

Connaissances générales utiles pour HyperFlex

Contenu

[Introduction](#)

[Acronymes couramment utilisés](#)

[Commande d'HyperFlex VMware VMNIC](#)

[Commande avant 3.5](#)

[Commande Post 3.5](#)

[SCVM sur noeud convergent par rapport au noeud de calcul](#)

[Noeud convergent](#)

[Réservations de ressources CPU](#)

[Réservations de ressources mémoire](#)

[Noeud de calcul](#)

[Scénarios de cluster malsains](#)

[Scénario 1 : Noeud désactivé](#)

[Scénario 2 : Disque dur](#)

[Scénario 3 : Ni Noeud Ni Disque Inactif](#)

[Comment rechercher un cluster SED à l'aide de l'interface de ligne de commande \(CLI\)](#)

[Mode de maintenance HX et mode de maintenance ESXi](#)

[Forum aux questions](#)

[Où sont installés les SCVM sur les serveurs Cisco HyperFlex M4 et M5 ?](#)

[Combien de noeuds défaillants un cluster peut-il tolérer ?](#)

[Clusters avec 5 noeuds ou plus](#)

[Clusters avec 3 et 4 noeuds](#)

[Que se passe-t-il si l'un des SCVM est arrêté ? Les machines virtuelles continuent-elles de fonctionner ?](#)

[La version matérielle VMware du SCVM a été mise à jour. Et maintenant ?](#)

Introduction

Ce document décrit les connaissances générales sur Cisco HyperFlex (HX) que les administrateurs doivent avoir à portée de main.

Acronymes couramment utilisés

SCVM = Machine virtuelle du contrôleur de stockage

VMNIC = Carte d'interface réseau de machine virtuelle

VNIC = Carte d'interface réseau virtuelle

SED = Lecteur de chiffrement automatique

VM = machine virtuelle

HX = HyperFlex

Commande d'HyperFlex VMware VMNIC

Le positionnement de VMNIC a été révisé dans HX version 3.5 et ultérieure.

Commande avant 3.5

Avant la version 3.5, les VNIC étaient attribuées en fonction des numéros de VNIC.

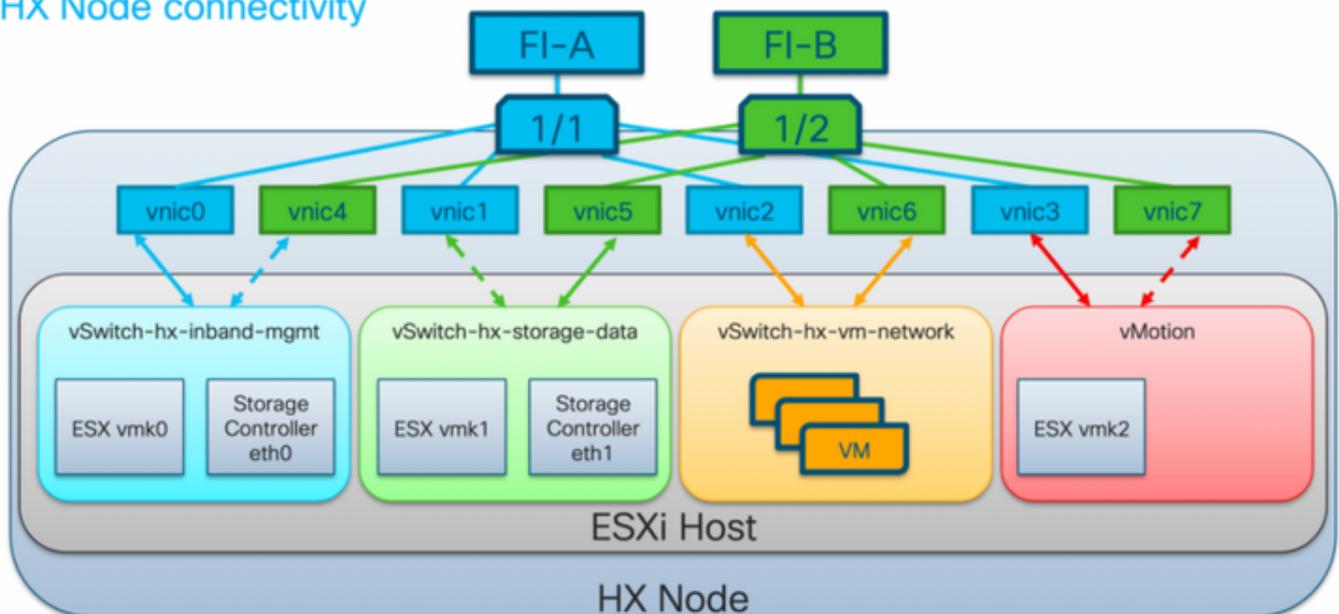
VNIC	Commutateur virtuel (vSwitch)
VNIC 0 et VNIC 1	vSwitch-hx-inband-mgmt
VNIC 2 et VNIC 3	vSwitch-hx-storage-data
VNIC 4 et VNIC 5	vSwitch-hx-vm-network
VNIC 6 et VNIC 7	vMotion

Commande Post 3.5

Dans les versions 3.5 et ultérieures, les cartes réseau virtuelles sont attribuées en fonction de l'adresse MAC (Media Access Control). Par conséquent, il n'existe pas d'ordre d'affectation particulier.

Si une mise à niveau d'une version antérieure à la version 3.5 vers la version 3.5 ou supérieure est effectuée, l'ordre VMNIC est maintenu.

HX Node connectivity



Note: Pour HX Hyper-V, cela ne s'applique pas, car Hyper-V utilise un nom de périphérique cohérent (CDN).

SCVM sur noeud convergent par rapport au noeud de calcul

Les SCVM résident à la fois sur les noeuds de calcul et convergents et il existe des différences entre eux.

Noeud convergent

Réservations de ressources CPU

Puisque les SCVM fournissent des fonctionnalités critiques de la plate-forme de données distribuées Cisco HX, le programme d'installation HyperFlex configurera les réservations de ressources CPU pour les machines virtuelles du contrôleur. Cette réservation garantit que les machines virtuelles du contrôleur disposeront au minimum de ressources d'unité de traitement centrale (UC), dans les situations où les ressources physiques du processeur de l'hôte d'hyperviseur ESXi sont fortement consommées par les machines virtuelles invitées. Il s'agit d'une garantie logicielle, ce qui signifie que dans la plupart des cas, les SCVM n'utilisent pas toutes les ressources CPU réservées, permettant ainsi aux machines virtuelles invitées de les utiliser. Le tableau suivant détaille la réservation des ressources CPU des machines virtuelles du contrôleur de stockage :

Nombre de vCPU	Partages	Réservation	Limite
8	Faible	1 0800 MHZ	Illimité

Réservations de ressources mémoire

Puisque les SCVM fournissent des fonctionnalités critiques de la plate-forme de données distribuées Cisco HX, le programme d'installation HyperFlex configurera les réservations de ressources de mémoire pour les machines virtuelles du contrôleur. Cette réservation garantit que les machines virtuelles du contrôleur disposeront de ressources de mémoire à un niveau minimal, dans les situations où les ressources de mémoire physique de l'hôte de l'hyperviseur ESXi sont fortement consommées par les machines virtuelles invitées. Le tableau suivant détaille la réservation des ressources de mémoire des machines virtuelles du contrôleur de stockage :

Modèles de serveur	Quantité de mémoire invité	Réserver toute la mémoire de
HX 220c-M5SX HXAF 220c-M5SX HX 220c-M4S HXAF220c-M4S	48 Go	Oui
HX 240c-M5SX HXAF 240c-M5SX HX240c-M4SX HXAF240c-M4SX HX240c-M5L	72 Go 78 Go	Oui Oui

Noeud de calcul

Les noeuds de calcul uniquement disposent d'un SCVM léger. Il est configuré avec seulement 1 vCPU de 1024 MHz et 512 Mo de réservation de mémoire.

L'objectif du noeud de calcul est principalement de maintenir les paramètres DRS (Distributed Resource Scheduler™) de vCluster, afin de s'assurer que DRS ne réoriente pas les machines virtuelles utilisateur vers des noeuds convergents.

Scénarios de cluster malsains

Un cluster HX peut être rendu malsain dans les scénarios suivants.

Scénario 1 : Noeud désactivé

Un cluster se trouve dans un état malsain lorsqu'un noeud tombe en panne. Un noeud est censé

être arrêté lors d'une mise à niveau de cluster ou lorsqu'un serveur est mis en mode maintenance.

```
root@SpringpathController:~# stcli cluster storage-summary --detail
<snip>
current ensemble size:3
# of caching failures before cluster shuts down:2
minimum cache copies remaining:2
minimum data copies available for some user data:2
current healing status:rebuilding/healing is needed, but not in progress yet. warning:
insufficient node or space resources may prevent healing. storage node 10.197.252.99is either
down or initializing disks.
minimum metadata copies available for cluster metadata:2
# of unavailable nodes:1
# of nodes failure tolerable for cluster to be available:0
health state reason:storage cluster is unhealthy.storage node 10.197.252.99 is unavailable.
# of node failures before cluster shuts down:2
# of node failures before cluster goes into readonly:2
# of persistent devices failures tolerable for cluster to be available:1
# of node failures before cluster goes to enospace warn trying to move the existing data:na
# of persistent devices failures before cluster shuts down:2
# of persistent devices failures before cluster goes into readonly:2
# of caching failures before cluster goes into readonly:na
# of caching devices failures tolerable for cluster to be available:1
resiliencyInfo:
messages:
-----
Storage cluster is unhealthy.
-----
Storage node 10.197.252.99 is unavailable.
-----
state: 2
nodeFailuresTolerable: 0
cachingDeviceFailuresTolerable: 1
persistentDeviceFailuresTolerable: 1
zoneResInfoList: None
spaceStatus: normal
totalCapacity: 3.0T
totalSavings: 5.17%
usedCapacity: 45.9G
zkHealth: online
clusterAccessPolicy: lenient
dataReplicationCompliance: non_compliant
dataReplicationFactor: 3
```

Scénario 2 : Disque dur

Un cluster se trouve dans un état malsain lorsqu'un disque n'est pas disponible. La condition doit être effacée lorsque les données sont distribuées à d'autres disques.

```
root@SpringpathController:~# stcli cluster storage-summary --detail
<snip>
current ensemble size:3
# of caching failures before cluster shuts down:2
minimum cache copies remaining:2
minimum data copies available for some user data:2
current healing status:rebuilding/healing is needed, but not in progress yet. warning:
insufficient node or space resources may prevent healing. storage node is either down or
initializing disks.
minimum metadata copies available for cluster metadata:2
# of unavailable nodes:1
```

```

# of nodes failure tolerable for cluster to be available:0
health state reason:storage cluster is unhealthy. persistent device disk [5000c5007e113d8b:0000000000000000] on node 10.197.252.99 is unavailable.
# of node failures before cluster shuts down:2
# of node failures before cluster goes into readonly:2
# of persistent devices failures tolerable for cluster to be available:1
# of node failures before cluster goes to enospace warn trying to move the existing data:na
# of persistent devices failures before cluster shuts down:2
# of persistent devices failures before cluster goes into readonly:2
# of caching failures before cluster goes into readonly:na
# of caching devices failures tolerable for cluster to be available:1
resiliencyInfo:
messages:
-----
Storage cluster is unhealthy.
-----
Persistent Device Disk [5000c5007e113d8b:0000000000000000] on node 10.197.252.99 is unavailable.
-----
state: 2
nodeFailuresTolerable: 0
cachingDeviceFailuresTolerable: 1
persistentDeviceFailuresTolerable: 1
zoneResInfoList: None
spaceStatus: normal
totalCapacity: 3.0T
totalSavings: 8.82%
usedCapacity: 45.9G
zkHealth: online
clusterAccessPolicy: lenient
dataReplicationCompliance: non_compliant
dataReplicationFactor: 3

```

Scénario 3 : Ni Noeud Ni Disque Inactif

Un cluster peut se retrouver dans un état malsain lorsqu'aucun noeud ni disque n'est hors service. Cette condition se produit si la reconstruction est en cours.

```

root@SpringpathController:~# stcli cluster storage-summary --detail
<snip>
resiliencyDetails:
    current ensemble size:5
    # of caching failures before cluster shuts down:3
    minimum cache copies remaining:3
    minimum data copies available for some user data:2
current healing status:rebuilding is in progress, 98% completed.           minimum metadata copies
available for cluster metadata:2
    time remaining before current healing operation finishes:7 hr(s), 15 min(s), and 34
sec(s)
    # of unavailable nodes:0
    # of nodes failure tolerable for cluster to be available:1
    health state reason:storage cluster is unhealthy.
    # of node failures before cluster shuts down:2
    # of node failures before cluster goes into readonly:2
    # of persistent devices failures tolerable for cluster to be available:1
    # of node failures before cluster goes to enospace warn trying to move the existing
data:na
    # of persistent devices failures before cluster shuts down:2
    # of persistent devices failures before cluster goes into readonly:2

```

```
# of caching failures before cluster goes into readonly:na
# of caching devices failures tolerable for cluster to be available:2
resiliencyInfo:
  messages:
    Storage cluster is unhealthy.
  state: 2
  nodeFailuresTolerable: 1
  cachingDeviceFailuresTolerable: 2
  persistentDeviceFailuresTolerable: 1
  zoneResInfoList: None
spaceStatus: normal
totalCapacity: 225.0T
totalSavings: 42.93%
usedCapacity: 67.7T
clusterAccessPolicy: lenient
dataReplicationCompliance: non_compliant
dataReplicationFactor: 3
```

Comment rechercher un cluster SED à l'aide de l'interface de ligne de commande (CLI)

Si l'accès à HX Connect n'est pas disponible, l'interface de ligne de commande peut être utilisée pour vérifier si le cluster est SED.

```
# Check if the cluster is SED capable
root@SpringpathController:~# cat /etc/springpath/sed_capability.conf
sed_capable_cluster=False
```

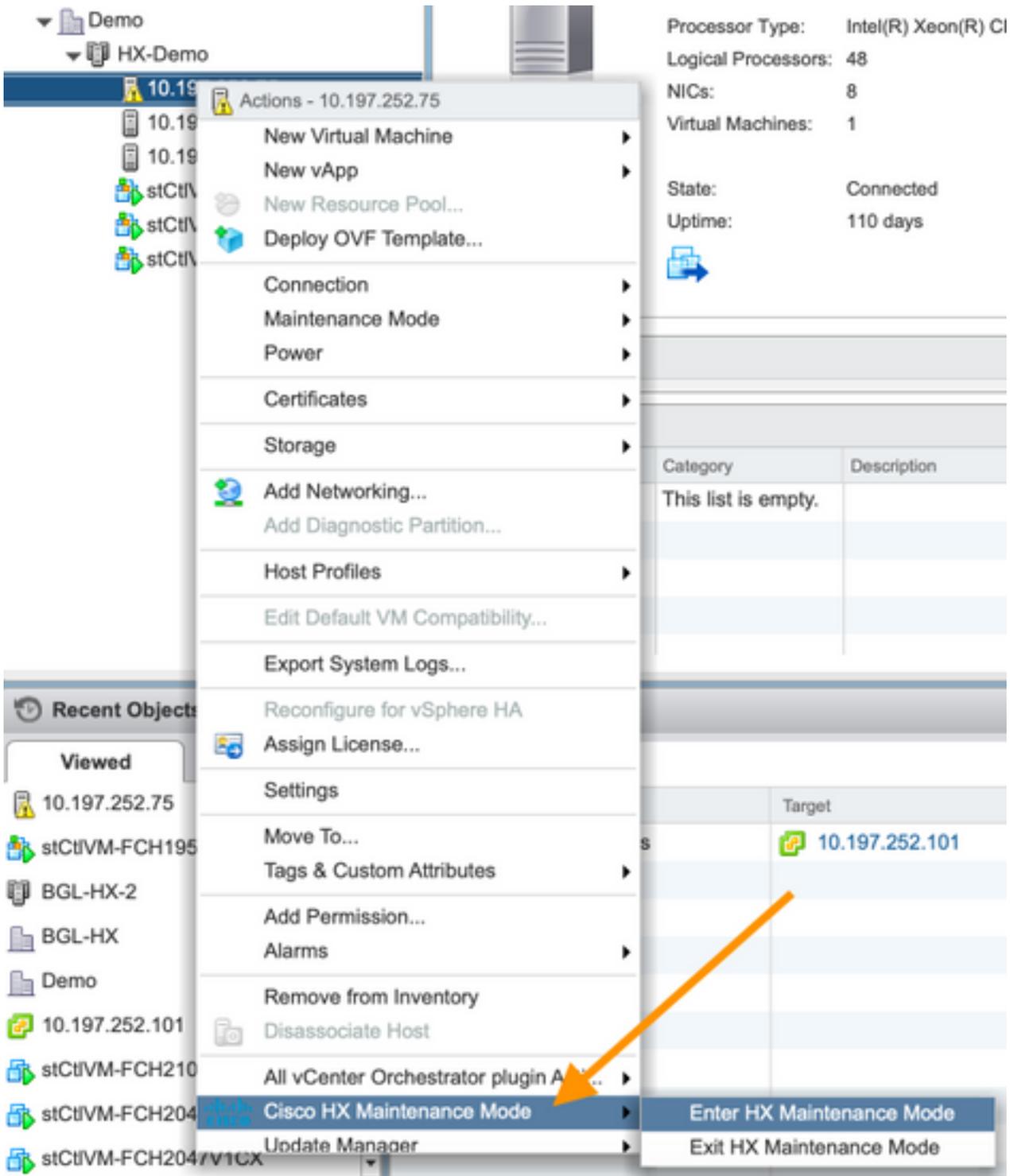
```
# Check if the cluster is SED enabled root@SpringpathController:~# cat /etc/springpath/sed.conf
sed_encryption_state=unknown
```

```
root@SpringpathController:~# /usr/share/springpath/storfs-appliance/sed-client.sh -l
WWN,Slot,Supported,Enabled,Locked,Vendor,Model,Serial,Size
5002538c40a42d38,1,0,0,0,Samsung,SAMSUNG_MZ7LM240HMHQ-00003,S3LKNX0K406548,228936
5000c50030278d83,25,1,1,0,MICRON,S650DC-800FIPS,ZAZ15QDM0000822150Z3,763097
500a07511d38cd36,2,1,1,0,MICRON,Micron_5100_MTFDDAK960TCB_SED,17261D38CD36,915715
500a07511d38efbe,4,1,1,0,MICRON,Micron_5100_MTFDDAK960TCB_SED,17261D38EFBE,915715
500a07511d38f350,7,1,1,0,MICRON,Micron_5100_MTFDDAK960TCB_SED,17261D38F350,915715
500a07511d38eaa6,3,1,1,0,MICRON,Micron_5100_MTFDDAK960TCB_SED,17261D38EAA6,915715
500a07511d38ce80,6,1,1,0,MICRON,Micron_5100_MTFDDAK960TCB_SED,17261D38CE80,915715
500a07511d38e4fc,5,1,1,0,MICRON,Micron_5100_MTFDDAK960TCB_SED,17261D38E4FC,915715
```

Mode de maintenance HX et mode de maintenance ESXi

Lorsque des activités de maintenance doivent être effectuées sur un serveur faisant partie d'un cluster HX, le mode de maintenance HX doit être utilisé au lieu du mode de maintenance ESXi. Le SCVM est arrêté avec grâce lorsque le mode de maintenance HX est utilisé alors qu'il est brusquement arrêté lorsque le mode de maintenance ESXi est utilisé.

Lorsqu'un noeud est en mode maintenance, il est considéré comme inactif, c'est-à-dire qu'il est défaillant pour un noeud.



Assurez-vous que le cluster est sain avant de passer en mode maintenance.

```

root@SpringpathController:~# stcli cluster storage-summary --detail
<snip>
current ensemble size:3
# of caching failures before cluster shuts down:3
minimum cache copies remaining:3
minimum data copies available for some user data:3
minimum metadata copies available for cluster metadata:3
# of unavailable nodes:0
# of nodes failure tolerable for cluster to be available:1
health state reason:storage cluster is healthy.
# of node failures before cluster shuts down:3
# of node failures before cluster goes into readonly:3
# of persistent devices failures tolerable for cluster to be available:2
  
```

```
# of node failures before cluster goes to enospace warn trying to move the existing data:na
# of persistent devices failures before cluster shuts down:3
# of persistent devices failures before cluster goes into readonly:3
# of caching failures before cluster goes into readonly:na
# of caching devices failures tolerable for cluster to be available:2
resiliencyInfo:
messages:
Storage cluster is healthy.
state: 1
nodeFailuresTolerable: 1
cachingDeviceFailuresTolerable: 2
<snip>
```

Forum aux questions

Où sont installés les SCVM sur les serveurs Cisco HyperFlex M4 et M5 ?

L'emplacement SCVM est différent entre les serveurs Cisco Hyperflex M4 et M5. Le tableau ci-dessous répertorie l'emplacement du SCVM et fournit d'autres informations utiles.

Serveur Cisco HX	ESXi	SCVM sda	Cache du disque dur SSD (Solid State Drive)	SSD ménager sdb1 et sdb2
HX 220 M4	Secure Digital (cartes SD)	3,5 G sur cartes SD	Slot 2	Emplacement 1
HX 240 M4	SD Cartes	Sur SSD contrôlé par PCH (esxi en a le contrôle)	Emplacement 1	Sur SSD contrôlé par PCH
HX 220 M5	M2 Lecteur	Lecteur M2	Slot 2	Emplacement 1
HX 240 M5	M2 Lecteur	Lecteur M2	Logement arrière SSD	Emplacement 1

Combien de noeuds défailants un cluster peut-il tolérer ?

Le nombre d'échecs qu'un cluster peut tolérer dépend du facteur de réplication et de la stratégie d'accès.

Clusters avec 5 noeuds ou plus

Lorsque le facteur de réplication (RF) est 3 et que la stratégie d'accès est définie sur Lénite, si 2 noeuds échouent, le cluster est toujours en lecture/écriture. Si 3 noeuds devaient échouer, le cluster s'arrêtera.

Facteur de réplication	Stratégie d'accès	Nombre de noeuds ayant échoué		Lecture	Arrêt
		Lecture/écriture	seule		
3	Lénite	2	seule	—	3

3	Strict	1	2	3
2	Lénite	1	—	2
2	Strict	—	1	2

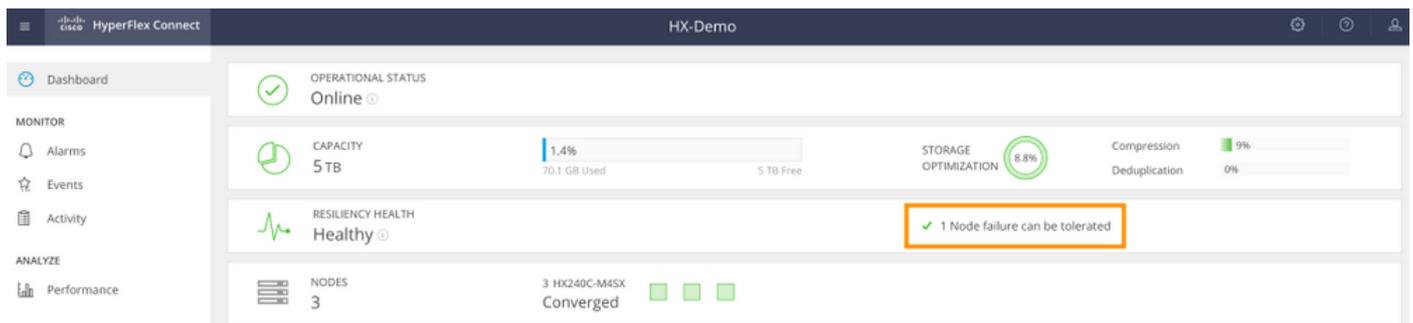
Clusters avec 3 et 4 noeuds

Lorsque la RF est 3 et que la stratégie d'accès est définie sur Lénite ou Strict, si un noeud unique échoue, le cluster est toujours en lecture/écriture. Si deux noeuds échouent, le cluster s'arrête.

Facteur de réplication	Stratégie d'accès	Nombre de noeuds ayant échoué		Lecture	Arrêt
		Lecture/écriture	seule		
3	Lénent ou strict	1	—	—	2
2	Lénite	1	—	—	2
2	Strict	—	—	1	2

Exemple de grappe de 3 noeuds (RF : 3, Stratégie d'accès : clémente)

Exemple d'interface utilisateur graphique



Exemple CLI

```

root@SpringpathController:~# stcli cluster storage-summary --detail
<snip>
current ensemble size:3
# of caching failures before cluster shuts down:3
minimum cache copies remaining:3
minimum data copies available for some user data:3
minimum metadata copies available for cluster metadata:3
# of unavailable nodes:0
# of nodes failure tolerable for cluster to be available:1
health state reason:storage cluster is healthy.
# of node failures before cluster shuts down:3
# of node failures before cluster goes into readonly:3
# of persistent devices failures tolerable for cluster to be available:2
# of node failures before cluster goes to enospace warn trying to move the existing data:na
# of persistent devices failures before cluster shuts down:3
# of persistent devices failures before cluster goes into readonly:3
# of caching failures before cluster goes into readonly:na
# of caching devices failures tolerable for cluster to be available:2
resiliencyInfo:
messages:
Storage cluster is healthy.
state: 1

```

<snip>

`clusterAccessPolicy: lenient`

Que se passe-t-il si l'un des SCVM est arrêté ? Les machines virtuelles continuent-elles de fonctionner ?

Avertissement : Il ne s'agit pas d'une opération prise en charge sur un SCVM. Ceci est uniquement à des fins de démonstration.

Note: Assurez-vous qu'un seul SCVM est arrêté à la fois. Vérifiez également que le cluster est sain avant d'arrêter un SCVM. Ce scénario est uniquement destiné à démontrer que les machines virtuelles et les magasins de données doivent fonctionner même si un SCVM est arrêté ou indisponible.

Les machines virtuelles continueront à fonctionner normalement. Voici un exemple de sortie dans lequel le SCVM a été arrêté, mais les data stores sont restés montés et disponibles.

```
[root@node1:~] vim-cmd vmsvc/getallvms
VmId Name File Guest OS Version Annotation
1 stCtlVM-F 9H [SpringpathDS-F 9H] stCtlVM-F 9H/stCtlVM-F 9H.vmx ubuntu64Guest vmx-13

[root@node1:~] vim-cmd vmsvc/power.off 1
Powering off VM:

[root@node1:~] vim-cmd vmsvc/power.getstate 1
Retrieved runtime info
Powered off

[root@node1:~] esxcfg-nas -l
Test is 10.197.252.106:Test from 3203172317343203629-5043383143428344954 mounted available
ReplSec is 10.197.252.106:ReplSec from 3203172317343203629-5043383143428344954 mounted available
New_DS is 10.197.252.106:New_DS from 3203172317343203629-5043383143428344954 mounted available
```

La version matérielle VMware du SCVM a été mise à jour. Et maintenant ?

Avertissement : Il ne s'agit pas d'une opération prise en charge sur un SCVM. Ceci est uniquement à des fins de démonstration.

La mise à niveau de la version matérielle de VMware en modifiant les paramètres de VM dans **Compatibility > Upgrade VM Compatibility** est le client Web vSphere qui n'est PAS pris en charge sur un SCVM. Le SCVM indique qu'il est hors connexion dans HX Connect.

```
root@SpringpathController0 UE:~# lsblk
NAME MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
sda 8:0 0 2.5G 0 disk
├─sda1 8:1 0 2.5G 0 part /
sdb 8:16 0 100G 0 disk
├─sdb1 8:17 0 64G 0 part /var/stv
└─sdb2 8:18 0 24G 0 part /var/zookeeper

root@SpringpathController0 UE:~# lsscsi
[2:0:0:0] disk VMware Virtual disk 2.0 /dev/sda
[2:0:1:0] disk VMware Virtual disk 2.0 /dev/sdb
```

```
root@SpringpathController0      UE:~# cat /var/log/springpath/diskslotmap-v2.txt
1.11.1:5002538a17221ab0:SAMSUNG:MZIES800HMHP/003:S1N2NY0J201389:EM19:SAS:SSD:763097:Inactive:/dev/sdc
1.11.2:5002538c405537e0:Samsung:SAMSUNG_MZ7LM3T8HMLP-00003:S
98:GXT51F3Q:SATA:SSD:3662830:Inactive:/dev/sdd
1.11.3:5002538c4055383a:Samsung:SAMSUNG_MZ7LM3T8HMLP-00003:S
88:GXT51F3Q:SATA:SSD:3662830:Inactive:/dev/sde
1.11.4:5002538c40553813:Samsung:SAMSUNG_MZ7LM3T8HMLP-00003:S
49:GXT51F3Q:SATA:SSD:3662830:Inactive:/dev/sdf
1.11.5:5002538c4055380e:Samsung:SAMSUNG_MZ7LM3T8HMLP-00003:S
44:GXT51F3Q:SATA:SSD:3662830:Inactive:/dev/sdg
1.11.6:5002538c40553818:Samsung:SAMSUNG_MZ7LM3T8HMLP-00003:S
54:GXT51F3Q:SATA:SSD:3662830:Inactive:/dev/sdh
1.11.7:5002538c405537d1:Samsung:SAMSUNG_MZ7LM3T8HMLP-00003:S
83:GXT51F3Q:SATA:SSD:3662830:Inactive:/dev/sdi
1.11.8:5002538c405537d8:Samsung:SAMSUNG_MZ7LM3T8HMLP-00003:S
90:GXT51F3Q:SATA:SSD:3662830:Inactive:/dev/sdj
1.11.9:5002538c4055383b:Samsung:SAMSUNG_MZ7LM3T8HMLP-00003:S
89:GXT51F3Q:SATA:SSD:3662830:Inactive:/dev/sdk
1.11.10:5002538c4055381f:Samsung:SAMSUNG_MZ7LM3T8HMLP-00003:S
61:GXT51F3Q:SATA:SSD:3662830:Inactive:/dev/sdl
1.11.11:5002538c40553823:Samsung:SAMSUNG_MZ7LM3T8HMLP-00003:S
65:GXT51F3Q:SATA:SSD:3662830:Inactive:/dev/sdm
```

Attention : Si cette opération a été accidentellement effectuée, contactez le support technique de Cisco pour obtenir de l'aide. Le SCVM devra être redéployé.