

# Les migrations

Il existe plusieurs types de migrations :

- La migration à froid (VM éteinte ou dans un état suspendu).
- Les migrations à chaud ou migration avec vMotion. Le vMotion est une fonctionnalité introduite par le vCenter, permettant de déplacer l'emprunte mémoire d'un hyperviseur à l'autre. Il existe plusieurs types de vMotion :
  - Les migrations inter vSwitch
  - Les migrations inter vCenter
  - Les migrations longues distances

Il existe quelques prérequis pour toutes ces migrations à chaud :

- Il faut au minimum 250 Mbps de bande passante dédiée par migration ce qui implique une limitation du nombre de migrations simultanées (4 pour un réseau gigabit, et 8 pour un réseau 10 gigabits). Attention, dans le cas où l'on utilise plusieurs interfaces réseaux pour le vMotion, nous pouvons dépasser cette limitation (<https://kb.vmware.com/kb/2007467>).
- Dans le cas des migrations longues distances, il est nécessaire d'avoir une latence inférieure à 150 ms sur l'aller-retour.

## 1. Migration à froid

Une migration à froid s'effectue lorsque la VM est soit éteinte, soit dans un état suspendu. Il est possible de migrer la VM entre hyperviseurs, datacenters et serveurs vCenter. Ce type de migration permet aussi de déplacer les fichiers constituant la VM entre datastores.

Le processus de migration à froid est le suivant :

- Validation d'un hyperviseur compatible.
- Sélection du choix du datastore cible, et copie des fichiers sur le datastore de l'hyperviseur cible. (Optionnel)
- Enregistrement de la VM sur le nouvel hyperviseur.
- Suppression de fichiers du datastore source s'il a été décidé de déplacer les fichiers. (Optionnel)

## 2. Migration à chaud

Il existe plusieurs types de migration à chaud.

- La migration à chaud d'un hyperviseur à l'autre est appelée vMotion.
- La migration à chaud des fichiers de machines virtuelles d'un datastore à l'autre est appelée storage vMotion.

Ces deux types de migration sont communément appelés simplement vMotion. Il est à noter que l'emplacement du fichier de Swap de la VM impacte le vMotion. Si le fichier de Swap est stocké sur le stockage local de l'hyperviseur source, il sera copié, ce qui peut ralentir le processus.

### a. vMotion

Le vMotion est une fonctionnalité qui autorise une VM à être déplacée d'un hyperviseur à l'autre. Ici nous ne parlons pas des fichiers constituant la VM, mais des processus (worlds, ou processus de machines virtuelles vu de l'hyperviseur) qui sont gérés par le VMKernel. La copie des informations de VM est effectuée via les interfaces vMkernel des hyperviseurs.

Le processus de vMotion est le suivant :

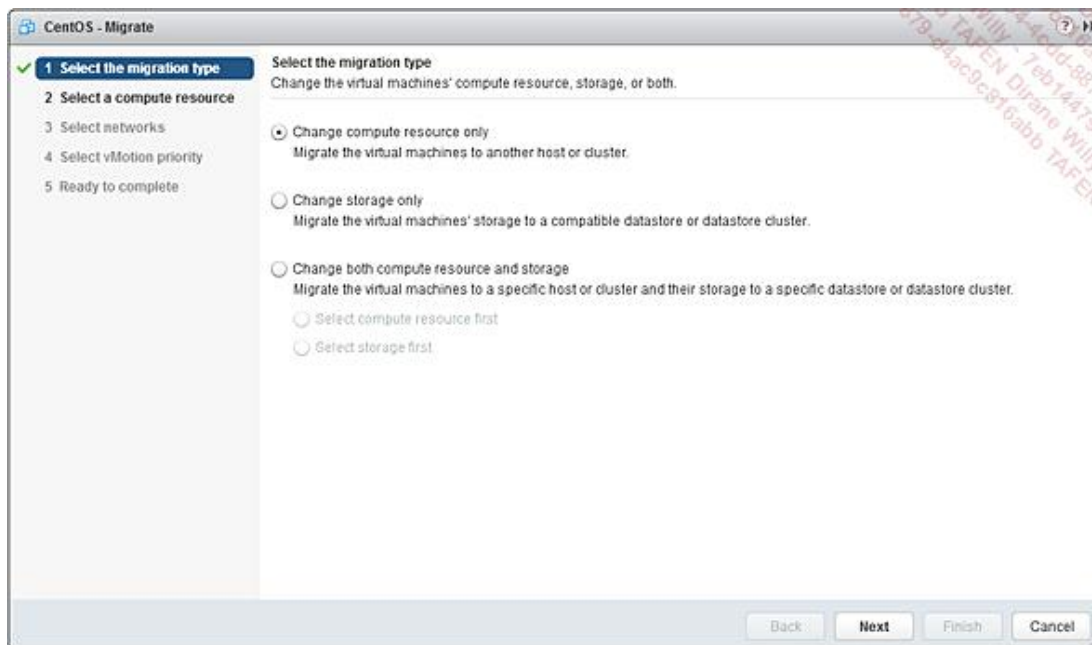
- Validation d'un hyperviseur compatible.
- Création d'une copie de la VM sur l'hyperviseur cible.
- Copie de chaque page mémoire depuis l'hyperviseur source vers l'hyperviseur cible, via le réseau vMotion. Cette étape est nommée précopie et elle s'exécute autant de fois que nécessaire tant qu'il y a des changements dans les pages mémoires.
- La VM sur l'hyperviseur source est figée et est relancée sur l'hyperviseur cible.

vMotion entre deux ESXi :

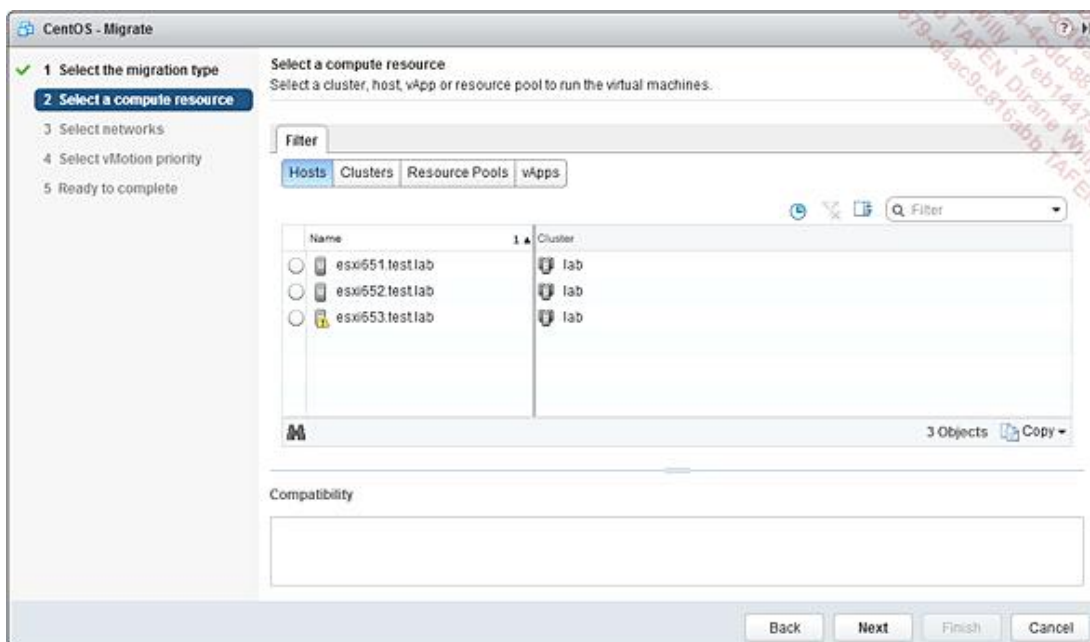
```
Get-VM [VM] | Move-VM -Destination (Get-VMHost [ESXHost])
```

Où [VM] représente le nom de la machine virtuelle et [ESXHost] le nom du serveur ESXi.

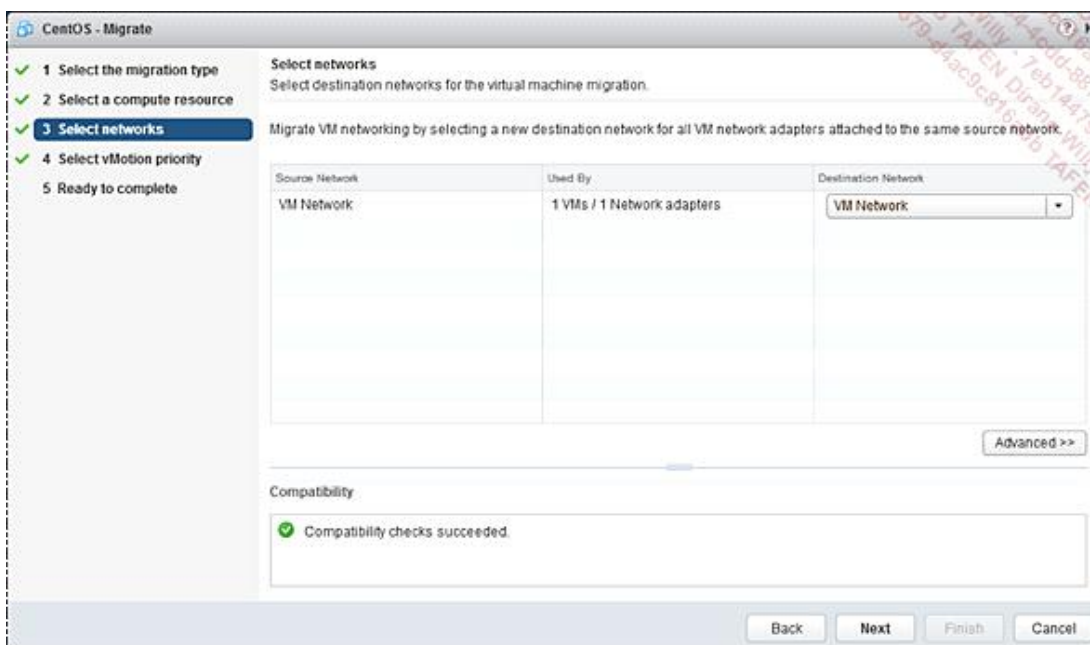
On sélectionne la VM que l'on souhaite migrer, puis le type de migration que l'on désire effectuer.



Sélectionnez la cible (ici cela sera un autre ESXi).



Choisissez le réseau cible et validez.



### **Enhanced vMotion Compatibility - EVC**

De base, la fonctionnalité vMotion requiert des processeurs disposant des mêmes instructions donc de même fondeur et de même famille. De ce fait, si deux hyperviseurs ont des processeurs physiques séparés de plusieurs générations (moins de deux ans suffisent pour que certaines instructions soient mises à jour), il est très probable que les migrations vMotion échouent car il ne sera pas garanti qu'une machine virtuelle puisse continuer à utiliser les mêmes types et versions d'instructions en passant d'un hyperviseur à l'autre. La fonctionnalité de compatibilité améliorée vMotion (EVC) permet d'« atténuer les différences entre les CPU physiques ».

L'EVC autorise des migrations vMotion entre CPU du même fondeur mais de génération différente. L'EVC permet de cacher des instructions CPU (CPU masking) sur les processeurs les plus récents. La prise en compte de la modification de la configuration EVC, qui est un paramètre de configuration du cluster, nécessite que les ESXi

soient en mode maintenance.

Génération de CPU Intel :

EVC Level	EVC Baseline	Description
L0	Intel® "Merom" Gen. (Intel® Xeon® Core™ 2)	Applies baseline feature set of Intel® "Merom" Generation (Intel® Xeon® Core™ 2) processors to all hosts in the cluster.
L1	Intel® "Penryn" Gen. (formerly Intel® Xeon® 45nm Core™ 2)	Applies baseline feature set of Intel® "Penryn" Generation (Intel® Xeon® 45nm Core™ 2) processors to all hosts in the cluster. Compared to the Intel® "Merom" Generation EVC mode, this EVC mode exposes additional CPU features including SSE4.1.
L2	Intel® "Nehalem" Gen. (formerly Intel® Xeon® Core™ i7)	Applies baseline feature set of Intel® "Nehalem" Generation (Intel® Xeon® Core™ i7) processors to all hosts in the cluster. Compared to the Intel® "Penryn" Generation EVC mode, this EVC mode exposes additional CPU features including SSE4.2 and POPCOUNT.
L3	Intel® "Westmere" Gen. (formerly Intel® Xeon® 32nm Core™ i7)	Applies baseline feature set of Intel® "Westmere" Generation (Intel® Xeon® 32nm Core™ i7) processors to all hosts in the cluster. Compared to the Intel® "Nehalem" Generation mode, this EVC mode exposes additional CPU features including AES and PCLMULQDQ.  <b>Note:</b> Intel® i3/i5 Xeon® Clarkdale Series processors that do not support AESNI and PCLMULQDQ cannot be admitted to EVC modes higher than the Intel® "Nehalem" Generation mode.  <b>Note:</b> Intel® Atom™ C2300-C2700 processors support the Intel® "Westmere" Gen. EVC baseline although their architecture is different from the architecture of the Intel® "Westmere" Generation processors.
L4	Intel® "Sandy Bridge" Generation	Applies baseline feature set of Intel® "Sandy Bridge" Generation processors to all hosts in the cluster. Compared to the Intel® "Westmere" Generation mode, this EVC mode exposes additional CPU features including AVX and XSAVE.  <b>Note:</b> Intel® "Sandy Bridge" processors that do not support AESNI and PCLMULQDQ cannot be admitted to EVC modes higher than the Intel® "Nehalem" Generation mode.
L5	Intel® "Ivy Bridge" Generation	Applies baseline feature set of Intel® "Ivy Bridge" Generation processors to all hosts in the cluster. Compared to the Intel® "Sandy Bridge" Generation EVC mode, this EVC mode exposes additional CPU features including RDRAND, ENFSTRG, FSGSBASE, SMEP, and F16C.  <b>Note:</b> Some Intel® "Ivy Bridge" processors do not provide the full "Ivy Bridge" feature set. Such processors cannot be admitted to EVC modes higher than the Intel® "Nehalem" Generation mode.
L6	Intel® "Haswell" Generation	Applies baseline feature set of Intel® "Haswell" Generation processors to all hosts in the cluster. Compared to the Intel® "Ivy Bridge" Generation EVC mode, this EVC mode exposes additional CPU features including ABMX2, MOVBE, FMA, PERMD, RORX/MULX, INVPCID, VMFUNC.

Génération de CPU AMD :

EVC Level	EVC Baseline	Description
A0	AMD Opteron™ Generation 1	Applies baseline feature set of AMD Opteron™ Generation 1 (Rev. E) processors to all hosts in the cluster.
A1	AMD Opteron™ Generation 2	Applies baseline feature set of AMD Opteron™ Generation 2 (Rev. F) processors to all hosts in the cluster. Compared to the AMD Opteron™ Generation 1 EVC mode, this EVC mode exposes additional CPU features including CPMXCHG16B and RDTSCP.
A3	AMD Opteron™ Generation 3	Applies baseline feature set of AMD Opteron™ Generation 3 (Greyhound) processors to all hosts in the cluster. Compared to the AMD Opteron™ Generation 2 EVC mode, this EVC mode exposes additional CPU features including SSE4A, MisAlignSSE, POPCOUNT and ABM (LZCNT).  <b>Note:</b> Due to 3DNow!™ support being removed from AMD processors after mid 2010, use AMD Opteron™ Generation 3 (no 3DNow!™) when possible to avoid compatibility issues with future processor generations.
A2, B0	AMD Opteron™ Generation 3 (no 3DNow!™)	Applies baseline feature set of AMD Opteron™ Generation 3 (Greyhound) processors with 3DNow!™ support removed, to all hosts in the cluster. This mode allows you to prepare clusters containing AMD hosts to accept AMD processors without 3DNow!™ support.
B1	AMD Opteron™ Generation 4	Applies baseline feature set of AMD Opteron™ Generation 4 (Bulldozer) processors to all hosts in the cluster. Compared to the AMD Opteron™ Generation 3 (no 3DNow!™) EVC mode, this EVC mode exposes additional CPU features including SSSE3, SSE4.1, AES, AVX, XSAVE, XOP, and FMA4.
B2	AMD Opteron™ "Piledriver" Generation	Applies baseline feature set of AMD Opteron™ "Piledriver" Generation processors to all hosts in the cluster. Compared to the AMD Opteron™ Generation 4 EVC mode, this EVC mode exposes additional CPU features including FMA, TBM, BMI1, and F16C.

Tous les types de CPU ne sont pas supportés par l'ensemble des versions de vCenter :

Génération de CPU Intel :



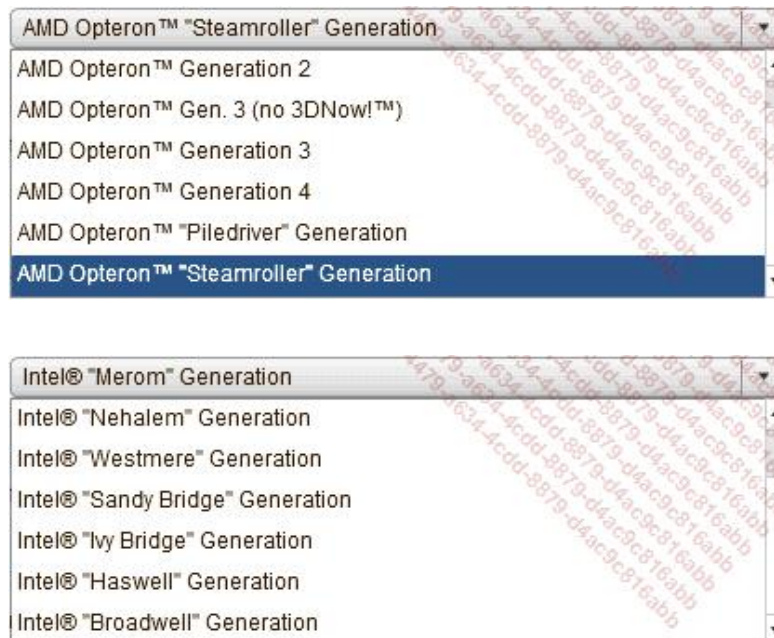
	EVC Cluster Baseline						
vCenter Server Release	Intel® "Merom" Generation	Intel® "Penryn" Generation	Intel® "Nehalem" Generation	Intel® "Westmere" Generation	Intel® "Sandy Bridge" Generation	Intel® "Ivy Bridge" Generation	Intel® "Haswell" Generation
VirtualCenter 2.5 U2 and later updates	Yes	No	No	No	No	No	No
vCenter Server 4.0	Yes	Yes	Yes	No	No	No	No
vCenter Server 4.0 U1 and later updates	Yes	Yes	Yes	Yes	No	No	No
vCenter Server 4.1	Yes	Yes	Yes	Yes	No	No	No
vCenter Server 5.0	Yes	Yes	Yes	Yes	Yes	No	No
vCenter Server 5.1	Yes	Yes	Yes	Yes	Yes	Yes	No
vCenter Server 5.5	Yes	Yes	Yes	Yes	Yes	Yes	No
vCenter Server 6.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Génération de CPU AMD :

	EVC Cluster Baseline					
vCenter Server Release	AMD Opteron™ Gen. 1	AMD Opteron™ Gen. 2	AMD Opteron™ Gen. 3	AMD Opteron™ Gen. 3 (no 3DNow!™)	AMD Opteron™ Gen. 4	AMD Opteron™ "Piledriver" Gen.
VirtualCenter 2.5 U2 and later updates	Yes	No	No	No	No	No
vCenter Server 4.0	Yes	Yes	Yes	No	No	No
vCenter Server 4.0 U1 and later updates	Yes	Yes	Yes	No	No	No
vCenter Server 4.1	Yes	Yes	Yes	Yes	No	No
vCenter Server 5.0	Yes	Yes	Yes	Yes	Yes	No
vCenter Server 5.1	Yes	Yes	Yes	Yes	Yes	Yes
vCenter Server 5.5	Yes	Yes	Yes	Yes	Yes	Yes

La KB1003212 (<https://kb.vmware.com/kb/1003212>) ne prend pas en compte les processeurs postérieurs à la génération des Haswell pour Intel et des Piledriver pour les AMD. Cependant, vSphere 6.5, lui, montre la génération Intel Broadwell et Stormroller pour AMD. Il s'agit d'un problème de mise à jour de la KB VMware au moment de l'écriture de ce livre.

Extrait des générations de CPU par fondeur depuis l'interface de configuration de l'EVC :



Activer EVC pour un cluster nécessite que tous les ESXi du cluster aient des processeurs du même fondeur. Il faut

donc choisir entre Intel et AMD.

## b. Storage vMotion

La migration vMotion permet de déplacer les processus de machines virtuelles tandis que les migrations Storage vMotion concernent les fichiers constitutifs des VM. La migration se fait de datastore(s) vers datastore(s).

Le SvMotion est un processus se déroulant en six étapes :

- L'agent VPXa copie le répertoire de la machine virtuelle depuis le datastore source vers le datastore cible.
- La machine virtuelle dite « shadow VM » est démarrée, mais reste en attente de la fin de copie des VMDKs.
- Le SvMotion demande au Mirror Driver de dupliquer chaque écriture pour chaque bloc déjà copié.
- En une seule passe la copie des VMDK est faite et complétée au niveau du datastore cible tout en continuant la duplication de chaque I/O.
- Le SvMotion fait appel aux fonctionnalités Fast Suspend et Resume de la machine virtuelle afin de transférer l'état actif de la machine virtuelle source vers la machine virtuelle dite shadow.
- Après que le Fast Suspend et Resume ont été complétés, le répertoire et les fichiers (VMX, VMDK) présents dans le datastore source sont supprimés.

Le Storage vMotion (<http://cgi.di.uoa.gr/~ad/M155/Papers/LiveMigrationESX.pdf>, [https://www.usenix.org/legacy/events/atc11/tech/final\\_files/Mashtizadeh.pdf](https://www.usenix.org/legacy/events/atc11/tech/final_files/Mashtizadeh.pdf) ou The Design and Evolution of Live Storage Migration in VMware ESX.) est une pratique existante depuis la version 3.0 des hyperviseurs connus à l'époque sous le nom de Virtual Infrastructure.

À ce moment a été introduit le snapshot des VMs, dans le cadre des « live Storage Migration », la technique de storage vMotion consistait à :

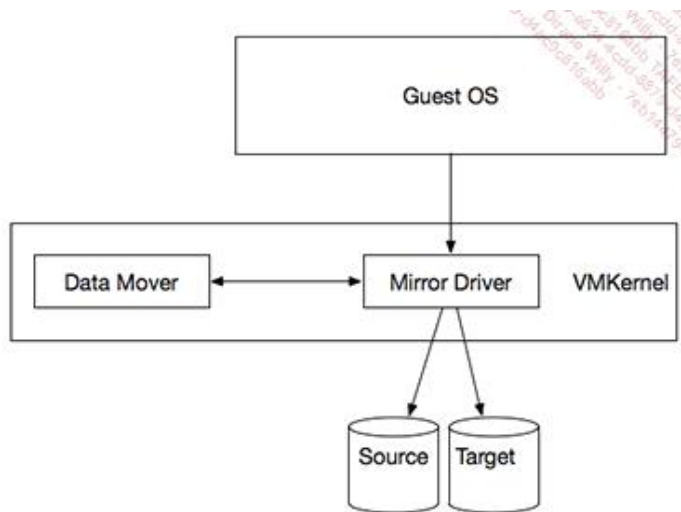
- Créer un snapshot afin de garder les I/O d'écritures et les modifications.
- Copier le VMDK de base sur le datastore cible.
- Créer un nouveau snapshot. Le premier snapshot est consolidé dans le disque de base.
- Ces opérations de création et consolidation se répètent et la taille du snapshot diminue de plus en plus.
- Lorsque la taille du snapshot atteint une taille minimale critique, la machine virtuelle passe en état suspendu et une dernière consolidation a lieu avant que la machine virtuelle soit de nouveau active.

La version 4.x d'ESX inclut le CBT (*Change Block Tracking*) ou Dirty Block Tracking. Tout comme la méthode du snapshot, le CBT est une méthode itérative, la différence avec le snapshot se trouve dans le fait que le CBT (<https://kb.vmware.com/kb/1020128>) travaille au niveau du bloc. (<https://blogs.vmware.com/vsphere/2011/07/new-vsphere-50-storage-features-part-2-storage-vmotion.html>)

Avec vSphere 5.1, la technique de l'I/O Mirror a été introduite. L'I/O Mirror étant un filtre, la machine virtuelle a besoin de passer rapidement dans mode dit figé ou « stun », afin d'activer l'I/O Mirror, une fois cela fait, la machine virtuelle retourne dans son état d'origine.

Pendant ce temps, le data mover fait la copie du disque de base tandis que l'I/O Mirror traque les accès en écritures en fonction des critères suivants :

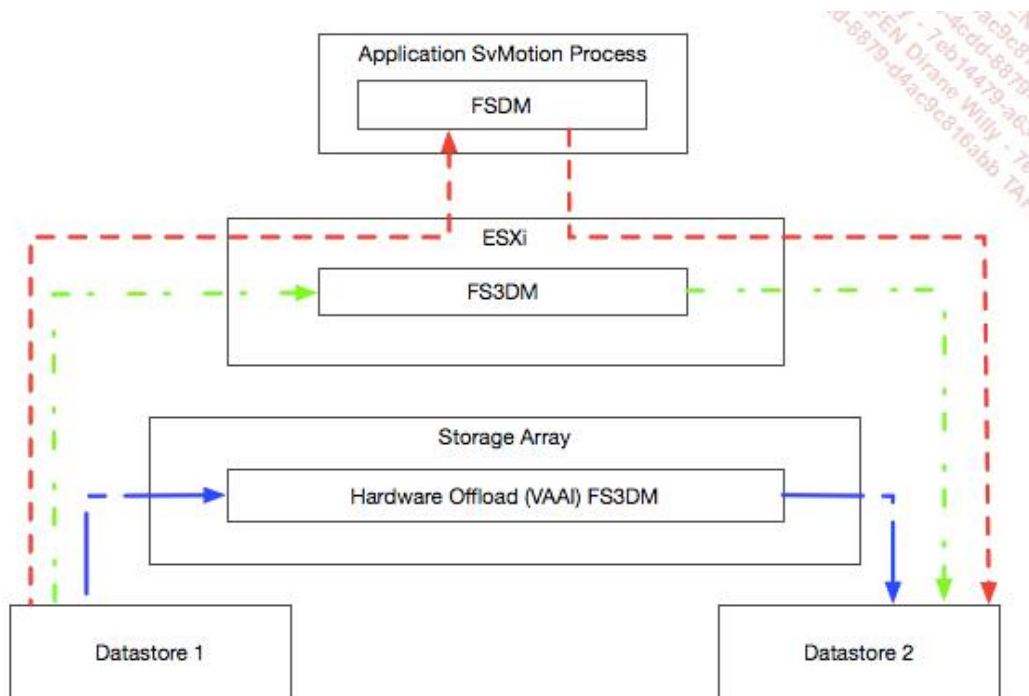
- Écriture sur une zone déjà copiée, et à mettre à jour sur la destination.
- Écriture sur une zone qui est en train d'être copiée, et à mettre à jour sur la destination.
- Écriture sur une zone qui n'est pas encore copiée.



Le technique CBT et I/O Mirror s'appuie sur le datamover.

Il existe trois types de datamovers :

- Le « legacy » datamover FSDM qui est la version la plus basique (et lente) car elle traverse l'ensemble des couches de la baie à la couche I/O de la machine virtuelle. Il s'appuie sur un tampon afin de faire la copie des fichiers. Il est utilisé dans le cas où les tailles des blocs au niveau des volumes VMFS ne sont pas identiques (VMFS3/5).
- Le datamover FS3DM qui a été introduit avec les versions 4.0, sera utilisé par défaut si la taille des blocs sur les volumes VMFS est identique et si les volumes sont sur la même baie.
- Le datamover FS3DM -hardware offload sera utilisé si le plug-in VAAI est présent. L'ESXi demandant à la baie d'effectuer les opérations.



Comme avec les versions précédentes, il est possible de faire jusqu'à 8 SvMotion simultanés par datastore et 2

par ESXi (<https://www.vmware.com/pdf/vsphere6/r65/vsphere-65-configuration-maximums.pdf>).

Object	5.5	6.0	6.5
Storage vMotion operations per host	2	2	2
Storage vMotion operations per datastore	8	8	8

Un storage vMotion en ligne de commande sans paramètre particulier est assez rapide à faire.

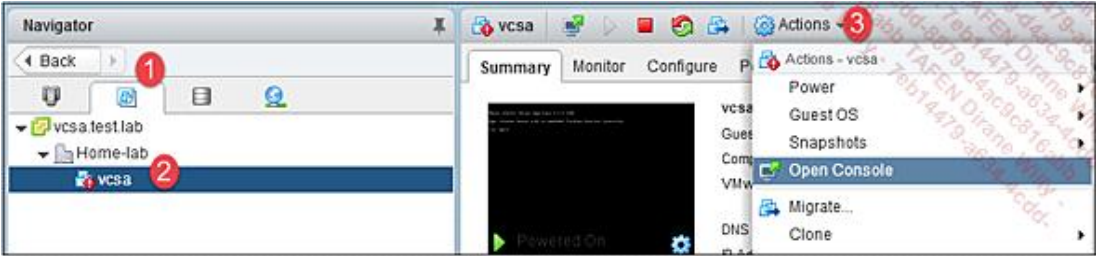
```
Get-VM [VM] | Move-VM -Datastore [DataStore]
```

[VM] et [DataStore] représentent respectivement une machine virtuelle et le datastore cible.

L'ensemble des paramètres que l'on utilise à partir du vCenter sont disponibles en powershell, avec la commande help move-vm.

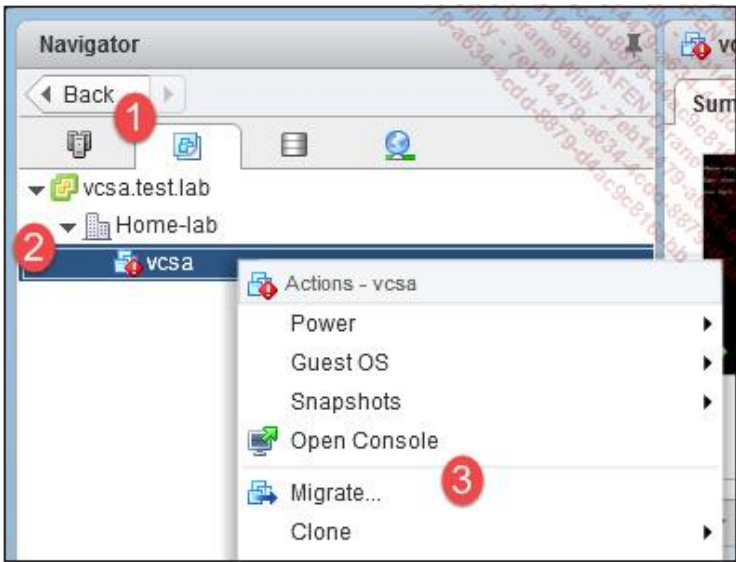
Plus d'informations en suivant ce lien : <https://www.vmware.com/support/developer/PowerCLI/PowerCLI65R1/html/index.html>

Ici, une opération de storage VMotion via le Web Client :

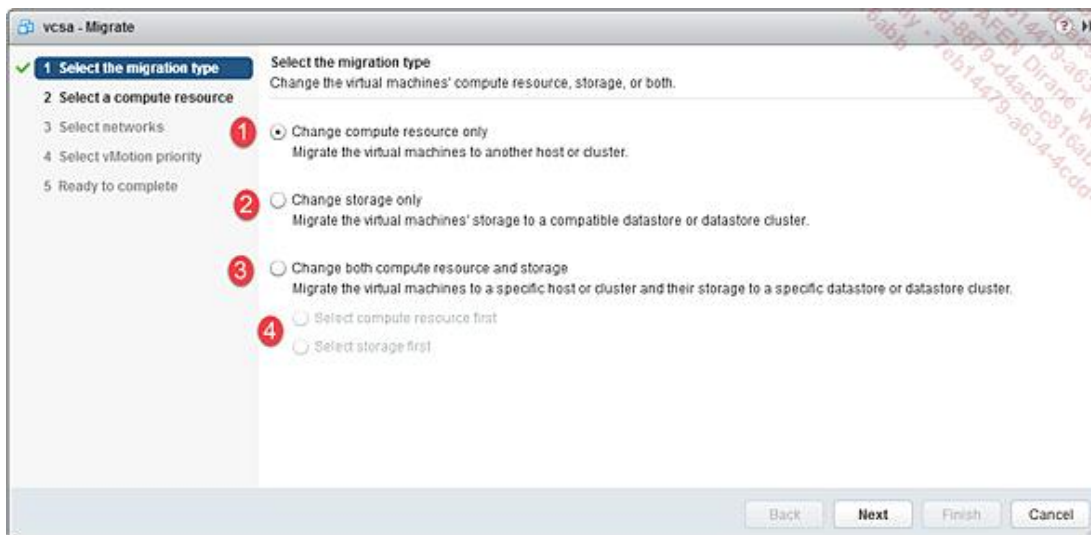


Sélectionnez la vue **Virtual machines et templates** (1).

Sélectionnez la machine virtuelle à migrer (2) et éventuellement ouvrez la console (3) pour vérifier son état (une machine virtuelle figée ne pourra pas migrer) :



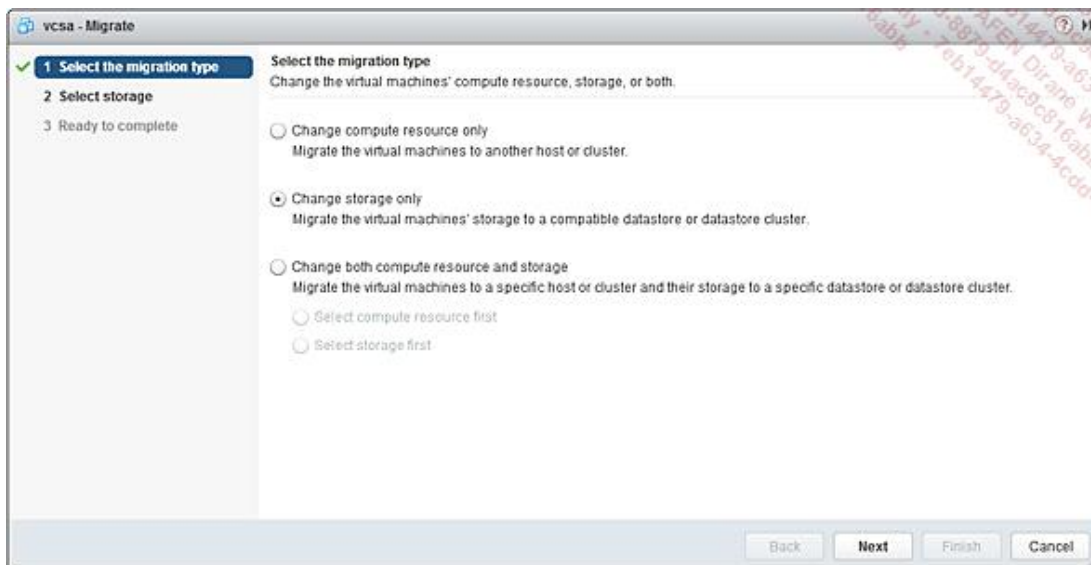




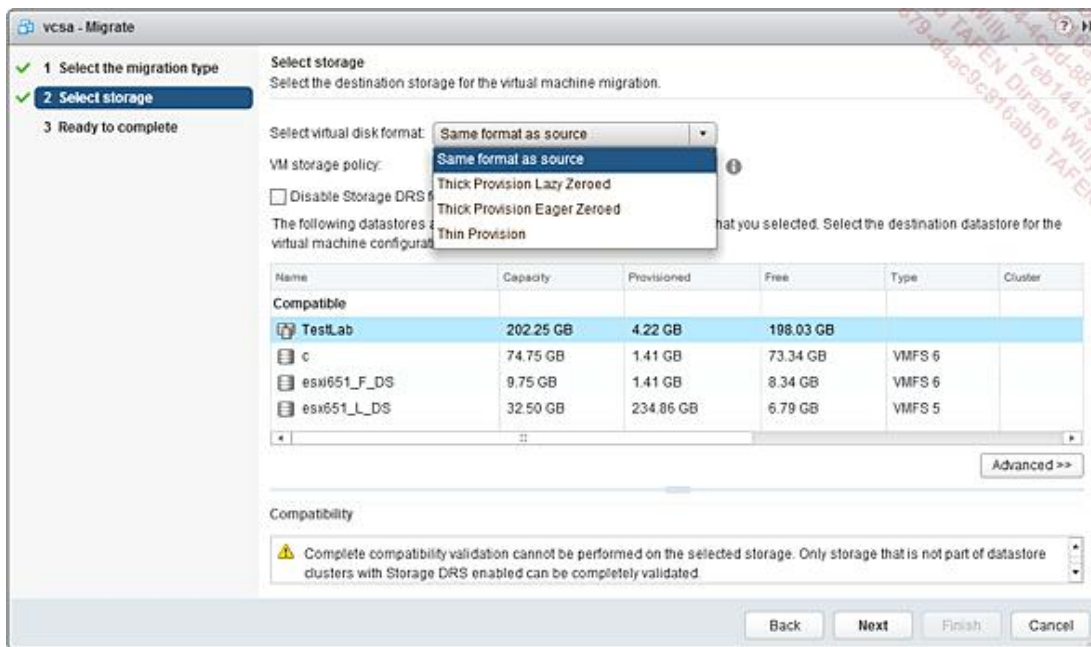
1 - vMotion

2 - Storage vMotion

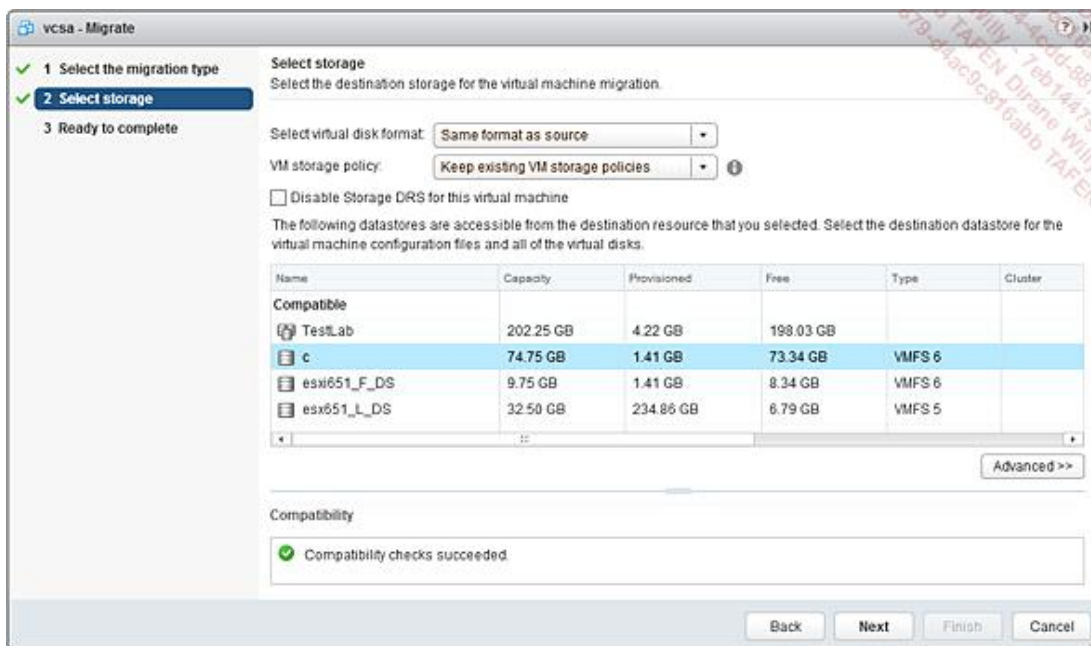
3 - vMotion shared-nothing. Le choix indiqué par la pastille 4 permet de sélectionner l'élément à migrer en premier (car vMotion shared-nothing permet de migrer compute et storage lors de la même opération)



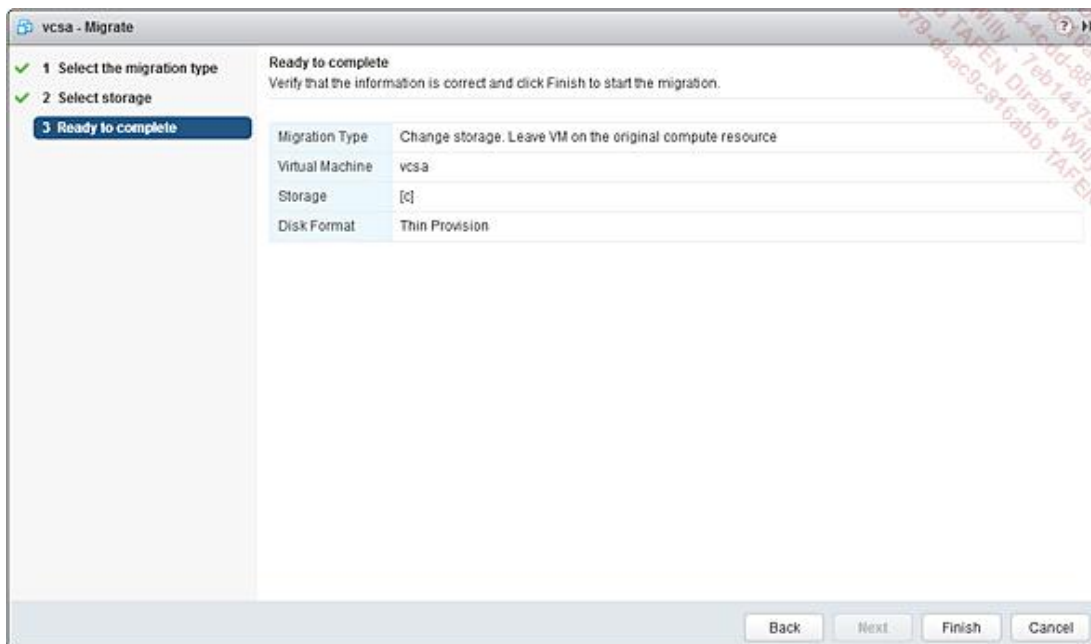
Choisissez donc l'option **Change Storage Only** pour effectuer un storage vMotion.



Conservez le format de stockage (il est possible de passer de thin à thickprovisionning et inversement).



Il est possible de modifier les stratégies de stockage associées à la machine virtuelle et de vérifier la conformité le cas échéant.

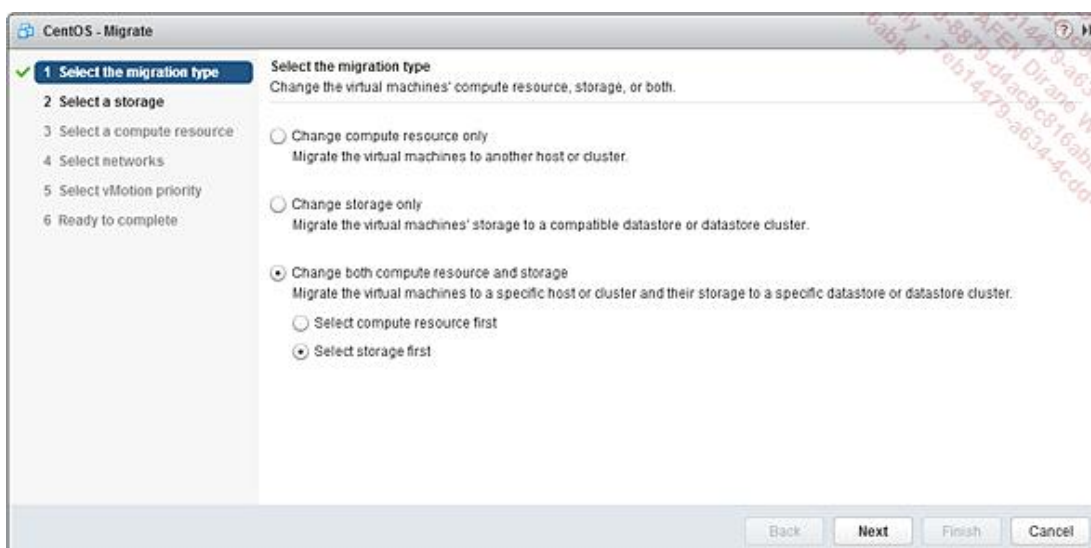


Avant la migration effective, l'interface propose un récapitulatif des paramètres choisis. En cliquant sur **Finish**, lancez la migration.

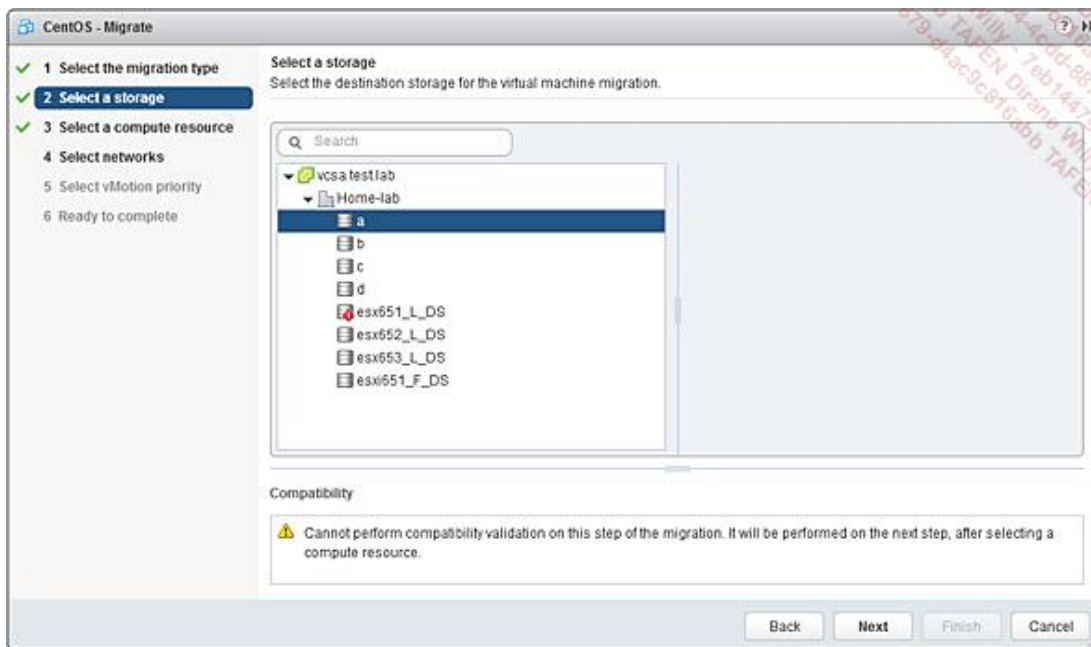
Le Storage vMotion permet aussi de placer les disques VMDK sur plusieurs datastores sous réserve que l'option Garder les VMDK ensemble (option d'affinités des VMDK) ne soit pas sélectionnée au niveau des paramètres de la machine virtuelle. Cette option est sélectionnée par défaut.

### c. Shared nothing vMotion

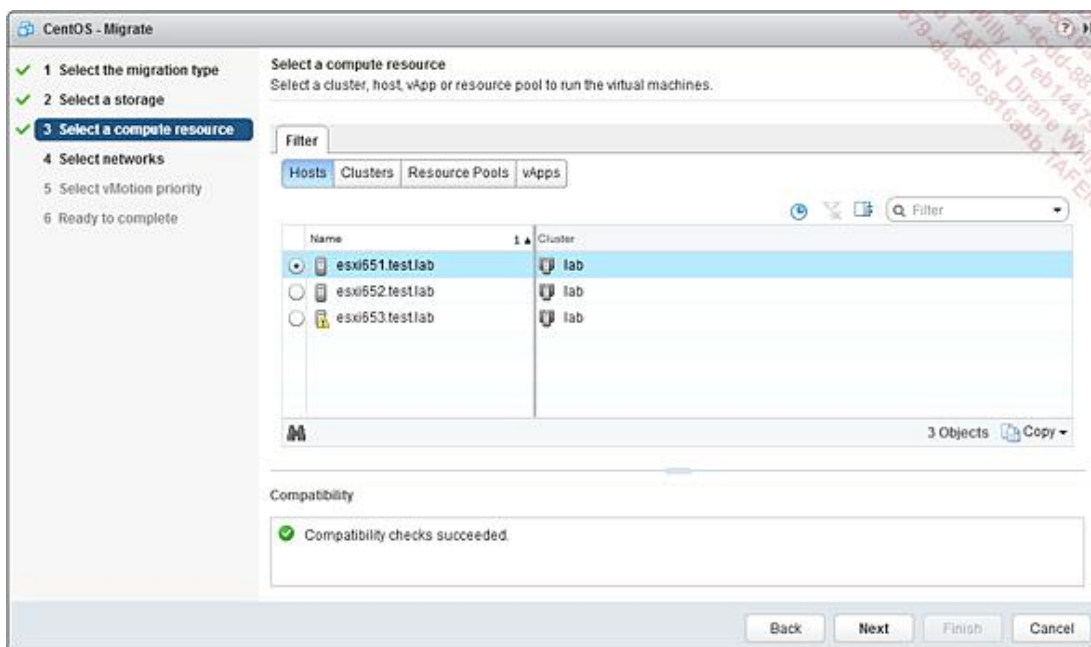
Le Shared Nothing vMotion est simplement le fait d'utiliser l'ensemble des autres types de vMotion en une seule et même opération. Dans l'exemple ci-dessous, nous choisissons de migrer le stockage avant les processus de VM.



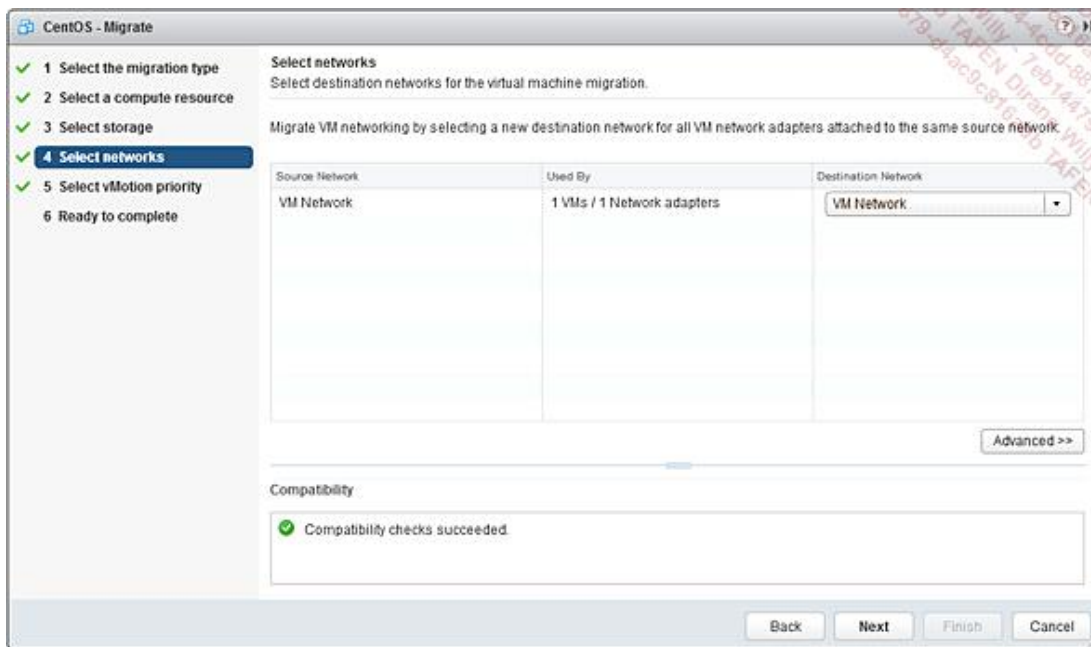
Sélectionnez le datastore cible.



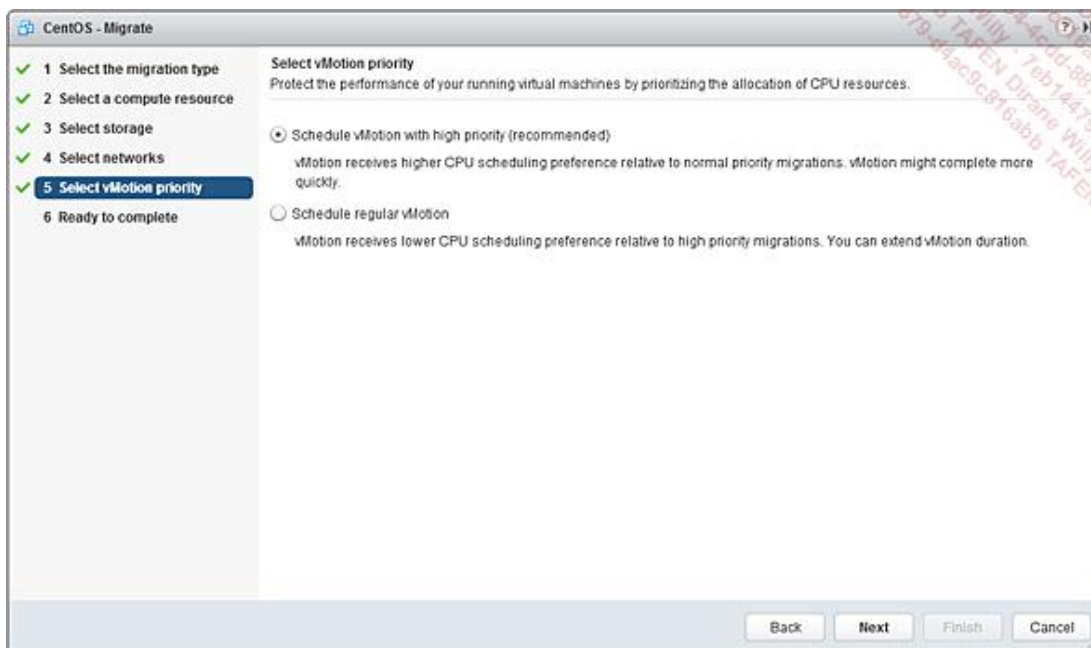
Vous pouvez utiliser les filtres automatiques afin de sélectionner un cluster, un pool de ressources ou bien une vApp.



Sélectionnez le réseau de destination.



Choisissez la priorité de l'opération du vMotion.



Puis validez avec **Finish**.



CentOS - Migrate

✓ 1 Select the migration type

✓ 2 Select a compute resource

✓ 3 Select storage

✓ 4 Select networks

✓ 5 Select vMotion priority

6 Ready to complete

Ready to complete

Verify that the information is correct and click Finish to start the migration.

Migration Type	Change compute resource and storage
Virtual Machine	CentOS
Cluster	lab
Host	esxi653.test.lab
Networks	No network reassignments
vMotion Priority	High
Storage	[a]
Disk Format	Same format as source

Back

Next

Finish

Cancel