

Le cluster DRS

DRS s'appuie sur le vCenter Agent (VPXa) qui autorise une communication dans les deux sens entre le vCenter et les serveurs ESXi. Le VPXa est présent sur chaque serveur de l'inventaire vCenter. Le serveur vCenter lance l'exécution DRS toutes les cinq minutes sur les serveurs ESXi du cluster.

DRS calcule dynamiquement les besoins et droits d'accès aux ressources (*dynamic entitlement*) de chaque VM en se basant sur les statistiques d'allocation de ressources et des métriques dynamiques. Dans le cas de fragmentation des ressources, DRS peut utiliser vMotion pour défragmenter les ressources. Cela crée une souplesse dans la gestion des ressources dans le placement initial et les déplacements des VMS sur les serveurs ESXi.

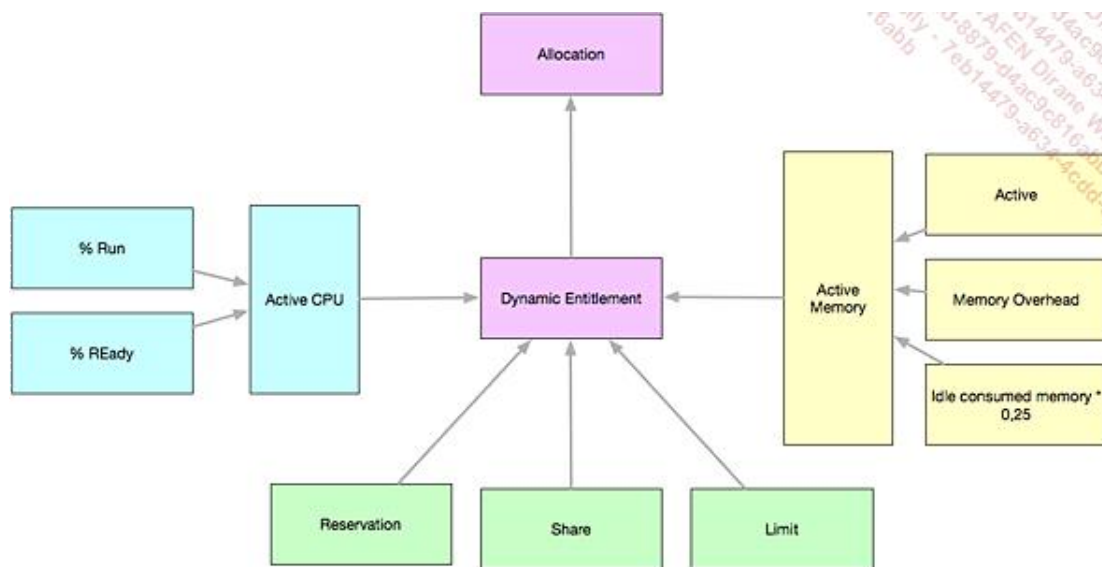
Au niveau du cluster, DRS est responsable des actions suivantes :

- Le placement initial : les recommandations du placement initial sont faites en respectant les contraintes d'utilisation CPU, mémoire et réseau.
- L'équilibrage de charge : est la répartition de la charge de manière équitable sur l'ensemble des serveurs du cluster. DRS essaie via ses algorithmes de faire cela.
- Gestion de la consommation électrique : lorsque la fonction DPM (*Distributed Power Management*) est utilisée, DRS peut recommander ou mettre en veille des serveurs ESXi, lorsqu'il y a un surplus de capacité, et remettre les hyperviseurs en services en cas de forte demande de ressources.
- Le mode maintenance du cluster : est une évaluation du cluster afin de déterminer quels serveurs ESXi peuvent être mis en maintenance en même temps afin de diminuer le temps de mise à jour via VUM (*VMware Update Manager*).
- Les contraintes de corrections : règles d'affinités VM/VM ou VM/Hyperviseur.

Cet ensemble de fonctionnalités nous offre une souplesse extrême dans la répartition des VMs sur l'ensemble du cluster ESXi.

1. Dynamic Entitlement et planificateur local

Derrière ces notions se cache le concept de la somme idéale des ressources utilisables. DRS et le planificateur local calculent cette somme de ressource afin de la mettre à disposition ou non des VMs ou des pools de ressources. Cet idéal est la résultante des données statiques et des données dynamiques. Les données statiques se basent sur les ressources allouées définies par les administrateurs, tandis que les données dynamiques se basent sur la demande et la congestion des ressources. Cette allocation dynamique ou Dynamic Entitlement se fait aussi bien sur les ressources CPU que mémoire. Les limites, shares et autres réservations influent dessus.



Le DRS s'appuie sur un composant propre à chaque hyperviseur le planificateur local ou Host Local Scheduler. Il est géré par le VMkernel. DRS travaille au niveau global via le vCenter, et est responsable pour l'allocation des ressources, tandis que le Host Local Scheduler travaillant au niveau de l'hyperviseur reproduit l'implémentation des pools de ressources et l'allocation des ressources au niveau des VMs. Cela nécessite une copie parfaite de l'arborescence du pool de ressources sur chaque Hyperviseur. Cette copie se trouve dans /host/user du planificateur local.

2. Les nouveautés de DRS

Lors du placement DRS prend en compte les réservations de capacité réseau, tout comme dans le déplacement des machines virtuelles dans la répartition de charge.

a. Predictive DRS

Avec vSphere 6.5, VMware nous a offert une nouvelle option pour le DRS : le predictive DRS.

L'idée du predictive DRS est de s'appuyer sur vRealize Operations Manager (vROps), qui est l'outil VMware pour la surveillance de l'infrastructure. vROps pour ses besoins propres fait de la collecte de métriques, cela nous permet de manière plus souple d'anticiper les risques et impacts liés à la charge des machines virtuelles et aux défaillances des ESXi. Cette anticipation est faite par le biais de l'analyse du comportement des objets constituant l'infrastructure, afin d'établir des seuils dynamiques d'activités.

D'une manière simple nous pourrions dire que vROps apprend de l'utilisation de l'infrastructure pour déterminer l'état de santé, le comportement normal ainsi que les exceptions d'utilisation dans notre environnement VMware. Cela permet à vROps de ne pas considérer une utilisation exceptionnelle des ressources comme normale : par exemple si pour un projet spécifique dix ESXi sont saturés, vROps pourra considérer cela comme une exception et ne vous fera pas un capacity planning demandant l'ajout d'une centaine de serveurs pour soutenir l'évolution de l'infrastructure.

b. Les options additionnelles

Ils existent trois options additionnelles :

- VM Distribution qui permet de répartir les machines virtuelles en les distribuant sur le cluster ESXi. Il se base sur la moyenne de machines virtuelles par ESXi. Le paramètre DRS qui est utilisé est LimitVMsperESXHostPercent. Le

placement initial peut violer ce paramètre.

- Memory Metric for Load Balancing. Par défaut, DRS utilise la mémoire active des machines virtuelles pour faire ses calculs. Cette fonctionnalité offre le choix d'utiliser la mémoire active plus un pourcentage de la mémoire inactive à la place de la mémoire active. Le paramètre DRS associé est PercentIdleMBinMemDemand. Le pourcentage de la mémoire inactive est calculé de la manière suivante : $100 - \text{mem.idleTax} = 25\%$. Mem.idleTax est un paramètre avancé de configuration de l'ESXi. Par défaut la valeur de Mem.idleTax est de 75.
- CPU Over-commitment ce paramètre permet de spécifier le ratio vCPU, pCPU au niveau du cluster. Le maximum étant 500 %, ce qui équivaut à un ratio de 5vCPU :1pCPU. Ce paramétrage peut être modifié manuellement : <https://kb.vmware.com/kb/1020121> Sachant que depuis vSphere 6.0, un serveur ESXi peut supporter jusqu'à 4096 vCPU.

3. Les niveaux d'automatisation de DRS

Il existe trois niveaux d'automatisation :

- Manuel : DRS gère le placement initial sous forme de recommandation. Dans le cas d'un cluster non équilibré, nous devons manuellement valider chacune des recommandations faites par DRS.
- Partiellement automatisé : DRS gère le placement initial, sans intervention d'un administrateur, sur le serveur ESXi le mieux à même de recevoir la VM. DRS fait des recommandations sur les migrations des VMs. Nous devons manuellement valider ou non chacune des recommandations faites par DRS.
- Entièrement automatisé : DRS s'occupe du placement et des migrations afin d'équilibrer la charge au mieux entre les serveurs du cluster ESXi.

4. Configuration du cluster DRS

La création d'un cluster est simple en ligne de commande Powercli :

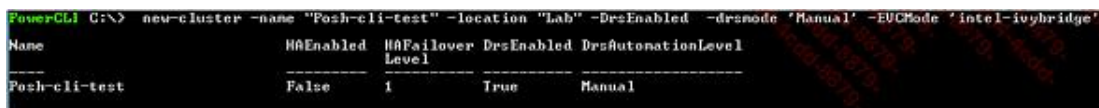
```
New-cluster -name '[Nom]' -location '[Datacenter]'  
-DrsEnabled -drsmode '[Option]' -EVCMode '[TypeCPU]'
```

[Nom] : nom du cluster que l'on crée.

[Datacenter] : nom du datacenter dans lequel on crée le cluster.

[Option] : option de migration des machines virtuelles.

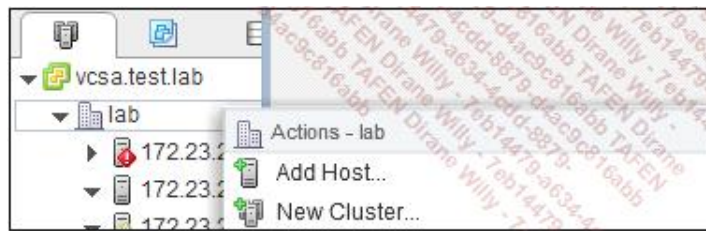
[TypeCPU] : type du CPU (Intel Ivy Bridge, AMD Streamroller...)



```
PowerCLI C:\> new-cluster -name "Pesh-cli-test" -location "Lab" -DrsEnabled -drsmode 'Manual' -EVCMode 'intel-ivybridge'
```

Name	HAEnabled	HAFailover Level	DrsEnabled	DrsAutomationLevel
Pesh-cli-test	False	1	True	Manual

En mode graphique, nous commençons par faire un clic droit sur le Datacenter où nous allons créer notre cluster vSphere (DRS et HA) via la commande New Cluster.



La fenêtre suivante permet de configurer le paramètre DRS du cluster. Nous donnons un nom au cluster, puis cochons pour le critère DRS Turn on. Connaissant notre matériel, nous en profitons pour configurer l'EVC. L'EVC sera expliqué dans la partie Migration. Attention, il est nécessaire de connaître le type de CPU que nous utilisons par sa génération (Ivy Bridge par exemple pour Intel, nous avons aussi les CPU AMD).

Name	Clu-Lab
Location	lab
DRS	<input checked="" type="checkbox"/> Turn ON
Automation Level	Fully automated
Migration Threshold	Conservative ——— Aggressive
vSphere HA	<input type="checkbox"/> Turn ON
EVC	
EVC Mode	Intel® "Broadwell" Generation
Description	<p>Applies the baseline feature set of Intel® "Broadwell" Generation processors to all hosts in the cluster.</p> <p>Hosts with the following processor types will be permitted to enter the cluster: Intel® "Broadwell" Generation Future Intel® processors</p> <p>Compared to the Intel® "Haswell" Generation EVC mode, this EVC mode exposes additional CPU features including Transactional Synchronization Extensions, Supervisor Mode Access Prevention, Multi-Precision Add-Carry Instruction Extensions, PREFETCHW and RDSEED.</p> <p>Note: Some "Broadwell" microarchitecture processors do not provide the full "Broadwell" feature set. Such processors do not support this EVC mode; they will only be admitted to the Intel® "Haswell" Generation mode or below.</p> <p>For more information, see Knowledge Base article 1003212.</p>
Virtual SAN	<input type="checkbox"/> Turn ON

OK Cancel

Notre cluster est paramétré au niveau DRS. Il ne reste qu'à faire glisser les ESXi que nous voulons voir travailler ensemble, ce qui implique que les machines virtuelles et les ressources pools présents sur ces ESXi sont également importés.

a. Les règles d'affinités entre machines virtuelles et ESXi

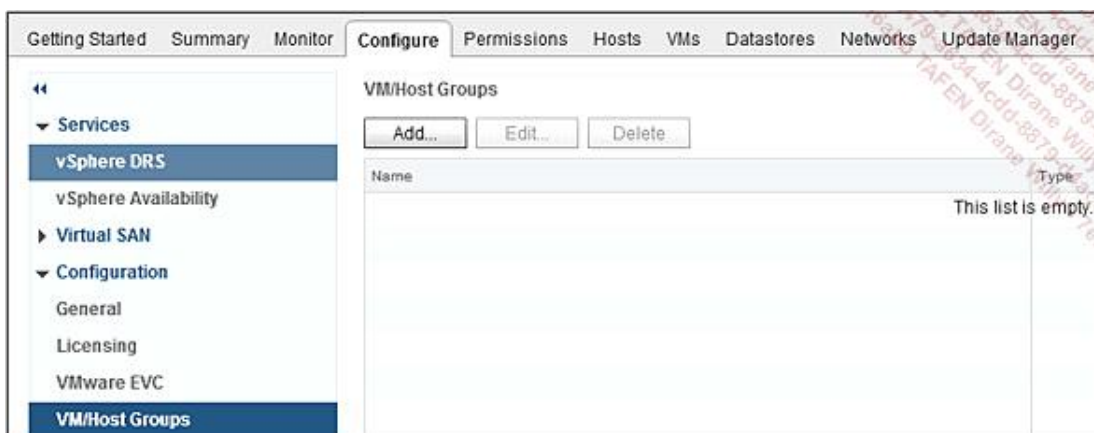
Les règles d'affinités permettent de créer des relations entre les machines virtuelles entre elles, et entre les machines virtuelles et les ESXi.

Cela peut être très utile lors que l'on virtualise des contrôleurs de domaines ou tout autre serveur d'infrastructure/application critique. Il serait dommage que l'ensemble des contrôleurs de domaine se retrouvent tous sur le même ESXi. En cas de perte de l'ESXi, nous perdrons aussi tout moyen d'authentification.

De même, dans le cas de cluster ESXi étendu (Stretched Cluster), il est intéressant de faire fonctionner les machines virtuelles d'un même site au plus près de leur stockage.

La première étape consiste à créer nos groupes d'ESXi et de machines virtuelles. La procédure étant la même pour ces deux entités, nous utiliserons la création d'un groupe d'ESXi.

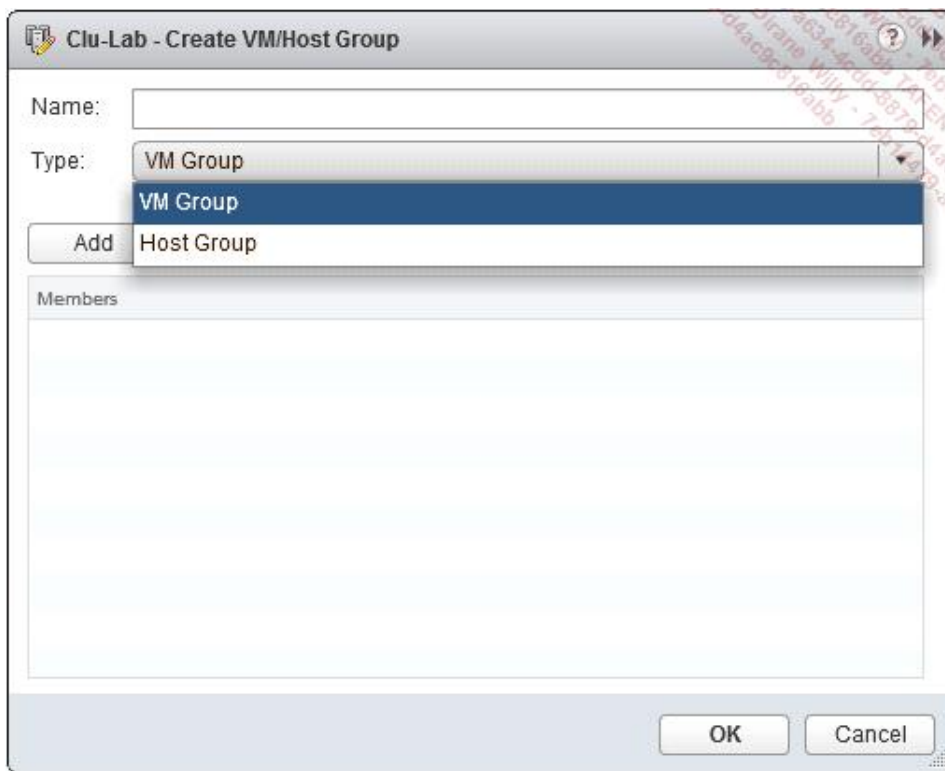
Lorsque le cluster est sélectionné, allez dans l'onglet **Configuration**, puis dans **VM/Host Groups**. En cliquant sur **Add** vous pouvez commencer la création d'un groupe.



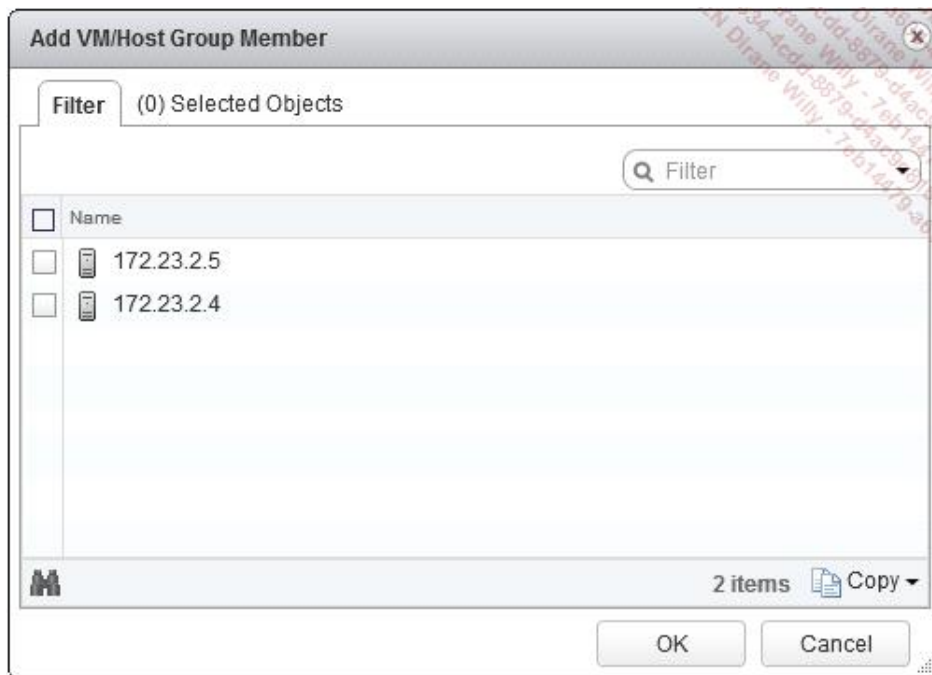
Sélectionnez le type de groupe à créer :

- Un groupe d'ESXi.
- Un groupe de machines virtuelles.

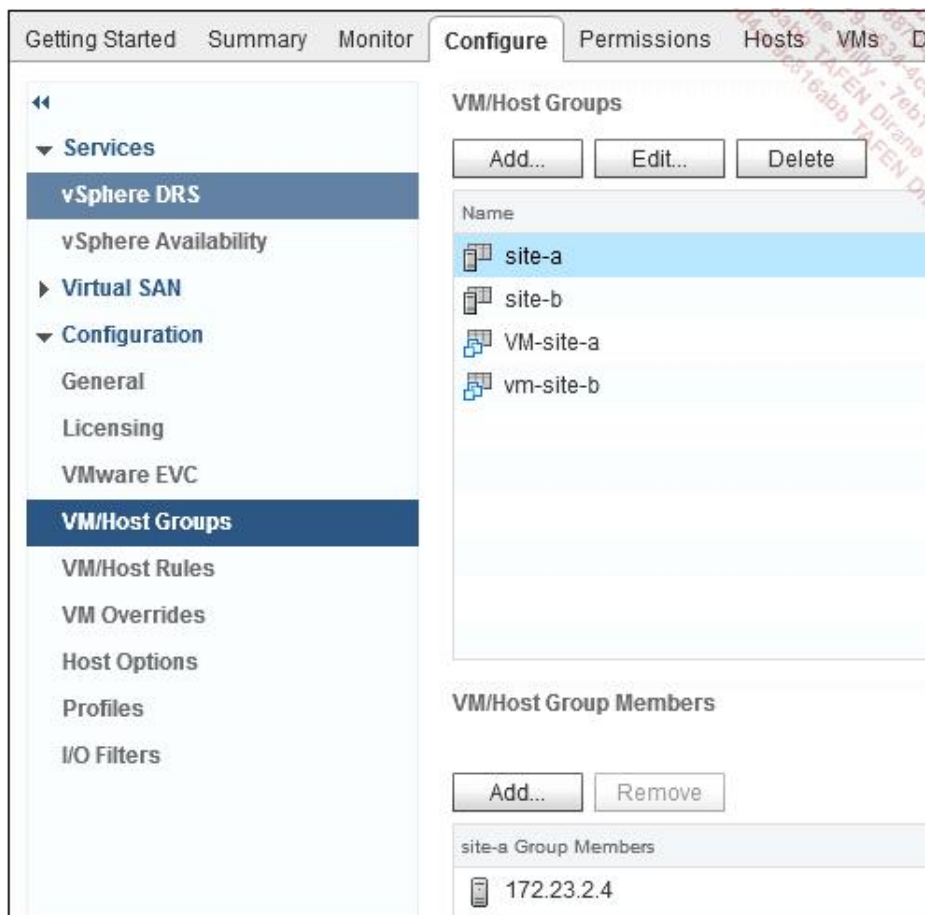
Définissez le nom du groupe. Il est préférable de donner des noms de groupe explicites afin de ne pas se perdre lorsqu'il y a plusieurs groupes présents.



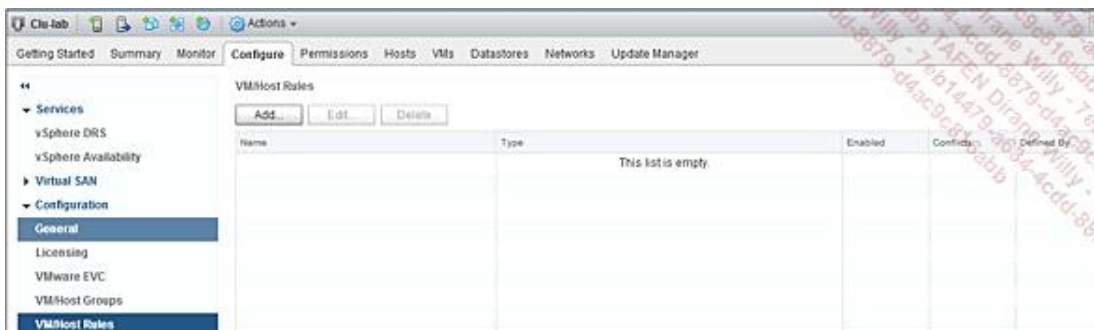
Lorsque vous cliquez sur **Add**, vous aurez au choix en fonction du type de groupe sélectionné : des machines virtuelles ou des ESXi présents dans le cluster.



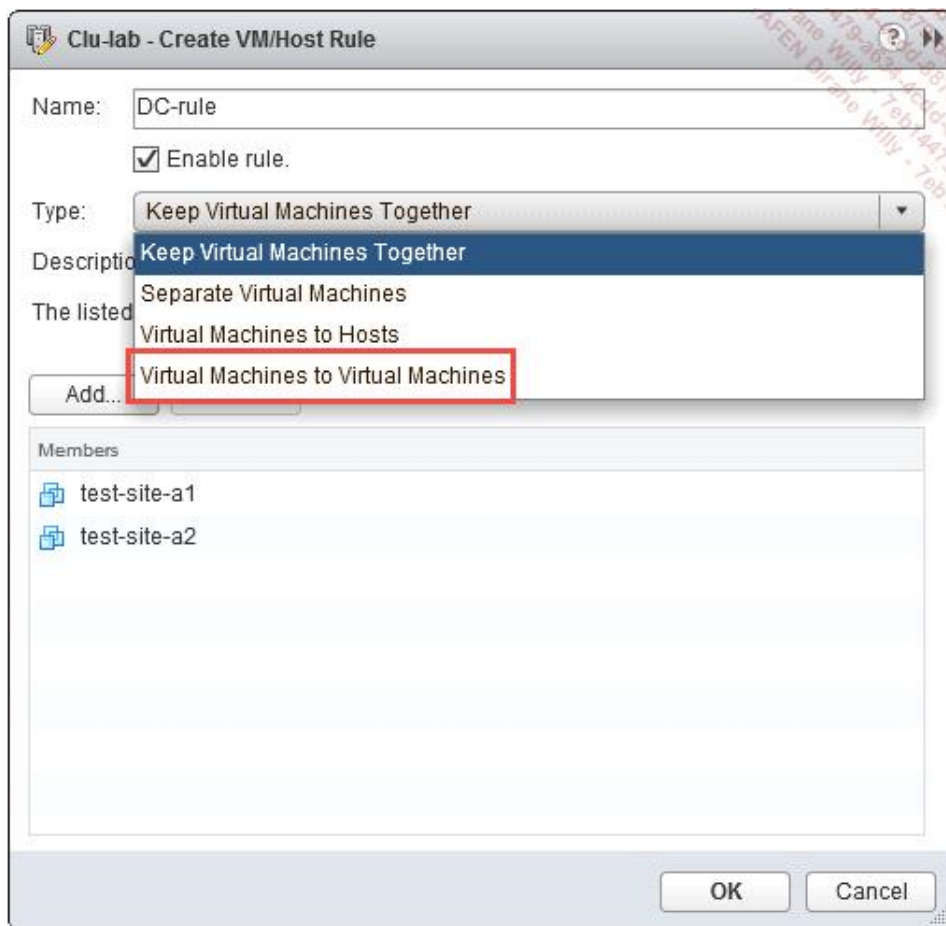
Une fois que cela est fait, il vous reste à cliquer sur **OK**. Répétez ces étapes en fonction de vos besoins en groupe d'ESXi et de machines virtuelles. Lorsque vous cliquez sur un groupe, vous pouvez voir ses membres.



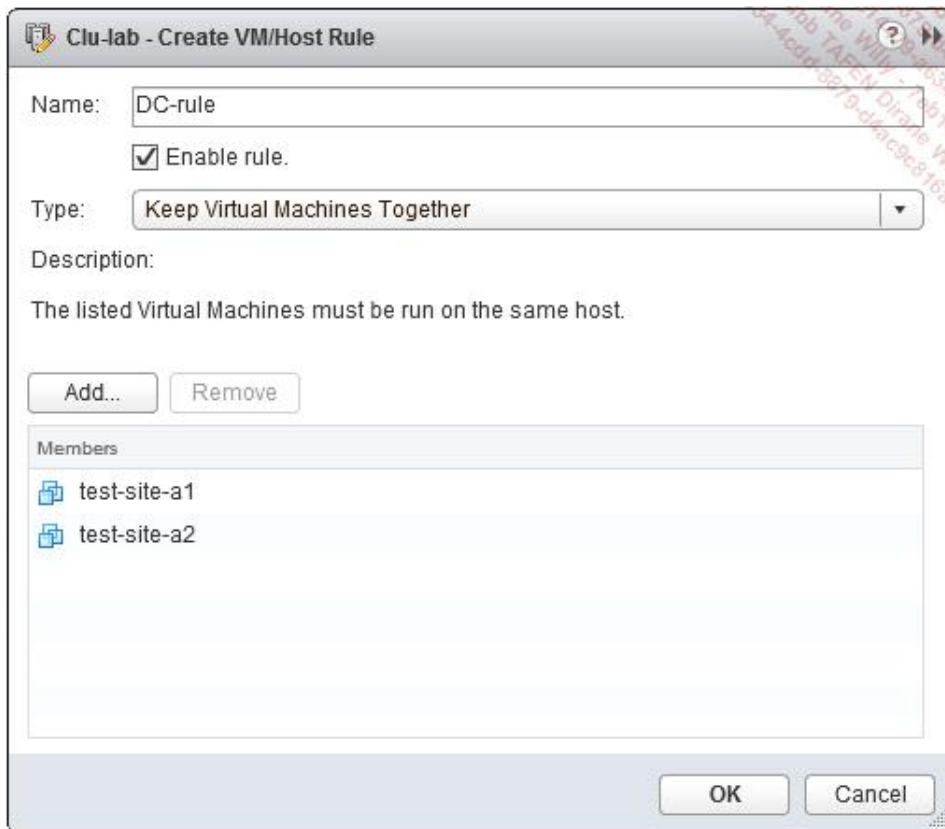
Vous allez maintenant passer à la partie configuration des affinités. Pour cela, toujours dans l'onglet **Configuration** du cluster, allez dans **VM/Host Rules** et cliquez sur **Add**.



Ici vous devez faire attention car il est possible de créer plusieurs types d'interactions. Notez que l'encadré rouge concerne la nouveauté liée à vSphere 6.5. Elle sera traitée dans les pages suivantes.



- Garder les machines virtuelles ensemble **Keep Virtual Machines Together** qui permet de faire en sorte que si une machine virtuelle est déplacée via une migration vMotion, les machines virtuelles associées seront aussi déplacées sur le même ESXi, si bien sûr nous ne violons pas les règles d'admission de HA.



- Séparer les machines virtuelles (**Separated Virtual Machines**), qui est l'exact opposé de la règle précédente. Les machines virtuelles dans cette règle ne pourront pas être exécutées sur le même ESXi. Sauf en cas de perte d'un ou plusieurs ESXi en fonction de la taille de notre cluster.
- Les règles de type groupe de machines virtuelles fonctionnant sur des groupes d'ESXi sont plus complexes que les deux types de règles précédemment abordées. Il existe quatre types de règles :
 - Doit s'exécuter sur des hosts en groupe (must run)
 - Devrait s'exécuter sur des hosts en groupe (should run)
 - Ne doit pas s'exécuter sur des hosts en groupe (must not run)
 - Ne devrait pas s'exécuter sur des hosts en groupe (should not run)

Dans le cas d'une règle de type must, les machines virtuelles ne pourront pas s'exécuter ailleurs que sur les ESXi configurés dans la règle.

Dans le cas d'une règle should, les machines virtuelles s'exécuteront de préférence sur le groupe d'ESXi associé, mais en cas de défaillance(s) d'ESXi elles pourront s'exécuter sur n'importe quel ESXi disponible.

Il reste à sélectionner le groupe de machines virtuelles et le groupe d'ESXi (caché par le menu des règles).

Clu-lab - Create VM/Host Rule

Name:

☒ Enable rule.

Type:

Description:

Virtual machines that are members of the VM Group vm-site-b should run on host group site-b.

VM group:

- Must run on hosts in group
- Should run on hosts in group**
- Must not run on hosts in group
- Should not run on hosts in group

- La version vSphere 6.5 ajoute l'interaction entre machines virtuelles (Virtual Machines to Virtual Machines). Tandis que les trois premières règles s'appuient sur l'exécution des machines virtuelles par les ESXi et la gestion du comportement des machines virtuelles lors de leur exécution, cette règle est différente car elle gère le redémarrage des groupes de machines virtuelles les unes par rapport aux autres, sous réserve que le HA (contrôle d'admission) autorise les redémarrages. Nous ne parlons pas ici de redémarrage des machines virtuelles dans un même groupe, cela se fait au niveau du paramétrage de HA.

Clu-lab - Create VM/Host Rule

Name:

☒ Enable rule.

Type:

Description:

Virtual machines in the VM group VM-site-a will be restarted first. Virtual machines in the VM group vm-site-b will be restarted afterwards, when the cluster dependency restart condition has been met.

First restart VMs in VM group:

Then restart VMs in VM group:

OK Cancel

Il est possible de désactiver une règle en décochant l'option **Enable rule**.

Lorsque les configurations sont terminées, vous pouvez voir quelles règles sont activées et quelles règles peuvent être en conflit (ce qu'il faudra corriger).

Getting Started Summary Monitor **Configure** Permissions Hosts VMs Datastores Networks Update Manager

Services

- vSphere DRS
- vSphere Availability

Virtual SAN

Configuration

- General
- Licensing
- VMware EVC
- VMHost Groups
- VMHost Rules**
- VM Overrides
- Host Options
- Profiles
- iO Filters

VMHost Rules

Add Edit Delete

Name	Type	Enabled	Conflicts	Defined By
DC-rule	Separate Virtual Machines	Yes	0	User
ESX-Local-Run	Run VMs on Hosts	Yes	0	User

VMHost Rule Details

The listed 2 Virtual Machines must run on different hosts.

Add Details Remove

Rule Members	Conflicts
test-site-a1	0
test-site-a2	0

Conflicts