mode maintenance : les machines virtuelles enregistrées dans l'inventaire seront déplacées vers un autre datastore. SDRS ne déplacera pas les fichiers de machines virtuelles non enregistrées ni les fichiers de templates.

À la création d'une VM, ou si l'administrateur clone ou déploie à partir d'un template, le choix de stockage ne se fait plus sur le datastore mais sur le cluster. SDRS utilisera le datastore selon l'utilisation et la charge en IOPS.

Par défaut, SDRS vérifie l'utilisation disque des datastores toutes les cinq minutes, et l'historique des IOPS toutes les huit heures afin de faire ses recommandations et éventuelles migrations.

3. Plus d'informations sur SIOC Storage I/O Control

a. Version 1

Avant l'existence de SIOC, le paramètre avancé `Disk.SchedNumReqOutstanding` (DSNRO) permettait de contrôler le nombre total de commandes (I/O) émis depuis l'ensemble des VMs d'un host vers un stockage partagé. La valeur par défaut du paramètre DSNRO est de 32.

Ce paramètre fonctionne avec le `Disk.SchedQControlVMSwitches`. La configuration par défaut de ce paramètre est 6. Cela signifie qu'il faut 6 I/O consécutif ne provenant pas de la même VM, avant que l'hyperviseur commence à réduire la valeur `Disk.SchedNumReqOutstanding`. Ces paramètres se configurent au niveau de l'hyperviseur.

Lorsque l'on ajoute le paramètre `Disk.SchedQuantum` qui correspond au nombre d'I/O consécutif dans les 2000 secteurs du disque qui est configuré à 8. Cela définit le nombre d'I/O indiqués comme séquentiel. Le paramètre `Disk.SchedQControlSeqReqs` permet à une VM d'envoyer jusqu'à 128 I/O avant qu'une autre VM ne puisse envoyer des I/O.

Afin que chaque machine virtuelle puisse travailler et avoir accès aux ressources de stockage, d'une manière plus juste, VMware fournit la fonction Storage I/O Control (SIOC). Le SIOC permet une gestion des I/O de manière plus juste entre les machines virtuelles basées sur la latence. Sur chaque LUN, toutes les 4 secondes s'effectue un test de latence. Le résultat de ce test se trouve dans le fichier `/vmfs/volumes/UUID/iormstats.sf`.

Il existe deux manières de configurer SIOC :

- L'utilisation d'un modèle injecteur : SIOC détermine le seuil de la meilleure latence par datastore. Ce résultat est dépendant de la baie et des disques. La valeur par défaut est de 90 % de la bande passante.
- La configuration manuelle : où l'on définit les seuils. La valeur par défaut est de 30 ms.

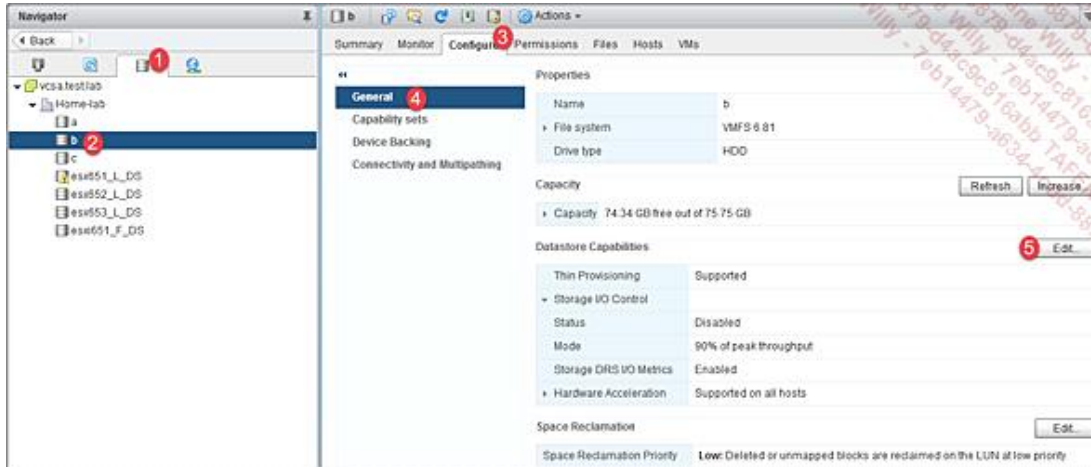
Nous ne devons pas confondre la latence SIOC et la latence SDRS. Tandis que le SDRS utilise la valeur de `vmObservedLatency`, qui correspond au temps que met un I/O entre la prise en charge par l'hyperviseur de la requête de l'I/O depuis la VM jusqu'à la réponse de la baie de stockage et du datastore. SIOC utilise la latence entre l'hyperviseur et le datastore. La finalité de SIOC diffère de celle de SDRS. SDRS va tenter de réduire et d'éviter les contentions, tandis que SIOC va réguler le flux des I/O lors de contention. Lorsque SIOC est configuré manuellement, il est recommandé de configurer la latence de SIOC de manière supérieure au seuil de latence de SDRS.

SIOC ne supporte pas les Raw Device Mapping, et les multiples extensions des datastores.

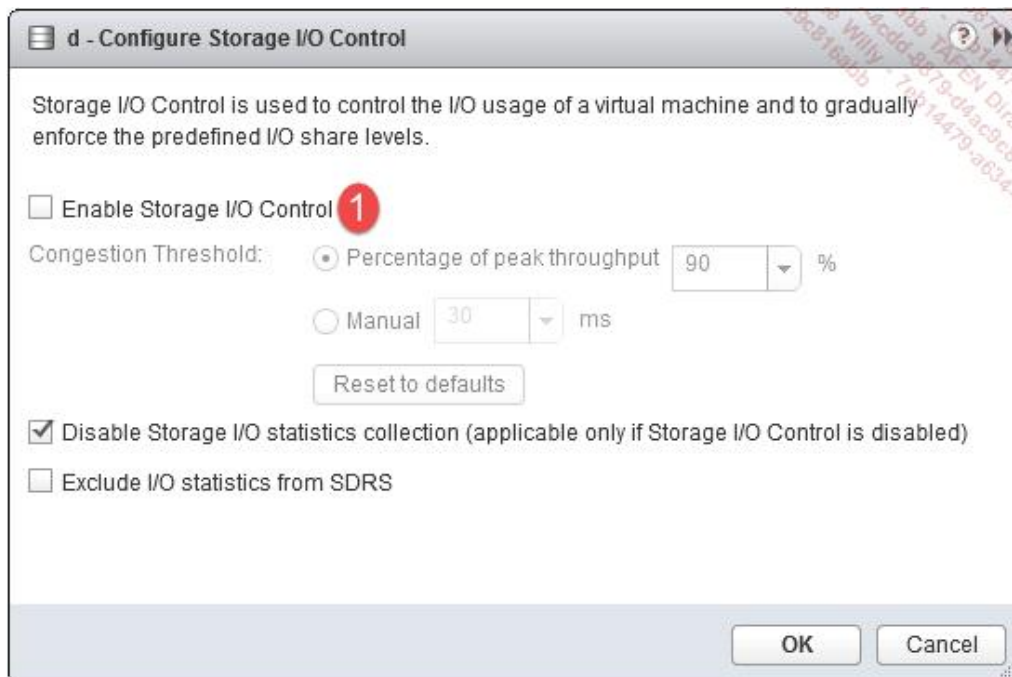
Le fichier de log de SIOC est propre à chaque serveur, il se trouve dans /var/log/vmkernel. Pour configurer ce fichier de log, nous devons configurer le paramètre Misc.SIOCControlLogLevel où 0 (inactif) et 7 (extrêmement verbeux). La prise en compte se fait en rebootant le service SIOC (/etc/init.d/storeRM restart).

L'activation de SIOC est très simple, elle s'effectue au niveau de chacun des datastores.

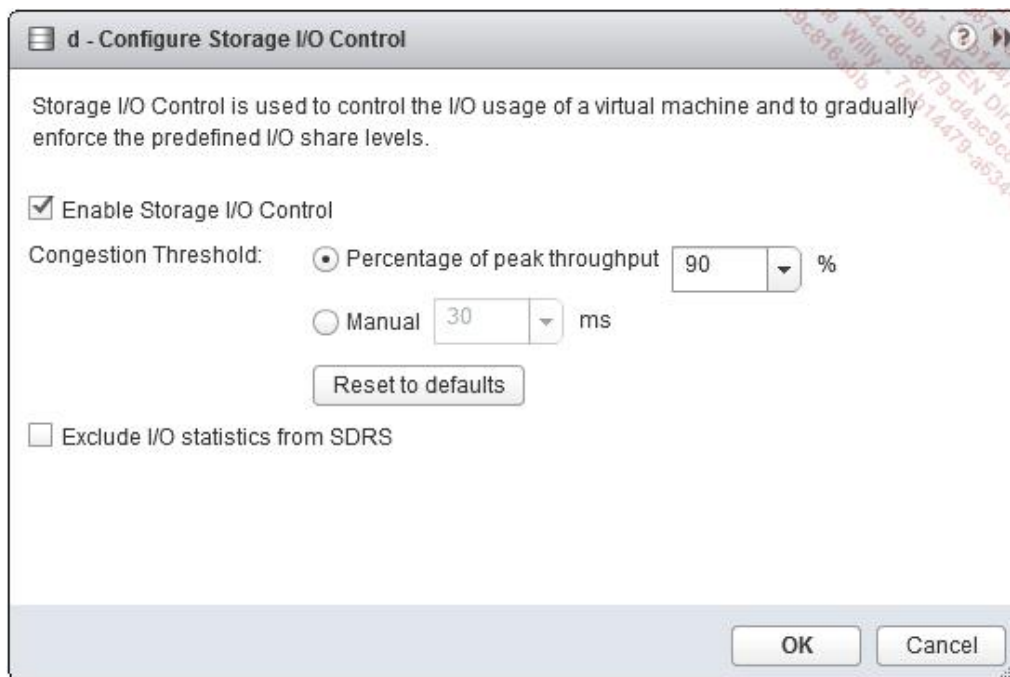
Éditez les propriétés du datastore.



Puis activez la fonctionnalité SIOC.



À ce moment-là, vous pouvez choisir la configuration du seuil de congestion (en pourcentage de bande passante ou latence).



Cliquez sur **OK**.

Datastore Capabilities		Edit...
Thin Provisioning	Supported	
▼ Storage I/O Control		
Status	Enabled	
Mode	90% of peak throughput	
Storage DRS I/O Metrics	Enabled	
► Hardware Acceleration	Supported on all hosts	

En cas de soucis avec SIOC, l'article suivant est un bon recueil des problèmes que l'on peut généralement rencontrer : <https://kb.vmware.com/kb/1022091>.

b. Version 2

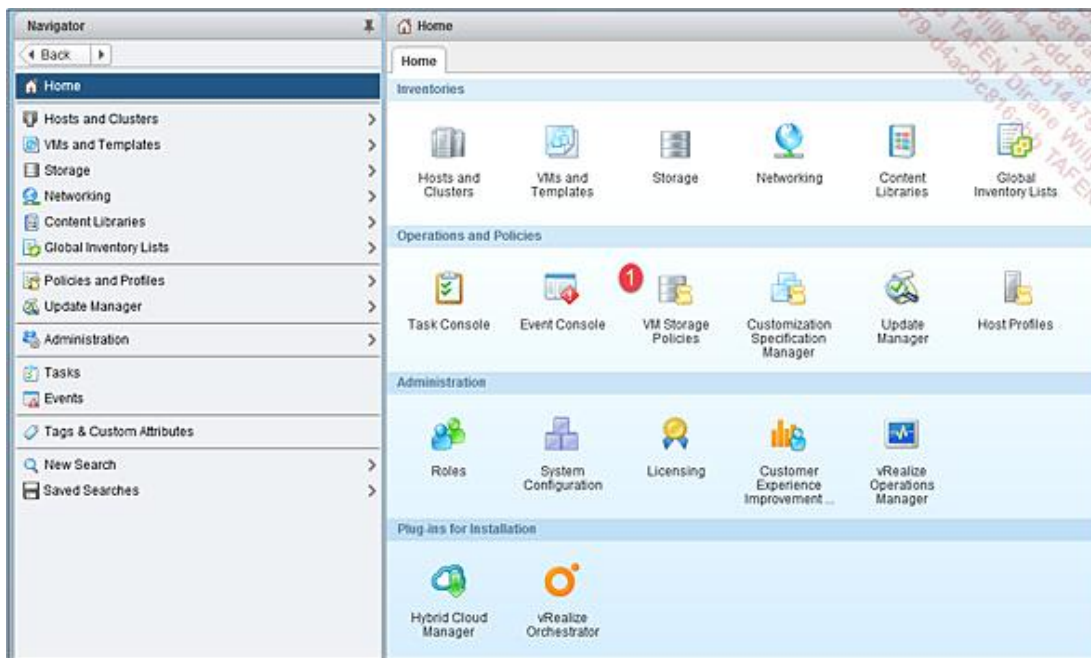
La version 2 (propre à la version vSphere 6.5) de SIOC travaille de pair avec VAAI ce qui permet d'inclure la configuration SIOC dans la politique de gestion de stockage. Il est utilisable sur tout type de stockage exception faite des RDM.

Il existe trois types de profils d'I/O qui sont définis par ce tableau.

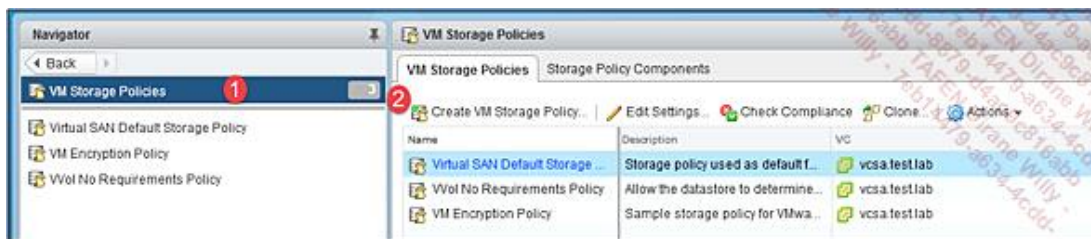
Type	Low	Normal	High
Limites	1000	10000	100000
Réservations	10	50	100
Shares	500	1000	2000

Ces valeurs pourront être configurées via la création d'une politique de stockage :

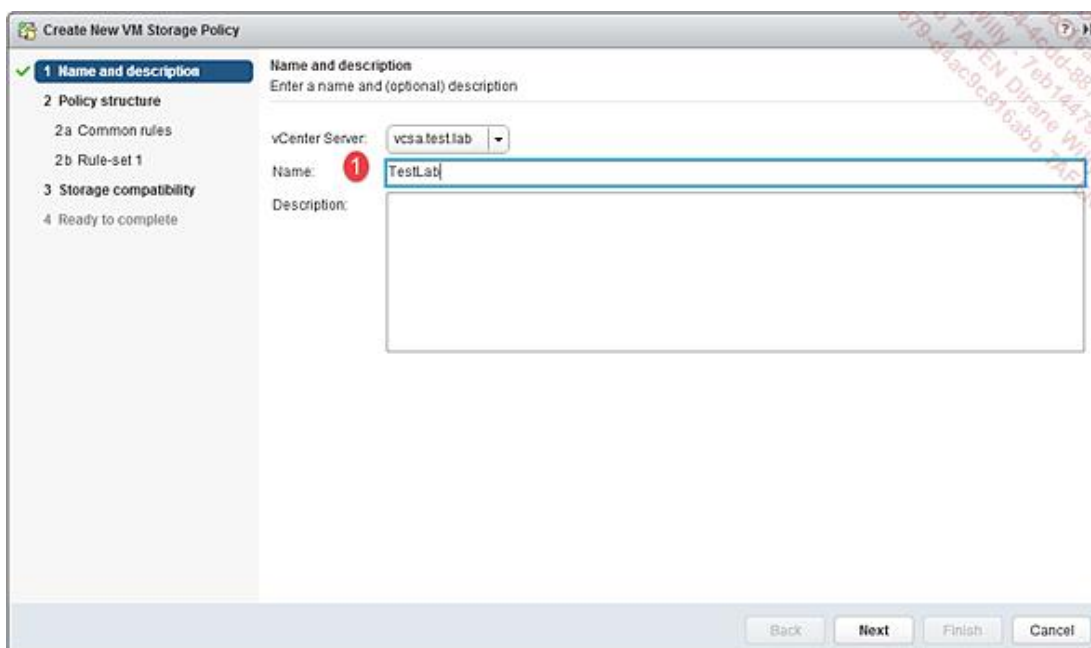
Dans le menu d'accueil, cliquez sur **VM Storage Policies** (1).



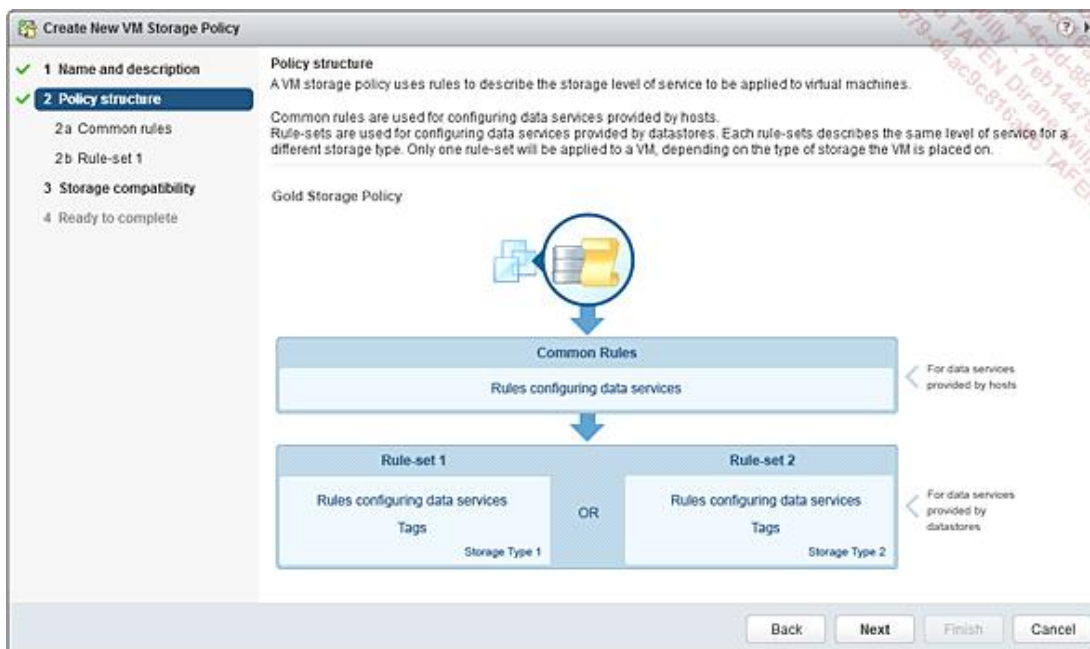
Cliquez sur **Create VM Storage Policy** pour créer une nouvelle stratégie de gestion du stockage (2).



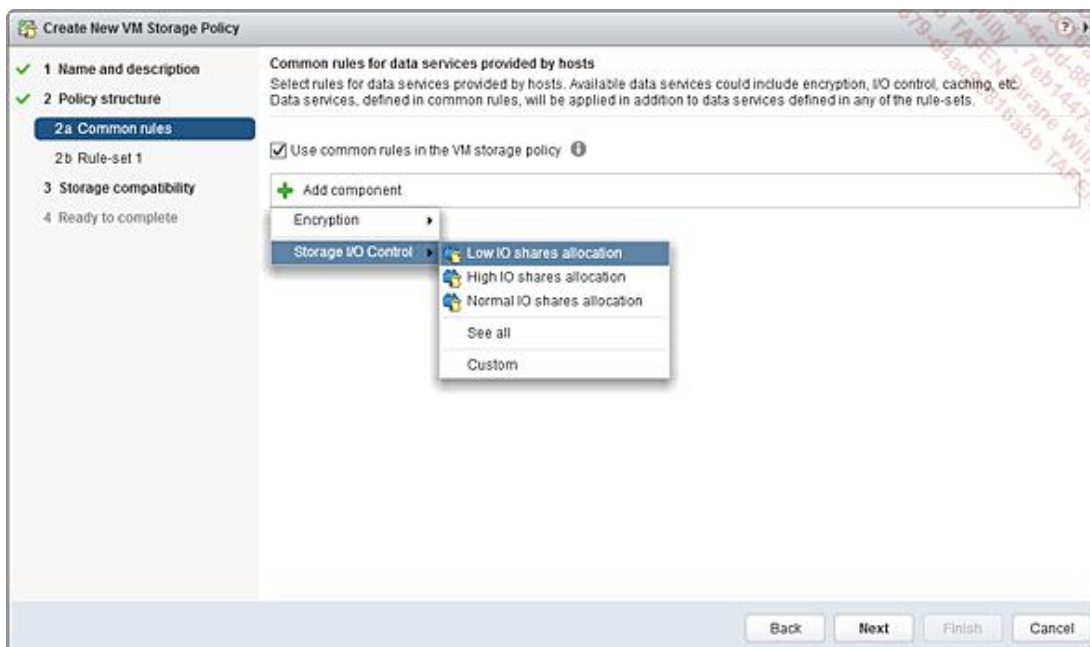
Nommez la stratégie de stockage.



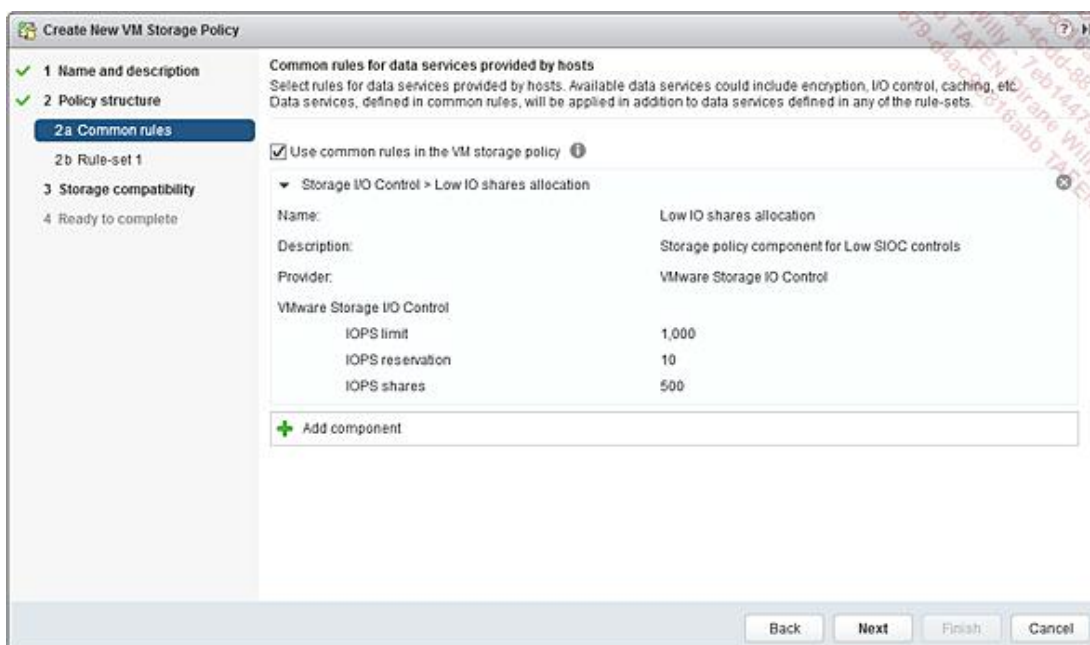
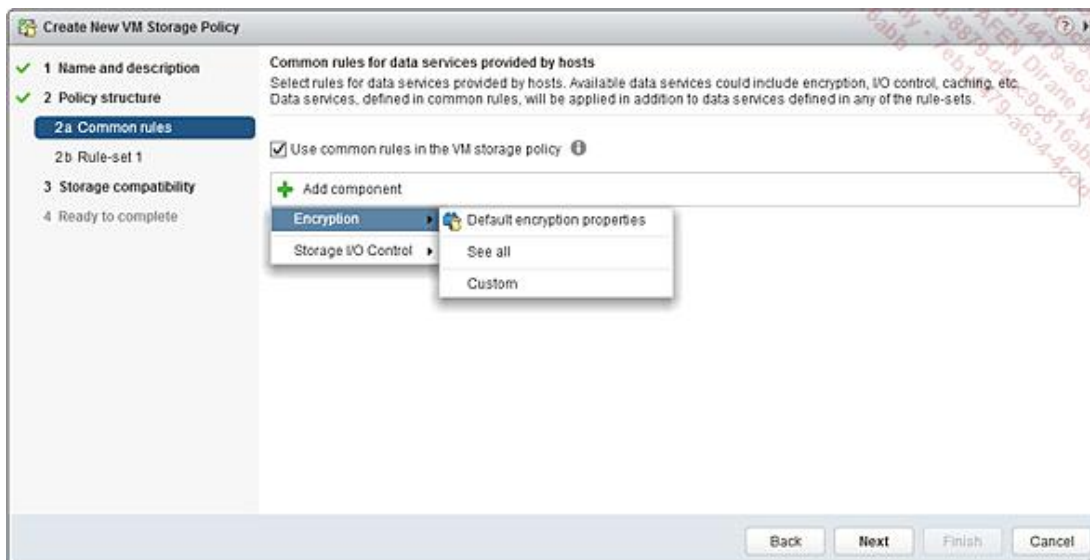
Ne modifiez pas la structure de la stratégie de stockage.



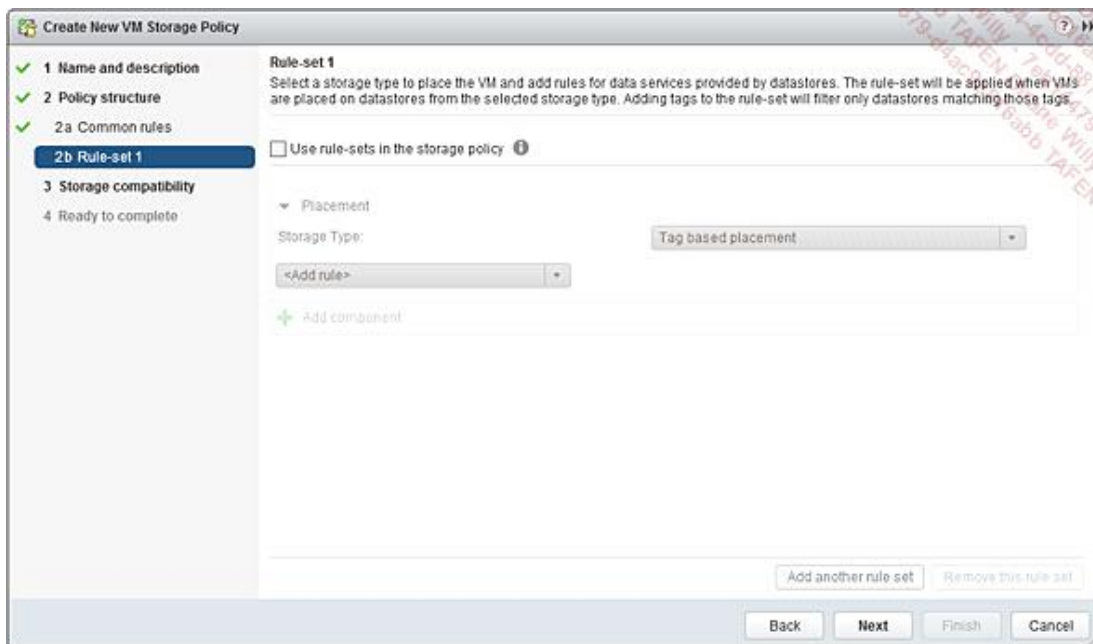
Explorez **Storage I/O Control** afin d'avoir accès aux différentes configurations de SIOC.



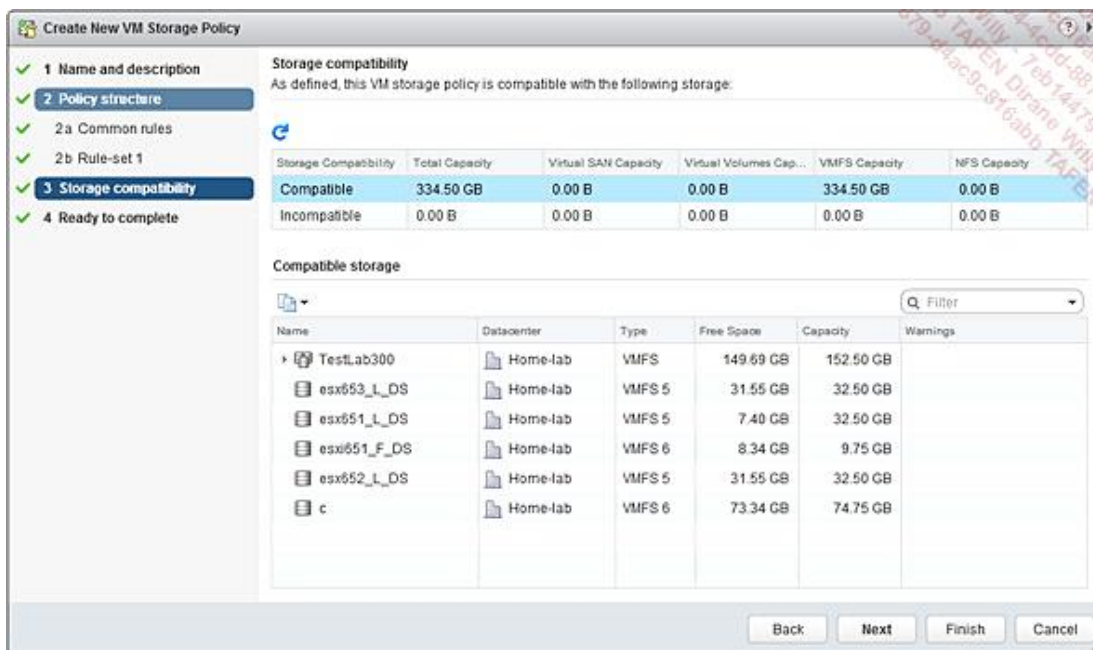
Sur le même écran, vous avez aussi la stratégie de chiffrement des disques des machines virtuelles.



Cliquez sur **Suivant**.



Sélectionnez le stockage compatible.



Cliquez sur **Finir** le résumé nous apparaît.

Create New VM Storage Policy

✓ 1 Name and description
✓ 2 Policy structure
✓ 2a Common rules
✓ 2b Rule-set 1
✓ 3 Storage compatibility
✓ 4 Ready to complete

Ready to complete
Confirm that the information below is correct and click Finish to create the VM storage policy.

General

Name	TestLab
Description	
vCenter Server	vcsa.testlab

Common rules

▼ Storage I/O Control > Low I/O shares allocation

Name:	Low I/O shares allocation
Description:	Storage policy component for Low SIOC controls
Provider:	VMware Storage I/O Control
VMware Storage I/O Control	
IOPS limit	1,000
IOPS reservation	10
IOPS shares	500

Back Next Finish Cancel