



Plongée en profondeur

ONTAP Select

NetApp
January 29, 2026

This PDF was generated from https://docs.netapp.com/fr-fr/ontap-select-9161/concept_stor_concepts_chars.html on January 29, 2026. Always check docs.netapp.com for the latest.

Sommaire

Plongée en profondeur	1
Stockage	1
Stockage ONTAP Select : concepts généraux et caractéristiques	1
Services RAID matériels pour le stockage local connecté ONTAP Select	6
Services de configuration RAID logiciel ONTAP Select pour le stockage local connecté	13
ONTAP Select VSAN et configurations de baies externes	22
Augmenter la capacité de stockage ONTAP Select	26
Prise en charge de l'efficacité du stockage ONTAP Select	29
Réseautage	32
Concepts et caractéristiques du réseau ONTAP Select	32
ONTAP Select les configurations de réseau à nœud unique et à nœuds multiples	34
ONTAP Select les réseaux internes et externes	39
Configurations réseau ONTAP Select en charge	42
ONTAP Select la configuration VMware vSphere vSwitch sur ESXi	43
Configuration du commutateur physique ONTAP Select	52
Séparation du trafic de données et de gestion ONTAP Select	54
Architecture haute disponibilité	56
Configurations de haute disponibilité ONTAP Select	56
ONTAP Select HA RSM et agrégats en miroir	59
ONTAP Select HA améliore la protection des données	62
Performances	65
Aperçu des performances ONTAP Select	65
Performances ONTAP Select 9.6 : stockage SSD à connexion directe HA haut de gamme	65

Plongée en profondeur

Stockage

Stockage ONTAP Select : concepts généraux et caractéristiques

Découvrez les concepts généraux de stockage qui s'appliquent à l'environnement ONTAP Select avant d'explorer les composants de stockage spécifiques.

Phases de configuration du stockage

Les principales phases de configuration du stockage hôte ONTAP Select incluent les suivantes :

- Prérequis avant le déploiement
 - Assurez-vous que chaque hôte hyperviseur est configuré et prêt pour un déploiement ONTAP Select .
 - La configuration implique les disques physiques, les contrôleurs et groupes RAID, les LUN, ainsi que la préparation réseau associée.
 - Cette configuration est effectuée en dehors d' ONTAP Select.
- Configuration à l'aide de l'utilitaire administrateur de l'hyperviseur
 - Vous pouvez configurer certains aspects du stockage à l'aide de l'utilitaire d'administration de l'hyperviseur (par exemple, vSphere dans un environnement VMware).
 - Cette configuration est effectuée en dehors d' ONTAP Select.
- Configuration à l'aide de l'utilitaire d'administration ONTAP Select Deploy
 - Vous pouvez utiliser l'utilitaire d'administration Déployer pour configurer les structures de stockage logique principales.
 - Cela est effectué soit explicitement via des commandes CLI, soit automatiquement par l'utilitaire dans le cadre d'un déploiement.
- Configuration post-déploiement
 - Une fois le déploiement ONTAP Select terminé, vous pouvez configurer le cluster à l'aide de l'interface de ligne de commande ONTAP ou du Gestionnaire système.
 - Cette configuration est effectuée en dehors d' ONTAP Select Deploy.

Stockage géré ou non géré

Le stockage accessible et directement contrôlé par ONTAP Select est un stockage géré. Tout autre stockage sur le même hôte hyperviseur est un stockage non géré.

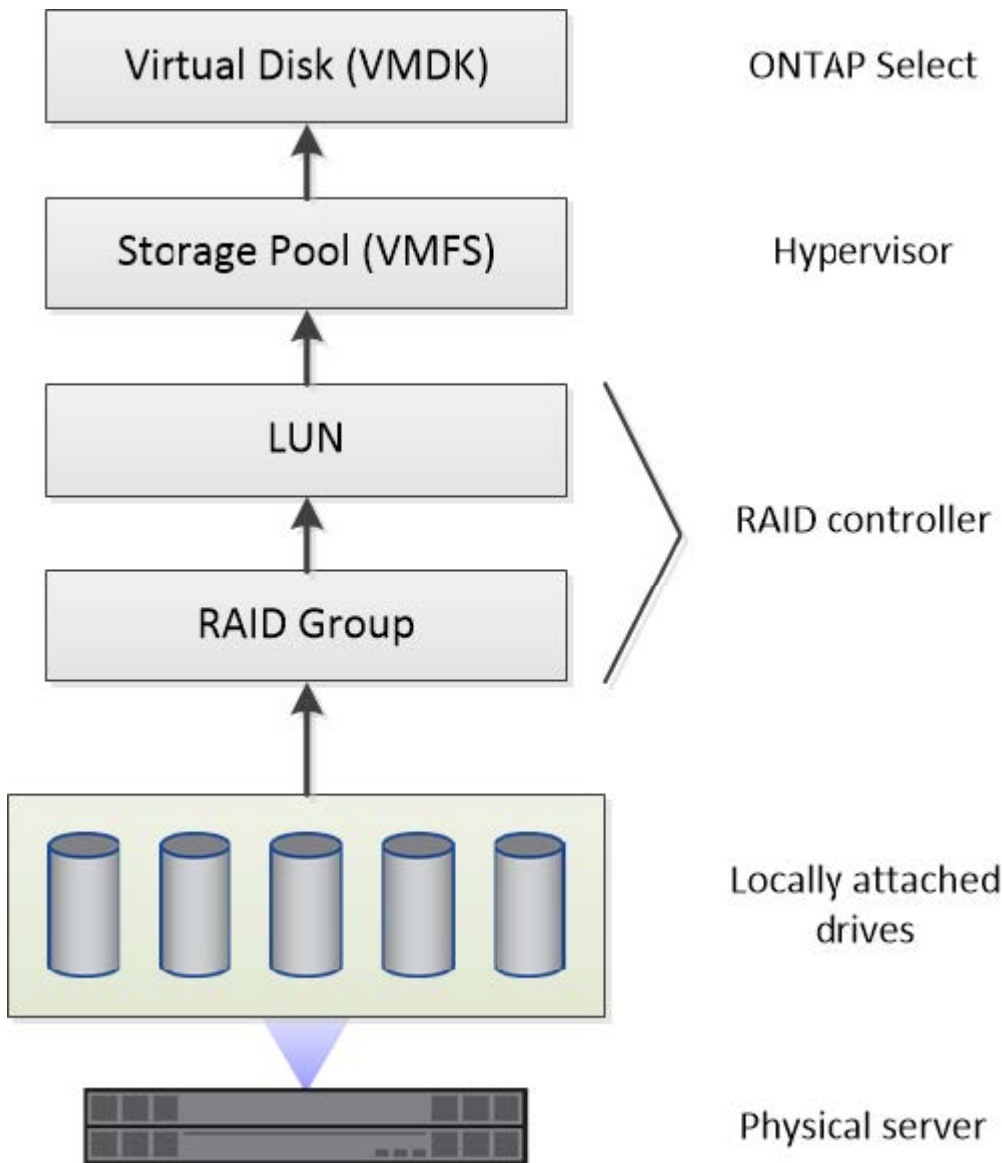
Stockage physique homogène

Tous les disques physiques composant le stockage géré ONTAP Select doivent être homogènes. Autrement dit, tout le matériel doit être identique sur les points suivants :

- Type (SAS, NL-SAS, SATA, SSD)
- Vitesse (tr/min)

Illustration de l'environnement de stockage local

Chaque hôte hyperviseur contient des disques locaux et d'autres composants de stockage logiques utilisables par ONTAP Select. Ces composants de stockage sont organisés en couches, à partir du disque physique.



Caractéristiques des composants de stockage local

Plusieurs concepts s'appliquent aux composants de stockage local utilisés dans un environnement ONTAP Select. Familiarisez-vous avec ces concepts avant de préparer un déploiement ONTAP Select. Ils sont classés par catégorie : groupes RAID et LUN, pools de stockage et disques virtuels.

Regroupement des disques physiques en groupes RAID et LUN

Un ou plusieurs disques physiques peuvent être connectés localement au serveur hôte et mis à disposition d'ONTAP Select. Ces disques physiques sont affectés à des groupes RAID, qui sont ensuite présentés au système d'exploitation hôte de l'hyperviseur sous la forme d'un ou plusieurs LUN. Chaque LUN est présenté au système d'exploitation hôte de l'hyperviseur comme un disque dur physique.

Lors de la configuration d'un hôte ONTAP Select, vous devez tenir compte des éléments suivants :

- Tout le stockage géré doit être accessible via un seul contrôleur RAID
- Selon le fournisseur, chaque contrôleur RAID prend en charge un nombre maximal de disques par groupe RAID

Un ou plusieurs groupes RAID

Chaque hôte ONTAP Select doit disposer d'un seul contrôleur RAID. Il est recommandé de créer un seul groupe RAID pour ONTAP Select. Cependant, dans certaines situations, il peut être judicieux d'en créer plusieurs. ["Résumé des meilleures pratiques"](#) .

Considérations relatives aux pools de stockage

Il existe plusieurs problèmes liés aux pools de stockage dont vous devez être conscient dans le cadre de la préparation du déploiement ONTAP Select.



Dans un environnement VMware, un pool de stockage est synonyme de banque de données VMware.

Pools de stockage et LUN

Chaque LUN est considéré comme un disque local sur l'hôte hyperviseur et peut faire partie d'un pool de stockage. Chaque pool de stockage est formaté avec un système de fichiers exploitable par le système d'exploitation hôte hyperviseur.

Vous devez vous assurer que les pools de stockage sont correctement créés dans le cadre d'un déploiement ONTAP Select . Vous pouvez créer un pool de stockage à l'aide de l'outil d'administration de l'hyperviseur. Par exemple, avec VMware, vous pouvez utiliser le client vSphere pour créer un pool de stockage. Ce pool est ensuite transmis à l'utilitaire d'administration ONTAP Select Deploy.

Gérer les disques virtuels sur ESXi

Il existe plusieurs problèmes liés aux disques virtuels dont vous devez être conscient dans le cadre de la préparation du déploiement ONTAP Select.

Disques virtuels et systèmes de fichiers

La machine virtuelle ONTAP Select se voit allouer plusieurs disques virtuels. Chaque disque virtuel est en réalité un fichier contenu dans un pool de stockage et géré par l'hyperviseur. ONTAP Select utilise plusieurs types de disques, principalement des disques système et des disques de données.

Vous devez également être conscient des points suivants concernant les disques virtuels :

- Le pool de stockage doit être disponible avant que les disques virtuels puissent être créés.
- Les disques virtuels ne peuvent pas être créés avant la création de la machine virtuelle.
- Vous devez vous fier à l'utilitaire d'administration ONTAP Select Deploy pour créer tous les disques virtuels (c'est-à-dire qu'un administrateur ne doit jamais créer un disque virtuel en dehors de l'utilitaire Deploy).

Configuration des disques virtuels

Les disques virtuels sont gérés par ONTAP Select. Ils sont créés automatiquement lors de la création d'un cluster à l'aide de l'utilitaire d'administration Deploy.

Illustration de l'environnement de stockage externe sur ESXi

La solution vNAS ONTAP Select permet à ONTAP Select d'utiliser des banques de données hébergées sur un stockage externe à l'hôte hyperviseur. Ces banques de données sont accessibles via le réseau via VMware vSAN ou directement sur une baie de stockage externe.

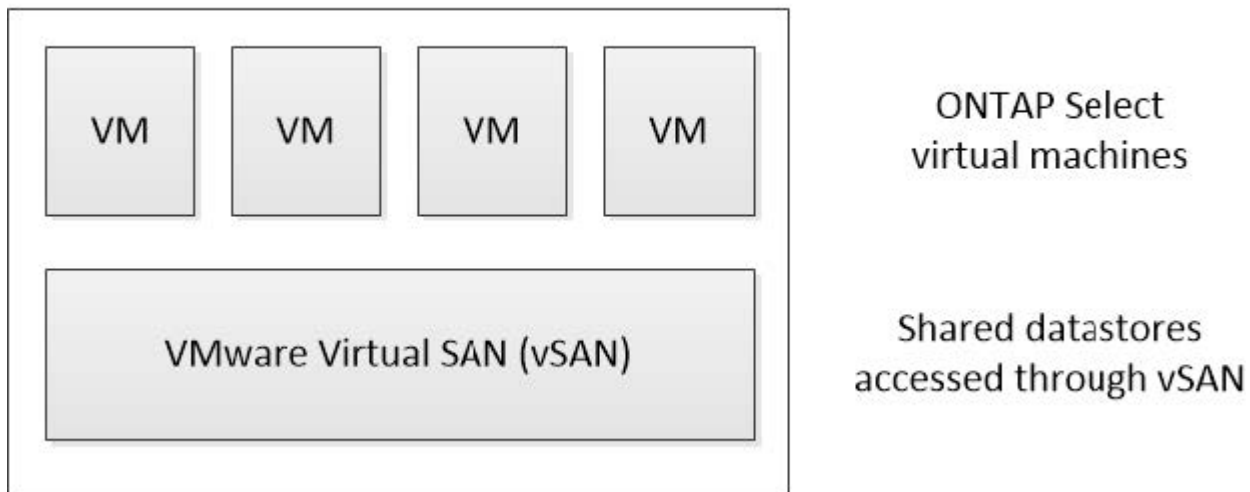
ONTAP Select peut être configuré pour utiliser les types suivants de banques de données réseau VMware ESXi qui sont externes à l'hôte hyperviseur :

- vSAN (SAN virtuel)
- VMFS
- NFS

magasins de données vSAN

Chaque hôte ESXi peut disposer d'un ou plusieurs datastores VMFS locaux. Normalement, ces datastores ne sont accessibles qu'à l'hôte local. Cependant, VMware vSAN permet à chaque hôte d'un cluster ESXi de partager tous les datastores du cluster comme s'ils étaient locaux. La figure suivante illustre comment vSAN crée un pool de datastores partagés entre les hôtes du cluster ESXi.

ESXi cluster

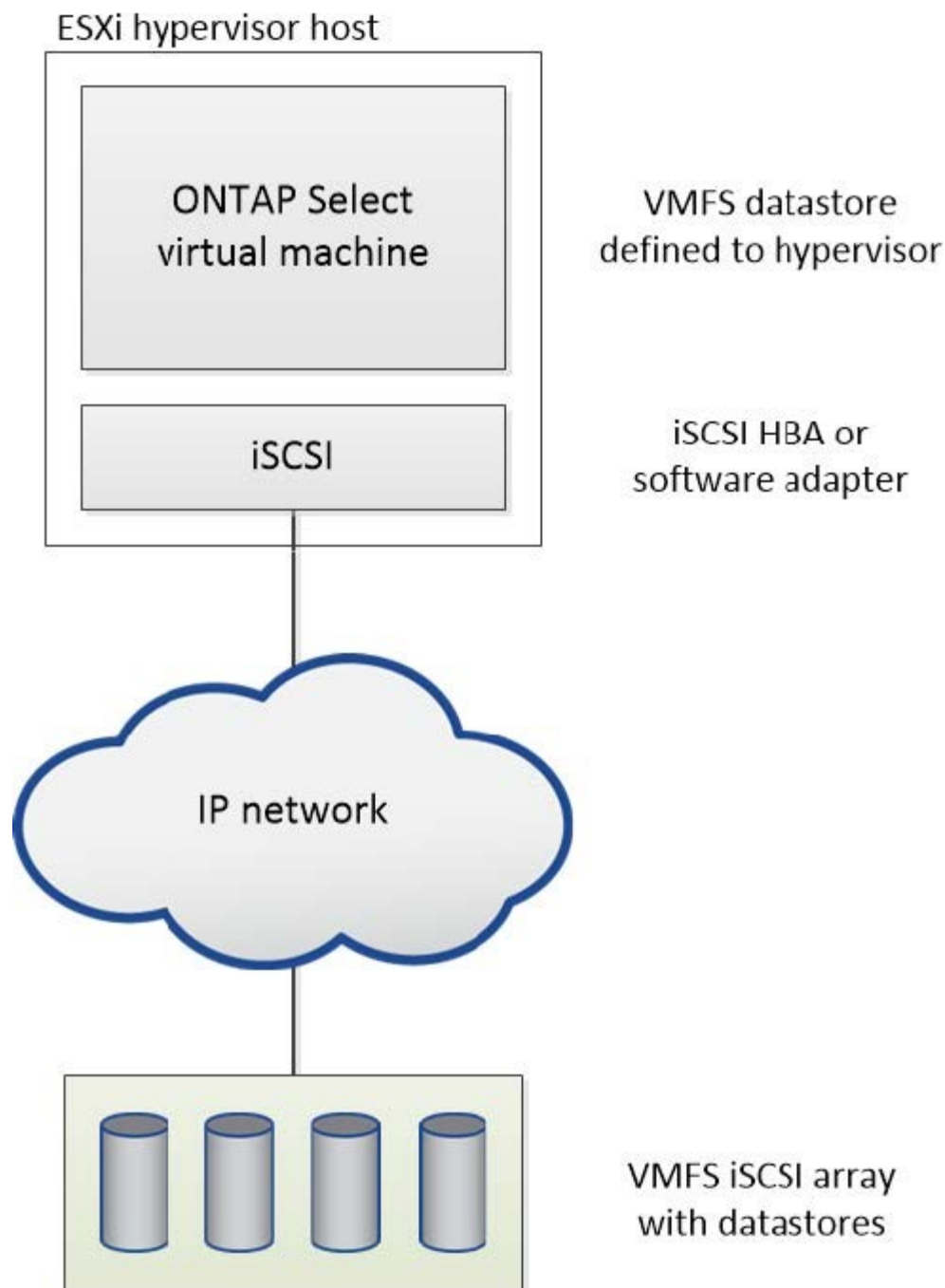


Banque de données VMFS sur une baie de stockage externe

Vous pouvez créer une banque de données VMFS résidant sur une baie de stockage externe. L'accès au stockage s'effectue via différents protocoles réseau. La figure suivante illustre une banque de données VMFS sur une baie de stockage externe accessible via le protocole iSCSI.

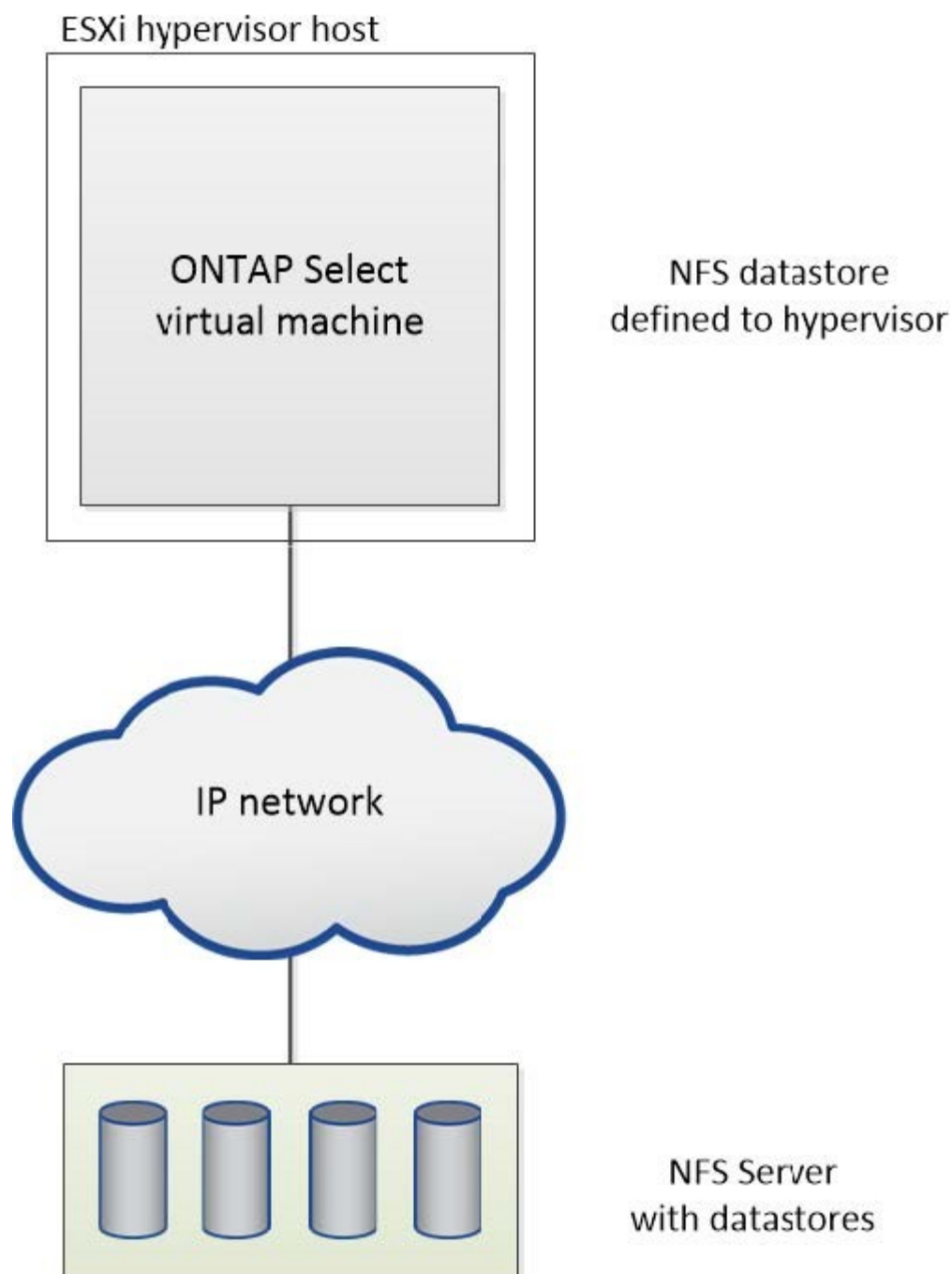


ONTAP Select prend en charge toutes les baies de stockage externes décrites dans la documentation de compatibilité VMware Storage/SAN, y compris iSCSI, Fiber Channel et Fiber Channel sur Ethernet.



Banque de données NFS sur une baie de stockage externe

Vous pouvez créer une banque de données NFS résidant sur une baie de stockage externe. L'accès au stockage s'effectue via le protocole réseau NFS. La figure suivante illustre une banque de données NFS sur un stockage externe accessible via le serveur NFS.



Services RAID matériels pour le stockage local connecté ONTAP Select

Lorsqu'un contrôleur RAID matériel est disponible, ONTAP Select peut déplacer les services RAID vers ce contrôleur afin d'optimiser les performances d'écriture et de protéger contre les pannes de disques physiques. Ainsi, la protection RAID de tous les nœuds du cluster ONTAP Select est assurée par le contrôleur RAID connecté localement, et non par le RAID logiciel ONTAP .



Les agrégats de données ONTAP Select sont configurés pour utiliser le niveau RAID 0, car le contrôleur RAID physique assure l'agrégation RAID sur les disques sous-jacents. Aucun autre niveau RAID n'est pris en charge.

Configuration du contrôleur RAID pour le stockage local connecté

Tous les disques connectés localement qui fournissent un stockage de secours à ONTAP Select doivent être placés derrière un contrôleur RAID. La plupart des serveurs grand public proposent plusieurs options de contrôleur RAID, à différents prix et avec des niveaux de fonctionnalités variés. L'objectif est de prendre en charge le plus grand nombre possible de ces options, à condition qu'elles répondent à certaines exigences minimales imposées au contrôleur.



Vous ne pouvez pas détacher de disques virtuels des machines virtuelles ONTAP Select utilisant la configuration RAID matérielle. La déconnexion de disques est uniquement prise en charge pour les machines virtuelles ONTAP Select utilisant la configuration RAID logicielle. Voir ["Remplacer un disque défectueux dans une configuration RAID logicielle ONTAP Select"](#) pour plus d'informations.

Le contrôleur RAID qui gère les disques ONTAP Select doit répondre aux exigences suivantes :

- Le contrôleur RAID matériel doit disposer d'une unité de secours sur batterie (BBU) ou d'un cache d'écriture flash (FBWC) et prendre en charge un débit de 12 Gbit/s.
- Le contrôleur RAID doit prendre en charge un mode capable de résister à au moins une ou deux pannes de disque (RAID 5 et RAID 6).
- Le cache du lecteur doit être désactivé.
- La politique d'écriture doit être configurée pour le mode d'écriture différée avec un repli vers l'écriture en cas de panne de la batterie de secours ou du flash.
- La politique d'E/S pour les lectures doit être définie sur mise en cache.

Tous les disques connectés localement qui fournissent un stockage de sauvegarde à ONTAP Select doivent être placés dans des groupes RAID exécutant RAID 5 ou RAID 6. Pour les disques SAS et SSD, l'utilisation de groupes RAID jusqu'à 24 disques permet à ONTAP de bénéficier de la répartition des requêtes de lecture entrantes sur un plus grand nombre de disques. Cela offre un gain de performances significatif. Avec les configurations SAS/SSD, des tests de performances ont été effectués sur des configurations à LUN unique et à LUN multiples. Aucune différence significative n'a été constatée ; par souci de simplicité, NetApp recommande donc de créer le moins de LUN possible pour répondre à vos besoins de configuration.

Les disques NL-SAS et SATA requièrent des bonnes pratiques différentes. Pour des raisons de performances, le nombre minimum de disques est toujours de huit, mais la taille du groupe RAID ne doit pas dépasser 12 disques. NetApp recommande également d'utiliser un disque de secours par groupe RAID ; toutefois, des disques de secours globaux peuvent être utilisés pour tous les groupes RAID. Par exemple, vous pouvez utiliser deux disques de secours pour trois groupes RAID, chaque groupe RAID étant composé de huit à douze disques.



La taille maximale de l'étendue et de la banque de données pour les anciennes versions d'ESX est de 64 To, ce qui peut affecter le nombre de LUN nécessaires pour prendre en charge la capacité brute totale fournie par ces disques de grande capacité.

Mode RAID

De nombreux contrôleurs RAID prennent en charge jusqu'à trois modes de fonctionnement, chacun représentant une différence significative dans le chemin emprunté par les requêtes d'écriture. Ces trois modes sont les suivants :

- Écriture directe. Toutes les requêtes d'E/S entrantes sont écrites dans le cache du contrôleur RAID, puis immédiatement vidées sur le disque avant de renvoyer la requête à l'hôte.

- Écriture de contournement. Toutes les requêtes d'E/S entrantes sont écrites directement sur le disque, contournant ainsi le cache du contrôleur RAID.
- Écriture différée. Toutes les requêtes d'E/S entrantes sont écrites directement dans le cache du contrôleur et immédiatement acquittées auprès de l'hôte. Les blocs de données sont vidés sur le disque de manière asynchrone via le contrôleur.

Le mode d'écriture différée offre le chemin de données le plus court, l'accusé de réception des E/S se produisant immédiatement après l'entrée des blocs en cache. Ce mode offre la latence la plus faible et le débit le plus élevé pour les charges de travail mixtes en lecture/écriture. Cependant, sans batterie de secours ni technologie flash non volatile, les utilisateurs courent le risque de perdre des données en cas de panne de courant du système.

ONTAP Select nécessite la présence d'une batterie de secours ou d'une unité flash ; nous pouvons donc être sûrs que les blocs mis en cache sont vidés sur le disque en cas de panne de ce type. C'est pourquoi le contrôleur RAID doit être configuré en mode écriture différée.

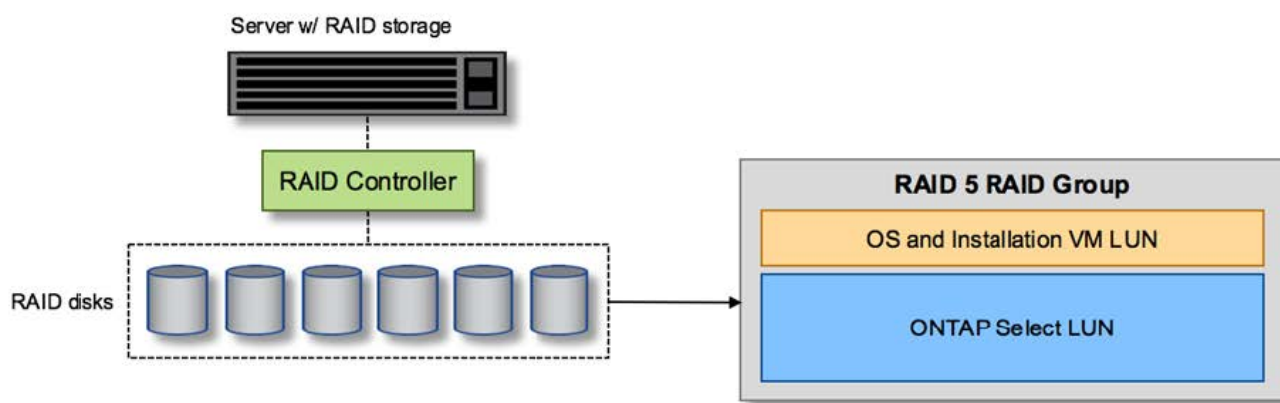
Disques locaux partagés entre ONTAP Select et le système d'exploitation

La configuration de serveur la plus courante est celle où tous les disques connectés localement sont placés derrière un seul contrôleur RAID. Vous devez provisionner au moins deux LUN : un pour l'hyperviseur et un pour la machine virtuelle ONTAP Select .

Prenons l'exemple d'un HP DL380 g8 équipé de six disques internes et d'un seul contrôleur RAID Smart Array P420i. Tous les disques internes sont gérés par ce contrôleur RAID, et aucun autre stockage n'est présent sur le système.

La figure suivante illustre ce type de configuration. Dans cet exemple, aucun autre stockage n'est présent sur le système ; par conséquent, l'hyperviseur doit partager le stockage avec le nœud ONTAP Select .

Configuration LUN du serveur avec uniquement des broches gérées par RAID



Le provisionnement des LUN du système d'exploitation à partir du même groupe RAID qu'ONTAP ONTAP Select permet au système d'exploitation de l'hyperviseur (et à toute machine virtuelle cliente également provisionnée à partir de ce stockage) de bénéficier de la protection RAID. Cette configuration empêche la panne d'un seul disque de mettre hors service l'ensemble du système.

Disques locaux répartis entre ONTAP Select et le système d'exploitation

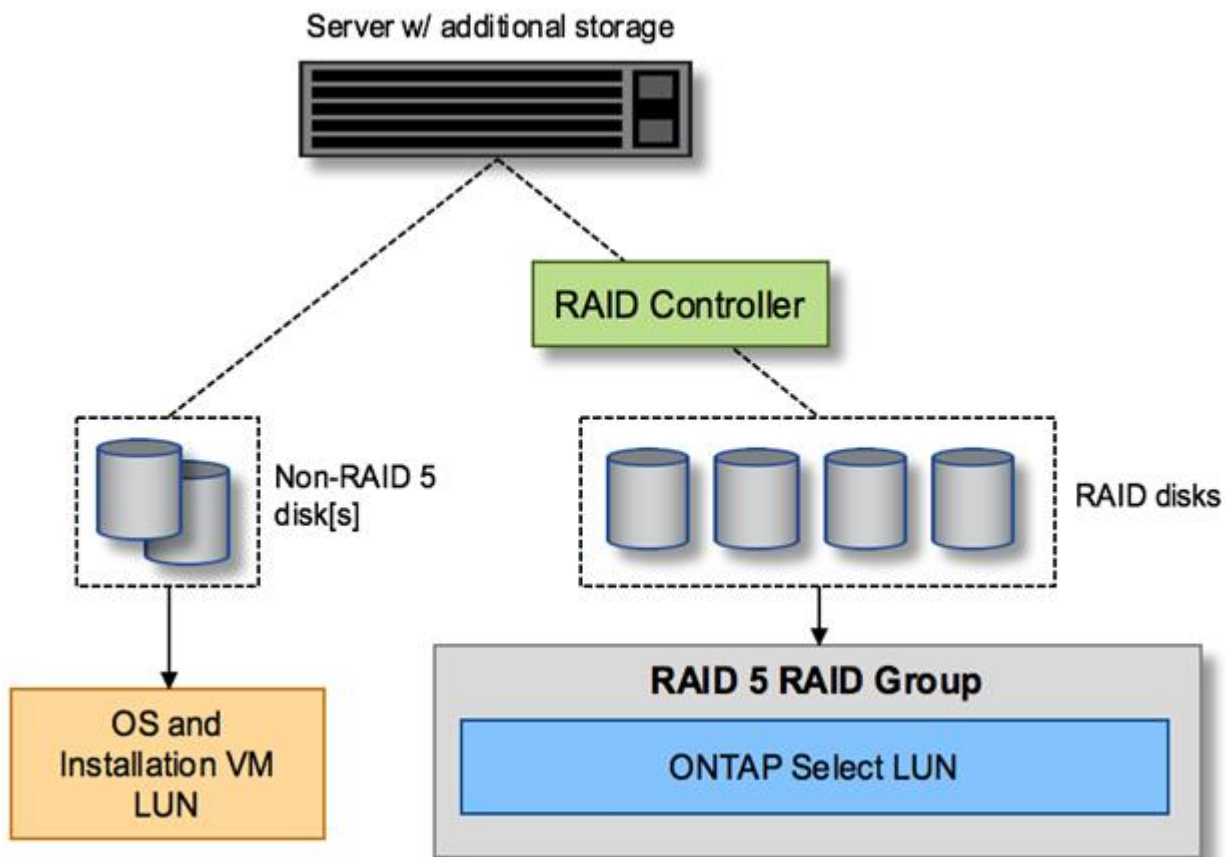
L'autre configuration possible proposée par les fournisseurs de serveurs consiste à configurer le système avec plusieurs contrôleurs RAID ou de disques. Dans cette configuration, un ensemble de disques est géré par un

contrôleur de disques, qui peut ou non offrir des services RAID. Un second ensemble de disques est géré par un contrôleur RAID matériel capable d'offrir des services RAID 5/6.

Avec ce type de configuration, l'ensemble des axes de disques situés derrière le contrôleur RAID capable de fournir les services RAID 5/6 doit être utilisé exclusivement par la machine virtuelle ONTAP Select. Selon la capacité de stockage totale gérée, vous devez configurer les axes de disques en un ou plusieurs groupes RAID et un ou plusieurs LUN. Ces LUN seront ensuite utilisés pour créer un ou plusieurs magasins de données, tous protégés par le contrôleur RAID.

Le premier ensemble de disques est réservé au système d'exploitation de l'hyperviseur et à toute machine virtuelle cliente qui n'utilise pas le stockage ONTAP, comme illustré dans la figure suivante.

Configuration LUN du serveur sur un système mixte RAID/non RAID



Plusieurs LUN

Il existe deux cas où les configurations de groupe RAID unique/LUN unique doivent être modifiées. Avec des disques NL-SAS ou SATA, la taille du groupe RAID ne doit pas dépasser 12 disques. De plus, un LUN unique peut dépasser les limites de stockage de l'hyperviseur sous-jacent, soit la taille maximale de l'extension du système de fichiers, soit la taille maximale du pool de stockage total. Le stockage physique sous-jacent doit alors être divisé en plusieurs LUN pour permettre la création réussie du système de fichiers.

Limites du système de fichiers de la machine virtuelle VMware vSphere

La taille maximale d'un magasin de données sur certaines versions d'ESX est de 64 To.

Si un serveur dispose de plus de 64 To de stockage, il peut être nécessaire de provisionner plusieurs LUN, chacun d'une taille inférieure à 64 To. La création de plusieurs groupes RAID pour améliorer le temps de reconstruction RAID des disques SATA/NL-SAS entraîne également le provisionnement de plusieurs LUN.

Lorsque plusieurs LUN sont nécessaires, il est essentiel de s'assurer que ces LUN présentent des performances similaires et cohérentes. Ceci est particulièrement important si tous les LUN doivent être utilisés dans un seul agrégat ONTAP. Par ailleurs, si un sous-ensemble d'un ou plusieurs LUN présente un profil de performances nettement différent, nous recommandons fortement d'isoler ces LUN dans un agrégat ONTAP distinct.

Plusieurs extensions de système de fichiers peuvent être utilisées pour créer un seul datastore jusqu'à sa taille maximale. Pour limiter la capacité nécessitant une licence ONTAP Select, veuillez à spécifier une limite de capacité lors de l'installation du cluster. Cette fonctionnalité permet à ONTAP Select d'utiliser (et donc de nécessiter une licence) uniquement une partie de l'espace d'un datastore.

Alternativement, il est possible de commencer par créer un seul datastore sur un seul LUN. Si de l'espace supplémentaire nécessitant une licence ONTAP Select de plus grande capacité est nécessaire, cet espace peut être ajouté au même datastore sous forme d'extension, jusqu'à la taille maximale du datastore. Une fois cette taille maximale atteinte, de nouveaux datastores peuvent être créés et ajoutés à ONTAP Select. Les deux types d'extension de capacité sont pris en charge et peuvent être réalisés grâce à la fonctionnalité d'ajout de stockage ONTAP Deploy. Chaque nœud ONTAP Select peut être configuré pour prendre en charge jusqu'à 400 To de stockage. Le provisionnement de capacité à partir de plusieurs datastores nécessite un processus en deux étapes.

La création initiale du cluster permet de créer un cluster ONTAP Select consommant tout ou partie de l'espace du datastore initial. Une deuxième étape consiste à effectuer une ou plusieurs opérations d'ajout de capacité en utilisant des datastores supplémentaires jusqu'à atteindre la capacité totale souhaitée. Cette fonctionnalité est détaillée dans la section "[Augmenter la capacité de stockage](#)".



La surcharge VMFS est différente de zéro (voir "[VMware KB 1001618](#)"), et la tentative d'utilisation de l'intégralité de l'espace signalé comme libre par une banque de données a entraîné des erreurs parasites lors des opérations de création de cluster.

Une mémoire tampon de 2 % est laissée inutilisée dans chaque banque de données. Cet espace ne nécessite pas de licence de capacité, car il n'est pas utilisé par ONTAP Select. ONTAP Deploy calcule automatiquement le nombre exact de gigaoctets pour la mémoire tampon, tant qu'aucune limite de capacité n'est spécifiée. Si une limite de capacité est spécifiée, cette taille est appliquée en premier. Si la taille de la limite de capacité est inférieure à la taille de la mémoire tampon, la création du cluster échoue et affiche un message d'erreur spécifiant la taille maximale correcte pouvant servir de limite de capacité :

```
"InvalidPoolCapacitySize: Invalid capacity specified for storage pool
"ontap-select-storage-pool", Specified value: 34334204 GB. Available
(after leaving 2% overhead space): 30948"
```

VMFS 6 est pris en charge pour les nouvelles installations et comme cible d'une opération Storage vMotion d'une machine virtuelle ONTAP Deploy ou ONTAP Select existante.

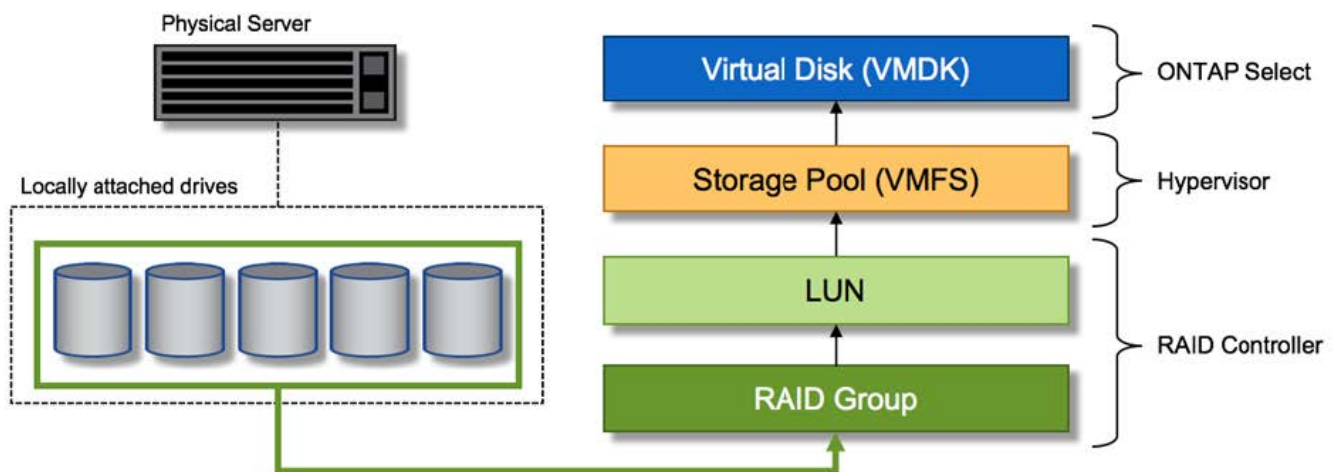
VMware ne prend pas en charge les mises à niveau sur place de VMFS 5 vers VMFS 6. Par conséquent, Storage vMotion est le seul mécanisme permettant à une machine virtuelle de passer d'une banque de données VMFS 5 à une banque de données VMFS 6. Cependant, la prise en charge de Storage vMotion avec ONTAP Select et ONTAP Deploy a été étendue pour couvrir d'autres scénarios que la transition de VMFS 5 vers VMFS 6.

ONTAP Select

ONTAP Select présente à ONTAP un ensemble de disques virtuels provisionnés à partir d'un ou plusieurs pools de stockage. ONTAP dispose d'un ensemble de disques virtuels qu'il traite comme physiques, tandis que le reste de la pile de stockage est abstrait par l'hyperviseur. La figure suivante illustre cette relation plus en détail, en mettant en évidence la relation entre le contrôleur RAID physique, l'hyperviseur et la machine virtuelle ONTAP Select .

- La configuration du groupe RAID et des LUN s'effectue depuis le logiciel du contrôleur RAID du serveur. Cette configuration n'est pas requise lors de l'utilisation de VSAN ou de baies externes.
- La configuration du pool de stockage s'effectue à partir de l'hyperviseur.
- Les disques virtuels sont créés et détenus par des machines virtuelles individuelles ; dans cet exemple, par ONTAP Select.

Mappage de disque virtuel vers disque physique



Provisionnement de disque virtuel

Pour une expérience utilisateur simplifiée, l'outil de gestion ONTAP Select , ONTAP Deploy, provisionne automatiquement les disques virtuels du pool de stockage associé et les associe à la machine virtuelle ONTAP Select . Cette opération s'effectue automatiquement lors de la configuration initiale et de l'ajout de stockage. Si le nœud ONTAP Select fait partie d'une paire HA, les disques virtuels sont automatiquement attribués à un pool de stockage local et miroir.

ONTAP Select divise le stockage sous-jacent en disques virtuels de taille égale, chacun ne dépassant pas 16 To. Si le nœud ONTAP Select fait partie d'une paire HA, au moins deux disques virtuels sont créés sur chaque nœud de cluster et affectés aux plex locaux et miroirs pour être utilisés dans un agrégat en miroir.

Par exemple, une ONTAP Select peut attribuer une banque de données ou un LUN de 31 To (l'espace restant après le déploiement de la machine virtuelle et le provisionnement des disques système et racine). Quatre disques virtuels d'environ 7,75 To sont ensuite créés et attribués au plex local et miroir ONTAP approprié.



L'ajout de capacité à une VM ONTAP Select génère probablement des VMDK de tailles différentes. Pour plus de détails, voir la section ["Augmenter la capacité de stockage"](#). Contrairement aux systèmes FAS , des VMDK de tailles différentes peuvent coexister dans le même agrégat. ONTAP Select utilise une bande RAID 0 sur ces VMDK, ce qui permet d'exploiter pleinement l'espace de chaque VMDK, quelle que soit sa taille

NVRAM virtualisée

Les systèmes NetApp FAS sont traditionnellement équipés d'une carte PCI NVRAM physique, une carte hautes performances contenant de la mémoire flash non volatile. Cette carte améliore considérablement les performances d'écriture en permettant à ONTAP d'accuser réception immédiate des écritures entrantes auprès du client. Elle peut également planifier le déplacement des blocs de données modifiés vers le support de stockage lent, grâce à un processus appelé « destaging ».

Les systèmes courants ne sont généralement pas équipés de ce type d'équipement. Par conséquent, la fonctionnalité de cette carte NVRAM a été virtualisée et placée dans une partition du disque de démarrage du système ONTAP Select . C'est pourquoi le placement du disque virtuel système de l'instance est extrêmement important. C'est également pourquoi le produit nécessite la présence d'un contrôleur RAID physique avec cache résilient pour les configurations de stockage local.

La NVRAM est placée sur son propre VMDK. Le fractionnement de la NVRAM dans son propre VMDK permet à la VM ONTAP Select d'utiliser le pilote vNVMe pour communiquer avec son VMDK NVRAM . Cela nécessite également que la VM ONTAP Select utilise la version matérielle 13, compatible avec ESX 6.5 et versions ultérieures.

Explication du chemin de données : NVRAM et contrôleur RAID

L'interaction entre la partition système NVRAM virtualisée et le contrôleur RAID peut être mieux mise en évidence en parcourant le chemin de données emprunté par une demande d'écriture lorsqu'elle entre dans le système.

Les requêtes d'écriture entrantes vers la machine virtuelle ONTAP Select ciblent la partition NVRAM de la machine virtuelle. Au niveau de la couche de virtualisation, cette partition se trouve dans un disque système ONTAP Select , un VMDK attaché à la machine virtuelle ONTAP Select . Au niveau de la couche physique, ces requêtes sont mises en cache dans le contrôleur RAID local, comme toutes les modifications de blocs ciblant les axes sous-jacents. À partir de là, l'écriture est confirmée à l'hôte.

À ce stade, le bloc réside physiquement dans le cache du contrôleur RAID, en attente d'être vidé sur le disque. Logiquement, il réside dans la NVRAM , en attente d'être transféré vers les disques de données utilisateur appropriés.

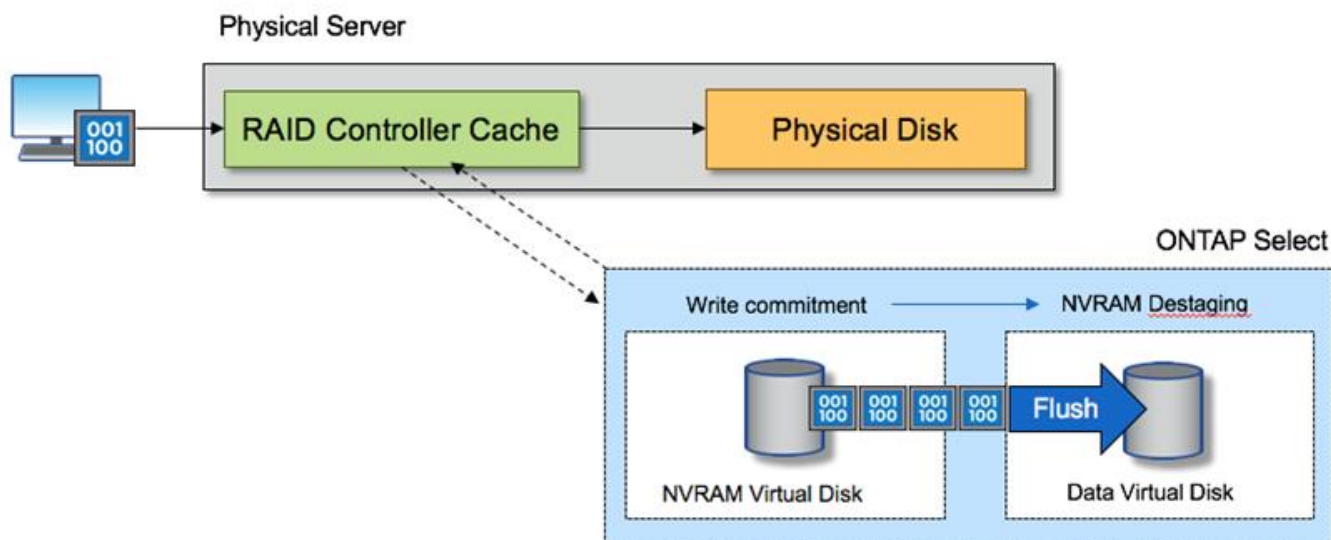
Les blocs modifiés étant automatiquement stockés dans le cache local du contrôleur RAID, les écritures entrantes sur la partition NVRAM sont automatiquement mises en cache et vidées périodiquement sur un support de stockage physique. Il ne faut pas confondre ce phénomène avec le vidage périodique du contenu de la NVRAM vers les disques de données ONTAP . Ces deux événements sont indépendants et se produisent à des moments et des fréquences différents.

La figure suivante illustre le chemin d'E/S emprunté par une écriture entrante. Elle met en évidence la différence entre la couche physique (représentée par le cache et les disques du contrôleur RAID) et la couche virtuelle (représentée par la NVRAM et les disques virtuels de données de la machine virtuelle).



Bien que les blocs modifiés sur la NVRAM VMDK soient mis en cache dans le cache du contrôleur RAID local, ce dernier ignore la structure de la VM ni ses disques virtuels. Il stocke tous les blocs modifiés sur le système, dont la NVRAM ne constitue qu'une partie. Cela inclut les requêtes d'écriture destinées à l'hyperviseur, si celui-ci est provisionné à partir des mêmes axes de sauvegarde.

*Écritures entrantes sur la machine ONTAP Select *



La partition NVRAM est séparée sur son propre VMDK. Ce VMDK est connecté via le pilote vNVME disponible dans les versions ESX 6.5 et ultérieures. Ce changement est particulièrement important pour les installations ONTAP Select avec RAID logiciel, qui ne bénéficient pas du cache du contrôleur RAID.

Services de configuration RAID logiciel ONTAP Select pour le stockage local connecté

Le RAID logiciel est une couche d'abstraction RAID implémentée au sein de la pile logicielle ONTAP . Il offre les mêmes fonctionnalités que la couche RAID d'une plateforme ONTAP traditionnelle telle que FAS. La couche RAID effectue les calculs de parité des disques et assure une protection contre les pannes de disques individuels au sein d'un nœud ONTAP Select .

Indépendamment des configurations RAID matérielles, ONTAP Select propose également une option RAID logiciel. Un contrôleur RAID matériel peut ne pas être disponible ou être indésirable dans certains environnements, par exemple lorsque ONTAP Select est déployé sur un matériel standard compact. Le RAID logiciel étend les options de déploiement disponibles à ces environnements. Pour activer le RAID logiciel dans votre environnement, voici quelques points à retenir :

- Il est disponible avec une licence Premium ou Premium XL.
- Il prend uniquement en charge les disques SSD ou NVMe (nécessite une licence Premium XL) pour les disques racine et de données ONTAP .
- Il nécessite un disque système distinct pour la partition de démarrage ONTAP Select VM.
 - Choisissez un disque séparé, soit un SSD, soit un lecteur NVMe, pour créer une banque de données pour les disques système (NVRAM, carte Boot/CF, Coredump et Mediator dans une configuration multi-nœuds).

Remarques

- Les termes disque de service et disque système sont utilisés de manière interchangeable.
 - Les disques de service sont les VMDK utilisés dans la machine virtuelle ONTAP Select pour gérer

divers éléments tels que le clustering, le démarrage, etc.

- Les disques de service sont physiquement situés sur un seul disque physique (appelé collectivement disque physique de service/système), vu depuis l'hôte. Ce disque physique doit contenir une banque de données DAS. ONTAP Deploy crée ces disques de service pour la machine virtuelle ONTAP Select lors du déploiement du cluster.
- Il n'est pas possible de séparer davantage les disques système ONTAP Select sur plusieurs banques de données ou sur plusieurs lecteurs physiques.
- Le RAID matériel n'est pas obsolète.

Configuration RAID logicielle pour le stockage local connecté

Lors de l'utilisation d'un RAID logiciel, l'absence de contrôleur RAID matériel est idéale, mais si un système dispose d'un contrôleur RAID existant, il doit respecter les exigences suivantes :

- Le contrôleur RAID matériel doit être désactivé pour que les disques puissent être présentés directement au système (JBOD). Cette modification peut généralement être effectuée dans le BIOS du contrôleur RAID.
- Le contrôleur RAID matériel doit également être en mode HBA SAS. Par exemple, certaines configurations du BIOS autorisent un mode « AHCI » en plus du RAID, qui pourrait être choisi pour activer le mode JBOD. Cela permet un transfert, permettant ainsi aux disques physiques d'être visibles tels quels sur l'hôte.

Selon le nombre maximal de disques pris en charge par le contrôleur, un contrôleur supplémentaire peut être nécessaire. Avec le mode HBA SAS, assurez-vous que le contrôleur d'E/S (HBA SAS) est pris en charge avec un débit minimal de 6 Gbit/s. Cependant, NetApp recommande un débit de 12 Gbit/s.

Aucun autre mode ou configuration de contrôleur RAID matériel n'est pris en charge. Par exemple, certains contrôleurs prennent en charge le RAID 0, ce qui permet artificiellement le passage des disques, mais les conséquences peuvent être indésirables. La taille prise en charge des disques physiques (SSD uniquement) est comprise entre 200 Go et 16 To.



Les administrateurs doivent suivre les lecteurs utilisés par la machine virtuelle ONTAP Select et empêcher l'utilisation accidentelle de ces lecteurs sur l'hôte.

ONTAP Select les disques virtuels et physiques

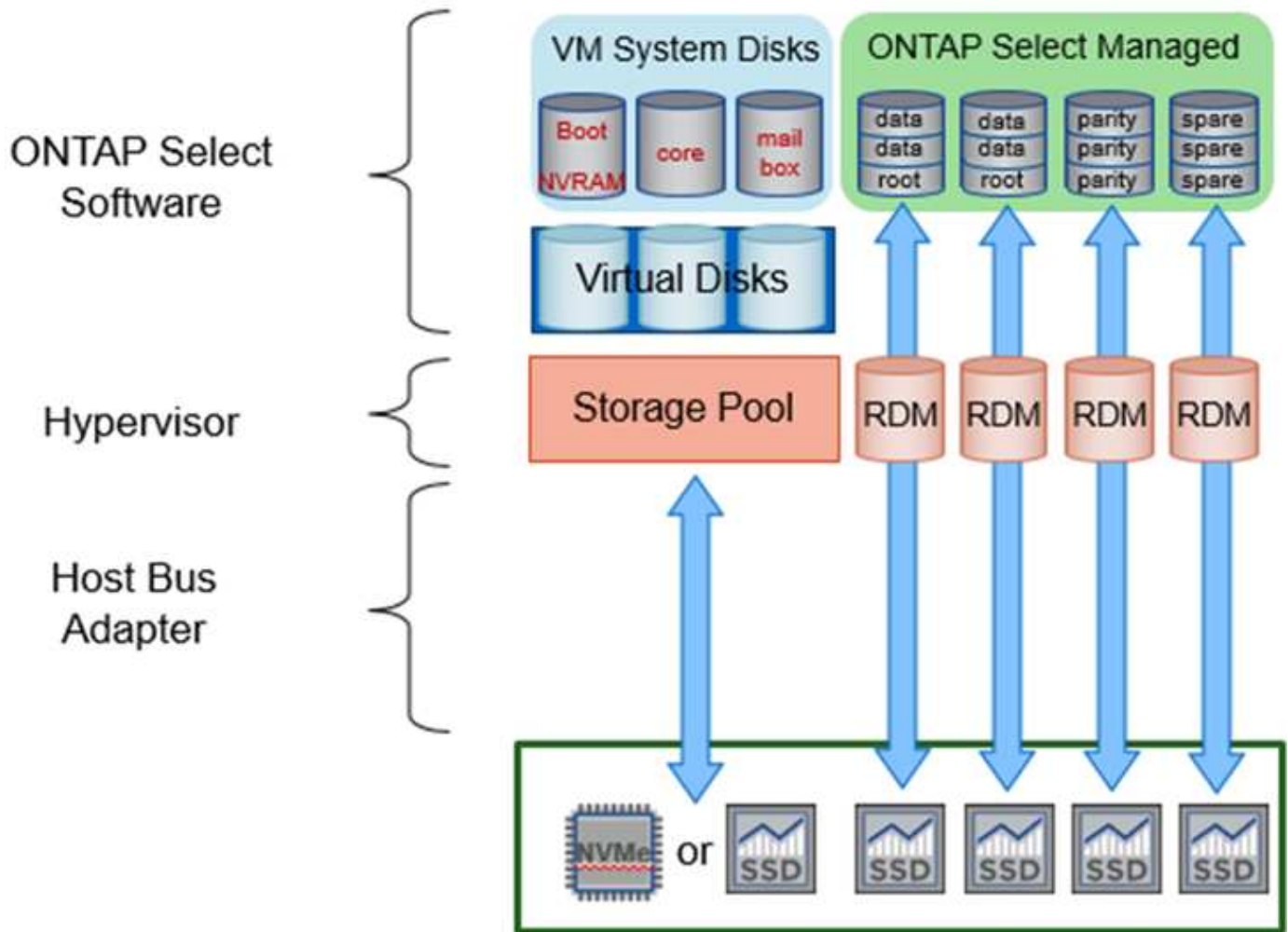
Pour les configurations avec contrôleurs RAID matériels, la redondance des disques physiques est assurée par le contrôleur RAID. ONTAP Select propose un ou plusieurs VMDK à partir desquels l'administrateur ONTAP peut configurer des agrégats de données. Ces VMDK sont répartis en RAID 0, car le RAID logiciel ONTAP est redondant et inefficace en raison de la résilience matérielle. De plus, les VMDK utilisés pour les disques système se trouvent dans le même magasin de données que ceux utilisés pour stocker les données utilisateur.

Lors de l'utilisation du RAID logiciel, ONTAP Deploy présente à ONTAP Select un ensemble de disques virtuels (VMDK) et de disques physiques, des mappages de périphériques bruts [RDM] pour les SSD et des périphériques d'E/S passthrough ou DirectPath pour les NVMe.

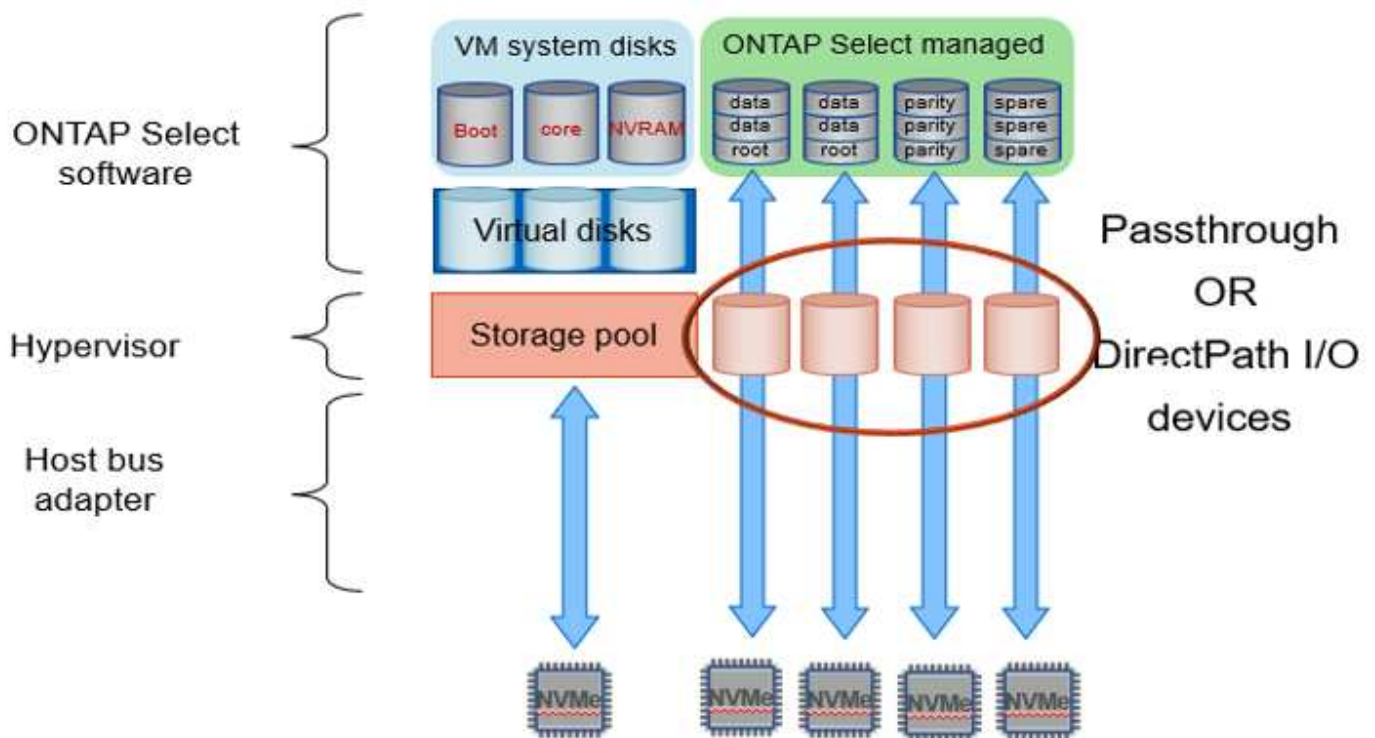
Les figures suivantes illustrent cette relation plus en détail, en soulignant la différence entre les disques virtualisés utilisés pour les composants internes de la machine ONTAP Select et les disques physiques utilisés pour stocker les données utilisateur.

- RAID logiciel ONTAP Select : utilisation de disques virtualisés et de RDM*

ONTAP Select with Software RAID



Les disques système (VMDK) résident dans le même magasin de données et sur le même disque physique. Le disque NVRAM virtuel nécessite un support rapide et durable. Par conséquent, seuls les magasins de données de type NVMe et SSD sont pris en charge.



Les disques système (VMDK) résident dans le même magasin de données et sur le même disque physique. Le disque NVRAM virtuel nécessite un support rapide et durable. Par conséquent, seuls les magasins de données de type NVMe et SSD sont pris en charge. Si vous utilisez des disques NVMe pour les données, le disque système doit également être un périphérique NVMe pour des raisons de performances. Une carte Intel Optane est un bon choix pour le disque système dans une configuration entièrement NVMe.

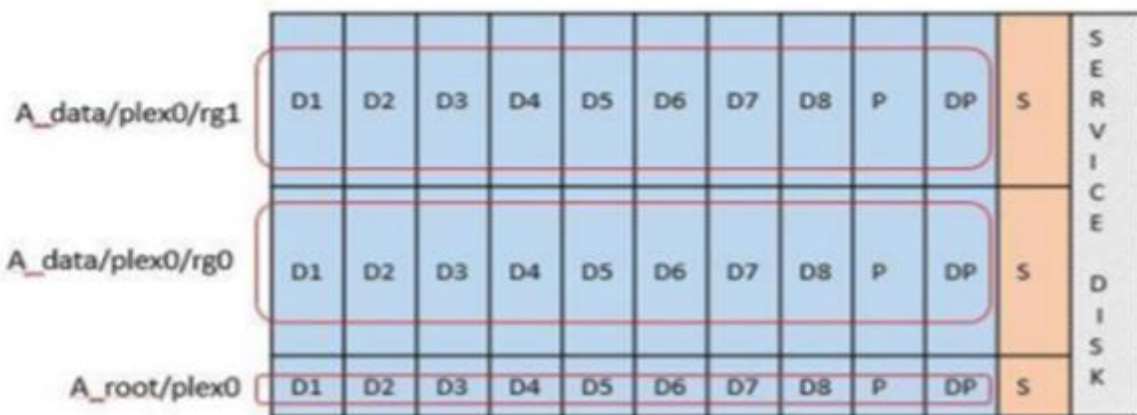


Avec la version actuelle, il n'est pas possible de séparer davantage les disques système ONTAP Select sur plusieurs banques de données ou plusieurs lecteurs physiques.

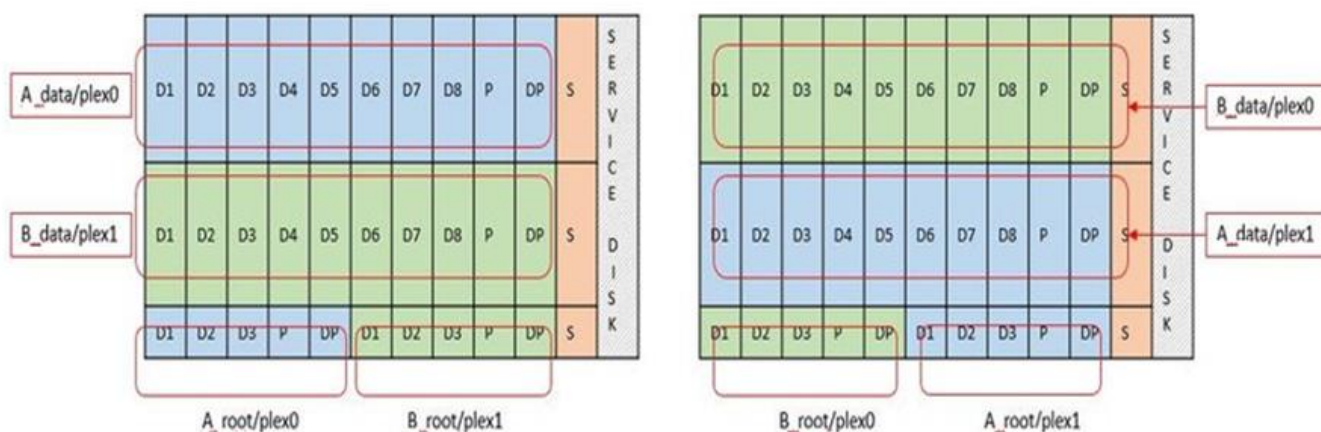
Chaque disque de données est divisé en trois parties : une petite partition racine (stripe) et deux partitions de taille égale pour créer deux disques de données visibles dans la machine virtuelle ONTAP Select. Les partitions utilisent le schéma Root Data Data (RD2) comme indiqué dans les figures suivantes pour un cluster à nœud unique et pour un nœud dans une paire HA.

P désigne un lecteur de parité. DP désigne un lecteur à double parité et S désigne un lecteur de secours.

Partitionnement de disque RDD pour les clusters à nœud unique



Partitionnement de disque RDD pour clusters multinœuds (paires HA)



Le RAID logiciel ONTAP prend en charge les types RAID suivants : RAID 4, RAID-DP et RAID-TEC. Il s'agit des mêmes structures RAID que celles utilisées par les plateformes FAS et AFF . Pour le provisionnement racine, ONTAP Select prend uniquement en charge RAID 4 et RAID-DP. Lorsque RAID-TEC est utilisé pour l'agrégation de données, la protection globale est RAID-DP. ONTAP Select HA utilise une architecture sans partage qui réplique la configuration de chaque nœud sur l'autre. Cela signifie que chaque nœud doit stocker sa partition racine et une copie de celle de son homologue. Étant donné qu'un disque de données possède une seule partition racine, le nombre minimal de disques de données varie selon que le nœud ONTAP Select fait partie ou non d'une paire HA.

Pour les clusters à nœud unique, toutes les partitions de données servent au stockage des données locales (actives). Pour les nœuds faisant partie d'une paire HA, une partition de données sert au stockage des données locales (actives) de ce nœud, tandis que la seconde partition de données sert à refléter les données actives de l'homologue HA.

Périphériques de transfert (DirectPath IO) vs. Cartes de périphériques bruts (RDM)

VMware ESX ne prend actuellement pas en charge les disques NVMe comme mappages de périphériques bruts. Pour ONTAP Select prenne directement le contrôle des disques NVMe, ces derniers doivent être configurés dans ESX comme périphériques de transfert. Veuillez noter que la configuration d'un périphérique NVMe comme périphérique de transfert nécessite la prise en charge du BIOS du serveur et constitue un processus perturbateur nécessitant un redémarrage de l'hôte ESX. De plus, le nombre maximal de

périphériques de transfert par hôte ESX est de 16. Cependant, ONTAP Deploy limite ce nombre à 14. Cette limite de 14 périphériques NVMe par nœud ONTAP Select signifie qu'une configuration 100 % NVMe offrira une densité d'IOPS très élevée (IOPS/To) au détriment de la capacité totale. Par ailleurs, si une configuration hautes performances avec une capacité de stockage plus importante est souhaitée, il est recommandé d'utiliser une VM ONTAP Select de grande taille, une carte Intel Optane pour le disque système et un nombre nominal de disques SSD pour le stockage des données.



Pour profiter pleinement des performances NVMe, tenez compte de la grande taille de la machine virtuelle ONTAP Select .

Il existe une différence supplémentaire entre les périphériques de transfert et les RDM. Les RDM peuvent être mappés à une machine virtuelle en cours d'exécution. Les périphériques de transfert nécessitent un redémarrage de la machine virtuelle. Cela signifie que tout remplacement de disque NVMe ou extension de capacité (ajout de disque) nécessite un redémarrage de la machine virtuelle ONTAP Select . Ces opérations sont gérées par un workflow dans ONTAP Deploy. ONTAP Deploy gère le redémarrage ONTAP Select pour les clusters à nœud unique et le basculement/la restauration automatique pour les paires haute disponibilité. Cependant, il est important de noter la différence entre le travail avec des lecteurs de données SSD (aucun redémarrage/basculement ONTAP Select n'est requis) et le travail avec des lecteurs de données NVMe (le redémarrage/basculement ONTAP Select est requis).

Provisionnement de disques physiques et virtuels

Pour une expérience utilisateur simplifiée, ONTAP Deploy provisionne automatiquement les disques système (virtuels) à partir du datastore spécifié (disque système physique) et les associe à la machine virtuelle ONTAP Select . Cette opération s'effectue automatiquement lors de la configuration initiale afin que la machine virtuelle ONTAP Select puisse démarrer. Les RDM sont partitionnés et l'agrégat racine est automatiquement créé. Si le nœud ONTAP Select fait partie d'une paire HA, les partitions de données sont automatiquement affectées à un pool de stockage local et à un pool de stockage miroir. Cette affectation s'effectue automatiquement lors des opérations de création de cluster et d'ajout de stockage.

Étant donné que les disques de données sur la machine virtuelle ONTAP Select sont associés aux disques physiques sous-jacents, la création de configurations avec un plus grand nombre de disques physiques a des conséquences sur les performances.



Le type de groupe RAID de l'agrégat racine dépend du nombre de disques disponibles. ONTAP Deploy sélectionne le type de groupe RAID approprié. Si le nombre de disques alloués au nœud est suffisant, il utilise RAID-DP ; sinon, il crée un agrégat racine RAID-4.

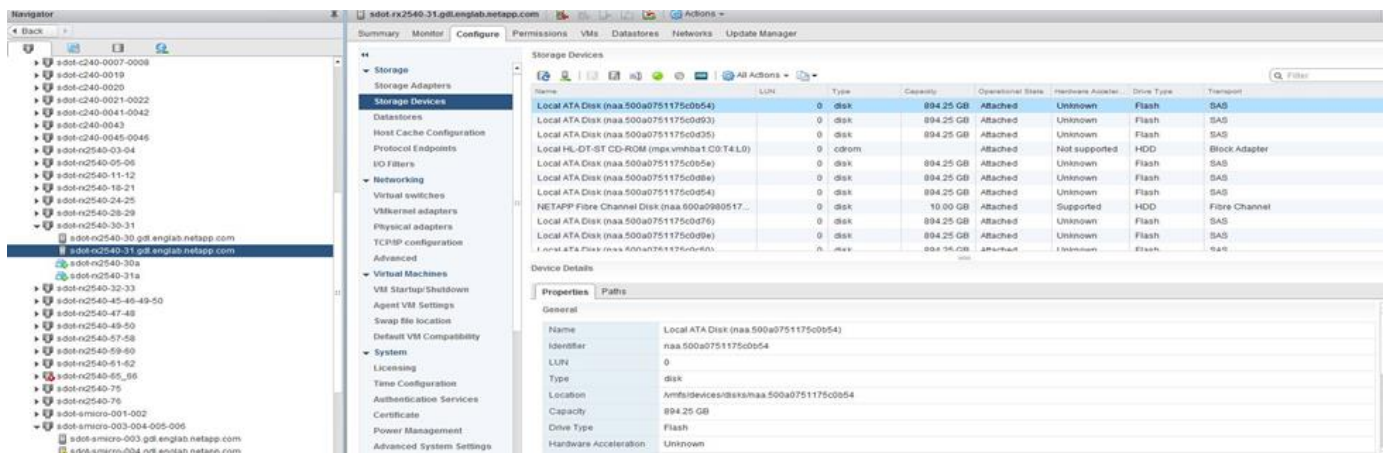
Lors de l'ajout de capacité à une machine virtuelle ONTAP Select via RAID logiciel, l'administrateur doit tenir compte de la taille du disque physique et du nombre de disques requis. Pour plus de détails, voir la section ["Augmenter la capacité de stockage"](#) .

Similaires aux systèmes FAS et AFF , seuls les disques de capacités égales ou supérieures peuvent être ajoutés à un groupe RAID existant. Les disques de plus grande capacité sont dimensionnés de manière optimale. Si vous créez de nouveaux groupes RAID, la taille du nouveau groupe RAID doit correspondre à celle du groupe RAID existant afin de garantir une performance globale optimale.

Associer un disque ONTAP Select au disque ESX correspondant

Les disques ONTAP Select sont généralement étiquetés NET xy. Vous pouvez utiliser la commande ONTAP suivante pour obtenir l'UUID du disque :

```
<system name>::> disk show NET-1.1
Disk: NET-1.1
Model: Micron_5100_MTFD
Serial Number: 1723175C0B5E
UID:
*500A0751:175C0B5E*:00000000:00000000:00000000:00000000:00000000:00000000:
00000000:00000000
BPS: 512
Physical Size: 894.3GB
Position: shared
Checksum Compatibility: advanced_zoned
Aggregate: -
Plex: -This UID can be matched with the device UID displayed in the
'storage devices' tab for the ESX host
```



Dans le shell ESXi, vous pouvez entrer la commande suivante pour faire clignoter la LED d'un disque physique donné (identifié par son naa.unique-id).

```
esxcli storage core device set -d <naa_id> -l=locator -L=<seconds>
```

Pannes de disques multiples lors de l'utilisation du RAID logiciel

Il est possible qu'un système soit confronté à une situation où plusieurs disques sont en panne simultanément. Le comportement du système dépend de la protection RAID globale et du nombre de disques défectueux.

Un agrégat RAID4 peut survivre à une panne de disque, un agrégat RAID-DP peut survivre à deux pannes de disque et un agrégat RAID-TEC peut survivre à trois pannes de disque.

Si le nombre de disques défectueux est inférieur au nombre maximal de pannes pris en charge par le type RAID et si un disque de secours est disponible, le processus de reconstruction démarre automatiquement. En l'absence de disques de secours, l'agrégat fournit les données dans un état dégradé jusqu'à l'ajout de disques de secours.

Si le nombre de disques défectueux est supérieur au nombre maximal de pannes pris en charge par le type RAID, le plex local est marqué comme défectueux et l'état de l'agrégat est dégradé. Les données sont servies

depuis le deuxième plex résidant sur le partenaire HA. Cela signifie que toutes les requêtes d'E/S du nœud 1 sont envoyées via le port d'interconnexion de cluster e0e (iSCSI) aux disques physiquement situés sur le nœud 2. Si le deuxième plex tombe également en panne, l'agrégat est marqué comme défaillant et les données sont indisponibles.

Un plex défaillant doit être supprimé et recréé pour que la mise en miroir correcte des données reprenne. Notez qu'une panne multidisque entraînant la dégradation d'un agrégat de données entraîne également la dégradation d'un agrégat racine. ONTAP Select utilise le schéma de partitionnement root-data-data (RDD) pour diviser chaque disque physique en une partition racine et deux partitions de données. Par conséquent, la perte d'un ou plusieurs disques peut affecter plusieurs agrégats, notamment la racine locale ou la copie de l'agrégat racine distant, ainsi que l'agrégat de données local et la copie de l'agrégat de données distant.

```
C3111E67::> storage aggregate plex delete -aggregate aggr1 -plex plex1
Warning: Deleting plex "plex1" of mirrored aggregate "aggr1" in a non-
shared HA configuration will disable its synchronous mirror protection and
disable
    negotiated takeover of node "sti-rx2540-335a" when aggregate
"aggr1" is online.
Do you want to continue? {y|n}: y
[Job 78] Job succeeded: DONE

C3111E67::> storage aggregate mirror -aggregate aggr1
Info: Disks would be added to aggregate "aggr1" on node "sti-rx2540-335a"
in the following manner:
    Second Plex
    RAID Group rg0, 5 disks (advanced_zoned checksum, raid_dp)
                                Usable
Physical
Size      Position  Disk                Type                Size
-----
-----
-          shared    NET-3.2             SSD                 -
-          shared    NET-3.3             SSD                 -
-          shared    NET-3.4             SSD                 208.4GB
208.4GB    shared    NET-3.5             SSD                 208.4GB
208.4GB    shared    NET-3.12            SSD                 208.4GB
208.4GB

    Aggregate capacity available for volume use would be 526.1GB.
    625.2GB would be used from capacity license.
Do you want to continue? {y|n}: y

C3111E67::> storage aggregate show-status -aggregate aggr1
```

Owner Node: sti-rx2540-335a

Aggregate: aggr1 (online, raid_dp, mirrored) (advanced_zoned checksums)

Plex: /aggr1/plex0 (online, normal, active, pool0)

RAID Group /aggr1/plex0/rg0 (normal, advanced_zoned checksums)

Usable

Physical

Position	Disk	Pool	Type	RPM	Size
----------	------	------	------	-----	------

Size Status

shared	NET-1.1	0	SSD	-	205.1GB
447.1GB (normal)					
shared	NET-1.2	0	SSD	-	205.1GB
447.1GB (normal)					
shared	NET-1.3	0	SSD	-	205.1GB
447.1GB (normal)					
shared	NET-1.10	0	SSD	-	205.1GB
447.1GB (normal)					
shared	NET-1.11	0	SSD	-	205.1GB
447.1GB (normal)					

Plex: /aggr1/plex3 (online, normal, active, pool1)

RAID Group /aggr1/plex3/rg0 (normal, advanced_zoned checksums)

Usable

Physical

Position	Disk	Pool	Type	RPM	Size
----------	------	------	------	-----	------

Size Status

shared	NET-3.2	1	SSD	-	205.1GB
447.1GB (normal)					
shared	NET-3.3	1	SSD	-	205.1GB
447.1GB (normal)					
shared	NET-3.4	1	SSD	-	205.1GB
447.1GB (normal)					
shared	NET-3.5	1	SSD	-	205.1GB
447.1GB (normal)					
shared	NET-3.12	1	SSD	-	205.1GB
447.1GB (normal)					

10 entries were displayed..



Afin de tester ou de simuler une ou plusieurs pannes de disque, utilisez le `storage disk fail -disk NET-x.y -immediate` Commande. Si le système dispose d'un disque de secours, l'agrégat commencera sa reconstruction. Vous pouvez vérifier l'état de la reconstruction à l'aide de la commande `storage aggregate show`. Vous pouvez supprimer le lecteur défaillant simulé à l'aide `ONTAP Deploy`. Notez qu'ONTAP a marqué le lecteur en tant que `Broken`. Le lecteur n'est pas endommagé et peut être réinstallé via `ONTAP Deploy`. Pour effacer l'étiquette « Cassé », saisissez les commandes suivantes dans l'interface de ligne de commande `ONTAP Select`

```
set advanced
disk unfail -disk NET-x.y -spare true
disk show -broken
```

La sortie de la dernière commande doit être vide.

NVRAM virtualisée

Les systèmes NetApp FAS sont traditionnellement équipés d'une carte PCI NVRAM physique. Cette carte hautes performances, dotée d'une mémoire flash non volatile, améliore considérablement les performances d'écriture. Elle permet à ONTAP d'accuser réception immédiate des écritures entrantes auprès du client. Elle peut également planifier le déplacement des blocs de données modifiés vers des supports de stockage plus lents, grâce à un processus appelé « destaging ».

Les systèmes courants ne sont généralement pas équipés de ce type d'équipement. Par conséquent, la fonctionnalité de la carte NVRAM a été virtualisée et placée dans une partition du disque de démarrage du système `ONTAP Select`. C'est pourquoi le placement du disque virtuel système de l'instance est extrêmement important.

ONTAP Select VSAN et configurations de baies externes

Les déploiements NAS virtuels (vNAS) prennent en charge les clusters `ONTAP Select` sur SAN virtuel (VSAN), certains produits HCI et les baies de stockage de données externes. L'infrastructure sous-jacente de ces configurations assure la résilience des banques de données.

L'exigence minimale est que la configuration sous-jacente soit prise en charge par VMware et doit être répertoriée dans les HCL VMware respectifs.

Architecture vNAS

La nomenclature vNAS est utilisée pour toutes les configurations qui n'utilisent pas DAS. Pour les clusters `ONTAP Select` multinœuds, cela inclut les architectures où les deux nœuds `ONTAP Select` d'une même paire HA partagent un même datastore (y compris les datastores vSAN). Les nœuds peuvent également être installés sur des datastores distincts à partir de la même baie externe partagée. Cela permet d'optimiser l'efficacité du stockage côté baie et de réduire l'empreinte globale de l'ensemble de la paire HA `ONTAP Select`. L'architecture des solutions vNAS `ONTAP Select` est très similaire à celle d' `ONTAP Select` sur DAS avec un contrôleur RAID local. Autrement dit, chaque nœud `ONTAP Select` conserve une copie des données de son partenaire HA. Les politiques d'efficacité du stockage `ONTAP` sont définies par nœud. Par conséquent, l'optimisation de l'efficacité du stockage côté baie est préférable, car elle peut potentiellement s'appliquer aux ensembles de données des deux nœuds `ONTAP Select`.

Il est également possible que chaque nœud ONTAP Select d'une paire HA utilise une baie externe distincte. C'est un choix courant lors de l'utilisation ONTAP Select Metrocluster SDS avec stockage externe.

Lorsque vous utilisez des baies externes distinctes pour chaque nœud ONTAP Select, il est très important que les deux baies fournissent des caractéristiques de performances similaires à la machine virtuelle ONTAP Select.

Architectures vNAS versus DAS local avec contrôleurs RAID matériels

L'architecture vNAS est logiquement très similaire à celle d'un serveur avec DAS et contrôleur RAID. Dans les deux cas, ONTAP Select consomme de l'espace de stockage. Cet espace est découpé en VMDK, qui constituent les agrégats de données ONTAP traditionnels. ONTAP Deploy s'assure que les VMDK sont correctement dimensionnés et affectés au bon plex (dans le cas de paires HA) lors des opérations de création de cluster et d'ajout de stockage.

Il existe deux différences majeures entre un vNAS et un DAS avec contrôleur RAID. La principale différence réside dans le fait que le vNAS ne nécessite pas de contrôleur RAID. Il suppose que la baie externe sous-jacente assure la persistance et la résilience des données qu'un DAS avec contrôleur RAID offrirait. La deuxième différence, plus subtile, concerne les performances de la NVRAM.

NVRAM vNAS

La NVRAM ONTAP Select est un VMDK. Autrement dit, ONTAP Select émule un espace adressable par octet (NVRAM traditionnelle) sur un périphérique adressable par bloc (VMDK). Cependant, les performances de la NVRAM sont essentielles aux performances globales du nœud ONTAP Select.

Pour les configurations DAS avec un contrôleur RAID matériel, le cache du contrôleur RAID matériel agit comme cache NVRAM de facto, car toutes les écritures dans le VMDK NVRAM sont d'abord hébergées dans le cache du contrôleur RAID.

Pour les architectures vNAS, ONTAP Deploy configure automatiquement les nœuds ONTAP Select avec un argument de démarrage appelé Single Instance Data Logging (SIDL). Lorsque cet argument est présent, ONTAP Select contourne la NVRAM et écrit la charge utile de données directement dans l'agrégat de données. La NVRAM sert uniquement à enregistrer l'adresse des blocs modifiés par l'opération d'écriture (WRITE). L'avantage de cette fonctionnalité est d'éviter une double écriture : une écriture dans la NVRAM et une seconde lors de la désactivation de la NVRAM. Cette fonctionnalité est uniquement activée pour les vNAS, car les écritures locales dans le cache du contrôleur RAID présentent une latence supplémentaire négligeable.

La fonctionnalité SIDL n'est pas compatible avec toutes les fonctionnalités d'optimisation du stockage ONTAP Select. Elle peut être désactivée au niveau agrégé à l'aide de la commande suivante :

```
storage aggregate modify -aggregate aggr-name -single-instance-data
-logging off
```

Notez que les performances d'écriture sont affectées si la fonctionnalité SIDL est désactivée. Il est possible de réactiver la fonctionnalité SIDL après avoir désactivé toutes les politiques d'efficacité de stockage sur tous les volumes de cet agrégat :

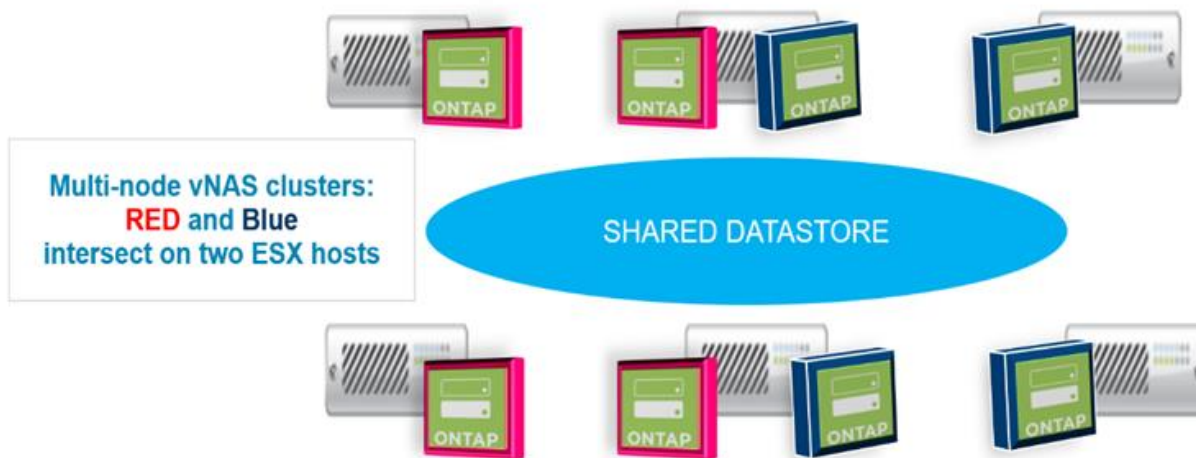
```
volume efficiency stop -all true -vserver * -volume * (all volumes in the
affected aggregate)
```

Colocaliser les nœuds ONTAP Select lors de l'utilisation de vNAS sur ESXi

ONTAP Select prend en charge les clusters ONTAP Select multinœuds sur un stockage partagé. ONTAP Deploy permet la configuration de plusieurs nœuds ONTAP Select sur le même hôte ESX, à condition qu'ils ne fassent pas partie du même cluster. Notez que cette configuration est uniquement valable pour les environnements VNAS (datastores partagés). L'utilisation de plusieurs instances ONTAP Select par hôte n'est pas prise en charge avec le stockage DAS, car elles sont en concurrence pour le même contrôleur RAID matériel.

ONTAP Deploy garantit que le déploiement initial du cluster VNAS multinœud ne place pas plusieurs instances ONTAP Select du même cluster sur le même hôte. L'illustration suivante illustre un exemple de déploiement correct de deux clusters à quatre nœuds qui se croisent sur deux hôtes.

Déploiement initial de clusters VNAS multinœuds



Après le déploiement, les nœuds ONTAP Select peuvent être migrés entre les hôtes. Cela pourrait entraîner des configurations non optimales et non prises en charge, où deux ou plusieurs nœuds ONTAP Select d'un même cluster partagent le même hôte sous-jacent. NetApp recommande la création manuelle de règles d'anti-affinité de VM afin que VMware maintienne automatiquement la séparation physique entre les nœuds d'un même cluster, et pas seulement entre les nœuds d'une même paire HA.



Les règles anti-affinité nécessitent que DRS soit activé sur le cluster ESX.

Consultez l'exemple suivant pour savoir comment créer une règle d'anti-affinité pour les machines virtuelles ONTAP Select. Si le cluster ONTAP Select contient plusieurs paires HA, tous les nœuds du cluster doivent être inclus dans cette règle.

Getting StartedSummaryMonitorConfigurePermissionsHostsVMsDatastoresNetworksUpdate Manager

◀

Services

vSphere DRS

vSphere Availability

vSAN

General

Disk Management

Fault Domains & Stretched Cluster

Health and Performance

iSCSI Targets

iSCSI Initiator Groups

Configuration Assist

Updates

Configuration

General

Licensing

VMware EVC

VM/Host Groups

VM/Host Rules

VM Overrides

Host Options

Profiles

I/O Filters

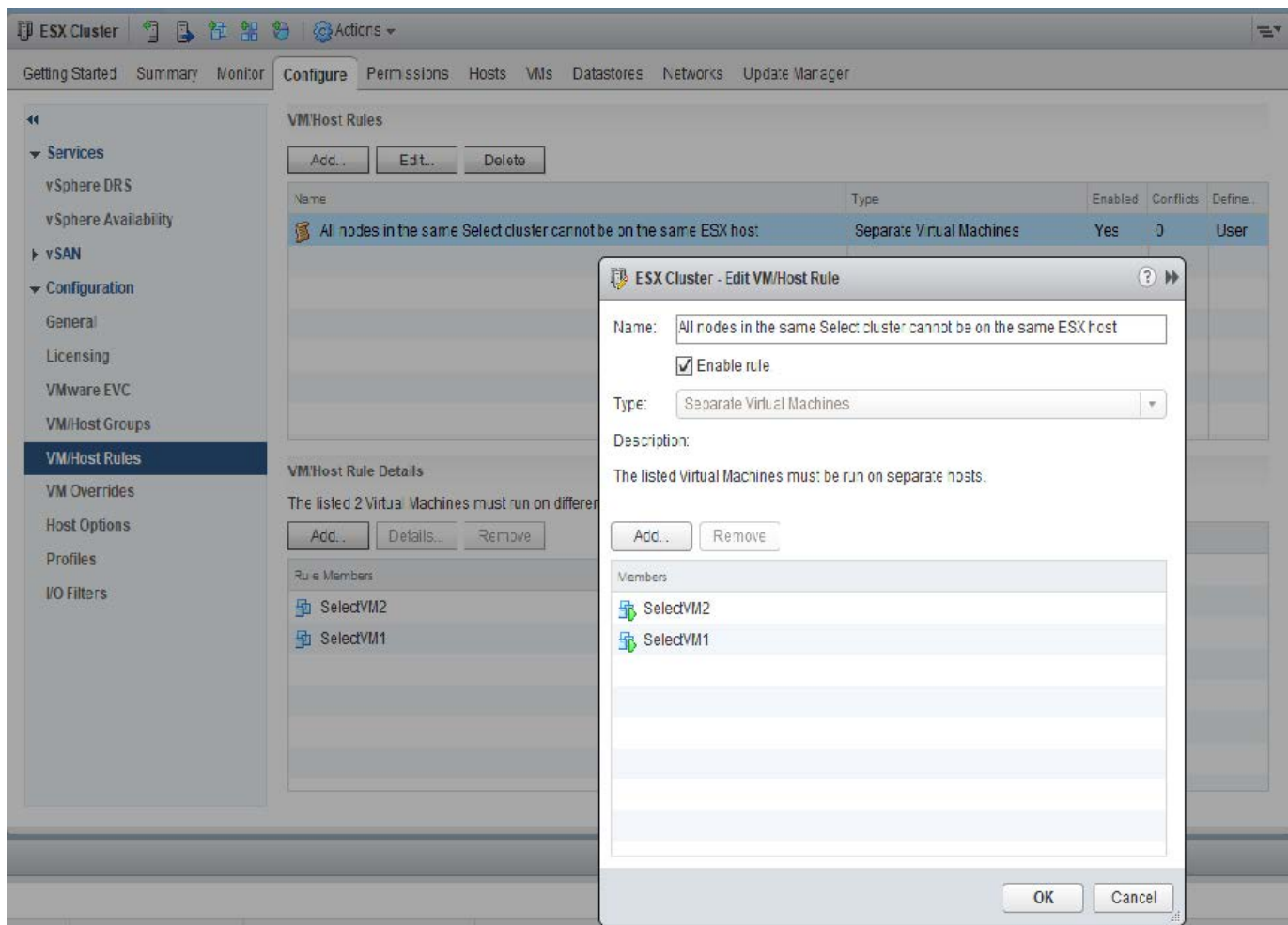
VM/Host Rules

AddEditDelete

Name	Type	Enabled	Conflicts	Defined By
This list is empty.				

No VM/Host rule selected

25



Deux ou plusieurs nœuds ONTAP Select du même cluster ONTAP Select peuvent potentiellement être trouvés sur le même hôte ESX pour l'une des raisons suivantes :

- DRS n'est pas présent en raison des limitations de licence VMware vSphere ou si DRS n'est pas activé.
- La règle anti-affinité DRS est contournée car une opération VMware HA ou une migration de machine virtuelle initiée par l'administrateur est prioritaire.

Notez ONTAP Deploy ne surveille pas proactivement les emplacements des machines virtuelles ONTAP Select . Cependant, une actualisation du cluster reflète cette configuration non prise en charge dans les journaux d' ONTAP Deploy :



Augmenter la capacité de stockage ONTAP Select

ONTAP Deploy peut être utilisé pour ajouter et octroyer une licence de stockage supplémentaire pour chaque nœud d'un cluster ONTAP Select .

La fonctionnalité d'ajout de stockage dans ONTAP Deploy est le seul moyen d'augmenter la capacité de stockage gérée. La modification directe de la VM ONTAP Select n'est pas prise en charge. L'illustration suivante montre l'icône « + » qui lance l'assistant d'ajout de stockage.

Cluster Details	
Name	onenode95IP15
ONTAP Image Version	9.5RC1
IPv4 Address	10.193.83.15
Netmask	255.255.255.128
Gateway	10.193.83.1
Last Refresh	-
Cluster Size	Single node cluster
Licensing	licensed
Domain Names	-
Server IP Addresses	-
NTP Server	216.239.35.0
Node Details	
Node	
Node	onenode95IP15-01 — 1.3 TB + # Host 10.193.39.54 — (Small (4 CPU, 16 GB Memory))

Les considérations suivantes sont importantes pour la réussite de l'extension de capacité. L'ajout de capacité nécessite que la licence existante couvre la quantité totale d'espace (existant et nouveau). Une opération d'ajout de stockage entraînant un dépassement de la capacité autorisée du nœud échoue. Une nouvelle licence d'une capacité suffisante doit être installée en premier.

Si la capacité supplémentaire est ajoutée à un agrégat ONTAP Select existant, le nouveau pool de stockage (datastore) doit présenter un profil de performances similaire à celui du pool de stockage existant. Notez qu'il est impossible d'ajouter du stockage non SSD à un nœud ONTAP Select installé avec une personnalité de type AFF(Flash activé). La combinaison de stockage DAS et de stockage externe n'est pas non plus prise en charge.

Si un stockage local est ajouté à un système pour fournir des pools de stockage locaux (DAS) supplémentaires, vous devez créer un groupe RAID et un ou plusieurs LUN supplémentaires. Comme pour les systèmes FAS, veillez à ce que les performances du nouveau groupe RAID soient similaires à celles du groupe RAID d'origine si vous ajoutez de l'espace au même agrégat. Si vous créez un nouvel agrégat, la configuration du nouveau groupe RAID peut être différente si les implications en termes de performances sont bien comprises.

Le nouvel espace peut être ajouté à ce même magasin de données en tant qu'extension si sa taille totale ne dépasse pas la taille maximale prise en charge. L'ajout d'une extension à un magasin de données sur lequel ONTAP Select est déjà installé peut être effectué dynamiquement et n'affecte pas le fonctionnement du nœud ONTAP Select.

Si le nœud ONTAP Select fait partie d'une paire HA, certains problèmes supplémentaires doivent être pris en compte.

Dans une paire HA, chaque nœud contient une copie miroir des données de son partenaire. L'ajout d'espace au nœud 1 nécessite l'ajout d'une quantité d'espace identique à son partenaire, le nœud 2, afin que toutes les données du nœud 1 soient répliquées sur le nœud 2. Autrement dit, l'espace ajouté au nœud 2 lors de l'opération d'ajout de capacité pour le nœud 1 n'est ni visible ni accessible sur le nœud 2. L'espace est ajouté au nœud 2 afin que les données du nœud 1 soient entièrement protégées lors d'un événement HA.

Un autre point à prendre en compte concerne les performances. Les données du nœud 1 sont répliquées de manière synchrone vers le nœud 2. Par conséquent, les performances du nouvel espace (magasin de données) du nœud 1 doivent être identiques à celles du nouvel espace (magasin de données) du nœud 2. Autrement dit, l'ajout d'espace sur les deux nœuds, mais avec des technologies de disque ou des tailles de groupe RAID différentes, peut entraîner des problèmes de performances. Cela est dû à l'opération RAID SyncMirror utilisée pour conserver une copie des données sur le nœud partenaire.

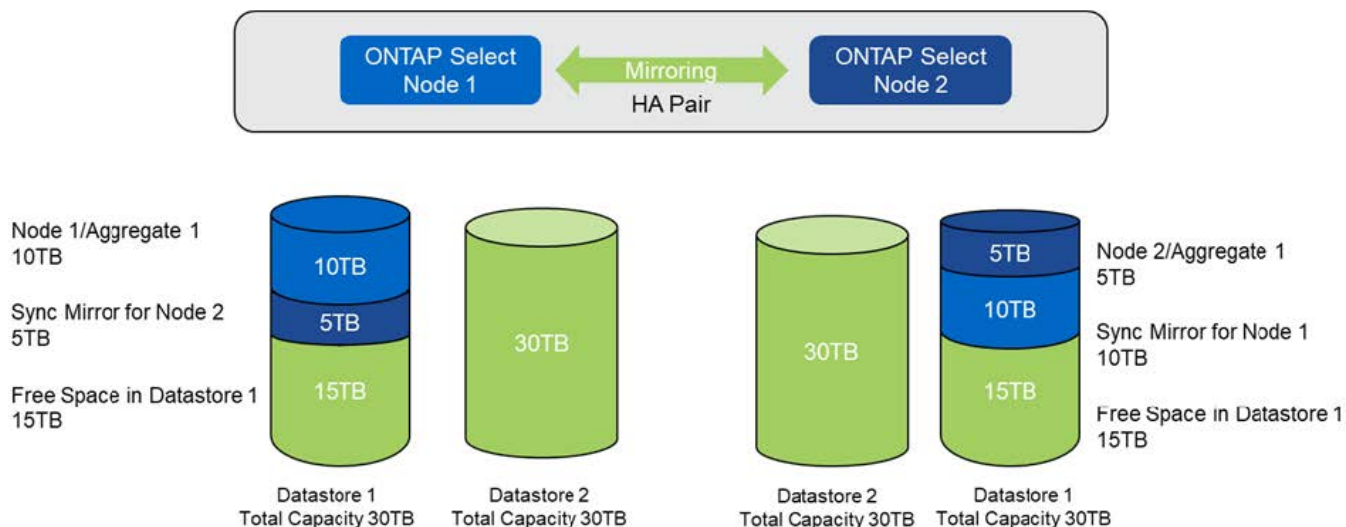
Pour augmenter la capacité accessible aux utilisateurs sur les deux nœuds d'une paire HA, deux ajouts de stockage doivent être effectués, un pour chaque nœud. Chaque ajout de stockage nécessite de l'espace

supplémentaire sur les deux nœuds. L'espace total requis sur chaque nœud est égal à l'espace requis sur le nœud 1 plus l'espace requis sur le nœud 2.

La configuration initiale comprend deux nœuds, chacun disposant de deux banques de données de 30 To chacune. ONTAP Deploy crée un cluster à deux nœuds, chaque nœud consommant 10 To d'espace de la banque de données 1. ONTAP Deploy configure chaque nœud avec 5 To d'espace actif par nœud.

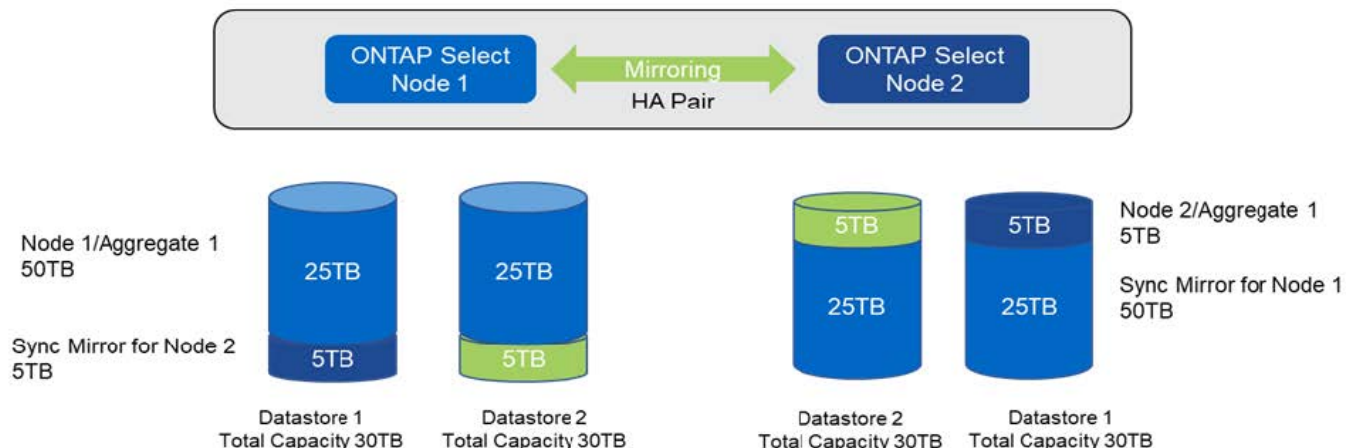
La figure suivante illustre les résultats d'une seule opération d'ajout de stockage pour le nœud 1. ONTAP Select utilise toujours la même quantité de stockage (15 To) sur chaque nœud. Cependant, le nœud 1 dispose de plus de stockage actif (10 To) que le nœud 2 (5 To). Les deux nœuds sont entièrement protégés, car chacun héberge une copie des données de l'autre. Il reste de l'espace libre dans le magasin de données 1, et le magasin de données 2 est toujours entièrement libre.

Répartition de la capacité : allocation et espace libre après une seule opération d'ajout de stockage



Deux opérations d'ajout de stockage supplémentaires sur le nœud 1 consomment le reste du datastore 1 et une partie du datastore 2 (en utilisant la capacité maximale). La première opération consomme les 15 To d'espace libre restant dans le datastore 1. La figure suivante illustre le résultat de la deuxième opération. À ce stade, le nœud 1 gère 50 To de données actives, tandis que le nœud 2 en gère les 5 To d'origine.

Répartition de la capacité : allocation et espace libre après deux opérations d'ajout de stockage supplémentaires pour le nœud 1



La taille maximale d'un VMDK utilisé lors des opérations d'ajout de capacité est de 16 To. La taille maximale d'un VMDK utilisé lors des opérations de création de cluster reste de 8 To. ONTAP Deploy crée des VMDK correctement dimensionnés en fonction de votre configuration (cluster à nœud unique ou multinœud) et de la capacité ajoutée. Cependant, la taille maximale de chaque VMDK ne doit pas dépasser 8 To lors des opérations de création de cluster et 16 To lors des opérations d'ajout de stockage.

Augmentez la capacité d' ONTAP Select avec le RAID logiciel

L'assistant d'ajout de stockage peut également être utilisé pour augmenter la capacité gérée des nœuds ONTAP Select utilisant le RAID logiciel. L'assistant affiche uniquement les disques SDD DAS disponibles et mappés en tant que RDM sur la machine virtuelle ONTAP Select .

Bien qu'il soit possible d'augmenter la capacité de la licence d'un seul To, avec un RAID logiciel, il est impossible d'augmenter physiquement la capacité d'un seul To. Comme pour l'ajout de disques à une matrice FAS ou AFF , certains facteurs déterminent la quantité minimale de stockage pouvant être ajoutée en une seule opération.

Notez que dans une paire HA, l'ajout de stockage au nœud 1 nécessite qu'un nombre identique de disques soit également disponible sur la paire HA du nœud (nœud 2). Les disques locaux et les disques distants sont utilisés par une opération d'ajout de stockage sur le nœud 1. Autrement dit, les disques distants servent à garantir que le nouveau stockage du nœud 1 est répliqué et protégé sur le nœud 2. Pour ajouter du stockage utilisable localement sur le nœud 2, une opération d'ajout de stockage distincte et un nombre égal de disques distincts doivent être disponibles sur les deux nœuds.

ONTAP Select partitionne tous les nouveaux disques dans les mêmes partitions racine, données et données que les disques existants. Le partitionnement a lieu lors de la création d'un nouvel agrégat ou de l'extension d'un agrégat existant. La taille de la bande de partition racine de chaque disque est définie pour correspondre à la taille de la partition racine existante sur les disques existants. Par conséquent, chacune des deux tailles de partition de données égales peut être calculée comme la capacité totale du disque moins la taille de la partition racine divisée par deux. La taille de la bande de partition racine est variable et est calculée lors de la configuration initiale du cluster comme suit : l'espace racine total requis (68 Go pour un cluster à nœud unique et 136 Go pour les paires HA) est divisé par le nombre initial de disques, moins les disques de secours et de parité. La taille de la bande de partition racine est maintenue constante sur tous les disques ajoutés au système.

Si vous créez un nouvel agrégat, le nombre minimum de disques requis varie en fonction du type de RAID et du fait que le nœud ONTAP Select fait partie d'une paire HA.

Si vous ajoutez du stockage à un agrégat existant, certaines considérations supplémentaires sont nécessaires. Il est possible d'ajouter des disques à un groupe RAID existant, à condition que celui-ci n'atteigne pas déjà sa limite maximale. Les bonnes pratiques FAS et AFF traditionnelles pour l'ajout de disques à des groupes RAID existants s'appliquent également, et la création d'un point chaud sur le nouveau disque peut poser problème. De plus, seuls des disques de tailles de partition de données égales ou supérieures peuvent être ajoutés à un groupe RAID existant. Comme expliqué précédemment, la taille de la partition de données diffère de la taille brute du disque. Si les partitions de données ajoutées sont plus grandes que les partitions existantes, les nouveaux disques sont dimensionnés correctement. Autrement dit, une partie de la capacité de chaque nouveau disque reste inutilisée.

Il est également possible d'utiliser les nouveaux disques pour créer un nouveau groupe RAID au sein d'un agrégat existant. Dans ce cas, la taille du groupe RAID doit correspondre à celle du groupe RAID existant.

Prise en charge de l'efficacité du stockage ONTAP Select

ONTAP Select fournit des options d'efficacité de stockage similaires aux options

d'efficacité de stockage présentes sur les baies FAS et AFF .

Les déploiements de NAS virtuels (vNAS) ONTAP Select utilisant des baies VSAN entièrement flash ou flash génériques doivent suivre les meilleures pratiques pour ONTAP Select avec un stockage à connexion directe (DAS) non SSD.

Une personnalité de type AFF est automatiquement activée sur les nouvelles installations tant que vous disposez d'un stockage DAS avec des disques SSD et d'une licence premium.

Avec une personnalité de type AFF, les fonctionnalités SE en ligne suivantes sont automatiquement activées lors de l'installation :

- Détection de motif zéro en ligne
- Déduplication en ligne du volume
- Déduplication d'arrière-plan du volume
- Compression adaptative en ligne
- Compactage des données en ligne
- Déduplication agrégée en ligne
- Déduplication d'arrière-plan agrégée

Pour vérifier qu'ONTAP ONTAP Select a activé toutes les stratégies d'efficacité de stockage par défaut, exécutez la commande suivante sur un volume nouvellement créé :

```
<system name>::> set diag
Warning: These diagnostic commands are for use by NetApp personnel only.
Do you want to continue? {y|n}: y
twonode95IP15::*> sis config
Vserver:                               SVM1
Volume:                               __export1_NFS_volume
Schedule:                             -
Policy:                               auto
Compression:                           true
Inline Compression:                    true
Compression Type:                      adaptive
Application IO Size:                   8K
Compression Algorithm:                 lzopro
Inline Dedupe:                         true
Data Compaction:                       true
Cross Volume Inline Deduplication:     true
Cross Volume Background Deduplication: true
```




Pour les mises à niveau ONTAP Select à partir de la version 9.6, vous devez installer ONTAP Select sur un stockage SSD DAS avec une licence premium. De plus, vous devez cocher la case « Activer l'efficacité du stockage » lors de l'installation initiale du cluster avec ONTAP Deploy. L'activation d'une personnalité de type AFF après la mise à niveau ONTAP lorsque les conditions préalables ne sont pas remplies nécessite la création manuelle d'un argument de démarrage et le redémarrage du nœud. Contactez le support technique pour plus d'informations.

Configurations d'efficacité de stockage ONTAP Select

Le tableau suivant résume les différentes options d'efficacité de stockage disponibles, activées par défaut ou non activées par défaut mais recommandées, selon le type de support et la licence du logiciel.

Fonctionnalités ONTAP Select	SSD DAS (premium ou premium XL ¹)	Disque dur DAS (toutes licences)	vNAS (toutes licences)
Détection du zéro en ligne	Oui (par défaut)	Oui Activé par l'utilisateur sur une base par volume	Oui Activé par l'utilisateur sur une base par volume
Déduplication en ligne du volume	Oui (par défaut)	Non disponible	Non pris en charge
Compression en ligne 32K (compression secondaire)	Oui Activé par l'utilisateur en fonction du volume.	Oui Activé par l'utilisateur sur une base par volume	Non pris en charge
Compression en ligne 8K (compression adaptative)	Oui (par défaut)	Oui Activé par l'utilisateur sur une base par volume	Non pris en charge
Compression d'arrière-plan	Non pris en charge	Oui Activé par l'utilisateur sur une base par volume	Oui Activé par l'utilisateur sur une base par volume
Scanner de compression	Oui	Oui	Oui Activé par l'utilisateur sur une base par volume
Compactage des données en ligne	Oui (par défaut)	Oui Activé par l'utilisateur sur une base par volume	Non pris en charge
Scanner de compactage	Oui	Oui	Non pris en charge
Déduplication agrégée en ligne	Oui (par défaut)	S/O	Non pris en charge
Déduplication d'arrière-plan du volume	Oui (par défaut)	Oui Activé par l'utilisateur sur une base par volume	Oui Activé par l'utilisateur sur une base par volume
Déduplication d'arrière-plan agrégée	Oui (par défaut)	S/O	Non pris en charge

¹ ONTAP Select 9.6 prend en charge une nouvelle licence (Premium XL) et une nouvelle taille de machine virtuelle (Large). Cependant, la grande machine virtuelle n'est prise en charge que pour les configurations DAS utilisant le RAID logiciel. Les configurations RAID matériel et vNAS ne sont pas prises en charge avec la grande machine virtuelle ONTAP Select de la version 9.6.

Remarques sur le comportement de mise à niveau pour les configurations SSD DAS

Après la mise à niveau vers ONTAP Select 9.6 ou une version ultérieure, attendez que le `system node upgrade-revert show` commande pour indiquer que la mise à niveau est terminée avant de vérifier les valeurs d'efficacité de stockage pour les volumes existants.

Sur un système mis à niveau vers ONTAP Select 9.6 ou version ultérieure, un nouveau volume créé sur un

agrégat existant ou nouvellement créé présente le même comportement qu'un volume créé lors d'un nouveau déploiement. Les volumes existants mis à niveau vers ONTAP Select appliquent la plupart des mêmes politiques d'efficacité de stockage qu'un volume nouvellement créé, avec quelques variations :

Scénario 1

Si aucune politique d'efficacité du stockage n'a été activée sur un volume avant la mise à niveau, alors :

- Volumes avec `space guarantee = volume` Les options de compactage des données en ligne, de déduplication agrégée en ligne et de déduplication agrégée en arrière-plan ne sont pas activées. Ces options peuvent être activées après la mise à niveau.
- Volumes avec `space guarantee = none` La compression d'arrière-plan n'est pas activée. Cette option peut être activée après la mise à niveau.
- La politique d'efficacité du stockage sur les volumes existants est définie sur automatique après la mise à niveau.

Scénario 2

Si certaines fonctionnalités de stockage sont déjà activées sur un volume avant la mise à niveau, alors :

- Volumes avec `space guarantee = volume` je ne vois aucune différence après la mise à niveau.
- Volumes avec `space guarantee = none` avoir activé la déduplication d'arrière-plan agrégée.
- Volumes avec `storage policy inline-only` avoir leur politique définie sur automatique.
- Les volumes avec des politiques d'efficacité de stockage définies par l'utilisateur ne subissent aucun changement de politique, à l'exception des volumes avec `space guarantee = none`. Ces volumes ont une déduplication d'arrière-plan agrégée activée.

Réseautage

Concepts et caractéristiques du réseau ONTAP Select

Familiarisez-vous d'abord avec les concepts réseau généraux applicables à l'environnement ONTAP Select. Explorez ensuite les caractéristiques et options spécifiques des clusters mono-nœud et multi-nœuds.

Réseau physique

Le réseau physique prend en charge le déploiement d'un cluster ONTAP Select principalement en fournissant l'infrastructure de commutation de couche 2 sous-jacente. La configuration liée au réseau physique comprend à la fois l'hôte hyperviseur et l'environnement réseau commuté plus large.

Options de la carte réseau hôte

Chaque hôte hyperviseur ONTAP Select doit être configuré avec deux ou quatre ports physiques. Le choix de la configuration dépend de plusieurs facteurs, notamment :

- Si le cluster contient un ou plusieurs hôtes ONTAP Select
- Quel système d'exploitation hyperviseur est utilisé
- Comment le commutateur virtuel est configuré
- Que LACP soit utilisé avec les liens ou non

Configuration du commutateur physique

Vous devez vous assurer que la configuration des commutateurs physiques prend en charge le déploiement ONTAP Select . Les commutateurs physiques sont intégrés aux commutateurs virtuels basés sur l'hyperviseur. Le choix de la configuration dépend de plusieurs facteurs, notamment :

- Comment allez-vous maintenir la séparation entre les réseaux internes et externes ?
- Allez-vous maintenir une séparation entre les réseaux de données et de gestion ?
- Comment les VLAN de couche deux seront-ils configurés ?

Réseau logique

ONTAP Select utilise deux réseaux logiques distincts, séparant le trafic selon son type. Plus précisément, le trafic peut circuler entre les hôtes du cluster, ainsi que vers les clients de stockage et d'autres machines externes. Les commutateurs virtuels gérés par les hyperviseurs contribuent à la prise en charge du réseau logique.

Réseau interne

Dans un déploiement de cluster multi-nœuds, les nœuds ONTAP Select communiquent via un réseau interne isolé. Ce réseau n'est ni exposé ni disponible en dehors des nœuds du cluster ONTAP Select .



Le réseau interne n'est présent qu'avec un cluster multi-nœuds.

Le réseau interne présente les caractéristiques suivantes :

- Utilisé pour traiter le trafic intra-cluster ONTAP , notamment :
 - Cluster
 - Interconnexion haute disponibilité (HA-IC)
 - Miroir de synchronisation RAID (RSM)
- Réseau monocouche à deux couches basé sur un VLAN
- Les adresses IP statiques sont attribuées par ONTAP Select:
 - IPv4 uniquement
 - DHCP non utilisé
 - Adresse de lien local
- La taille MTU est de 9 000 octets par défaut et peut être ajustée dans une plage de 7 500 à 9 000 (inclus)

Réseau externe

Le réseau externe gère le trafic entre les nœuds d'un cluster ONTAP Select et les clients de stockage externes, ainsi que les autres machines. Il fait partie intégrante de tout déploiement de cluster et présente les caractéristiques suivantes :

- Utilisé pour traiter le trafic ONTAP , notamment :
 - Données (NFS, CIFS, iSCSI)
 - Gestion (cluster et nœud ; SVM en option)
 - Intercluster (facultatif)

- Prend en charge les VLAN en option :
 - Groupe de ports de données
 - Groupe de gestion portuaire
- Adresses IP attribuées en fonction des choix de configuration de l'administrateur :
 - IPv4 ou IPv6
- La taille MTU est de 1 500 octets par défaut (peut être ajustée)

Le réseau externe est présent avec des clusters de toutes tailles.

Environnement réseau de machines virtuelles

L'hôte hyperviseur fournit plusieurs fonctionnalités réseau.

ONTAP Select s'appuie sur les fonctionnalités suivantes exposées via la machine virtuelle :

Ports de machines virtuelles

Plusieurs ports sont disponibles pour ONTAP Select. Leur attribution et leur utilisation dépendent de plusieurs facteurs, notamment la taille du cluster.

Commutateur virtuel

Le logiciel de commutation virtuelle de l'environnement hyperviseur, qu'il s'agisse de vSwitch (VMware) ou d'Open vSwitch (KVM), relie les ports exposés par la machine virtuelle aux ports physiques de la carte réseau Ethernet. Vous devez configurer un vSwitch pour chaque hôte ONTAP Select , en fonction de votre environnement.

ONTAP Select les configurations de réseau à nœud unique et à nœuds multiples

ONTAP Select prend en charge les configurations de réseau à nœud unique et à nœuds multiples.

Configuration de réseau à nœud unique

Les configurations ONTAP Select à nœud unique ne nécessitent pas le réseau interne ONTAP , car il n'y a pas de trafic de cluster, de haute disponibilité ou de miroir.

Contrairement à la version multinœud du produit ONTAP Select , chaque machine virtuelle ONTAP Select contient trois adaptateurs réseau virtuels, présentés aux ports réseau ONTAP e0a, e0b et e0c.

Ces ports sont utilisés pour fournir les services suivants : gestion, données et LIF intercluster.

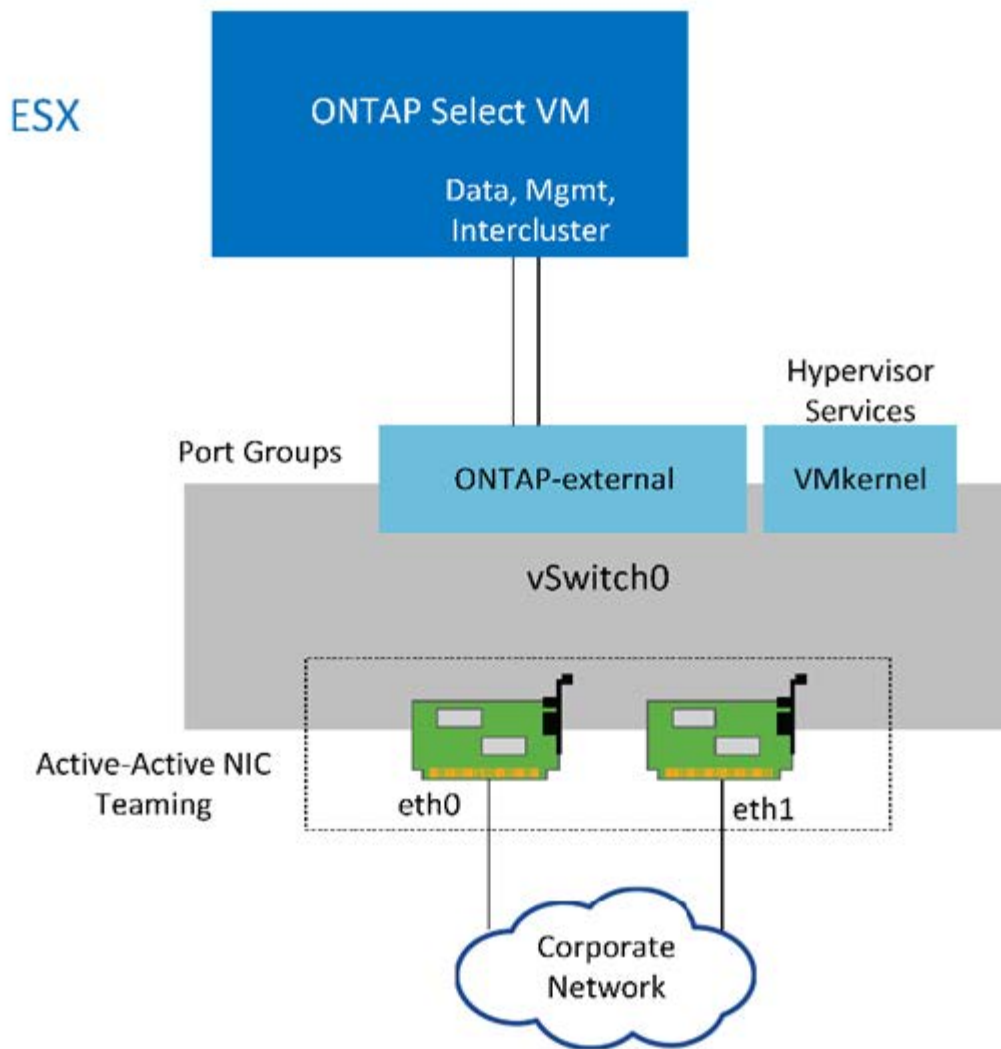
KVM

ONTAP Select peut être déployé sous forme de cluster à nœud unique. L'hôte hyperviseur comprend un commutateur virtuel donnant accès au réseau externe.

ESXi

La relation entre ces ports et les adaptateurs physiques sous-jacents peut être observée dans la figure suivante, qui représente un nœud de cluster ONTAP Select sur l'hyperviseur ESX.

Configuration réseau du cluster ONTAP Select à nœud unique



Même si deux adaptateurs suffisent pour un cluster à nœud unique, l'association de cartes réseau est toujours nécessaire.

Affectation LIF

Comme expliqué dans la section Affectation de LIF multinœuds de ce document, les espaces IP sont utilisés par ONTAP Select pour séparer le trafic réseau du cluster du trafic de données et de gestion. La variante à nœud unique de cette plateforme ne contient pas de réseau de cluster. Par conséquent, aucun port n'est présent dans l'espace IP du cluster.



Les LIF de gestion des clusters et des nœuds sont automatiquement créés lors de la configuration du cluster ONTAP Select. Les LIF restants peuvent être créés après le déploiement.

Gestion et données LIF (e0a, e0b et e0c)

Les ports ONTAP e0a, e0b et e0c sont délégués comme ports candidats pour les LIF qui transportent les types de trafic suivants :

- Trafic du protocole SAN/NAS (CIFS, NFS et iSCSI)

- Trafic de gestion des clusters, des nœuds et des SVM
- Trafic intercluster (SnapMirror et SnapVault)

Configuration du réseau multinœud

La configuration réseau multinœud ONTAP Select se compose de deux réseaux.

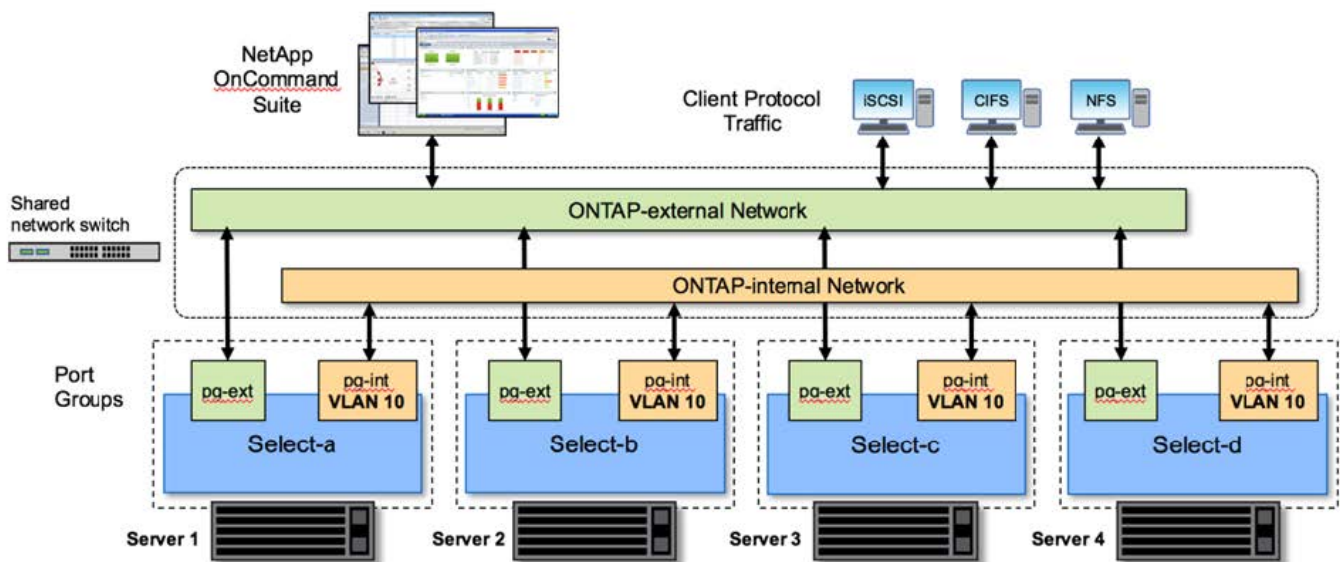
Il s'agit d'un réseau interne, chargé de fournir des services de cluster et de réplication interne, et d'un réseau externe, chargé de fournir des services d'accès et de gestion des données. L'isolation de bout en bout du trafic circulant au sein de ces deux réseaux est essentielle pour créer un environnement propice à la résilience des clusters.

Ces réseaux sont représentés dans la figure suivante, qui montre un cluster ONTAP Select à quatre nœuds exécuté sur une plateforme VMware vSphere. Les clusters à six et huit nœuds présentent une configuration réseau similaire.



Chaque instance ONTAP Select réside sur un serveur physique distinct. Le trafic interne et externe est isolé à l'aide de groupes de ports réseau distincts, attribués à chaque interface réseau virtuelle, permettant aux nœuds du cluster de partager la même infrastructure de commutation physique.

*Présentation d'une configuration réseau de cluster multinœud ONTAP Select *



Chaque machine virtuelle ONTAP Select contient sept cartes réseau virtuelles présentées à ONTAP sous la forme d'un ensemble de sept ports réseau, de e0a à e0g. ONTAP qu'ONTAP traite ces cartes comme des cartes réseau physiques, elles sont en réalité virtuelles et mappées à un ensemble d'interfaces physiques via une couche réseau virtualisée. Par conséquent, chaque serveur d'hébergement ne nécessite pas six ports réseau physiques.



L'ajout d'adaptateurs réseau virtuels à la machine virtuelle ONTAP Select n'est pas pris en charge.

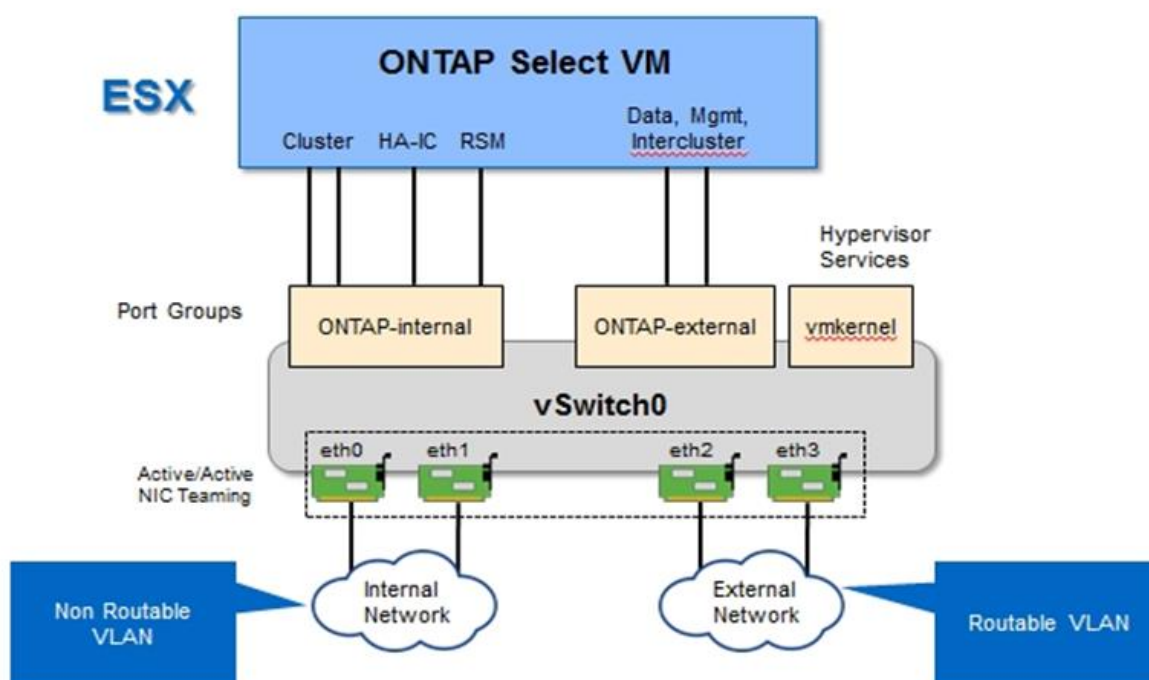
Ces ports sont préconfigurés pour fournir les services suivants :

- e0a, e0b et e0g. Gestion et données LIF
- e0c, e0d. Réseaux de clusters LIF
- e0e. RSM
- e0f. Interconnexion HA

Les ports e0a, e0b et e0g résident sur le réseau externe. Bien que les ports e0c à e0f remplissent plusieurs fonctions différentes, ils constituent collectivement le réseau Select interne. Lors de la conception du réseau, il est conseillé de placer ces ports sur un seul réseau de couche 2. Il n'est pas nécessaire de séparer ces adaptateurs virtuels sur différents réseaux.

La relation entre ces ports et les adaptateurs physiques sous-jacents est illustrée dans la figure suivante, qui représente un nœud de cluster ONTAP Select sur l'hyperviseur ESX.

Configuration réseau d'un nœud unique faisant partie d'un cluster ONTAP Select multi-nœuds



La séparation du trafic interne et externe sur différentes cartes réseau physiques évite les latences dues à un accès insuffisant aux ressources réseau. De plus, l'agrégation par regroupement de cartes réseau garantit qu'une défaillance d'une seule carte réseau n'empêche pas le nœud du cluster ONTAP Select d'accéder au réseau concerné.

Notez que les groupes de ports réseau externes et internes contiennent les quatre cartes réseau de manière symétrique. Les ports actifs du groupe de ports réseau externes sont les ports de secours du réseau interne. Inversement, les ports actifs du groupe de ports réseau internes sont les ports de secours du groupe de ports réseau externes.

Affectation LIF

Avec l'introduction des espaces IP, les rôles de port ONTAP sont devenus obsolètes. Comme les baies FAS, les clusters ONTAP Select contiennent un espace IP par défaut et un espace IP de cluster. En plaçant les ports réseau e0a, e0b et e0g dans l'espace IP par défaut et les ports e0c et e0d dans l'espace IP de cluster,

ces ports ont été isolés de l'hébergement de LIF non autorisées. Les ports restants du cluster ONTAP Select sont utilisés via l'affectation automatique d'interfaces fournissant des services internes. Ils ne sont pas exposés via le shell ONTAP , contrairement aux interfaces d'interconnexion RSM et HA.



Tous les LIF ne sont pas visibles via l'interpréteur de commandes ONTAP . Les interfaces d'interconnexion HA et RSM sont masquées par ONTAP et utilisées en interne pour fournir leurs services respectifs.

Les ports réseau et les LIF sont expliqués en détail dans les sections suivantes.

Gestion et données LIF (e0a, e0b et e0g)

Les ports ONTAP e0a, e0b et e0g sont délégués comme ports candidats pour les LIF qui transportent les types de trafic suivants :

- Trafic du protocole SAN/NAS (CIFS, NFS et iSCSI)
- Trafic de gestion des clusters, des nœuds et des SVM
- Trafic intercluster (SnapMirror et SnapVault)



Les LIF de gestion des clusters et des nœuds sont automatiquement créés lors de la configuration du cluster ONTAP Select . Les LIF restants peuvent être créés après le déploiement.

Réseau de clusters LIF (e0c, e0d)

Les ports ONTAP e0c et e0d sont délégués comme ports d'accueil pour les interfaces de cluster. Au sein de chaque nœud de cluster ONTAP Select , deux interfaces de cluster sont automatiquement générées lors de la configuration ONTAP à l'aide des adresses IP locales (169.254.xx).



Ces interfaces ne peuvent pas se voir attribuer d'adresses IP statiques et des interfaces de cluster supplémentaires ne doivent pas être créées.

Le trafic réseau du cluster doit transiter par un réseau de couche 2 non routé et à faible latence. En raison des exigences de débit et de latence du cluster, le cluster ONTAP Select doit être physiquement situé à proximité (par exemple, dans un cluster multipack ou un centre de données unique). La création de configurations de cluster extensibles à quatre, six ou huit nœuds en séparant les nœuds HA sur un WAN ou sur des distances géographiques importantes n'est pas prise en charge. Une configuration étendue à deux nœuds avec un médiateur est prise en charge.

Pour plus de détails, voir la section ["Bonnes pratiques pour la haute disponibilité étendue à deux nœuds \(MetroCluster SDS\)"](#) .



Pour garantir un débit maximal pour le trafic réseau du cluster, ce port réseau est configuré pour utiliser des trames jumbo (7 500 à 9 000 MTU). Pour un fonctionnement optimal du cluster, vérifiez que les trames jumbo sont activées sur tous les commutateurs virtuels et physiques en amont fournissant des services réseau internes aux nœuds du cluster ONTAP Select .

Trafic RAID SyncMirror (e0e)

La réplication synchrone des blocs entre les nœuds partenaires HA s'effectue via une interface réseau interne résidant sur le port réseau e0e. Cette fonctionnalité est automatique, grâce aux interfaces réseau configurées par ONTAP lors de la configuration du cluster, et ne nécessite aucune configuration de la part de

l'administrateur.



Le port e0e est réservé par ONTAP pour le trafic de réplication interne. Par conséquent, ni le port ni la LIF hébergée ne sont visibles dans l'interface de ligne de commande ONTAP ni dans System Manager. Cette interface est configurée pour utiliser une adresse IP locale de liaison générée automatiquement, et la réattribution d'une autre adresse IP n'est pas prise en charge. Ce port réseau nécessite l'utilisation de trames jumbo (7 500 à 9 000 MTU).

Interconnexion HA (e0f)

Les baies NetApp FAS utilisent du matériel spécialisé pour transmettre les informations entre les paires HA d'un cluster ONTAP . Cependant, les environnements définis par logiciel (SSD) ne disposent généralement pas de ce type d'équipement (comme les périphériques InfiniBand ou iWARP), une solution alternative est donc nécessaire. Bien que plusieurs possibilités aient été envisagées, les exigences ONTAP imposées au transport d'interconnexion nécessitaient que cette fonctionnalité soit émulée par logiciel. Par conséquent, au sein d'un cluster ONTAP Select , la fonctionnalité d'interconnexion HA (traditionnellement assurée par le matériel) a été intégrée au système d'exploitation, utilisant Ethernet comme mécanisme de transport.

Chaque nœud ONTAP Select est configuré avec un port d'interconnexion haute disponibilité (e0f). Ce port héberge l'interface réseau d'interconnexion haute disponibilité, responsable de deux fonctions principales :

- Mise en miroir du contenu de la NVRAM entre les paires HA
- Envoi/réception d'informations d'état HA et de messages de pulsation réseau entre les paires HA

Le trafic d'interconnexion HA circule via ce port réseau à l'aide d'une interface réseau unique en superposant des trames d'accès direct à la mémoire à distance (RDMA) dans des paquets Ethernet.



Comme pour le port RSM (e0e), ni le port physique ni l'interface réseau hébergée ne sont visibles par les utilisateurs depuis l'interface de ligne de commande ONTAP ou le Gestionnaire système. Par conséquent, l'adresse IP de cette interface et l'état du port ne peuvent pas être modifiés. Ce port réseau nécessite l'utilisation de trames jumbo (7 500 à 9 000 MTU).

ONTAP Select les réseaux internes et externes

Caractéristiques d' ONTAP Select les réseaux internes et externes.

ONTAP Select le réseau interne

Le réseau interne ONTAP Select , présent uniquement dans la version multinœud du produit, assure la communication entre les clusters ONTAP Select , l'interconnexion haute disponibilité et les services de réplication synchrone. Ce réseau comprend les ports et interfaces suivants :

- **e0c, e0d.** Hébergement de LIF de réseau de cluster
- **e0e.** Hébergement du RSM LIF
- **e0f.** Hébergement du LIF d'interconnexion HA

Le débit et la latence de ce réseau sont essentiels pour déterminer les performances et la résilience du cluster ONTAP Select . L'isolation du réseau est nécessaire pour la sécurité du cluster et pour garantir la séparation des interfaces système du reste du trafic réseau. Par conséquent, ce réseau doit être utilisé exclusivement par le cluster ONTAP Select .



L'utilisation du réseau interne Select pour un trafic autre que celui du cluster Select, comme le trafic d'application ou de gestion, n'est pas prise en charge. Aucune autre machine virtuelle ni aucun autre hôte ne peut être présent sur le VLAN interne ONTAP .

Les paquets réseau transitant par le réseau interne doivent transiter par un réseau de couche 2 dédié et étiqueté VLAN. Pour ce faire, effectuez l'une des tâches suivantes :

- Attribution d'un groupe de ports étiquetés VLAN aux cartes réseau virtuelles internes (e0c à e0f) (mode VST)
- Utilisation du VLAN natif fourni par le commutateur en amont où le VLAN natif n'est utilisé pour aucun autre trafic (attribuer un groupe de ports sans ID VLAN, c'est-à-dire en mode EST)

Dans tous les cas, le balisage VLAN pour le trafic réseau interne est effectué en dehors de la machine virtuelle ONTAP Select .



Seuls les commutateurs virtuels ESX standard et distribués sont pris en charge. Les autres commutateurs virtuels ou la connectivité directe entre hôtes ESX ne sont pas pris en charge. Le réseau interne doit être entièrement ouvert ; la NAT et les pare-feu ne sont pas pris en charge.

Au sein d'un cluster ONTAP Select , le trafic interne et le trafic externe sont séparés par des objets réseau virtuels de couche 2, appelés groupes de ports. L'affectation correcte de ces groupes de ports par vSwitch est essentielle, notamment pour le réseau interne, chargé de fournir les services de cluster, d'interconnexion haute disponibilité et de réplication miroir. Une bande passante réseau insuffisante vers ces ports peut entraîner une dégradation des performances, voire affecter la stabilité du nœud du cluster. Par conséquent, les clusters à quatre, six et huit nœuds nécessitent une connectivité 10 Gbit/s sur le réseau ONTAP Select interne ; les cartes réseau 1 Gbit/s ne sont pas prises en charge. Des compromis peuvent toutefois être trouvés sur le réseau externe, car limiter le flux de données entrantes vers un cluster ONTAP Select n'affecte pas sa fiabilité de fonctionnement.

Un cluster à deux nœuds peut utiliser soit quatre ports de 1 Gbit/s pour le trafic interne, soit un seul port de 10 Gbit/s au lieu des deux ports de 10 Gbit/s requis par le cluster à quatre nœuds. Si les conditions ne permettent pas d'équiper le serveur de quatre cartes réseau de 10 Gbit/s, deux cartes réseau de 10 Gbit/s peuvent être utilisées pour le réseau interne et deux cartes réseau de 1 Gbit/s pour le réseau ONTAP externe.

Validation et dépannage du réseau interne

Le réseau interne d'un cluster multinœud peut être validé grâce au vérificateur de connectivité réseau. Cette fonction peut être invoquée depuis l'interface de ligne de commande (CLI) de déploiement exécutant la commande `network connectivity-check start` commande.

Exécutez la commande suivante pour afficher la sortie du test :

```
network connectivity-check show --run-id X (X is a number)
```

Cet outil est uniquement utile pour le dépannage du réseau interne d'un cluster Select multinœud. Il ne doit pas être utilisé pour résoudre les problèmes de clusters à nœud unique (y compris les configurations vNAS), de connectivité ONTAP Deploy vers ONTAP Select ou de connectivité côté client.

L'assistant de création de cluster (intégré à l'interface graphique ONTAP Deploy) inclut le vérificateur de réseau interne comme étape facultative lors de la création de clusters multinœuds. Compte tenu du rôle important du réseau interne dans les clusters multinœuds, l'intégration de cette étape au workflow de création

de cluster améliore le taux de réussite des opérations de création.

À partir d'ONTAP Deploy 2.10, la taille MTU utilisée par le réseau interne peut être définie entre 7 500 et 9 000. Le vérificateur de connectivité réseau permet également de tester une taille MTU comprise entre 7 500 et 9 000. La valeur MTU par défaut correspond à celle du commutateur réseau virtuel. Cette valeur par défaut doit être remplacée par une valeur inférieure si une superposition réseau telle que VXLAN est présente dans l'environnement.

ONTAP Select un réseau externe

Le réseau externe ONTAP Select est responsable de toutes les communications sortantes du cluster et est donc présent dans les configurations mono-nœud et multi-nœud. Bien que ce réseau ne réponde pas aux exigences de débit strictes du réseau interne, l'administrateur doit veiller à ne pas créer de goulots d'étranglement entre le client et la machine virtuelle ONTAP, car des problèmes de performances pourraient être interprétés à tort comme des problèmes ONTAP Select.



De la même manière que pour le trafic interne, le trafic externe peut être étiqueté au niveau de la couche vSwitch (VST) et de la couche de commutation externe (EST). De plus, le trafic externe peut être étiqueté par la machine virtuelle ONTAP Select elle-même, via un processus appelé VGT. Voir la section "[Séparation du trafic de données et de gestion](#)" pour plus de détails.

Le tableau suivant met en évidence les principales différences entre les réseaux internes et externes ONTAP Select.

Référence rapide réseau interne et réseau externe

Description	Réseau interne	Réseau externe
Services réseau	Cluster HA/IC RAID SyncMirror (RSM)	Gestion des données intercluster (SnapMirror et SnapVault)
Isolation du réseau	Obligatoire	Facultatif
Taille de la trame (MTU)	7 500 à 9 000	1 500 (par défaut) 9 000 (pris en charge)
Attribution d'adresse IP	Généré automatiquement	Défini par l'utilisateur
Prise en charge DHCP	Non	Non

Regroupement de cartes réseau

Pour garantir que les réseaux internes et externes disposent de la bande passante et des caractéristiques de résilience nécessaires pour offrir des performances élevées et une tolérance aux pannes, il est recommandé de regrouper les cartes réseau physiques. Les configurations de cluster à deux nœuds avec une seule liaison 10 Gbit/s sont prises en charge. Cependant, la meilleure pratique recommandée par NetApp consiste à utiliser le regroupement de cartes réseau sur les réseaux internes et externes du cluster ONTAP Select.

génération d'adresses MAC

Les adresses MAC attribuées à tous les ports réseau ONTAP Select sont générées automatiquement par l'utilitaire de déploiement inclus. Cet utilitaire utilise un identifiant unique d'organisation (OUI) propre à la plateforme NetApp afin d'éviter tout conflit avec les systèmes FAS. Une copie de cette adresse est ensuite stockée dans une base de données interne de la machine virtuelle d'installation d'ONTAP Select (ONTAP Deploy), afin d'éviter toute réaffectation accidentelle lors des déploiements ultérieurs de nœuds. L'administrateur ne doit en aucun cas modifier l'adresse MAC attribuée à un port réseau.

Configurations réseau ONTAP Select en charge

Sélectionnez le meilleur matériel et configurez votre réseau pour optimiser les performances et la résilience.

Les fournisseurs de serveurs comprennent que les besoins des clients sont variés et que le choix est essentiel. Par conséquent, lors de l'achat d'un serveur physique, de nombreuses options s'offrent à vous pour choisir la connectivité réseau. La plupart des systèmes grand public sont livrés avec différentes cartes réseau, monoport ou multiport, offrant différentes possibilités de débit et de vitesse. Cela inclut la prise en charge des adaptateurs réseau 25 Gbit/s et 40 Gbit/s avec VMware ESX.

Les performances de la machine virtuelle ONTAP Select étant directement liées aux caractéristiques du matériel sous-jacent, l'augmentation du débit vers la machine virtuelle grâce à des cartes réseau plus rapides permet d'obtenir un cluster plus performant et une meilleure expérience utilisateur. Quatre cartes réseau 10 Gbit/s ou deux cartes réseau plus rapides (25/40 Gbit/s) peuvent être utilisées pour obtenir une configuration réseau hautes performances. Plusieurs autres configurations sont également prises en charge. Pour les clusters à deux nœuds, quatre ports 1 Gbit/s ou un port 10 Gbit/s sont pris en charge. Pour les clusters à un seul nœud, deux ports 1 Gbit/s sont pris en charge.

Configurations réseau minimales et recommandées

Il existe plusieurs configurations Ethernet prises en charge en fonction de la taille du cluster.

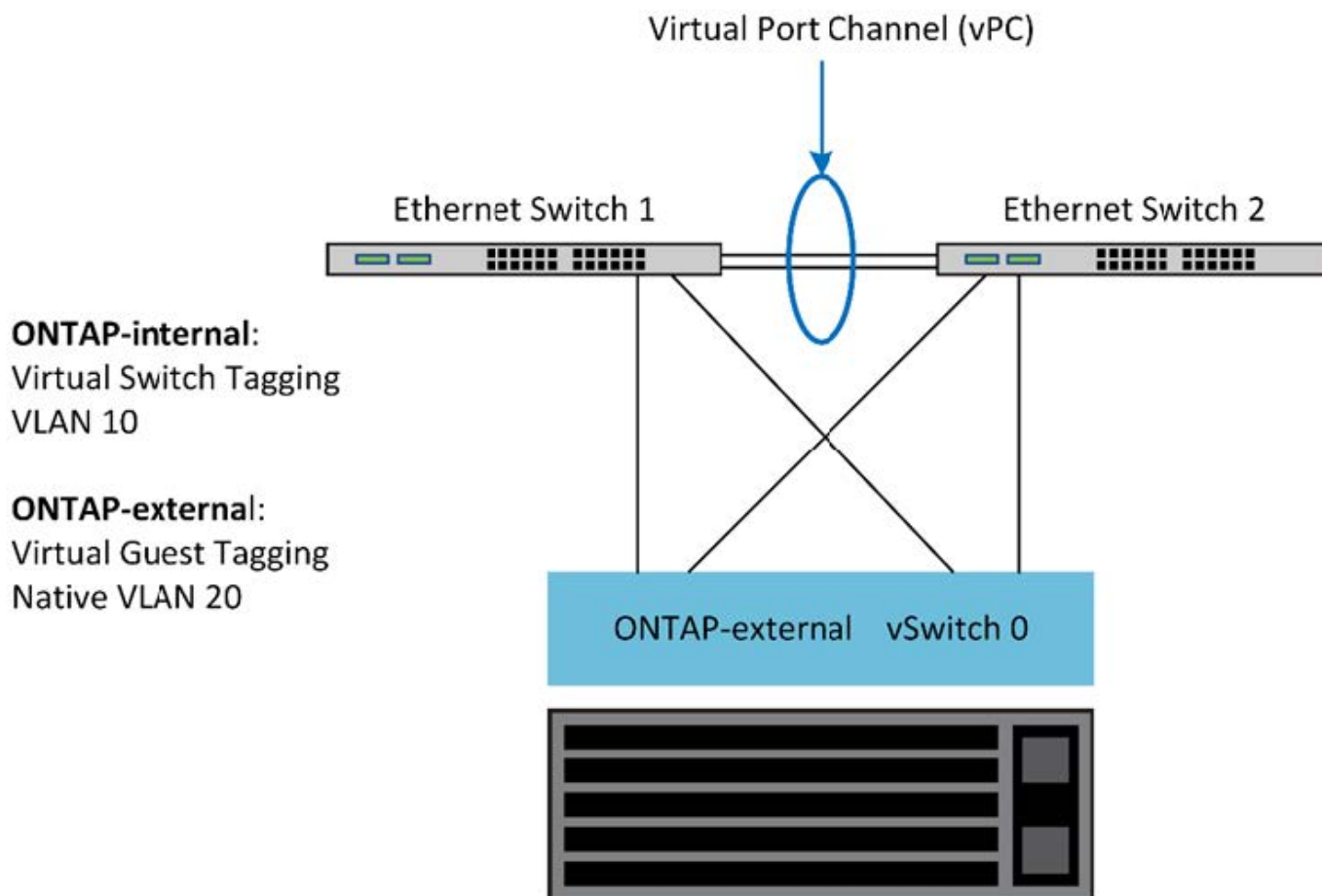
Taille du cluster	Exigences minimales	Recommandation
Cluster à nœud unique	2 x 1 GbE	2 x 10 GbE
Cluster à deux nœuds ou MetroCluster SDS	4 x 1 GbE ou 1 x 10 GbE	2 x 10 GbE
Cluster de 4/6/8 nœuds	2 x 10 GbE	4 x 10 GbE ou 2 x 25/40 GbE



La conversion entre les topologies à liaison unique et à liaisons multiples sur un cluster en cours d'exécution n'est pas prise en charge en raison de la nécessité éventuelle de convertir entre différentes configurations d'association de cartes réseau requises pour chaque topologie.

Configuration du réseau à l'aide de plusieurs commutateurs physiques

Lorsque suffisamment de matériel est disponible, NetApp recommande d'utiliser la configuration multi-commutateurs illustrée dans la figure suivante, en raison de la protection supplémentaire contre les pannes de commutateur physique.



ONTAP Select la configuration VMware vSphere vSwitch sur ESXi

ONTAP Select pour les configurations à deux et quatre cartes réseau.

ONTAP Select prend en charge les configurations de commutateurs virtuels standard et distribués. Les commutateurs virtuels distribués prennent en charge les structures d'agrégation de liens (LACP). L'agrégation de liens est une structure réseau courante utilisée pour agréger la bande passante entre plusieurs adaptateurs physiques. LACP est une norme indépendante du fournisseur qui fournit un protocole ouvert pour les points de terminaison réseau regroupant des groupes de ports réseau physiques dans un seul canal logique. ONTAP Select peut fonctionner avec des groupes de ports configurés en tant que groupe d'agrégation de liens (LAG). Cependant, NetApp recommande d'utiliser les ports physiques individuels comme de simples ports de liaison montante (trunk) afin d'éviter la configuration LAG. Dans ce cas, les bonnes pratiques pour les commutateurs virtuels standard et distribués sont identiques.

Cette section décrit la configuration vSwitch et les stratégies d'équilibrage de charge qui doivent être utilisées dans les configurations à deux et quatre cartes réseau.

Lors de la configuration des groupes de ports à utiliser par ONTAP Select, les bonnes pratiques suivantes doivent être respectées : la stratégie d'équilibrage de charge au niveau du groupe de ports est basée sur l'itinéraire basé sur l'ID de port virtuel d'origine. VMware recommande de définir STP sur Portfast sur les ports de commutateur connectés aux hôtes ESXi.

Toutes les configurations vSwitch nécessitent au moins deux cartes réseau physiques regroupées dans une même équipe de cartes réseau. ONTAP Select prend en charge une seule liaison 10 Gbit/s pour les clusters à deux nœuds. Cependant, il est recommandé chez NetApp de garantir la redondance matérielle grâce à l'agrégation de cartes réseau.

Sur un serveur vSphere, les équipes de cartes réseau constituent la structure d'agrégation utilisée pour regrouper plusieurs cartes réseau physiques dans un seul canal logique, permettant ainsi de répartir la charge réseau entre tous les ports membres. Il est important de noter que les équipes de cartes réseau peuvent être créées sans l'aide du commutateur physique. Les politiques d'équilibrage de charge et de basculement peuvent être appliquées directement à une équipe de cartes réseau, qui ignore la configuration du commutateur en amont. Dans ce cas, les politiques ne s'appliquent qu'au trafic sortant.



Les canaux de port statiques ne sont pas pris en charge avec ONTAP Select. Les canaux compatibles LACP sont pris en charge avec les vSwitches distribués, mais l'utilisation de LAG LACP peut entraîner une répartition de charge inégale entre les membres du LAG.

Pour les clusters à nœud unique, ONTAP Deploy configure la machine virtuelle ONTAP Select pour utiliser un groupe de ports pour le réseau externe et soit le même groupe de ports, soit, éventuellement, un groupe de ports différent pour le trafic de gestion du cluster et des nœuds. Pour les clusters à nœud unique, le nombre souhaité de ports physiques peut être ajouté au groupe de ports externes en tant qu'adaptateurs actifs.

Pour les clusters multinœuds, ONTAP Deploy configure chaque machine virtuelle ONTAP Select pour utiliser un ou deux groupes de ports pour le réseau interne et, séparément, un ou deux groupes de ports pour le réseau externe. Le trafic de gestion du cluster et des nœuds peut utiliser le même groupe de ports que le trafic externe, ou éventuellement un groupe de ports distinct. Le trafic de gestion du cluster et des nœuds ne peut pas partager le même groupe de ports que le trafic interne.

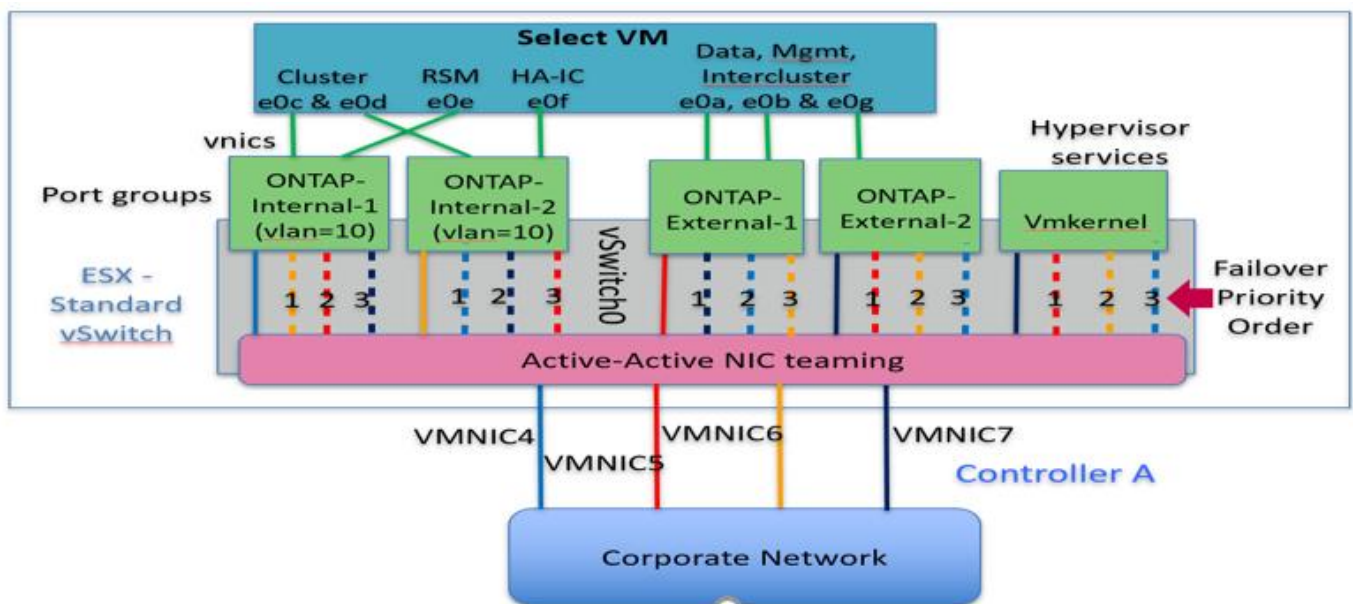


ONTAP Select prend en charge un maximum de quatre VMNIC.

vSwitch standard ou distribué et quatre ports physiques par nœud

Quatre groupes de ports peuvent être attribués à chaque nœud d'un cluster multinœud. Chaque groupe de ports possède un port physique actif et trois ports physiques de secours, comme illustré ci-dessous.

vSwitch avec quatre ports physiques par nœud



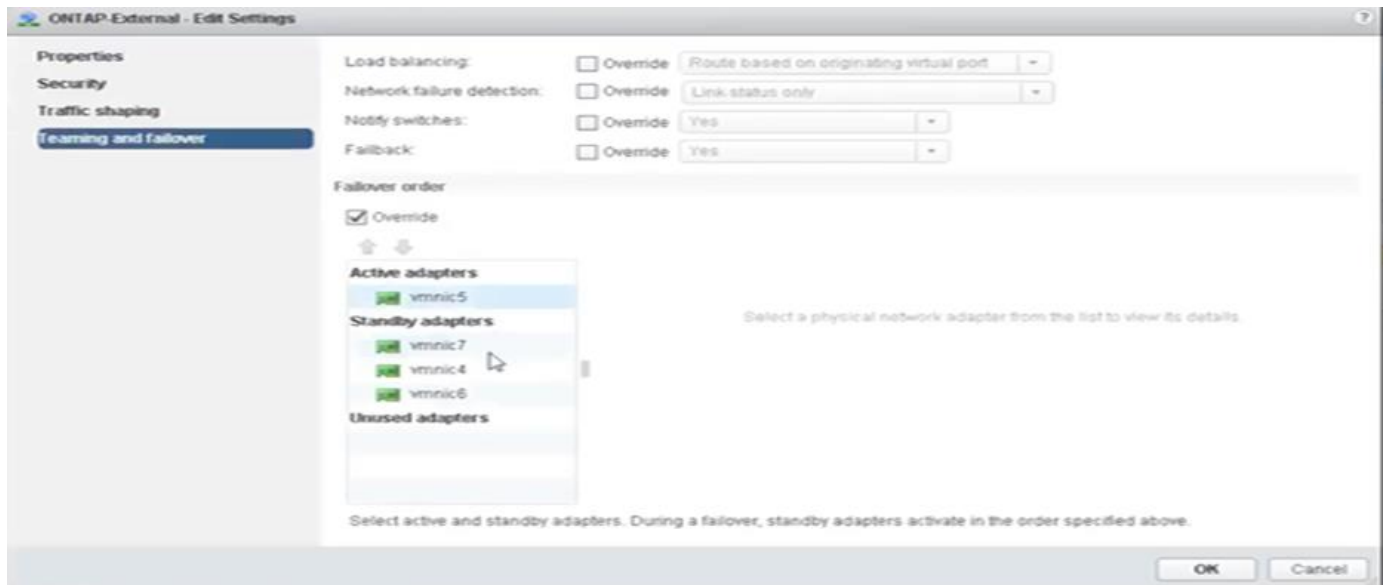
L'ordre des ports dans la liste de secours est important. Le tableau suivant présente un exemple de répartition physique des ports entre les quatre groupes de ports.

Configurations réseau minimales et recommandées

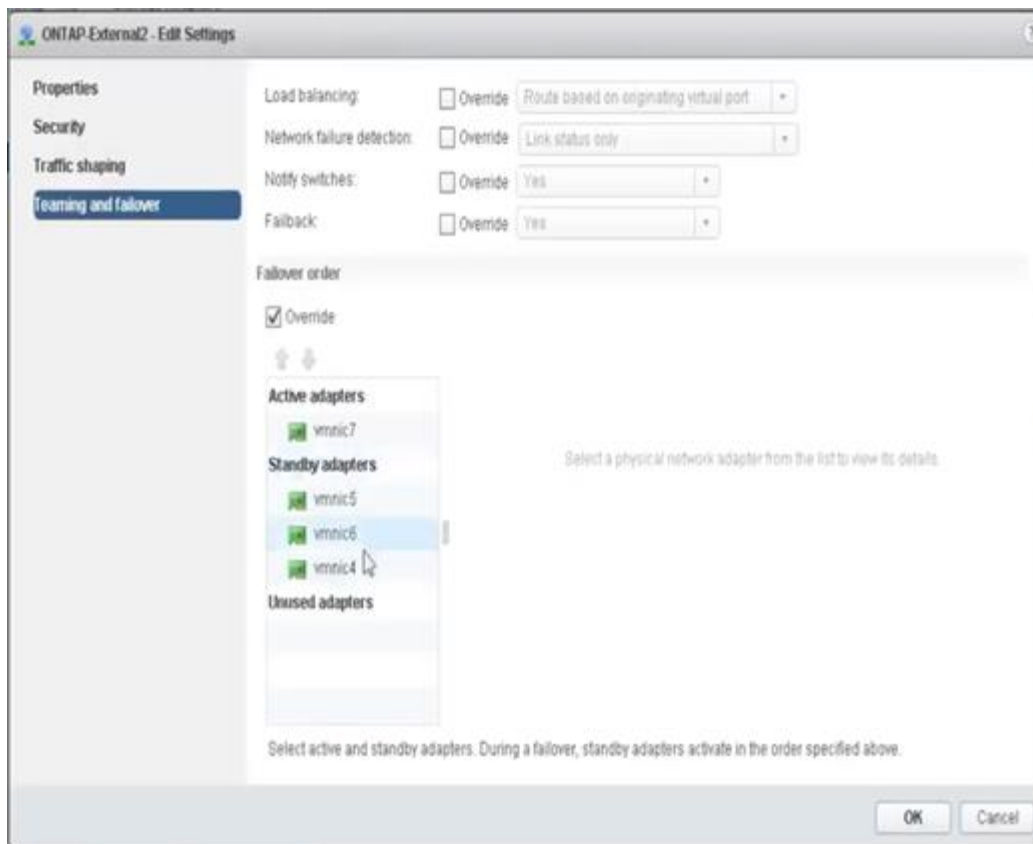
Groupe Port	Externe 1	Externe 2	Interne 1	Interne 2
Actif	vmnic0	vmnic1	vmnic2	vmnic3
Veille 1	vmnic1	vmnic0	vmnic3	vmnic2
Veille 2	vmnic2	vmnic3	vmnic0	vmnic1
Veille 3	vmnic3	vmnic2	vmnic1	vmnic0

Les figures suivantes illustrent les configurations des groupes de ports réseau externes depuis l'interface graphique de vCenter (ONTAP-External et ONTAP-External2). Notez que les adaptateurs actifs proviennent de cartes réseau différentes. Dans cette configuration, vmnic 4 et vmnic 5 sont des ports doubles sur la même carte réseau physique, tandis que vmnic 6 et vmnic 7 sont également des ports doubles sur une carte réseau distincte (les vmnic 0 à 3 ne sont pas utilisés dans cet exemple). L'ordre des adaptateurs de secours permet un basculement hiérarchique, les ports du réseau interne étant les derniers. L'ordre des ports internes dans la liste de secours est également inversé entre les deux groupes de ports externes.

*Partie 1 : Configurations de groupe de ports externes ONTAP Select *



*Partie 2 : Configurations de groupes de ports externes ONTAP Select *

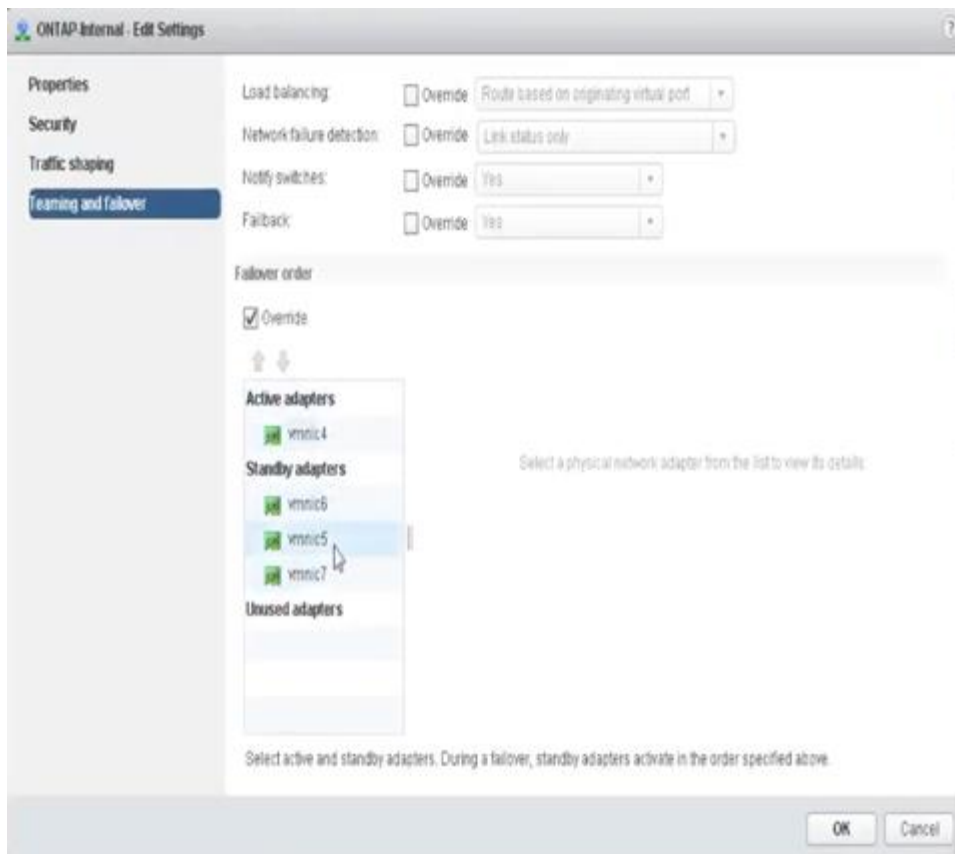


Pour plus de lisibilité, les devoirs sont les suivants :

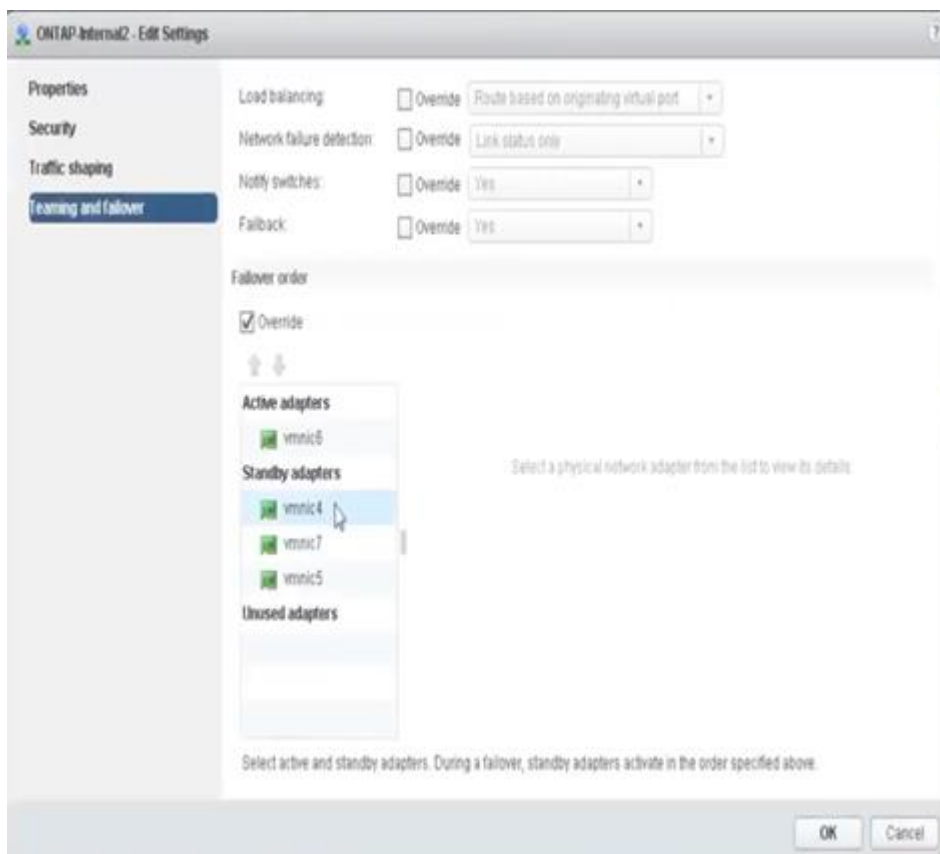
ONTAP- Externe	ONTAP-External2
Adaptateurs actifs : vmnic5 Adaptateurs de secours : vmnic7, vmnic4, vmnic6	Adaptateurs actifs : vmnic7 Adaptateurs de secours : vmnic5, vmnic6, vmnic4

Les figures suivantes illustrent la configuration des groupes de ports réseau internes (ONTAP-Internal et ONTAP-Internal2). Notez que les adaptateurs actifs proviennent de cartes réseau différentes. Dans cette configuration, vmnic 4 et vmnic 5 sont des ports doubles sur le même ASIC physique, tandis que vmnic 6 et vmnic 7 sont également des ports doubles sur un ASIC distinct. L'ordre des adaptateurs de secours assure un basculement hiérarchique, les ports du réseau externe étant les derniers. L'ordre des ports externes dans la liste de secours est également inversé entre les deux groupes de ports internes.

*Partie 1 : Configurations de groupe de ports internes ONTAP Select *



Partie 2 : ONTAP Select groupes de ports internes



Pour plus de lisibilité, les devoirs sont les suivants :

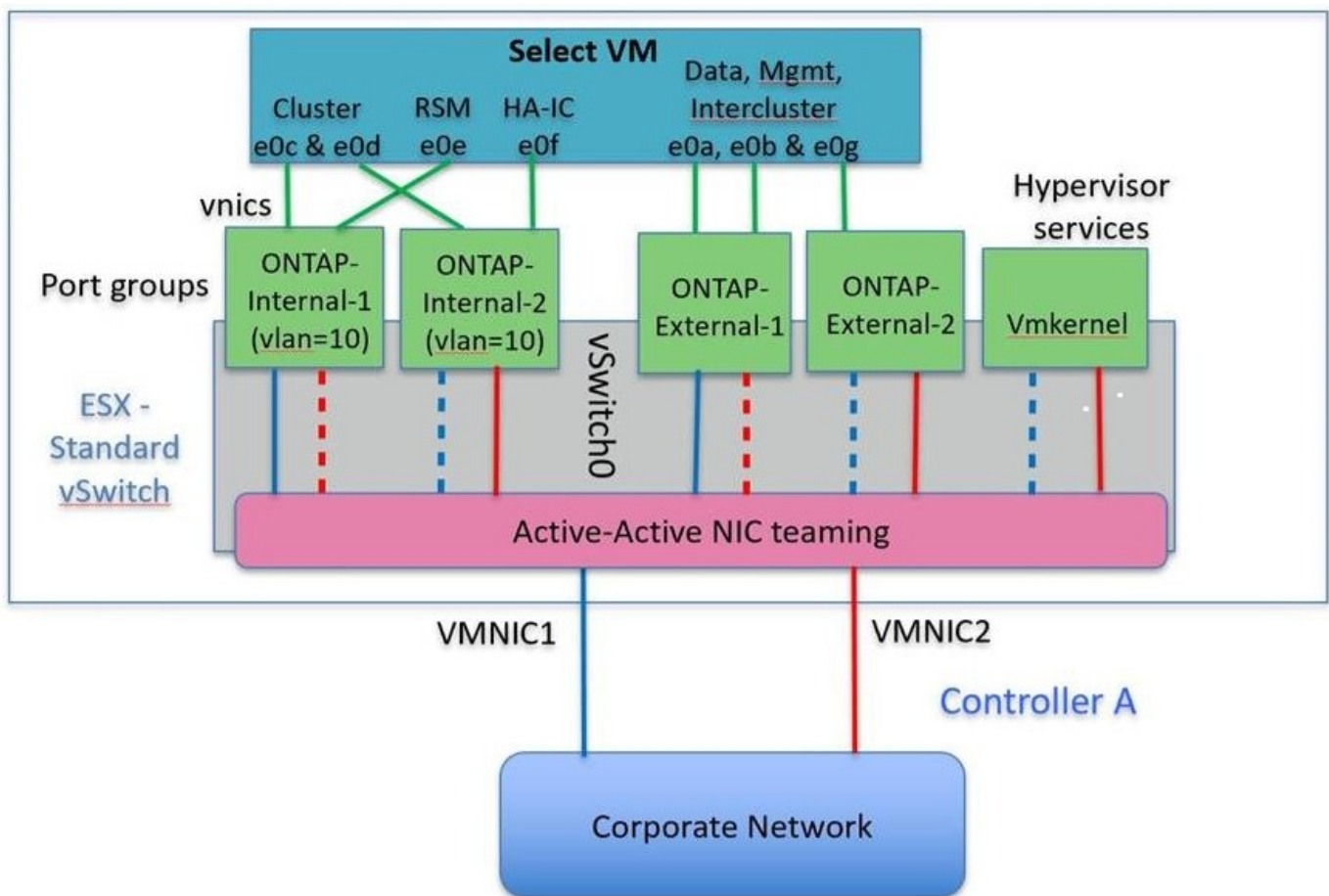
ONTAP- Interne	ONTAP-Internal2
Adaptateurs actifs : vmnic4 Adaptateurs de secours : vmnic6, vmnic5, vmnic7	Adaptateurs actifs : vmnic6 Adaptateurs de secours : vmnic4, vmnic7, vmnic5

vSwitch standard ou distribué et deux ports physiques par nœud

Lors de l'utilisation de deux cartes réseau haut débit (25/40 Gbit/s), la configuration recommandée des groupes de ports est très similaire à celle avec quatre adaptateurs 10 Gbit/s. Il est recommandé d'utiliser quatre groupes de ports, même avec seulement deux adaptateurs physiques. Les affectations des groupes de ports sont les suivantes :

Groupe Port	Externe 1 (e0a,e0b)	Interne 1 (e0c,e0e)	Interne 2 (e0d,e0f)	Externe 2 (e0g)
Actif	vmnic0	vmnic0	vmnic1	vmnic1
Attendre	vmnic1	vmnic1	vmnic0	vmnic0

vSwitch avec deux ports physiques haut débit (25/40 Go) par nœud

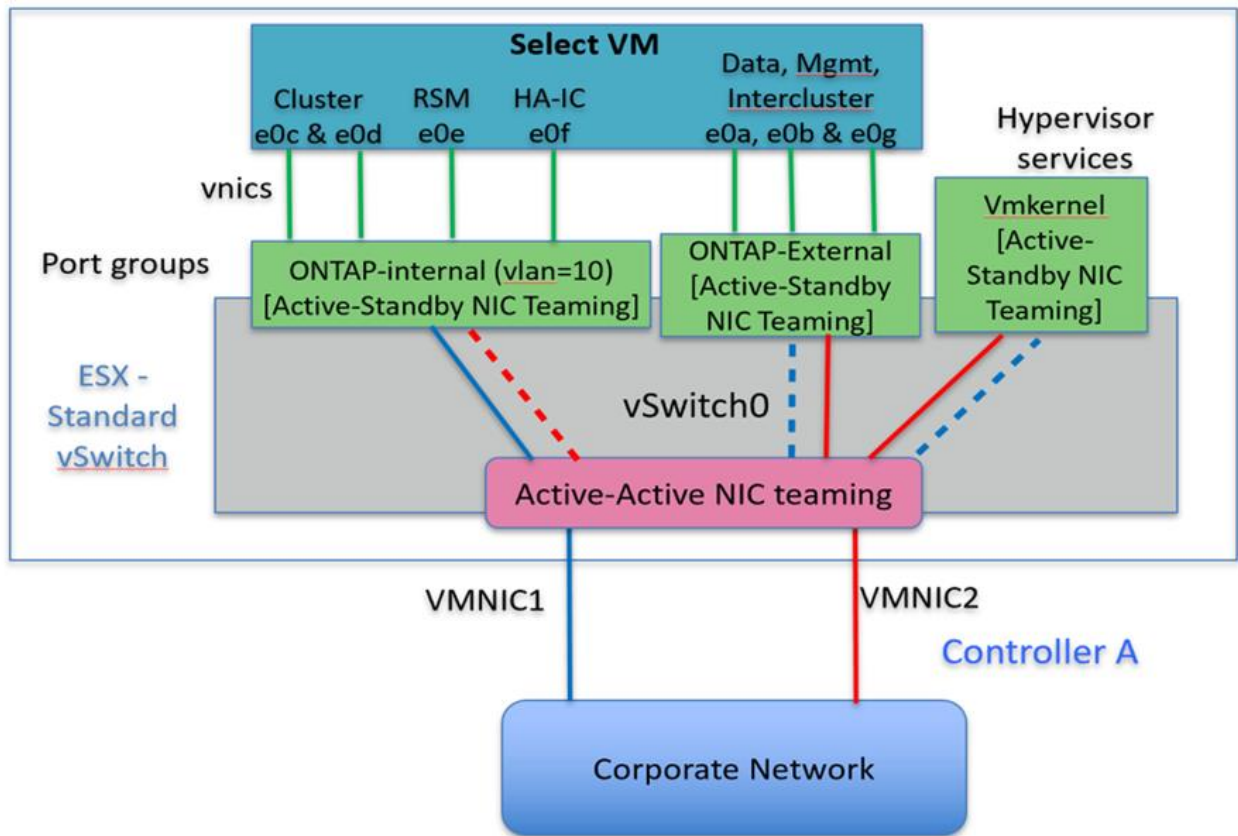


Lors de l'utilisation de deux ports physiques (10 Gbit/s ou moins), chaque groupe de ports doit disposer d'un adaptateur actif et d'un adaptateur de secours configurés en vis-à-vis. Le réseau interne est uniquement présent pour les clusters ONTAP Select multinœuds. Pour les clusters mononœuds, les deux adaptateurs peuvent être configurés comme actifs dans le groupe de ports externes.

L'exemple suivant illustre la configuration d'un vSwitch et des deux groupes de ports responsables de la gestion des services de communication internes et externes pour un cluster ONTAP Select multinœud. Le

réseau externe peut utiliser la VMNIC du réseau interne en cas de panne réseau, car ces dernières font partie de ce groupe de ports et sont configurées en mode veille. L'inverse est vrai pour le réseau externe. L'alternance des VMNIC actives et de secours entre les deux groupes de ports est essentielle au basculement correct des VM ONTAP Select en cas de panne réseau.

vSwitch avec deux ports physiques (10 Go ou moins) par nœud

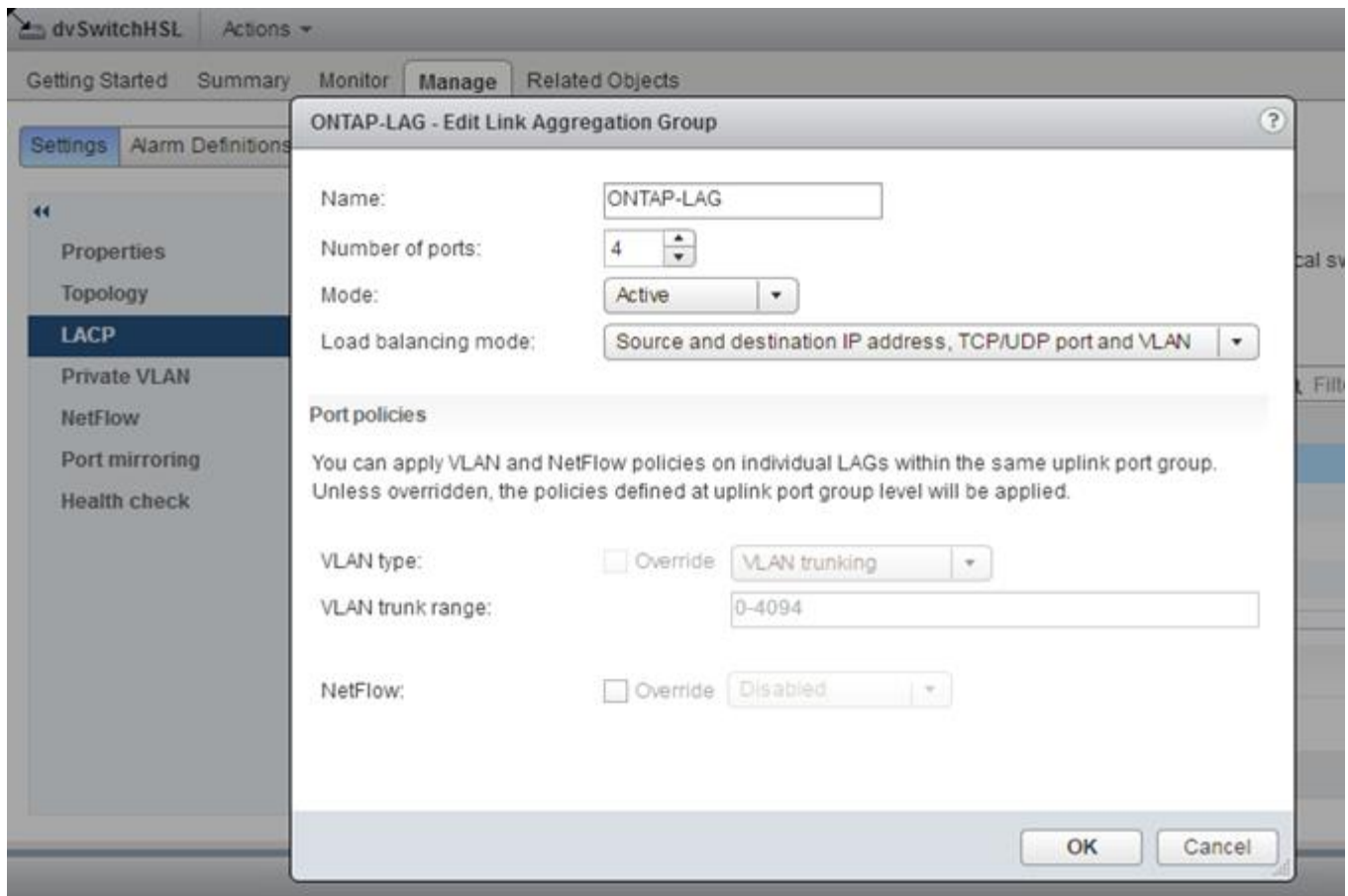


vSwitch distribué avec LACP

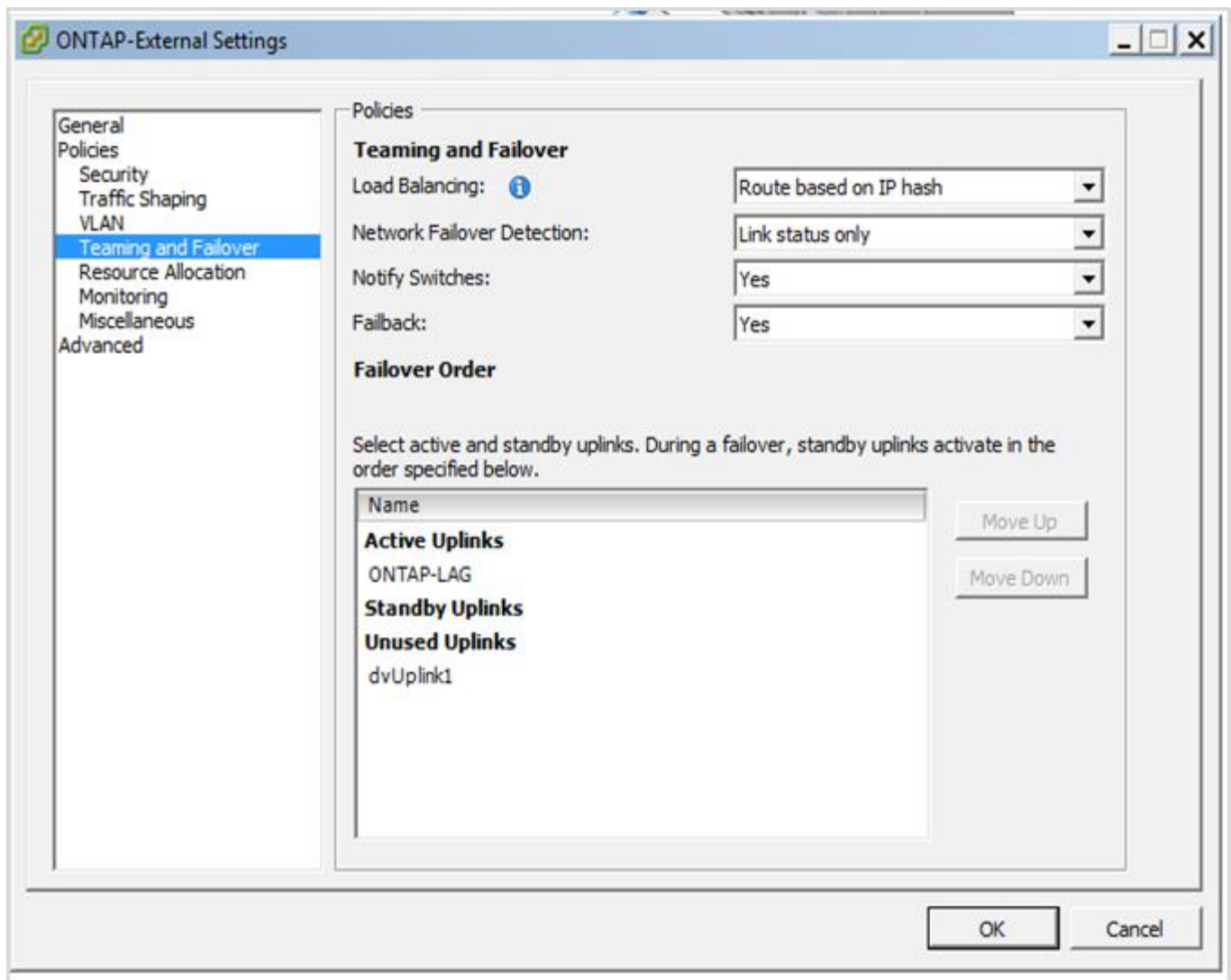
Lors de l'utilisation de commutateurs virtuels distribués dans votre configuration, le protocole LACP peut être utilisé (bien que ce ne soit pas une bonne pratique) afin de simplifier la configuration réseau. La seule configuration LACP prise en charge nécessite que toutes les VMNIC soient dans un seul LAG. Le commutateur physique de liaison montante doit prendre en charge une taille de MTU comprise entre 7 500 et 9 000 sur tous les ports du canal. Les réseaux ONTAP Select internes et externes doivent être isolés au niveau du groupe de ports. Le réseau interne doit utiliser un VLAN non routable (isolé). Le réseau externe peut utiliser VST, EST ou VGT.

Les exemples suivants montrent la configuration vSwitch distribuée à l'aide de LACP.

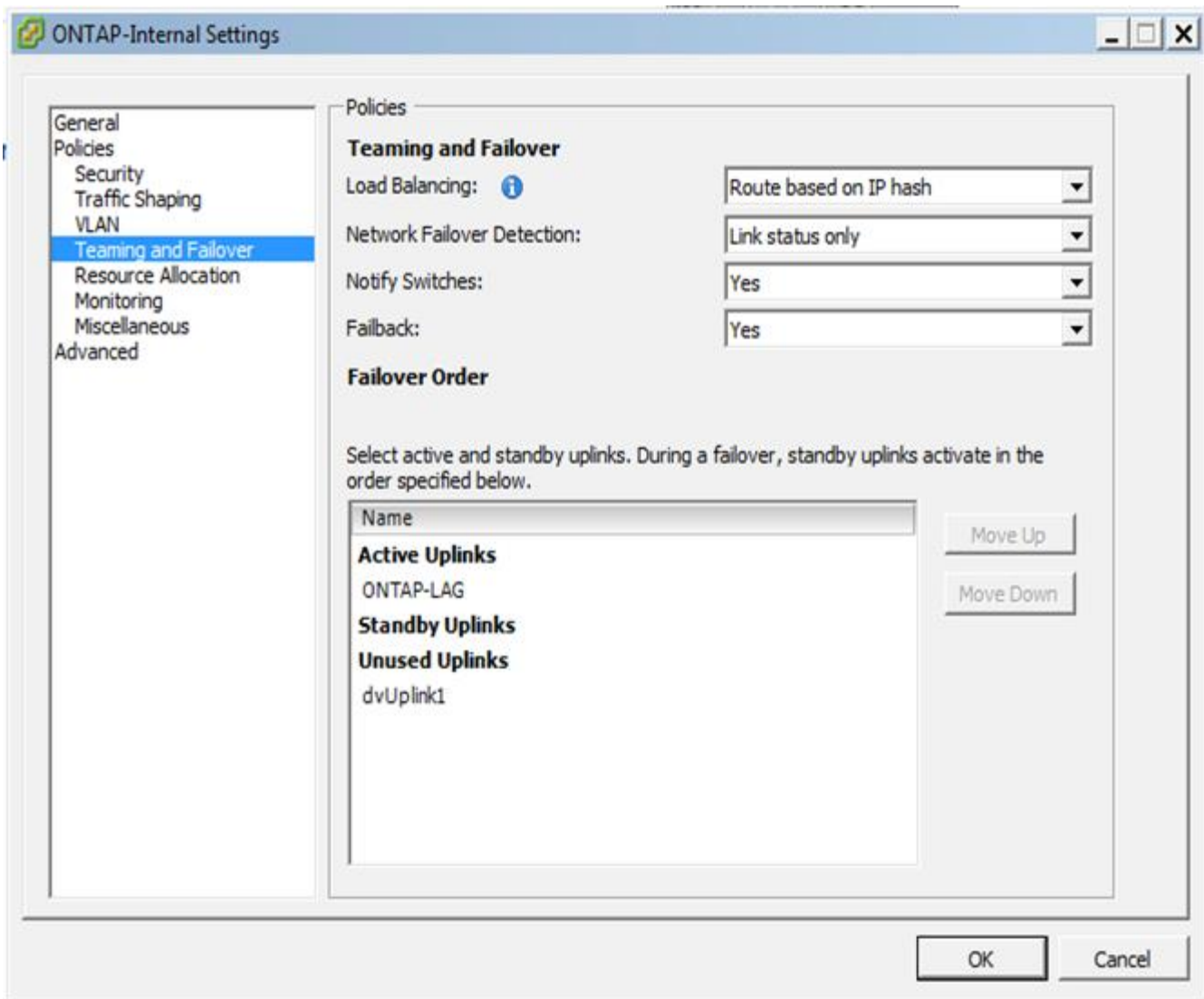
Propriétés LAG lors de l'utilisation de LACP



Configurations de groupes de ports externes utilisant un vSwitch distribué avec LACP activé



Configurations de groupes de ports internes utilisant un vSwitch distribué avec LACP activé



LACP nécessite la configuration des ports du commutateur en amont comme canal de port. Avant d'activer cette fonctionnalité sur le vSwitch distribué, assurez-vous qu'un canal de port compatible LACP est correctement configuré.

Configuration du commutateur physique ONTAP Select

Détails de configuration du commutateur physique en amont basés sur des environnements à commutateur unique et à commutateurs multiples.

Une attention particulière doit être portée aux décisions de connectivité entre la couche de commutation virtuelle et les commutateurs physiques. La séparation du trafic interne du cluster et des services de données externes doit s'étendre à la couche réseau physique en amont grâce à l'isolation assurée par les VLAN de couche 2.

Les ports de commutation physiques doivent être configurés comme ports trunk. Le trafic externe ONTAP Select peut être réparti sur plusieurs réseaux de couche 2 de deux manières : utiliser des ports virtuels ONTAP VLAN étiquetés avec un seul groupe de ports. L'autre méthode consiste à attribuer des groupes de ports distincts en mode VST au port de gestion e0a. Vous devez également attribuer des ports de données à e0b et e0c/e0g, selon la version ONTAP Select et la configuration mono-nœud ou multi-nœud. Si le trafic externe est réparti sur plusieurs réseaux de couche 2, les ports de commutation physiques de liaison montante doivent

inclure ces VLAN dans leur liste de VLAN autorisés.

Le trafic réseau interne ONTAP Select s'effectue via des interfaces virtuelles définies avec des adresses IP locales de liaison. Ces adresses IP n'étant pas routables, le trafic interne entre les nœuds du cluster doit transiter par un seul réseau de couche 2. Les sauts de routage entre les nœuds du cluster ONTAP Select ne sont pas pris en charge.

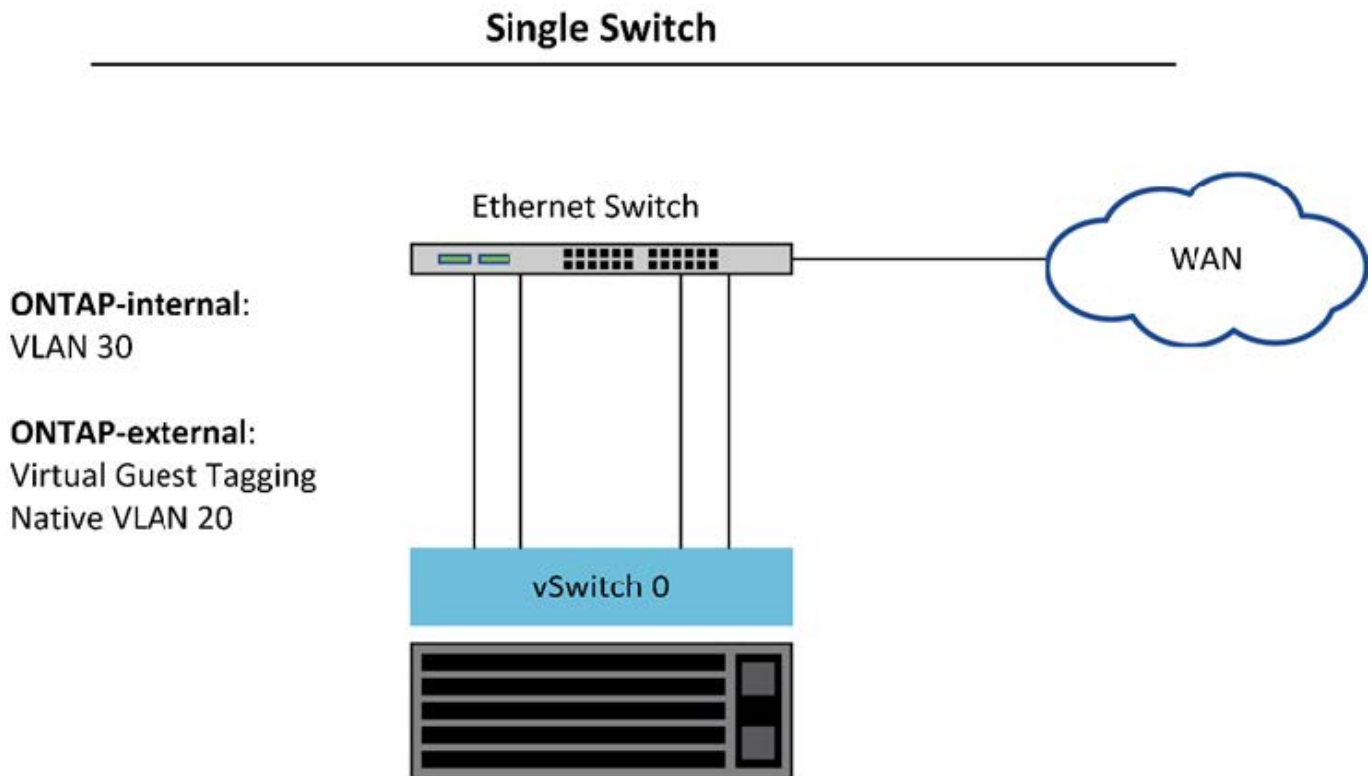
Commutateur physique partagé

La figure suivante illustre une configuration de commutateur possible utilisée par un nœud d'un cluster ONTAP Select multinœud. Dans cet exemple, les cartes réseau physiques utilisées par les commutateurs virtuels hébergeant les groupes de ports réseau internes et externes sont câblées au même commutateur en amont. Le trafic du commutateur est isolé grâce à des domaines de diffusion contenus dans des VLAN distincts.



Pour le réseau interne ONTAP Select, le marquage est effectué au niveau du groupe de ports. Bien que l'exemple suivant utilise VGT pour le réseau externe, VGT et VST sont tous deux pris en charge sur ce groupe de ports.

Configuration réseau à l'aide d'un commutateur physique partagé



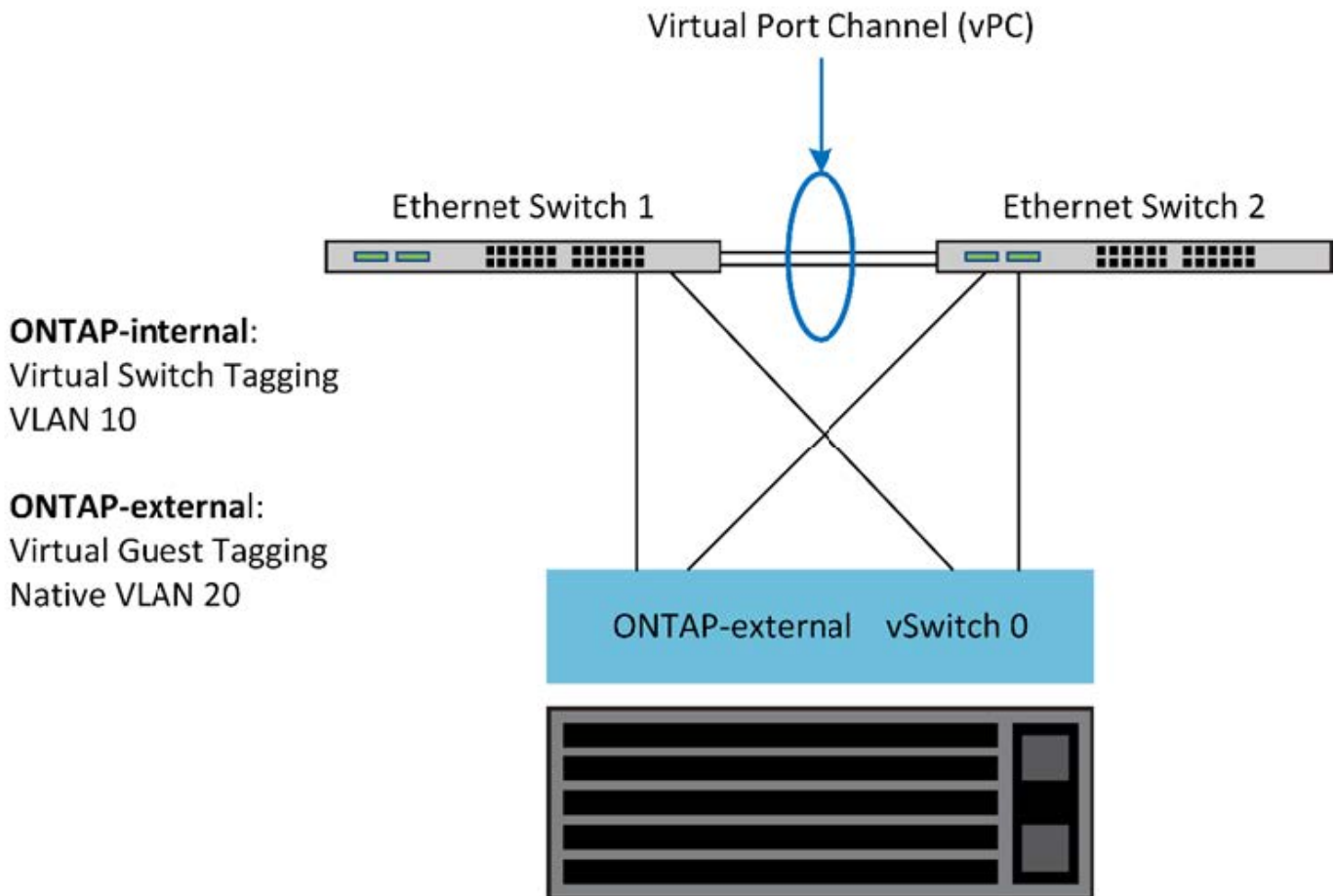
Dans cette configuration, le commutateur partagé devient un point de défaillance unique. Si possible, il est conseillé d'utiliser plusieurs commutateurs pour éviter qu'une panne matérielle physique n'entraîne une panne du réseau du cluster.

Plusieurs commutateurs physiques

Lorsque la redondance est nécessaire, il est conseillé d'utiliser plusieurs commutateurs réseau physiques. La figure suivante illustre une configuration recommandée pour un nœud d'un cluster ONTAP Select multinœud. Les cartes réseau des groupes de ports internes et externes sont câblées sur des commutateurs physiques

distincts, protégeant ainsi l'utilisateur d'une panne matérielle. Un canal de port virtuel est configuré entre les commutateurs pour éviter les problèmes de spanning tree.

Configuration du réseau utilisant plusieurs commutateurs physiques



Séparation du trafic de données et de gestion ONTAP Select

Isolez le trafic de données et le trafic de gestion dans des réseaux de couche 2 distincts.

Le trafic réseau externe ONTAP Select comprend le trafic de données (CIFS, NFS et iSCSI), de gestion et de réplication (SnapMirror). Au sein d'un cluster ONTAP, chaque type de trafic utilise une interface logique distincte, hébergée sur un port réseau virtuel. Dans la configuration multinœud d' ONTAP Select, ces ports sont désignés par les ports e0a et e0b/e0g. Dans la configuration mononœud, ils sont désignés par les ports e0a et e0b/e0c, les autres étant réservés aux services internes du cluster.

NetApp recommande d'isoler le trafic de données et le trafic de gestion dans des réseaux de couche 2 distincts. Dans l'environnement ONTAP Select, cela s'effectue à l'aide de balises VLAN. Pour ce faire, attribuez un groupe de ports VLAN à la carte réseau 1 (port e0a) pour le trafic de gestion. Vous pouvez ensuite attribuer un ou plusieurs groupes de ports distincts aux ports e0b et e0c (clusters mono-nœuds) et e0b et e0g (clusters multi-nœuds) pour le trafic de données.

Si la solution VST décrite précédemment dans ce document ne suffit pas, il peut être nécessaire de regrouper les LIF de données et de gestion sur le même port virtuel. Pour ce faire, utilisez un processus appelé VGT, dans lequel le marquage VLAN est effectué par la machine virtuelle.

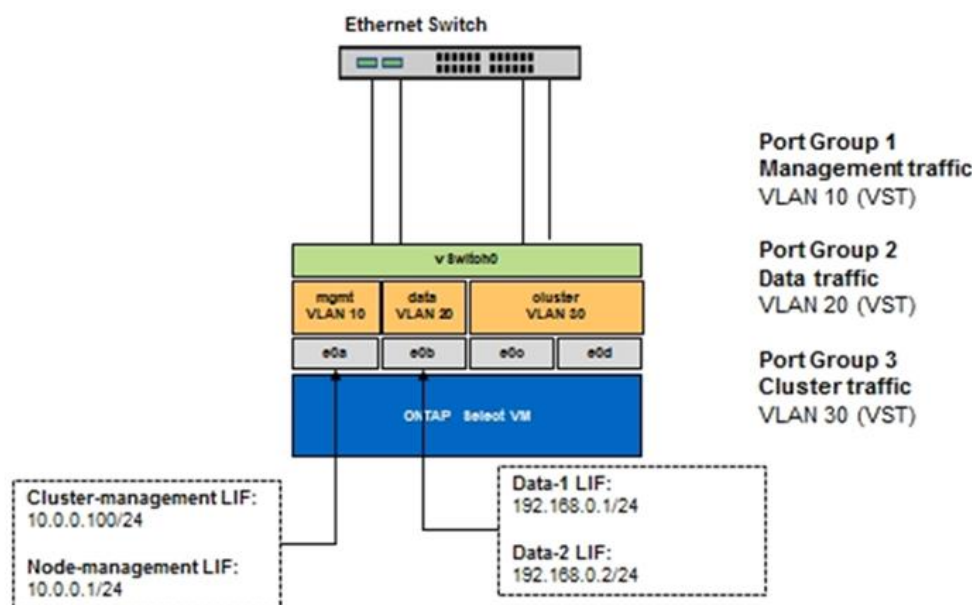


La séparation des données et du réseau de gestion via VGT n'est pas disponible avec l'utilitaire ONTAP Deploy. Cette procédure doit être effectuée une fois la configuration du cluster terminée.

Il existe une mise en garde supplémentaire lors de l'utilisation de VGT et de clusters à deux nœuds. Dans les configurations de cluster à deux nœuds, l'adresse IP de gestion du nœud est utilisée pour établir la connectivité au médiateur avant ONTAP ne soit entièrement disponible. Par conséquent, seuls les marquages EST et VST sont pris en charge sur le groupe de ports mappé à la LIF de gestion du nœud (port e0a). De plus, si le trafic de gestion et le trafic de données utilisent le même groupe de ports, seuls les protocoles EST/VST sont pris en charge pour l'ensemble du cluster à deux nœuds.

Les deux options de configuration, VST et VGT, sont prises en charge. La figure suivante illustre le premier scénario, VST, dans lequel le trafic est balisé au niveau de la couche vSwitch via le groupe de ports attribué. Dans cette configuration, les LIF de gestion de cluster et de nœud sont affectées au port ONTAP e0a et balisées avec l'ID VLAN 10 via le groupe de ports attribué. Les LIF de données sont affectées au port e0b et soit à e0c, soit à e0g, et reçoivent l'ID VLAN 20 via un deuxième groupe de ports. Les ports de cluster utilisent un troisième groupe de ports et se trouvent sur l'ID VLAN 30.

Séparation des données et de la gestion à l'aide de VST



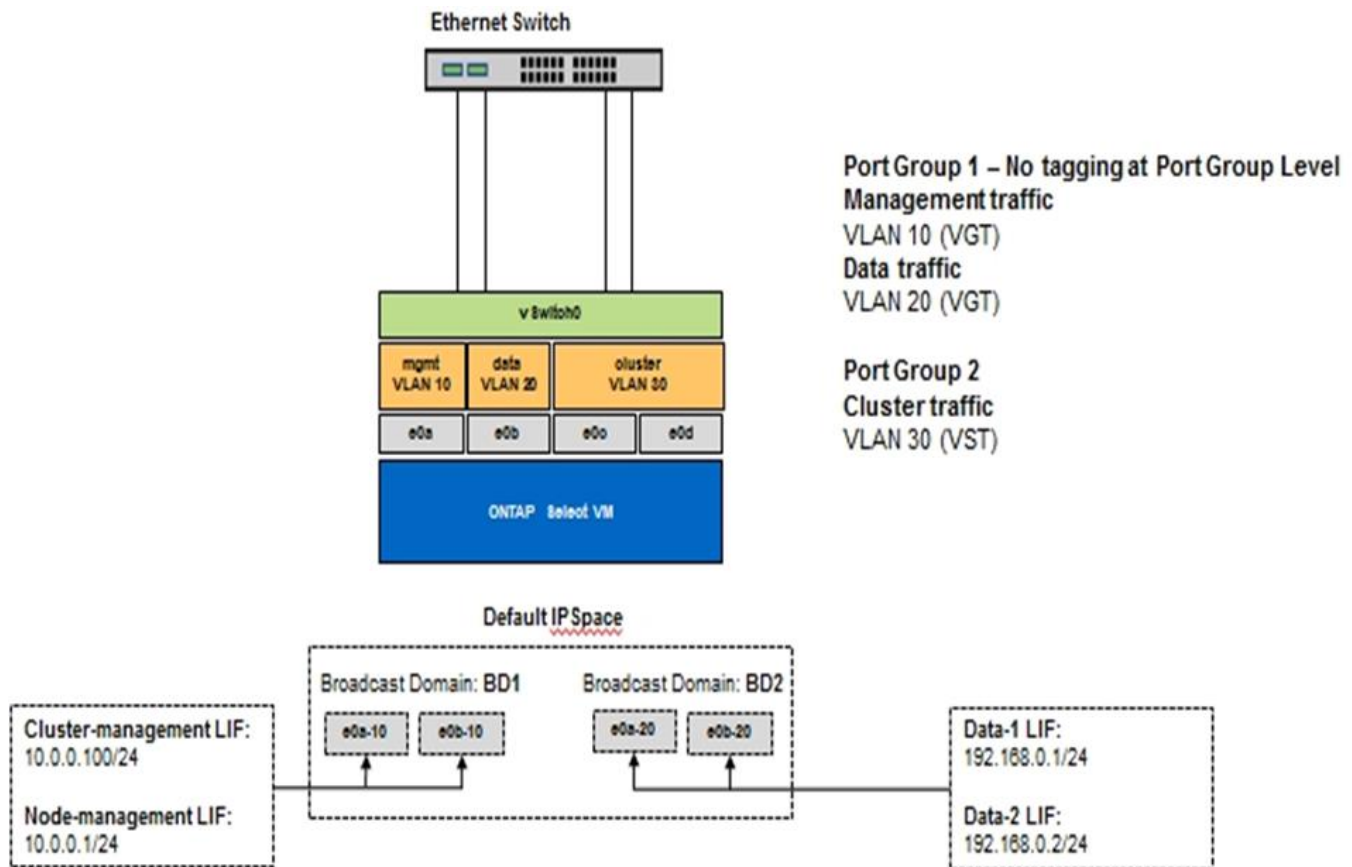
La figure suivante illustre le deuxième scénario, VGT, dans lequel le trafic est balisé par la machine virtuelle ONTAP à l'aide de ports VLAN placés dans des domaines de diffusion distincts. Dans cet exemple, les ports virtuels e0a-10/e0b-10/(e0c ou e0g)-10 et e0a-20/e0b-20 sont placés au-dessus des ports e0a et e0b de la machine virtuelle. Cette configuration permet d'effectuer le balisage réseau directement dans ONTAP, plutôt qu'au niveau de la couche vSwitch. Les LIF de gestion et de données sont placées sur ces ports virtuels, permettant une subdivision supplémentaire de la couche 2 au sein d'un même port de machine virtuelle. Le VLAN du cluster (ID VLAN 30) est toujours balisé au niveau du groupe de ports.

Remarques :

- Ce type de configuration est particulièrement utile lors de l'utilisation de plusieurs espaces IP. Regroupez les ports VLAN dans des espaces IP personnalisés distincts si vous souhaitez une isolation logique et une mutualisation renforcées.
- Pour prendre en charge VGT, les adaptateurs réseau hôtes ESXi/ESX doivent être connectés aux ports

trunk du commutateur physique. L'ID VLAN des groupes de ports connectés au commutateur virtuel doit être défini sur 4095 pour activer le trunking sur le groupe de ports.

Séparation des données et de la gestion à l'aide de VGT



Architecture haute disponibilité

Configurations de haute disponibilité ONTAP Select

Découvrez les options de haute disponibilité pour sélectionner la meilleure configuration HA pour votre environnement.

Bien que les clients commencent à migrer leurs applications des appliances de stockage professionnelles vers des solutions logicielles fonctionnant sur du matériel standard, les attentes et les besoins en matière de résilience et de tolérance aux pannes restent inchangés. Une solution haute disponibilité avec un objectif de point de récupération zéro (RPO) protège le client contre la perte de données due à une défaillance de l'un des composants de l'infrastructure.

Une grande partie du marché des SDS repose sur le concept de stockage sans partage, la réplication logicielle assurant la résilience des données en stockant plusieurs copies des données utilisateur sur différents silos de stockage. ONTAP Select s'appuie sur ce principe en utilisant les fonctionnalités de réplication synchrone (RAID SyncMirror) fournies par ONTAP pour stocker une copie supplémentaire des données utilisateur au sein du cluster. Cela se produit dans le cadre d'une paire HA. Chaque paire HA stocke deux copies des données utilisateur : une sur le stockage fourni par le nœud local et une sur le stockage fourni par le partenaire HA. Au sein d'un cluster ONTAP Select, la haute disponibilité et la réplication synchrone sont liées, et leurs fonctionnalités ne peuvent être ni dissociées ni utilisées indépendamment. Par conséquent, la fonctionnalité de

réplication synchrone est uniquement disponible dans l'offre multinœud.

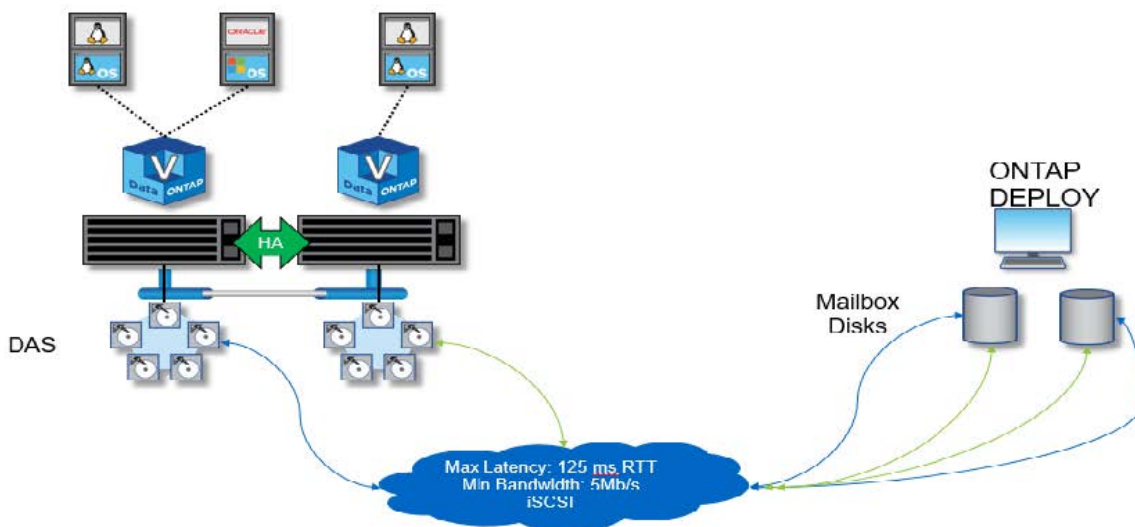


Dans un cluster ONTAP Select, la réplication synchrone est une fonction de l'implémentation HA, et non un remplacement des moteurs de réplication asynchrone SnapMirror ou SnapVault. La réplication synchrone ne peut pas être utilisée indépendamment de la haute disponibilité.

Il existe deux modèles de déploiement ONTAP Select HA : les clusters multi-nœuds (quatre, six ou huit nœuds) et les clusters à deux nœuds. La principale caractéristique d'un cluster ONTAP Select à deux nœuds est l'utilisation d'un service de médiation externe pour résoudre les scénarios de split-brain. La machine virtuelle ONTAP Deploy sert de médiateur par défaut pour toutes les paires HA à deux nœuds qu'elle configure.

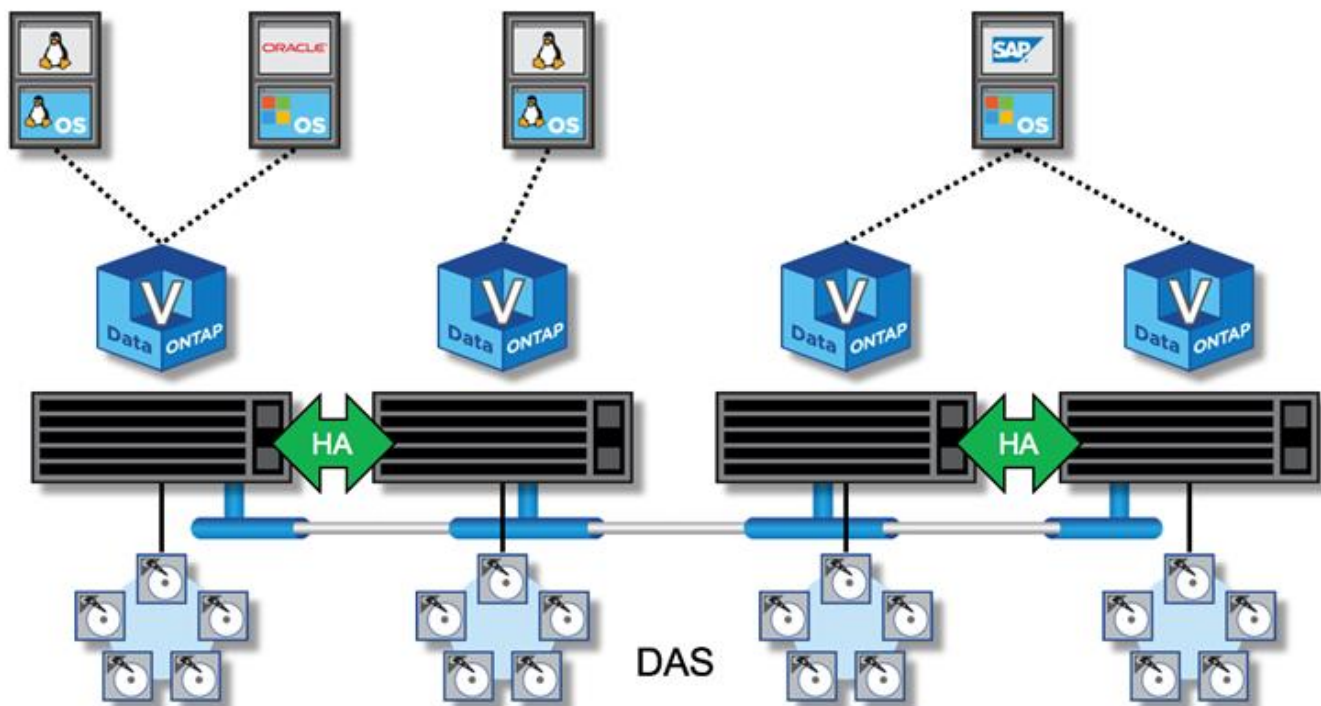
Les deux architectures sont représentées dans les figures suivantes.

Cluster ONTAP Select à deux nœuds avec médiateur distant et utilisant un stockage local



Le cluster ONTAP Select à deux nœuds est composé d'une paire HA et d'un médiateur. Au sein de cette paire HA, les agrégats de données de chaque nœud du cluster sont mis en miroir de manière synchrone, évitant ainsi toute perte de données en cas de basculement.

Cluster ONTAP Select à quatre nœuds utilisant un stockage local



- Le cluster ONTAP Select à quatre nœuds est composé de deux paires HA. Les clusters à six et huit nœuds sont composés respectivement de trois et quatre paires HA. Au sein de chaque paire HA, les agrégats de données de chaque nœud sont mis en miroir de manière synchrone, ce qui évite toute perte de données en cas de basculement.
- Une seule instance ONTAP Select peut être présente sur un serveur physique lors de l'utilisation du stockage DAS. ONTAP Select nécessite un accès non partagé au contrôleur RAID local du système et est conçu pour gérer les disques connectés localement, ce qui serait impossible sans connectivité physique au stockage.

HA à deux nœuds contre HA à plusieurs nœuds

Contrairement aux baies FAS, les nœuds ONTAP Select d'une paire HA communiquent exclusivement via le réseau IP. Cela signifie que le réseau IP constitue un point de défaillance unique (SPOF), et la protection contre les partitions réseau et les scénarios de « split-brain » devient un aspect important de la conception. Le cluster multinœud peut supporter les pannes d'un seul nœud, car le quorum du cluster peut être établi par les trois nœuds survivants ou plus. Le cluster à deux nœuds s'appuie sur le service médiateur hébergé par la machine virtuelle ONTAP Deploy pour obtenir le même résultat.

Le trafic réseau de pulsation entre les nœuds ONTAP Select et le service médiateur ONTAP Deploy est minimal et résilient, de sorte que la machine virtuelle ONTAP Deploy peut être hébergée dans un centre de données différent de celui du cluster à deux nœuds ONTAP Select.



La machine virtuelle ONTAP Deploy devient partie intégrante d'un cluster à deux nœuds lorsqu'elle sert de médiateur pour ce cluster. Si le service de médiateur n'est pas disponible, le cluster à deux nœuds continue de fournir des données, mais les fonctionnalités de basculement de stockage du cluster ONTAP Select sont désactivées. Par conséquent, le service de médiateur ONTAP Deploy doit maintenir une communication constante avec chaque nœud ONTAP Select de la paire HA. Une bande passante minimale de 5 Mbit/s et une latence aller-retour maximale (RTT) de 125 ms sont requises pour assurer le bon fonctionnement du quorum du cluster.

Si la machine virtuelle ONTAP Deploy servant de médiateur est temporairement ou potentiellement définitivement indisponible, une machine virtuelle ONTAP Deploy secondaire peut être utilisée pour restaurer le quorum du cluster à deux nœuds. Cela crée une configuration dans laquelle la nouvelle machine virtuelle ONTAP Deploy ne peut pas gérer les nœuds ONTAP Select, mais participe avec succès à l'algorithme de quorum du cluster. La communication entre les nœuds ONTAP Select et la machine virtuelle ONTAP Deploy s'effectue via le protocole iSCSI sur IPv4. L'adresse IP de gestion du nœud ONTAP Select est l'initiateur et celle de la machine virtuelle ONTAP Deploy est la cible. Par conséquent, il n'est pas possible de prendre en charge les adresses IPv6 pour les adresses IP de gestion des nœuds lors de la création d'un cluster à deux nœuds. Les disques de boîtes aux lettres hébergés par ONTAP Deploy sont automatiquement créés et masqués par les adresses IP de gestion des nœuds ONTAP Select appropriées lors de la création du cluster à deux nœuds. La configuration complète est effectuée automatiquement lors de l'installation, et aucune autre action administrative n'est requise. L'instance ONTAP Deploy créant le cluster est le médiateur par défaut pour ce cluster.

Une action administrative est requise si l'emplacement d'origine du médiateur doit être modifié. Il est possible de récupérer le quorum d'un cluster même en cas de perte de la machine virtuelle ONTAP Deploy d'origine. Cependant, NetApp recommande de sauvegarder la base de données ONTAP Deploy après chaque instantiation d'un cluster à deux nœuds.

HA à deux nœuds contre HA étendu à deux nœuds (MetroCluster SDS)

Il est possible d'étendre un cluster HA actif/actif à deux nœuds sur de plus grandes distances et de placer chaque nœud dans un centre de données différent. La seule différence entre un cluster à deux nœuds et un cluster étendu à deux nœuds (également appelé MetroCluster SDS) réside dans la distance de connectivité réseau entre les nœuds.

Un cluster à deux nœuds est défini comme un cluster dont les deux nœuds sont situés dans le même centre de données, à une distance de 300 m. En général, les deux nœuds disposent de liaisons montantes vers le même commutateur réseau ou ensemble de commutateurs ISL (Interswitch Link).

Un SDS MetroCluster à deux nœuds est défini comme un cluster dont les nœuds sont physiquement séparés (salles, bâtiments et centres de données différents) de plus de 300 m. De plus, les connexions montantes de chaque nœud sont connectées à des commutateurs réseau distincts. Le SDS MetroCluster ne nécessite pas de matériel dédié. Cependant, l'environnement doit respecter les exigences de latence (5 ms maximum pour le RTT et 5 ms pour la gigue, soit un total de 10 ms) et de distance physique (10 km maximum).

MetroCluster SDS est une fonctionnalité premium nécessitant une licence Premium ou Premium XL. La licence Premium prend en charge la création de machines virtuelles de petite et moyenne taille, ainsi que de disques durs et SSD. La licence Premium XL prend également en charge la création de disques NVMe.



MetroCluster SDS est compatible avec le stockage local (DAS) et le stockage partagé (vNAS). Notez que les configurations vNAS présentent généralement une latence intrinsèque plus élevée en raison du réseau entre la machine virtuelle ONTAP Select et le stockage partagé. Les configurations MetroCluster SDS doivent offrir une latence maximale de 10 ms entre les nœuds, latence du stockage partagé comprise. Autrement dit, mesurer uniquement la latence entre les machines virtuelles Select n'est pas suffisant, car la latence du stockage partagé n'est pas négligeable pour ces configurations.

ONTAP Select HA RSM et agrégats en miroir

Empêchez la perte de données à l'aide de RAID SyncMirror (RSM), des agrégats en miroir et du chemin d'écriture.

Réplication synchrone

Le modèle ONTAP HA repose sur le concept de partenaires HA. ONTAP Select étend cette architecture aux serveurs standard non partagés en utilisant la fonctionnalité RAID SyncMirror (RSM) présente dans ONTAP pour répliquer les blocs de données entre les nœuds du cluster, fournissant ainsi deux copies des données utilisateur réparties sur une paire HA.

Un cluster à deux nœuds avec médiateur peut couvrir deux centres de données. Pour plus d'informations, consultez la section ["Bonnes pratiques pour la haute disponibilité étendue à deux nœuds \(MetroCluster SDS\)"](#).

Agrégats en miroir

Un cluster ONTAP Select est composé de deux à huit nœuds. Chaque paire HA contient deux copies des données utilisateur, mises en miroir de manière synchrone entre les nœuds via un réseau IP. Cette mise en miroir est transparente pour l'utilisateur et constitue une propriété de l'agrégat de données, configurée automatiquement lors de sa création.

Tous les agrégats d'un cluster ONTAP Select doivent être mis en miroir pour garantir la disponibilité des données en cas de basculement d'un nœud et éviter un SPOF en cas de panne matérielle. Les agrégats d'un cluster ONTAP Select sont créés à partir de disques virtuels fournis par chaque nœud de la paire HA et utilisent les disques suivants :

- Un ensemble local de disques (fourni par le nœud ONTAP Select actuel)
- Un ensemble de disques en miroir (fourni par le partenaire HA du nœud actuel)



Les disques locaux et miroirs utilisés pour créer un agrégat en miroir doivent être de même taille. Ces agrégats sont appelés plex 0 et plex 1 (pour indiquer respectivement les paires de miroirs locaux et distants). Les numéros de plex réels peuvent être différents dans votre installation.

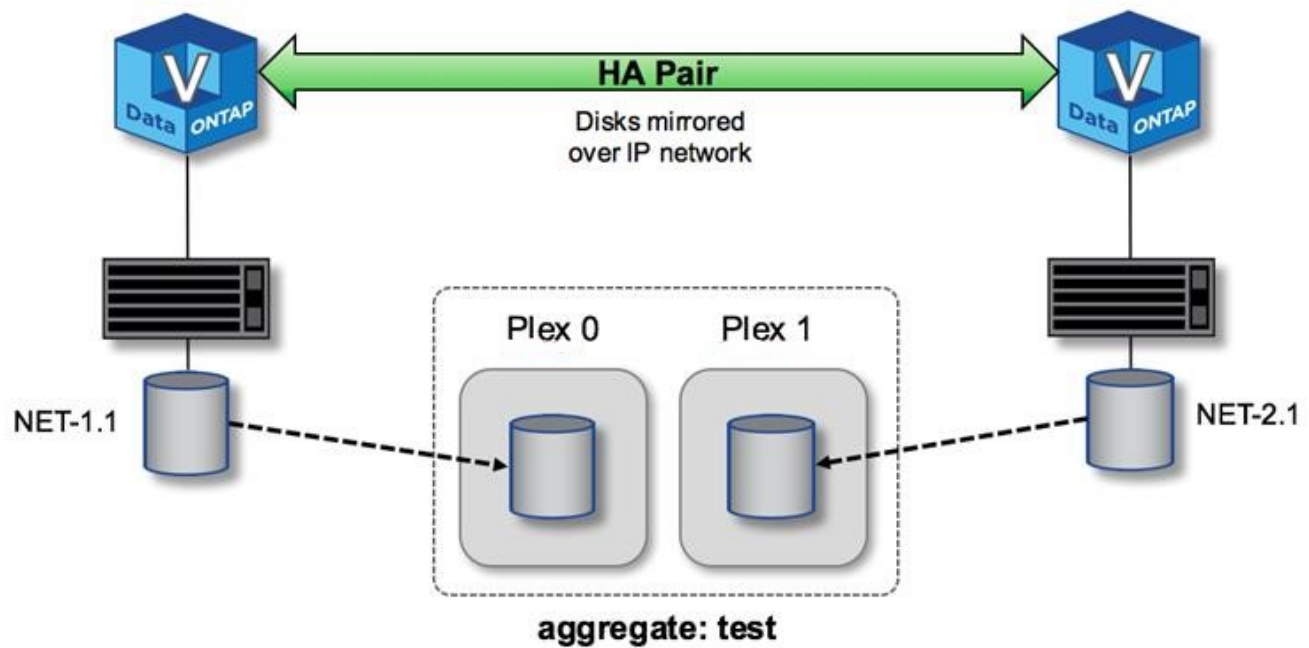
Cette approche est fondamentalement différente du fonctionnement des clusters ONTAP standard. Elle s'applique à tous les disques racine et de données du cluster ONTAP Select. L'agrégat contient des copies locales et miroir des données. Par conséquent, un agrégat contenant N disques virtuels offre un stockage unique équivalent à N/2 disques, car la seconde copie des données réside sur ses propres disques.

La figure suivante illustre une paire HA au sein d'un cluster ONTAP Select à quatre nœuds. Ce cluster contient un agrégat unique (test) qui utilise le stockage des deux partenaires HA. Cet agrégat de données est composé de deux ensembles de disques virtuels : un ensemble local, fourni par le nœud de cluster propriétaire d'ONTAP Select (Plex 0), et un ensemble distant, fourni par le partenaire de basculement (Plex 1).

Le Plex 0 est le compartiment qui contient tous les disques locaux. Le Plex 1 est le compartiment qui contient les disques miroirs, c'est-à-dire les disques chargés de stocker une seconde copie répliquée des données utilisateur. Le nœud propriétaire de l'agrégat fournit des disques au Plex 0, et le partenaire HA de ce nœud fournit des disques au Plex 1.

Dans la figure suivante, un agrégat en miroir avec deux disques est utilisé. Son contenu est mis en miroir sur nos deux nœuds de cluster : le disque local NET-1.1 est placé dans le compartiment Plex 0 et le disque distant NET-2.1 dans le compartiment Plex 1. Dans cet exemple, le test d'agrégat appartient au nœud de cluster de gauche et utilise le disque local NET-1.1 et le disque miroir partenaire HA NET-2.1.

- ONTAP Select agrégat en miroir*



Lors du déploiement d'un cluster ONTAP Select, tous les disques virtuels présents sur le système sont automatiquement attribués au plex approprié, sans intervention supplémentaire de l'utilisateur. Cela évite l'attribution accidentelle de disques à un plex incorrect et assure une configuration optimale des disques miroirs.

Chemin d'écriture

La mise en miroir synchrone des blocs de données entre les nœuds du cluster et l'obligation d'éviter toute perte de données en cas de panne système ont un impact significatif sur le chemin emprunté par une écriture entrante lors de sa propagation dans un cluster ONTAP Select. Ce processus se déroule en deux étapes :

- Reconnaissance
- Démantèlement

Les écritures sur un volume cible s'effectuent via une LIF de données et sont validées sur la partition NVRAM virtualisée, présente sur un disque système du nœud ONTAP Select, avant d'être confirmées au client. Dans une configuration haute disponibilité, une étape supplémentaire est nécessaire : ces écritures NVRAM sont immédiatement mises en miroir sur le partenaire haute disponibilité du propriétaire du volume cible avant d'être confirmées. Ce processus garantit la cohérence du système de fichiers sur le nœud partenaire haute disponibilité en cas de panne matérielle sur le nœud d'origine.

Une fois l'écriture validée dans la NVRAM, ONTAP déplace périodiquement le contenu de cette partition vers le disque virtuel approprié, un processus appelé désactivation. Ce processus n'a lieu qu'une seule fois, sur le nœud de cluster propriétaire du volume cible, et non sur le partenaire haute disponibilité.

La figure suivante montre le chemin d'écriture d'une demande d'écriture entrante vers un nœud ONTAP Select

- Flux de travail de ONTAP Select *

cluster, généralement causé par des pannes réseau, où chaque côté croit que l'autre est en panne et tente de s'approprier les ressources du cluster.

Les implémentations HA d'entreprise doivent gérer ce type de scénario avec élégance. ONTAP y parvient grâce à une méthode de pulsation personnalisée sur disque. Cette tâche incombe à la boîte aux lettres HA, un emplacement sur le stockage physique utilisé par les nœuds du cluster pour transmettre les pulsations. Cela permet au cluster de déterminer la connectivité et donc de définir le quorum en cas de basculement.

Sur les baies FAS, qui utilisent une architecture HA de stockage partagé, ONTAP résout les problèmes de split-brain des manières suivantes :

- Réservations persistantes SCSI
- Métadonnées HA persistantes
- État HA envoyé via l'interconnexion HA

Cependant, dans l'architecture « shared-nothing » d'un cluster ONTAP Select, un nœud ne peut voir que son propre stockage local et non celui du partenaire HA. Par conséquent, lorsque le partitionnement réseau isole chaque côté d'une paire HA, les méthodes précédentes de détermination du quorum du cluster et du comportement de basculement ne sont pas disponibles.

Bien que la méthode actuelle de détection et d'évitement du split-brain ne soit pas utilisable, une méthode de médiation, compatible avec les contraintes d'un environnement sans partage, reste nécessaire. ONTAP Select étend l'infrastructure de boîtes aux lettres existante en lui permettant d'agir comme méthode de médiation en cas de partitionnement du réseau. Le stockage partagé étant indisponible, la médiation s'effectue via l'accès aux disques de boîtes aux lettres via le NAS. Ces disques sont répartis dans le cluster, y compris le médiateur dans un cluster à deux nœuds, via le protocole iSCSI. Ainsi, des décisions de basculement intelligentes peuvent être prises par un nœud de cluster en fonction de l'accès à ces disques. Si un nœud peut accéder aux disques de boîtes aux lettres d'autres nœuds en dehors de son partenaire haute disponibilité, il est probablement opérationnel et en bon état.



L'architecture de la boîte aux lettres et la méthode de pulsation basée sur le disque pour résoudre les problèmes de quorum de cluster et de split-brain sont les raisons pour lesquelles la variante multinœud d'ONTAP Select nécessite soit quatre nœuds distincts, soit un médiateur pour un cluster à deux nœuds.

Publication dans la boîte aux lettres HA

L'architecture de boîte aux lettres HA utilise un modèle de publication de messages. À intervalles réguliers, les nœuds du cluster envoient des messages à tous les autres disques de boîte aux lettres du cluster, y compris le médiateur, indiquant que le nœud est opérationnel. Dans un cluster sain, à tout moment, un seul disque de boîte aux lettres sur un nœud du cluster reçoit des messages de tous les autres nœuds du cluster.

Un disque virtuel est attaché à chaque nœud du cluster Select, spécifiquement utilisé pour l'accès aux boîtes aux lettres partagées. Ce disque est appelé disque de boîte aux lettres médiateur, car sa fonction principale est de servir de méthode de médiation du cluster en cas de panne de nœud ou de partitionnement réseau. Ce disque contient des partitions pour chaque nœud du cluster et est monté sur un réseau iSCSI par d'autres nœuds du cluster Select. Ces nœuds publient régulièrement des états d'intégrité sur la partition correspondante du disque de boîte aux lettres. L'utilisation de disques de boîte aux lettres accessibles par le réseau, répartis sur l'ensemble du cluster, permet de déduire l'état de santé des nœuds grâce à une matrice d'accessibilité. Par exemple, les nœuds A et B du cluster peuvent publier des messages sur la boîte aux lettres du nœud D, mais pas sur celle du nœud C. De plus, le nœud D ne peut pas publier de messages sur la boîte aux lettres du nœud C ; il est donc probable que le nœud C soit en panne ou isolé du réseau et doive être pris en charge.

HA battements de cœur

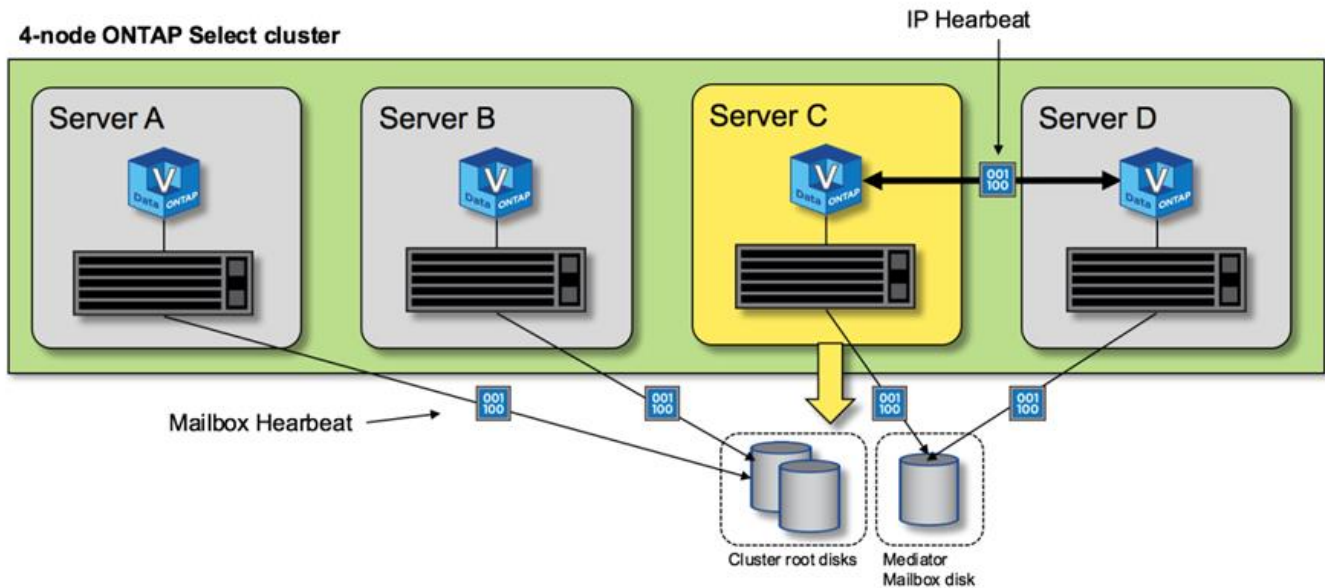
Comme avec les plateformes NetApp FAS , ONTAP Select envoie régulièrement des messages de pulsation HA via l'interconnexion HA. Au sein du cluster ONTAP Select , cette opération s'effectue via une connexion réseau TCP/IP entre les partenaires HA. De plus, les messages de pulsation sur disque sont transmis à tous les disques de boîtes aux lettres HA, y compris les disques de boîtes aux lettres médiatrices. Ces messages sont transmis toutes les quelques secondes et relus périodiquement. La fréquence d'envoi et de réception permet au cluster ONTAP Select de détecter les pannes HA en environ 15 secondes, soit le même délai que sur les plateformes FAS . Lorsque les messages de pulsation ne sont plus lus, un basculement est déclenché.

La figure suivante montre le processus d'envoi et de réception de messages de pulsation sur les disques d'interconnexion et de médiateur HA du point de vue d'un seul nœud de cluster ONTAP Select , le nœud C.



Les pulsations du réseau sont envoyées via l'interconnexion HA au partenaire HA, le nœud D, tandis que les pulsations du disque utilisent des disques de boîte aux lettres sur tous les nœuds du cluster, A, B, C et D.

Haute fréquence cardiaque dans un cluster à quatre nœuds : état stable



Basculement et retour HA

Lors d'un basculement, le nœud survivant assume les responsabilités de service de données pour son nœud homologue en utilisant la copie locale des données de son partenaire HA. Les E/S client peuvent se poursuivre sans interruption, mais les modifications apportées à ces données doivent être répliquées avant la restitution. Notez qu'ONTAP Select ne prend pas en charge la restitution forcée, car cela entraîne la perte des modifications stockées sur le nœud survivant.

La synchronisation est automatiquement déclenchée lorsque le nœud redémarré rejoint le cluster. Le temps nécessaire à la synchronisation dépend de plusieurs facteurs, notamment le nombre de modifications à répliquer, la latence réseau entre les nœuds et la vitesse des sous-systèmes de disque sur chaque nœud. Il est possible que le temps nécessaire à la synchronisation dépasse la fenêtre de retour automatique de 10 minutes. Dans ce cas, une synchronisation manuelle est nécessaire après la synchronisation. La progression de la synchronisation peut être surveillée à l'aide de la commande suivante :

```
storage aggregate status -r -aggregate <aggregate name>
```

Performances

Aperçu des performances ONTAP Select

Les performances d'un cluster ONTAP Select peuvent varier considérablement en fonction des caractéristiques du matériel et de la configuration sous-jacents. La configuration matérielle spécifique est le facteur le plus important pour les performances d'une instance ONTAP Select donnée. Voici quelques facteurs qui influencent les performances d'une instance ONTAP Select spécifique :

- **Fréquence principale.** En général, une fréquence plus élevée est préférable.
- **Socket unique ou multisocket.** ONTAP Select n'utilise pas les fonctionnalités multisocket, mais la surcharge de l'hyperviseur pour la prise en charge des configurations multisocket explique une certaine variation des performances globales.
- **Configuration de la carte RAID et pilote d'hyperviseur associé.** Le pilote par défaut fourni par l'hyperviseur devra peut-être être remplacé par le pilote du fournisseur du matériel.
- **Type de lecteur et nombre de lecteurs dans le(s) groupe(s) RAID.**
- **Version de l'hyperviseur et niveau de correctif.**

Performances ONTAP Select 9.6 : stockage SSD à connexion directe HA haut de gamme

Informations sur les performances de la plateforme de référence.

Plateforme de référence

Matériel ONTAP Select (Premium XL) (par nœud)

- FUJITSU PRIMERGY RX2540 M4 :
 - Processeur Intel® Xeon® Gold 6142b à 2,6 GHz
 - 32 cœurs physiques (16 x 2 sockets), 64 cœurs logiques
 - 256 Go de RAM
 - Disques par hôte : 24 SSD de 960 Go
 - ESX 6.5U1

Matériel client

- 5 clients NFSv3 IBM 3550m4

Informations de configuration

- SW RAID 1 x 9 + 2 RAID-DP (11 disques)
- 22+1 RAID-5 (RAID-0 dans ONTAP) / cache RAID NVRAM

- Aucune fonctionnalité d'efficacité de stockage n'est utilisée (compression, déduplication, copies Snapshot, SnapMirror, etc.)

Le tableau suivant répertorie le débit mesuré par rapport aux charges de travail en lecture/écriture sur une paire de nœuds ONTAP Select haute disponibilité (HA) utilisant à la fois le RAID logiciel et le RAID matériel. Les mesures de performances ont été réalisées à l'aide de l'outil de génération de charge SIO.



Ces chiffres de performances sont basés sur ONTAP Select 9.6.

Résultats de performances pour un cluster ONTAP Select à nœud unique (partie d'une instance moyenne à quatre nœuds) sur un SSD de stockage à connexion directe (DAS), avec RAID logiciel et RAID matériel

Description	Lecture séquentielle 64 Ko	Écriture séquentielle 64 Ko	Lecture aléatoire 8 Ko	Écriture aléatoire 8 Ko	WR/RD aléatoire (50/50) 8 Ko
ONTAP Select grande instance avec logiciel RAID DAS (SSD)	2171 Mbit/s	559 Mbit/s	954 Mbit/s	394 Mbit/s	564 Mbit/s
Instance moyenne ONTAP Select avec logiciel RAID DAS (SSD)	2090 Mbit/s	592 Mbit/s	677 Mbit/s	335 Mbit/s	441 3 Mbit/s
Instance moyenne ONTAP Select avec RAID matériel DAS (SSD)	2038 Mbit/s	520 Mbit/s	578 Mbit/s	325 Mbit/s	399 Mbit/s

Lecture séquentielle de 64 K

Détails:

- E/S directes SIO activées
- 2 nœuds
- 2 cartes réseau de données par nœud
- 1 x agrégat de données par nœud (RAID matériel 2 To), (RAID logiciel 8 To)
- 64 processus SIO, 1 thread par processus
- 32 volumes par nœud
- 1 x fichiers par processus ; les fichiers font 12 000 Mo chacun

Écriture séquentielle de 64 Ko

Détails:

- E/S directes SIO activées
- 2 nœuds
- 2 cartes d'interface réseau de données (NIC) par nœud
- 1 x agrégat de données par nœud (RAID matériel 2 To), (RAID logiciel 4 To)
- 128 processus SIO, 1 thread par processus
- Volumes par nœud : 32 (RAID matériel), 16 (RAID logiciel)
- 1 x fichiers par processus ; les fichiers font 30 720 Mo chacun

Lecture aléatoire de 8 Ko

Détails:

- E/S directes SIO activées
- 2 nœuds
- 2 cartes réseau de données par nœud
- 1 x agrégat de données par nœud (RAID matériel 2 To), (RAID logiciel 4 To)
- 64 processus SIO, 8 threads par processus
- Volumes par nœud : 32
- 1 x fichiers par processus ; les fichiers font 12 228 Mo chacun

Écriture aléatoire de 8 Ko

Détails:

- E/S directes SIO activées
- 2 nœuds
- 2 cartes réseau de données par nœud
- 1 x agrégat de données par nœud (RAID matériel 2 To), (RAID logiciel 4 To)
- 64 processus SIO, 8 threads par processus
- Volumes par nœud : 32
- 1 x fichiers par processus ; les fichiers font 8 192 Mo chacun

8K aléatoire 50% écriture 50% lecture

Détails:

- E/S directes SIO activées
- 2 nœuds
- 2 cartes réseau de données par nœud
- 1 x agrégat de données par nœud (RAID matériel 2 To), (RAID logiciel 4 To)
- 64 SIO proc208 threads par proc
- Volumes par nœud : 32
- 1 x fichiers par processus ; les fichiers font 12 228 Mo chacun

Informations sur le copyright

Copyright © 2026 NetApp, Inc. Tous droits réservés. Imprimé aux États-Unis. Aucune partie de ce document protégé par copyright ne peut être reproduite sous quelque forme que ce soit ou selon quelque méthode que ce soit (graphique, électronique ou mécanique, notamment par photocopie, enregistrement ou stockage dans un système de récupération électronique) sans l'autorisation écrite préalable du détenteur du droit de copyright.

Les logiciels dérivés des éléments NetApp protégés par copyright sont soumis à la licence et à l'avis de non-responsabilité suivants :

CE LOGICIEL EST FOURNI PAR NETAPP « EN L'ÉTAT » ET SANS GARANTIES EXPRESSES OU TACITES, Y COMPRIS LES GARANTIES TACITES DE QUALITÉ MARCHANDE ET D'ADÉQUATION À UN USAGE PARTICULIER, QUI SONT EXCLUES PAR LES PRÉSENTES. EN AUCUN CAS NETAPP NE SERA TENU POUR RESPONSABLE DE DOMMAGES DIRECTS, INDIRECTS, ACCESSOIRES, PARTICULIERS OU EXEMPLAIRES (Y COMPRIS L'ACHAT DE BIENS ET DE SERVICES DE SUBSTITUTION, LA PERTE DE JOUISSANCE, DE DONNÉES OU DE PROFITS, OU L'INTERRUPTION D'ACTIVITÉ), QUELLES QU'EN SOIENT LA CAUSE ET LA DOCTRINE DE RESPONSABILITÉ, QU'IL S'AGISSE DE RESPONSABILITÉ CONTRACTUELLE, STRICTE OU DÉLICTELLE (Y COMPRIS LA NÉGLIGENCE OU AUTRE) DÉCOULANT DE L'UTILISATION DE CE LOGICIEL, MÊME SI LA SOCIÉTÉ A ÉTÉ INFORMÉE DE LA POSSIBILITÉ DE TELS DOMMAGES.

NetApp se réserve le droit de modifier les produits décrits dans le présent document à tout moment et sans préavis. NetApp décline toute responsabilité découlant de l'utilisation des produits décrits dans le présent document, sauf accord explicite écrit de NetApp. L'utilisation ou l'achat de ce produit ne concède pas de licence dans le cadre de droits de brevet, de droits de marque commerciale ou de tout autre droit de propriété intellectuelle de NetApp.

Le produit décrit dans ce manuel peut être protégé par un ou plusieurs brevets américains, étrangers ou par une demande en attente.

LÉGENDE DE RESTRICTION DES DROITS : L'utilisation, la duplication ou la divulgation par le gouvernement sont sujettes aux restrictions énoncées dans le sous-paragraphe (b)(3) de la clause Rights in Technical Data-Noncommercial Items du DFARS 252.227-7013 (février 2014) et du FAR 52.227-19 (décembre 2007).

Les données contenues dans les présentes se rapportent à un produit et/ou service commercial (tel que défini par la clause FAR 2.101). Il s'agit de données propriétaires de NetApp, Inc. Toutes les données techniques et tous les logiciels fournis par NetApp en vertu du présent Accord sont à caractère commercial et ont été exclusivement développés à l'aide de fonds privés. Le gouvernement des États-Unis dispose d'une licence limitée irrévocable, non exclusive, non cessible, non transférable et mondiale. Cette licence lui permet d'utiliser uniquement les données relatives au contrat du gouvernement des États-Unis d'après lequel les données lui ont été fournies ou celles qui sont nécessaires à son exécution. Sauf dispositions contraires énoncées dans les présentes, l'utilisation, la divulgation, la reproduction, la modification, l'exécution, l'affichage des données sont interdits sans avoir obtenu le consentement écrit préalable de NetApp, Inc. Les droits de licences du Département de la Défense du gouvernement des États-Unis se limitent aux droits identifiés par la clause 252.227-7015(b) du DFARS (février 2014).

Informations sur les marques commerciales

NETAPP, le logo NETAPP et les marques citées sur le site <http://www.netapp.com/TM> sont des marques déposées ou des marques commerciales de NetApp, Inc. Les autres noms de marques et de produits sont des marques commerciales de leurs propriétaires respectifs.