Analyse approfondie

ONTAP Select

NetApp
November 09, 2026

This PDF was generated from https://docs.netapp.com/fr-fr/ontap-select-9191/concept_stor_concepts_chars.html on November 09, 2026. Always check docs.netapp.com for the latest.

Sommaire

Analyse approfondie . . . 1
- Stockage . . . 1
  - Découvrez les caractéristiques générales de stockage d'ONTAP Select . . . 1
  - Services RAID matériels pour ONTAP Select avec stockage local connecté . . . 6
  - Services de configuration RAID du logiciel ONTAP Select pour le stockage local connecté . . . 13
  - ONTAP Select vSAN et configurations de baies externes . . . 22
  - Considérations relatives à l'augmentation de la capacité de stockage d'ONTAP Select . . . 26
  - Prise en charge de l'efficacité du stockage ONTAP Select . . . 30
- Réseautique . . . 32
  - Concepts et caractéristiques de mise en réseau ONTAP Select . . . 32
  - Configurations réseau ONTAP Select à nœud unique et à nœuds multiples . . . 34
  - Réseaux internes et externes ONTAP Select . . . 39
  - Configurations réseau ONTAP Select prises en charge . . . 42
  - Configuration ONTAP Select VMware vSphere vSwitch sur ESXi . . . 43
  - Configuration du commutateur physique ONTAP Select . . . 53
  - Séparation du trafic de données et de gestion ONTAP Select . . . 55
- architecture à haute disponibilité . . . 57
  - Configurations à haute disponibilité ONTAP Select . . . 57
  - ONTAP Select HA RSM et agrégats en miroir . . . 60
  - Améliorations de la protection des données avec ONTAP Select HA . . . 63
- Découvrez les performances d'ONTAP Select . . . 66

Analyse approfondie

Stockage

Découvrez les caractéristiques générales de stockage d'ONTAP Select

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 comprennent les éléments suivants :

- Prérequis avant déploiement
  - Assurez-vous que chaque hôte hyperviseur est configuré et prêt pour un déploiement ONTAP Select.
  - La configuration comprend les disques physiques, les contrôleurs RAID et groupes RAID, les LUN, ainsi que la préparation du réseau associé.
  - Cette configuration est effectuée en dehors d'ONTAP Select.
- Configuration à l'aide de l'utilitaire d'administration 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 Deploy pour configurer les structures de stockage logique de base.
  - Cela se fait soit explicitement via des commandes CLI, soit automatiquement par l'utilitaire dans le cadre d'un déploiement.
- Configuration post-déploiement
  - Une fois le déploiement d'ONTAP Select terminé, vous pouvez configurer le cluster à l'aide de l'interface de ligne de commande ONTAP ou de System Manager.
  - Cette configuration est effectuée en dehors d'ONTAP Select Deploy.

Stockage géré versus stockage non géré

Le stockage auquel ONTAP Select accède et qu'il contrôle directement est un stockage géré. Tout autre stockage sur le même hôte hyperviseur est un stockage non géré.

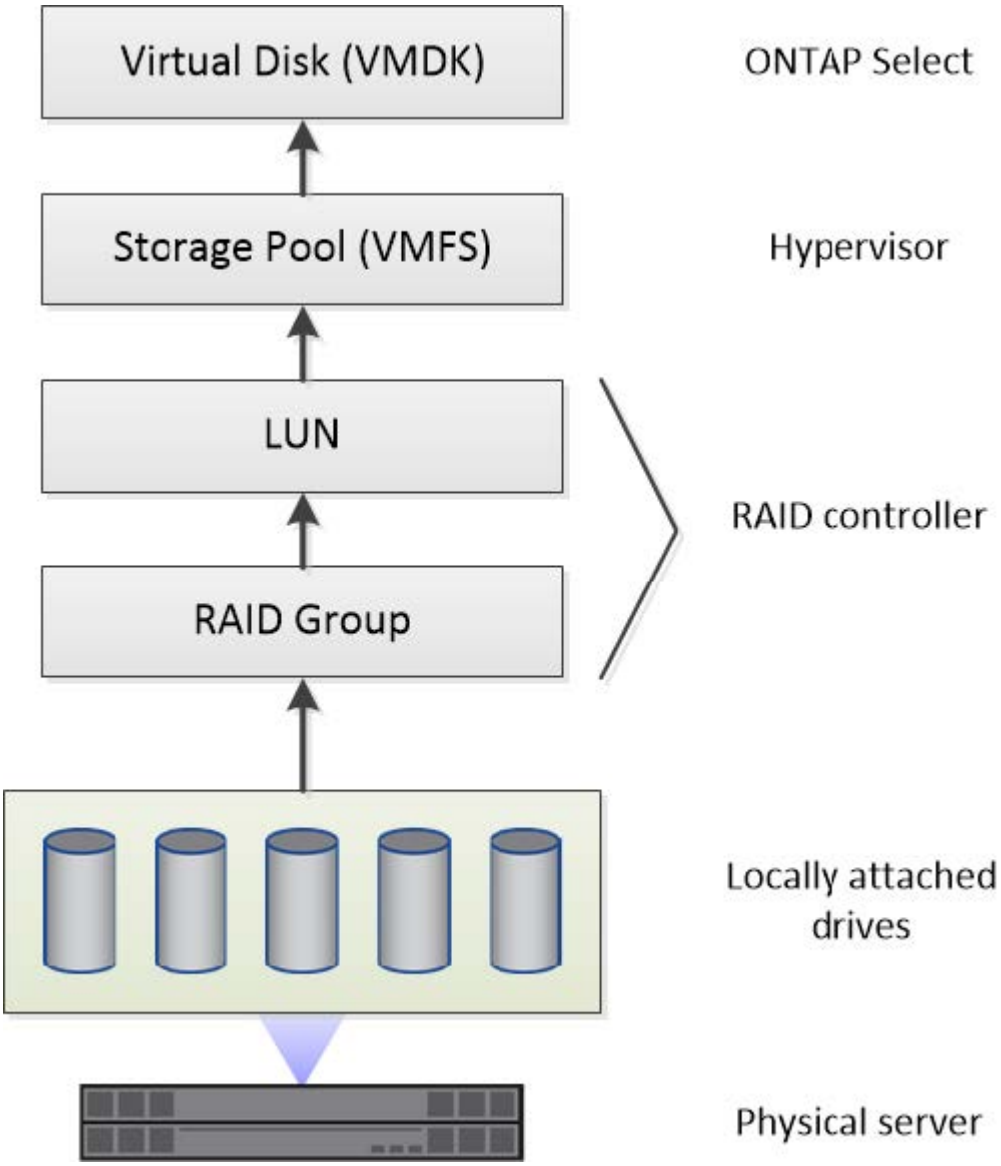
stockage physique homogène

Tous les disques physiques composant le stockage géré ONTAP Select doivent être homogènes. Autrement dit, tout le matériel doit être identique en ce qui concerne les caractéristiques suivantes :

- 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 logique pouvant être utilisés par ONTAP Select. Ces composants de stockage sont organisés en une structure 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. Vous devez être familiarisé avec ces concepts avant de préparer un déploiement ONTAP Select. Ces concepts sont classés par catégorie : groupes RAID et LUN, pools de stockage et disques virtuels.

Regroupement des disques physiques en groupes RAID et en LUN

Un ou plusieurs disques physiques peuvent être connectés localement au serveur hôte et mis à disposition pour ONTAP Select. Les disques physiques sont affectés à des groupes RAID, qui sont ensuite présentés au système d'exploitation hôte de l'hyperviseur comme une ou plusieurs LUN. Chaque LUN est présentée 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 points 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, vous pouvez envisager de créer plus d'un groupe RAID. Reportez-vous à "Résumé des meilleures pratiques".

Considérations relatives au pool de stockage

Plusieurs points relatifs aux pools de stockage doivent être pris en compte lors de la préparation du déploiement ONTAP Select.

Dans un environnement VMware, un pool de stockage est synonyme de datastore VMware.

Pools de stockage et LUNs

Chaque LUN est considérée 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 que le système d'exploitation de l'hôte hyperviseur peut utiliser.

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. Le pool de stockage est ensuite transmis à l'utilitaire d'administration ONTAP Select Deploy.

Gérer les disques virtuels sur ESXi

Plusieurs points relatifs aux disques virtuels doivent être pris en compte lors 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 correspond en réalité à un fichier contenu dans un pool de stockage et est géré par l'hyperviseur. Il existe plusieurs types de disques utilisés par ONTAP Select, principalement des disques système et des disques de données.

Vous devez également prendre en compte les 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 utiliser l'utilitaire d'administration ONTAP Select Deploy pour créer tous les disques virtuels (c'est-à-dire qu'un administrateur ne doit jamais créer de 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

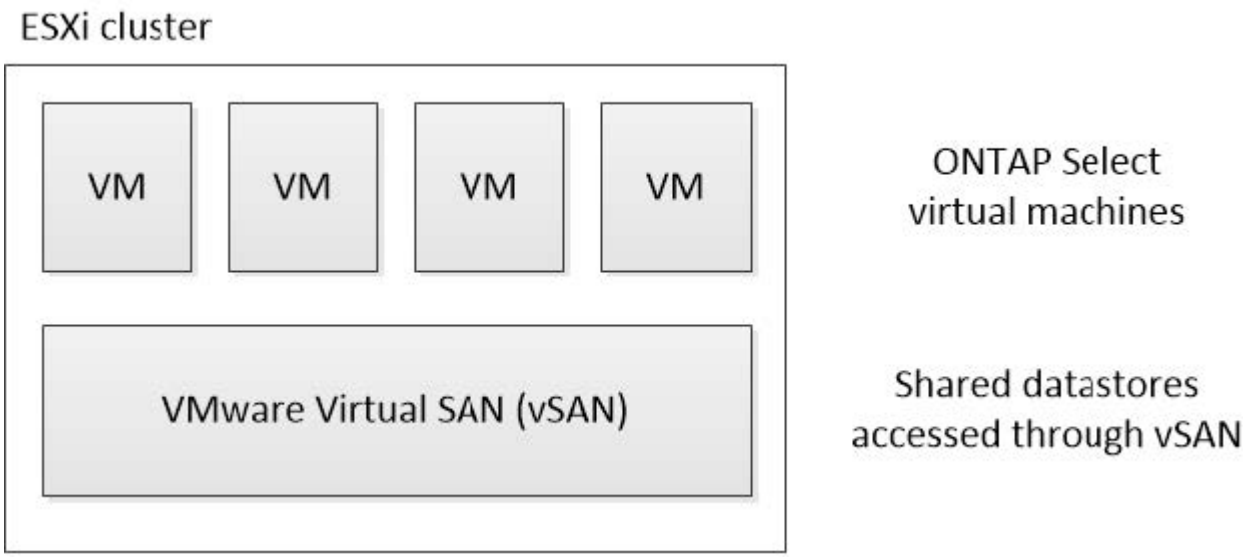
La solution ONTAP Select vNAS permet à ONTAP Select d'utiliser des banques de données résidant sur un stockage externe à l'hôte de l'hyperviseur. Les banques de données peuvent être accessibles via le réseau à l'aide de VMware vSAN ou directement sur une baie de stockage externe.

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

- vSAN (Virtual SAN)
- VMFS
- NFS

banques de données vSAN

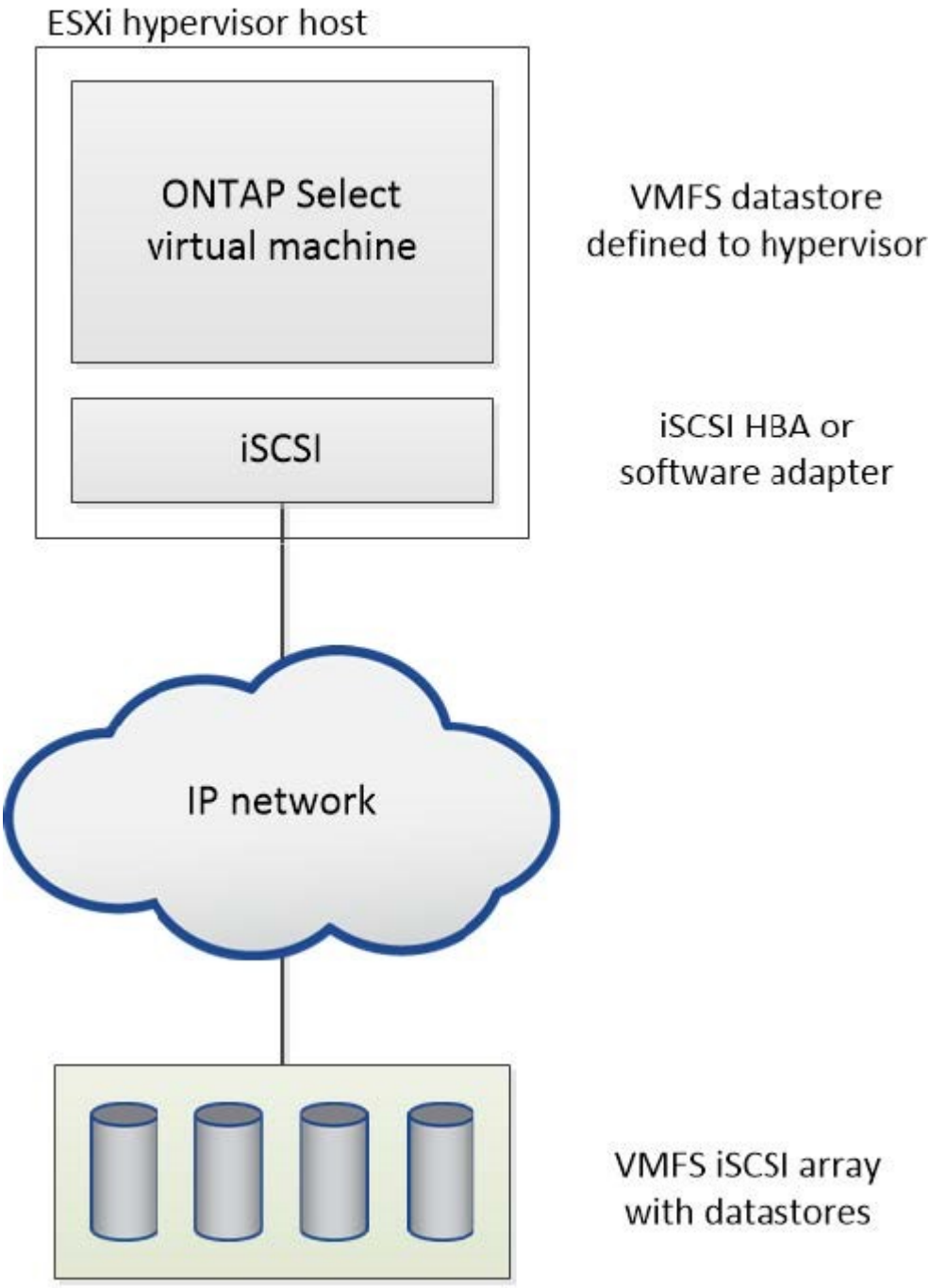
Chaque hôte ESXi peut disposer d'une ou plusieurs banques de données VMFS locales. Normalement, ces banques de données ne sont accessibles qu'à l'hôte local. Cependant, VMware vSAN permet à chaque hôte d'un cluster ESXi de partager toutes les banques de données du cluster comme s'il s'agissait de banques locales. La figure suivante illustre comment vSAN crée un pool de banques de données partagées entre les hôtes du cluster ESXi.



Stockage de données VMFS sur baie de stockage externe

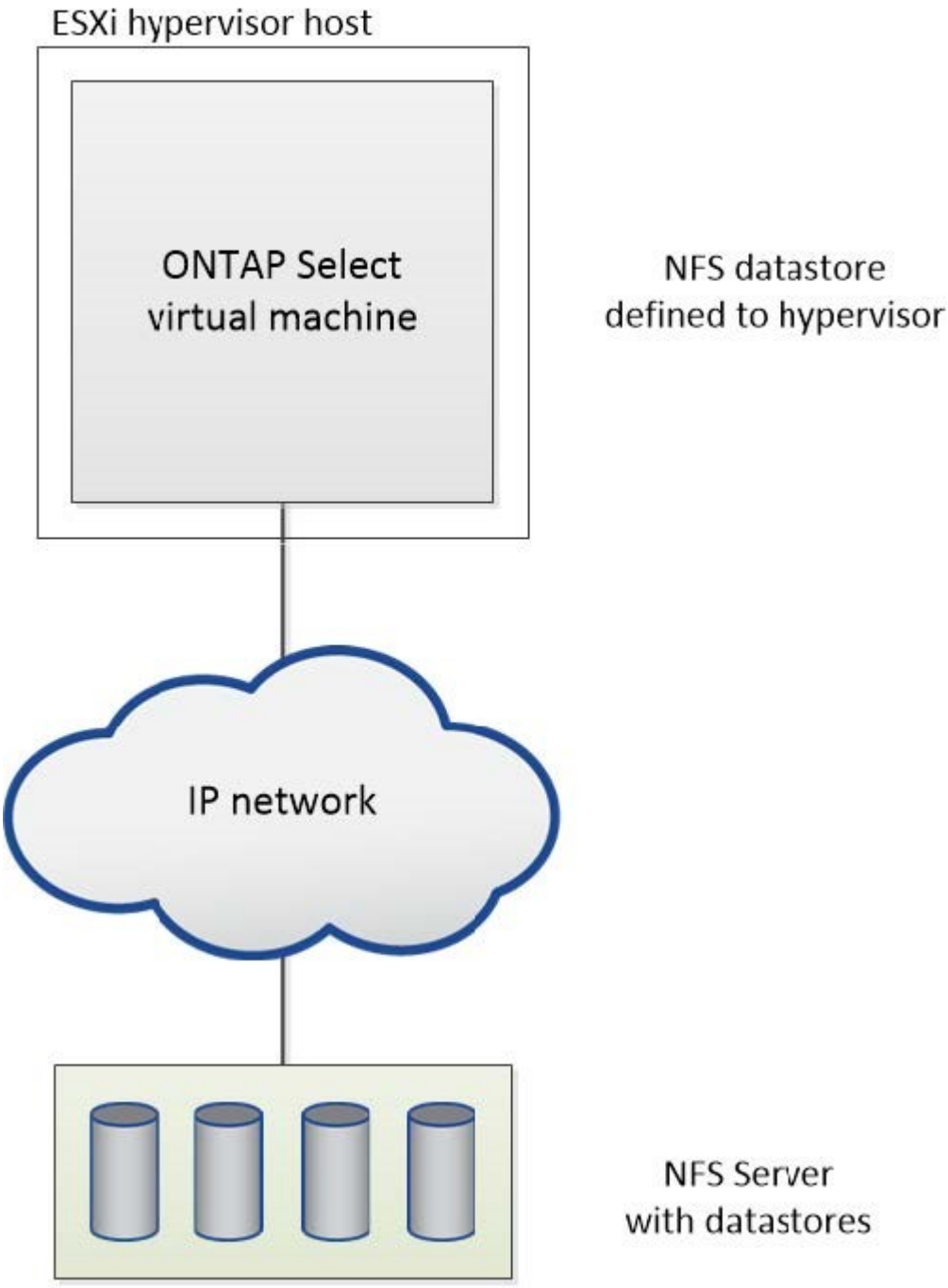
Vous pouvez créer une banque de données VMFS sur une baie de stockage externe. Le stockage est accessible à l'aide de l'un des différents protocoles réseau. La figure suivante illustre une banque de données VMFS sur une baie de stockage externe accessible via le protocole iSCSI.

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 over Ethernet.



Stockage de données NFS sur baie de stockage externe

Vous pouvez créer un datastore NFS sur une baie de stockage externe. Le stockage est accessible à l'aide du protocole réseau NFS. La figure suivante illustre un datastore NFS sur un stockage externe, accessible via l'appliance serveur NFS.



Services RAID matériels pour ONTAP Select avec stockage local connecté

Lorsqu'un contrôleur RAID matériel est disponible, ONTAP Select peut transférer les services RAID vers ce contrôleur afin d'améliorer les performances d'écriture et d'assurer une protection contre les pannes de disques physiques. Par conséquent, la protection RAID de tous les nœuds au sein du cluster ONTAP Select est assurée par le contrôleur RAID local et non par le RAID logiciel d'ONTAP.

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

Configuration du contrôleur RAID pour le stockage local connecté

Tous les disques locaux attachés qui fournissent à ONTAP Select le stockage de support doivent être placés derrière un contrôleur RAID. La plupart des serveurs de matériel ordinaire sont livrés avec plusieurs options de contrôleurs RAID à différents niveaux de prix, chacune offrant des niveaux variés de fonctionnalité. L'intention est de prendre en charge autant de ces options que possible, à condition qu'elles répondent à certaines exigences minimales imposées au contrôleur.

> Il est impossible de détacher les disques virtuels des machines virtuelles ONTAP Select utilisant une configuration RAID matérielle. Le détachement des disques est uniquement pris en charge pour les machines virtuelles ONTAP Select utilisant une configuration RAID logicielle. Consultez "Remplacer un disque défaillant 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 par batterie (BBU) ou d'un cache d'écriture sauvegardé par flash (FBWC) et prendre en charge un débit de 12Gbps.
- Le contrôleur RAID doit prendre en charge un mode capable de supporter au moins une ou deux pannes de disque (RAID 5 et RAID 6).
- Le cache du disque doit être désactivé.
- La politique d'écriture doit être configurée en mode d'écriture différée avec un basculement en mode d'écriture directe en cas de défaillance de la BBU ou de la mémoire flash.
- La stratégie d'E/S pour les lectures doit être définie sur mise en cache.

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

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

> La taille maximale des volumes et des banques de données pour les anciennes versions d'ESXi 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 d'accès aux données emprunté par les requêtes d'écriture. Ces trois modes sont les suivants :

- Écriture immédiate. Toutes les requêtes d'E/S entrantes sont écrites dans le cache du contrôleur RAID,

puis immédiatement vidées sur le disque avant que la requête ne soit renvoyée à l'hôte.

- Technique d'écriture alternative. 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 un accusé de réception est immédiatement renvoyé à l'hôte. Les blocs de données sont vidés sur le disque de manière asynchrone par le contrôleur.

Le mode d'écriture différée offre le chemin d'accès aux données le plus court, l'accusé de réception des E/S intervenant immédiatement après l'entrée des blocs dans le cache. Ce mode garantit la latence la plus faible et le débit le plus élevé pour les charges de travail mixtes lecture/écriture. Toutefois, en l'absence d'une BBU ou de technologie Flash non volatile, les utilisateurs risquent de perdre des données en cas de coupure de courant lors du fonctionnement dans ce mode.

ONTAP Select requiert la présence d'une batterie de secours ou d'une unité de mémoire flash ; ainsi, nous pouvons être certains que les blocs mis en cache sont écrits sur disque en cas de ce type de panne. Pour cette raison, il est nécessaire que le contrôleur RAID soit configuré en mode writeback.

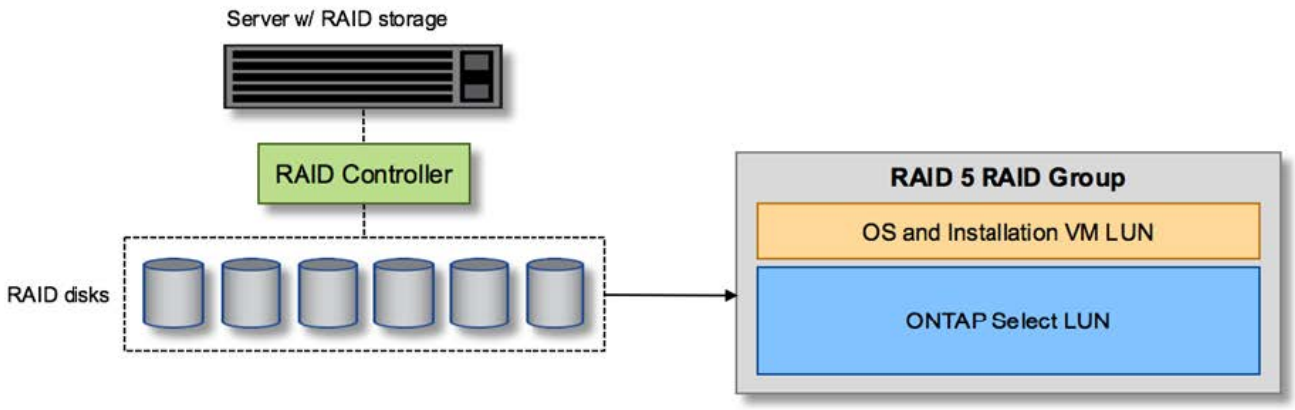
Disques locaux partagés entre ONTAP Select et le système d'exploitation

La configuration serveur la plus courante consiste à placer tous les disques durs locaux derrière un seul contrôleur RAID. Vous devez prévoir au minimum deux LUN : une pour l'hyperviseur et une pour la machine virtuelle ONTAP Select.

Prenons l'exemple d'un HP DL380 g8 équipé de six disques internes et d'un 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 disques gérés par RAID



Le provisionnement des LUN du système d'exploitation à partir du même groupe RAID que 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 défaillance d'un seul disque d'entraîner l'arrêt complet du système.

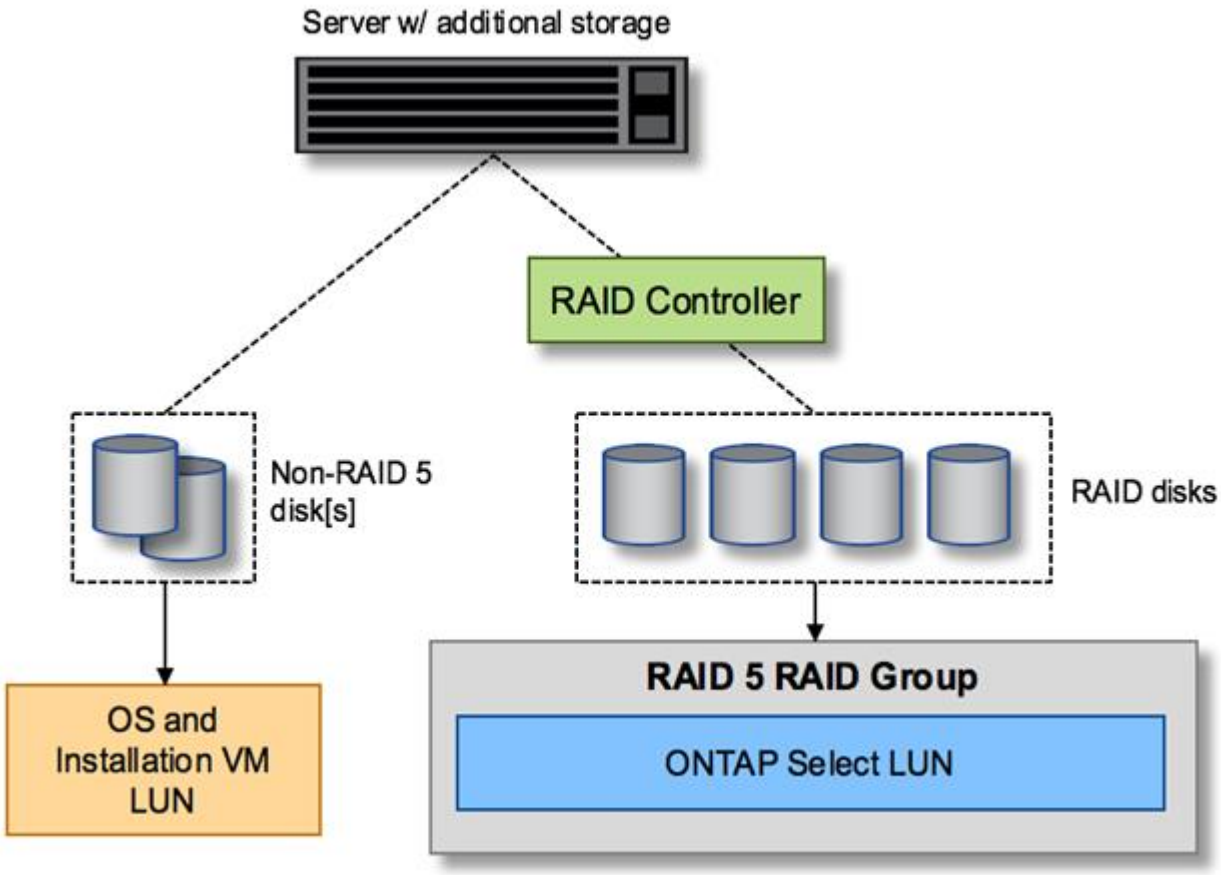
Disques locaux répartis entre ONTAP Select et le système d'exploitation

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 cette configuration, l'ensemble des disques situés derrière le contrôleur RAID, capables de fournir des services RAID 5/6, doit être utilisé exclusivement par la machine virtuelle ONTAP Select. En fonction de la capacité de stockage totale gérée, vous devez configurer les disques en un ou plusieurs groupes RAID et un ou plusieurs LUN. Ces LUN seraient ensuite utilisés pour créer un ou plusieurs datastores, tous les datastores étant 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 RAID/non-RAID mixte



LUN multiples

Il existe deux cas où les configurations à groupe RAID unique/LUN unique doivent être modifiées. Lors de l'utilisation de disques NL-SAS ou SATA, la taille du groupe RAID ne doit pas dépasser 12 disques. De plus, un LUN unique peut devenir plus grand que les limites de stockage de l'hyperviseur sous-jacent, soit la taille maximale d'une extension de système de fichiers individuelle, soit la taille maximale totale du pool de stockage. Dans ce cas, le stockage physique sous-jacent doit être divisé en plusieurs LUN pour permettre la création réussie du système de fichiers.

Limites du système de fichiers des machines virtuelles VMware vSphere

La taille maximale d'un datastore sur certaines versions d'ESXi est de 64TB.

Si un serveur dispose de plus de 64 To de stockage connecté, il peut être nécessaire de provisionner plusieurs LUN, chacune d'une capacité 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 primordial de s'assurer que ces LUN offrent des performances similaires et homogènes. Ceci est particulièrement important si toutes les LUN sont utilisées dans un seul agrégat ONTAP. En revanche, si un sous-ensemble d'une ou plusieurs LUN présente un profil de performances nettement différent, nous recommandons vivement d'isoler ces LUN dans un agrégat ONTAP distinct.

Il est possible d'utiliser plusieurs extensions de système de fichiers pour créer un seul datastore, jusqu'à la taille maximale de celui-ci. Pour limiter la capacité nécessitant une licence ONTAP Select, veillez 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 pour) uniquement une partie de l'espace dans un datastore.

Il est également possible de commencer par créer un seul datastore sur un seul LUN. Lorsque de l'espace supplémentaire nécessitant une licence de capacité ONTAP Select plus importante est requis, cet espace peut être ajouté au même datastore sous forme d'extension, jusqu'à la taille maximale du datastore. Une fois la taille maximale atteinte, de nouveaux datastores peuvent être créés et ajoutés à ONTAP Select. Les deux types d'opérations d'extension de capacité sont pris en charge et peuvent être réalisés à l'aide de la fonctionnalité storage-add d'ONTAP Deploy. Chaque nœud ONTAP Select peut être configuré pour prendre en charge jusqu'à 400 To de stockage. L'allocation de capacité à partir de plusieurs datastores nécessite une procédure en deux étapes.

La création initiale d'un cluster peut être utilisée pour créer un cluster ONTAP Select consommant une partie ou la totalité de l'espace dans le datastore initial. La deuxième étape consiste à effectuer une ou plusieurs opérations d'ajout de capacité à l'aide de datastores supplémentaires jusqu'à atteindre la capacité totale souhaitée. Cette fonctionnalité est détaillée dans la section "Augmenter la capacité de stockage".



La surcharge VMFS n'est pas nulle (voir l'article 1001618 de la base de connaissances VMware), et la tentative d'utiliser l'intégralité de l'espace indiqué comme libre par un datastore a entraîné des erreurs parasites lors des opérations de création de cluster.

Un espace tampon de 2 % est laissé inutilisé dans chaque datastore. 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 le tampon, tant qu'une limite de capacité n'est pas spécifiée. Si une limite de capacité est spécifiée, cette taille est appliquée en premier. Si la taille de la limite de capacité se situe dans la taille du tampon, la création du cluster échoue avec un message d'erreur précisant le paramètre de taille maximale correct pouvant être utilisé comme 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 à la fois pour les nouvelles installations et comme cible d'une opération de Storage vMotion d'une VM ONTAP Deploy ou ONTAP Select existante.

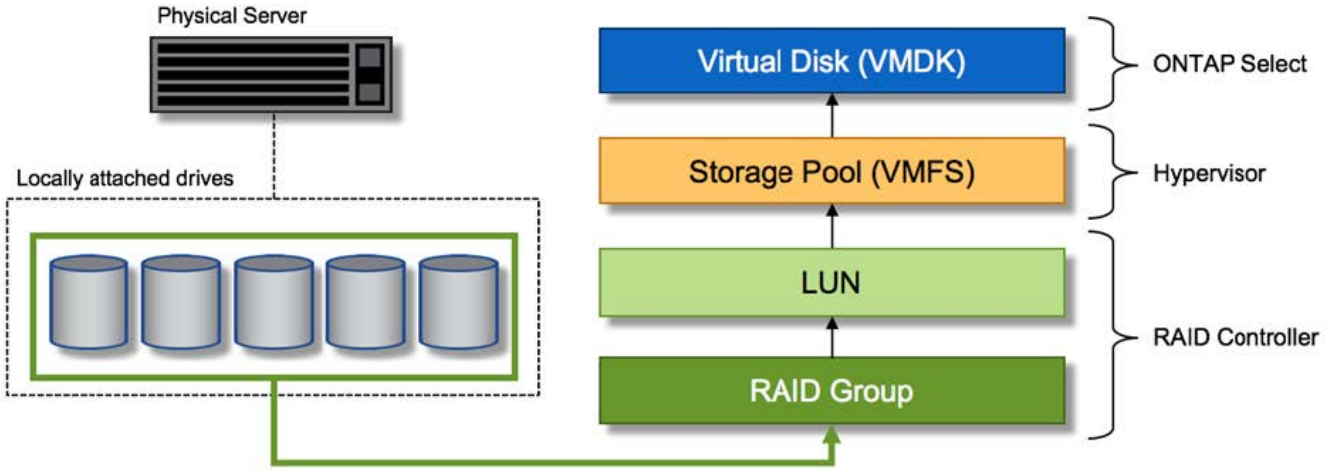
VMware ne prend pas en charge les mises à niveau sur place de VMFS 5 vers VMFS 6. Par conséquent, Storage vMotion est le seul mécanisme permettant à une machine virtuelle de migrer d'une banque de données VMFS 5 vers une banque de données VMFS 6. Cependant, la prise en charge de Storage vMotion avec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

Disques virtuels ONTAP Select

En résumé, ONTAP Select présente à ONTAP un ensemble de disques virtuels provisionnés à partir d'un ou plusieurs pools de stockage. ONTAP se voit présenter un ensemble de disques virtuels qu'il traite comme physiques, et la partie restante de la pile de stockage est abstraite par l'hyperviseur. La figure suivante montre 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 des groupes RAID et des LUN s'effectue au sein du logiciel du contrôleur RAID du serveur. Cette configuration n'est pas nécessaire lors de l'utilisation de VSAN ou de baies externes.
- La configuration du pool de stockage s'effectue au sein de l'hyperviseur.
- Les disques virtuels sont créés et détenus par des machines virtuelles individuelles ; dans cet exemple, par ONTAP Select.

Mappage d'un disque virtuel avec un disque physique



Provisionnement de disques virtuels

Pour offrir une expérience utilisateur plus fluide, l'outil de gestion ONTAP Select, ONTAP Deploy, provisionne automatiquement les disques virtuels à partir du pool de stockage associé et les attache à la machine virtuelle ONTAP Select. Cette opération se produit automatiquement lors de la configuration initiale ainsi que lors des opérations d'ajout de stockage. Si le nœud ONTAP Select fait partie d'une paire haute disponibilité, les disques virtuels sont automatiquement affectés à un pool de stockage local et à un pool de stockage miroir.

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

Par exemple, un ONTAP Select peut se voir attribuer un datastore ou un LUN de 31 To (l'espace restant après le déploiement de la VM et le provisionnement des disques système et racine). Ensuite, quatre disques virtuels

d'environ 7,75 To sont créés et affectés au plex local ONTAP et au plex miroir approprié.

> L'ajout de capacité à une machine virtuelle ONTAP Select peut entraîner la création de VMDK de tailles différentes. Pour plus de détails, consultez la section "Augmenter la capacité de stockage". Contrairement aux systèmes FAS, des VMDK de tailles différentes peuvent exister dans le même agrégat. ONTAP Select utilise une répartition RAID 0 entre ces VMDK, ce qui permet d'utiliser pleinement tout l'espace dans chaque VMDK, quelle que soit sa taille.

NVRAM virtualisée

NetApp FAS systèmes sont traditionnellement équipés d'une carte NVRAM PCI physique, une carte haute performance contenant de la mémoire flash non volatile. Cette carte offre une amélioration significative des performances d'écriture en donnant à ONTAP la capacité d'accuser immédiatement réception des écritures entrantes auprès du client. Elle peut également planifier le déplacement des blocs de données modifiés vers le support de stockage plus lent, dans un processus appelé destaging.

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

La NVRAM est placée sur son propre VMDK. Séparer la NVRAM dans son propre VMDK permet à la machine virtuelle ONTAP Select d'utiliser le pilote vNVMe pour communiquer avec son VMDK NVRAM. Cela nécessite également que la machine virtuelle ONTAP Select utilise la version matérielle 13, qui est compatible avec ESXi 8.0 et versions ultérieures.

Explication du chemin d'accès aux 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 d'accès aux données emprunté par une requête d'écriture lorsqu'elle entre dans le système.

Les requêtes d'écriture entrantes vers la machine virtuelle ONTAP Select sont dirigées vers la partition NVRAM de la VM. Au niveau de la virtualisation, cette partition existe dans un disque système ONTAP Select, un VMDK attaché à la machine virtuelle ONTAP Select. Au niveau physique, ces requêtes sont mises en cache dans le contrôleur RAID local, comme toutes les modifications de blocs destinées aux disques sous-jacents. À partir de là, l'écriture est accusée de réception auprès de l'hôte.

Physiquement, le bloc se trouve alors dans le cache du contrôleur RAID, en attente d'être écrit sur le disque. Logiquement, le bloc réside dans NVRAM en attente d'être transféré vers les disques de données utilisateur appropriés.

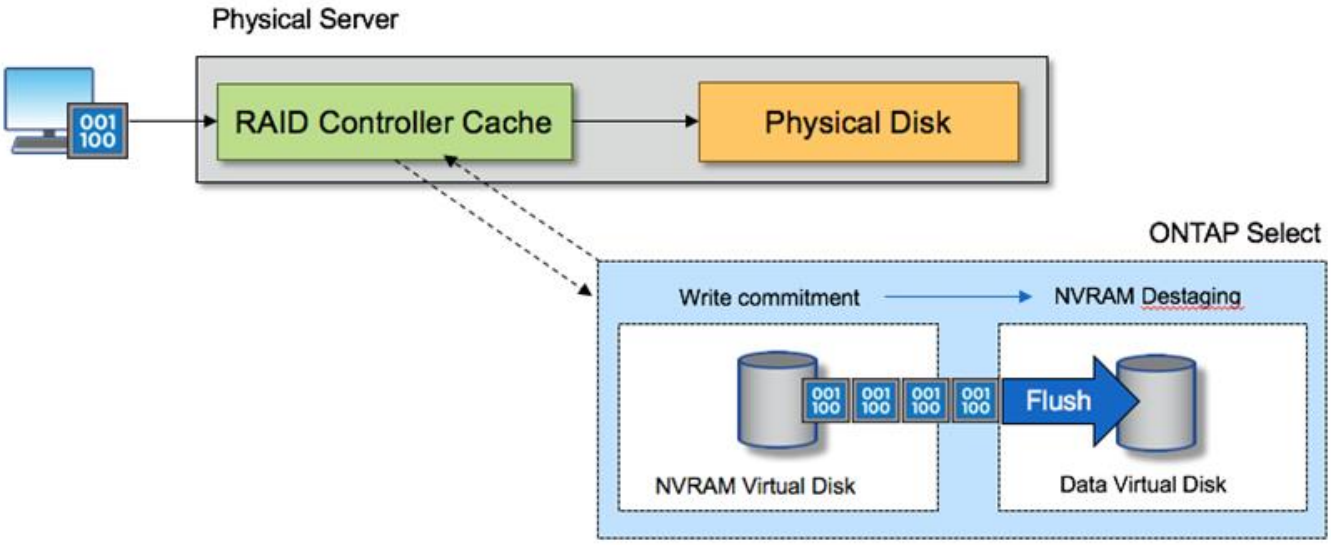
Les blocs modifiés étant automatiquement stockés dans le cache local du contrôleur RAID, les écritures entrantes sur la partition NVRAM sont automatiquement mises en cache et périodiquement transférées vers le support de stockage physique. Il ne faut pas confondre ce processus avec le transfert périodique du contenu de la NVRAM vers les disques de données ONTAP. Ces deux opérations sont indépendantes 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 le VMDK NVRAM soient mis en cache dans le cache du

contrôleur RAID local, ce dernier ignore la structure de la machine virtuelle et ses disques virtuels. Il stocke tous les blocs modifiés sur le système, dont 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 disques sous-jacents.

Écritures entrantes vers la machine virtuelle ONTAP Select



La partition NVRAM est séparée sur son propre VMDK. Ce VMDK est attaché à l'aide du pilote vNVME disponible dans les versions ESXi 8.0 ou 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 du 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 la même fonctionnalité que la couche RAID d'une plateforme ONTAP traditionnelle telle que FAS. La couche RAID effectue les calculs de parité des disques et assure la protection contre les défaillances individuelles de disques 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 être indisponible ou non souhaitable dans certains environnements, notamment lorsque ONTAP Select est déployé sur un matériel ordinaire. 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 de la machine virtuelle ONTAP Select.

- Choisissez un disque séparé, soit un SSD, soit un disque NVMe, pour créer un datastore pour les disques système (NVRAM, carte de démarrage/CF, Coredump et Mediator dans une configuration multi-nœuds).

> - Les termes « disque de service » et « disque système » sont utilisés indifféremment.
>   - Les disques de service sont les disques virtuels (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 de l'hôte. Ce disque physique doit contenir une banque de données DAS. ONTAP Deploy crée ces disques de service pour la VM 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 disques 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 possède déjà un contrôleur RAID, celui-ci doit respecter les exigences suivantes :

- Vous devez désactiver le contrôleur RAID matériel pour que les disques soient présentés directement au système (un JBOD). Vous pouvez généralement effectuer cette modification dans le BIOS du contrôleur RAID.
- Ou le contrôleur RAID matériel doit être en mode SAS HBA. Par exemple, certaines configurations BIOS proposent un mode « AHCI » en plus du RAID, que vous pouvez choisir pour activer le mode JBOD. Cela permet un accès direct, de sorte que les disques physiques soient visibles tels quels sur l'hôte.

Selon le nombre maximal de disques pris en charge par le contrôleur, un contrôleur supplémentaire peut être nécessaire. En mode SAS HBA, assurez-vous que le contrôleur d'E/S (SAS HBA) est pris en charge avec une vitesse minimale de 6Gbps. Cependant, NetApp recommande une vitesse de 12Gbps.

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

> Les administrateurs doivent surveiller quels disques sont utilisés par la machine virtuelle ONTAP Select et empêcher toute utilisation involontaire de ces disques sur l'hôte.

Disques virtuels et physiques ONTAP Select

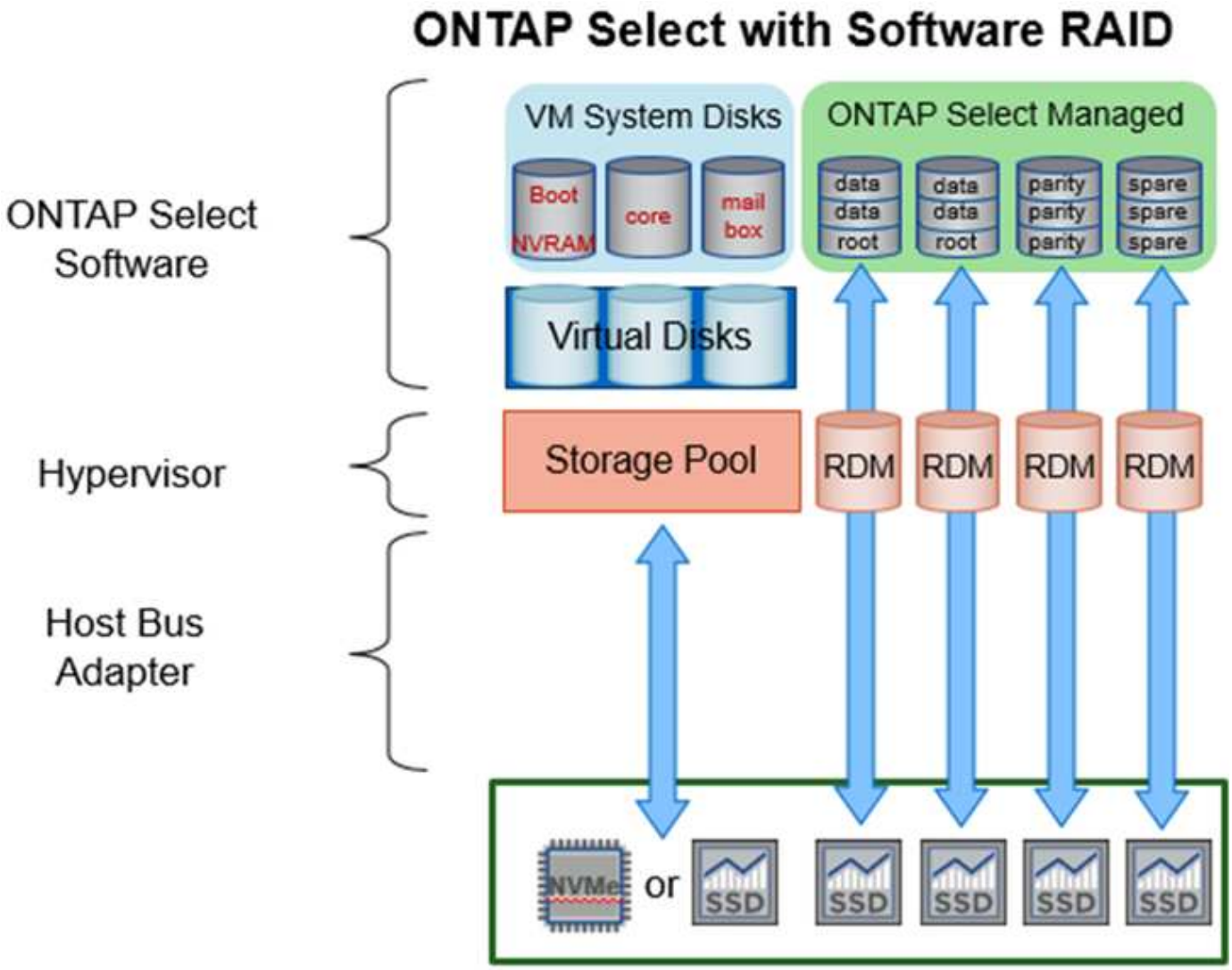
Pour les configurations avec contrôleurs RAID matériels, la redondance physique des disques est assurée par le contrôleur RAID. ONTAP Select est présenté avec un ou plusieurs VMDK à partir desquels l'administrateur ONTAP peut configurer des agrégats de données. Ces VMDK sont configurés en RAID 0, car l'utilisation du RAID logiciel ONTAP est redondante, inefficace et peu performante en raison de la résilience fournie au niveau matériel. De plus, les VMDK utilisés pour les disques système se trouvent dans le même datastore que les VMDK utilisés pour stocker les données utilisateur.

Lors de l'utilisation du RAID logiciel, ONTAP Deploy présente à ONTAP Select un ensemble de VMDK et de mappages de périphériques bruts [RDM] pour les SSD et des périphériques passthrough ou DirectPath IO

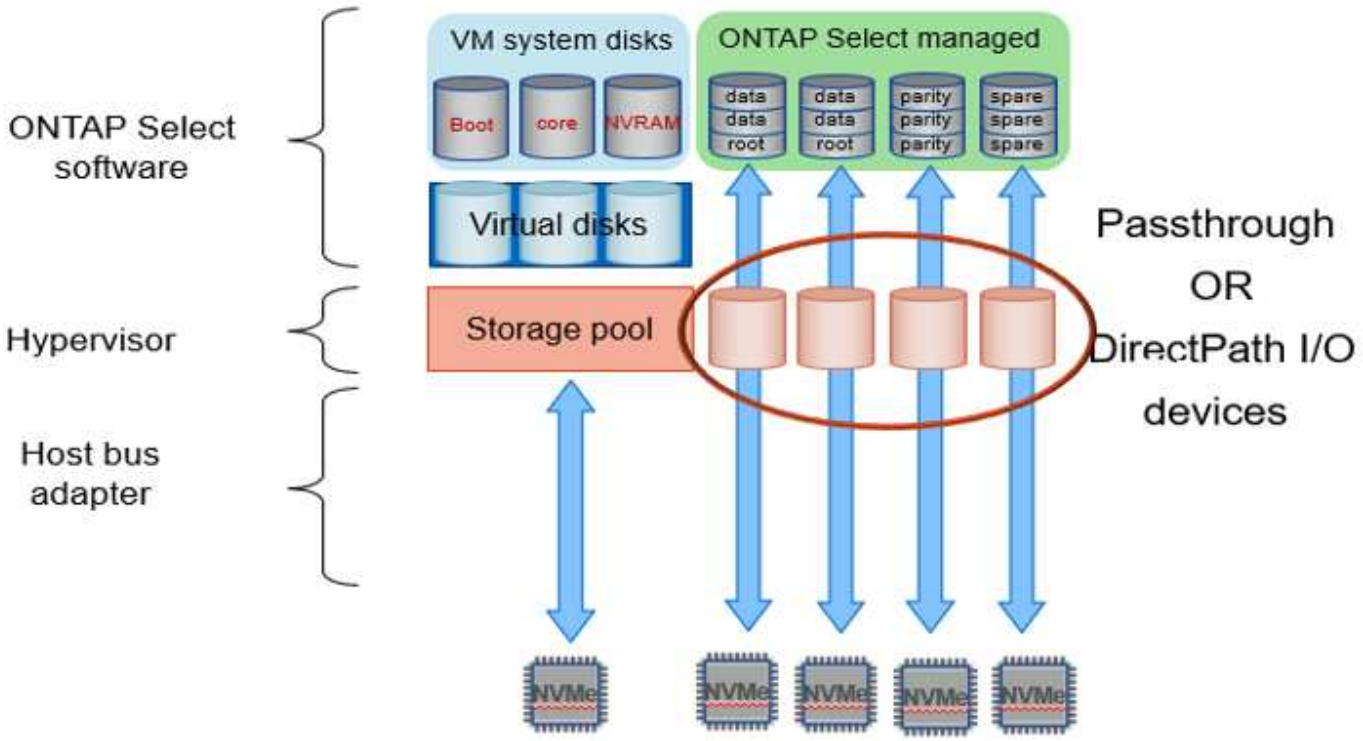
pour NVMe.

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 virtuelle ONTAP Select et les disques physiques utilisés pour stocker les données utilisateur.

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



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



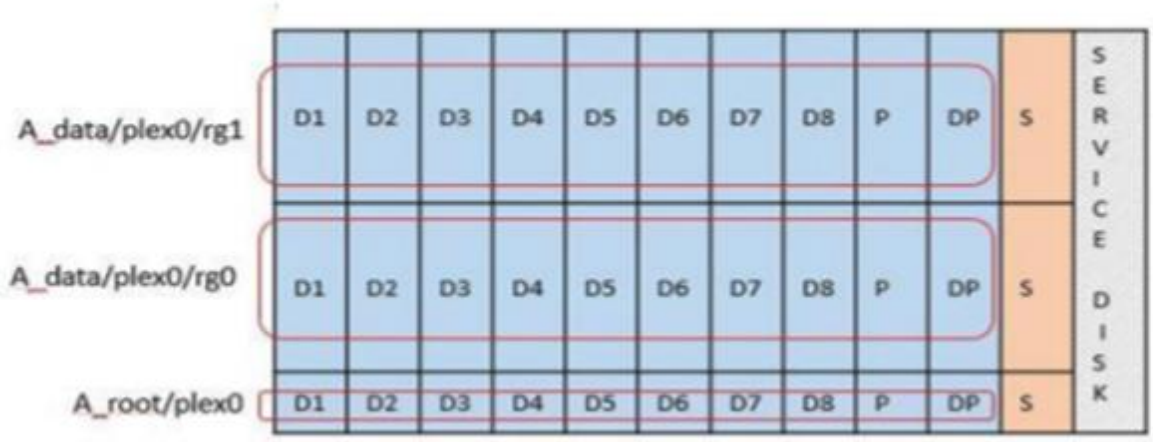
Les disques système (VMDK) résident dans le même datastore et sur le même disque physique. Le disque NVRAM virtuel nécessite un support rapide et durable. Par conséquent, seuls les datastores de type NVMe et SSD sont pris en charge. Lors de l'utilisation de disques NVMe pour les données, le disque système doit également être un périphérique NVMe pour des raisons de performances. Une carte INTEL Optane constitue un excellent candidat 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 disques physiques.

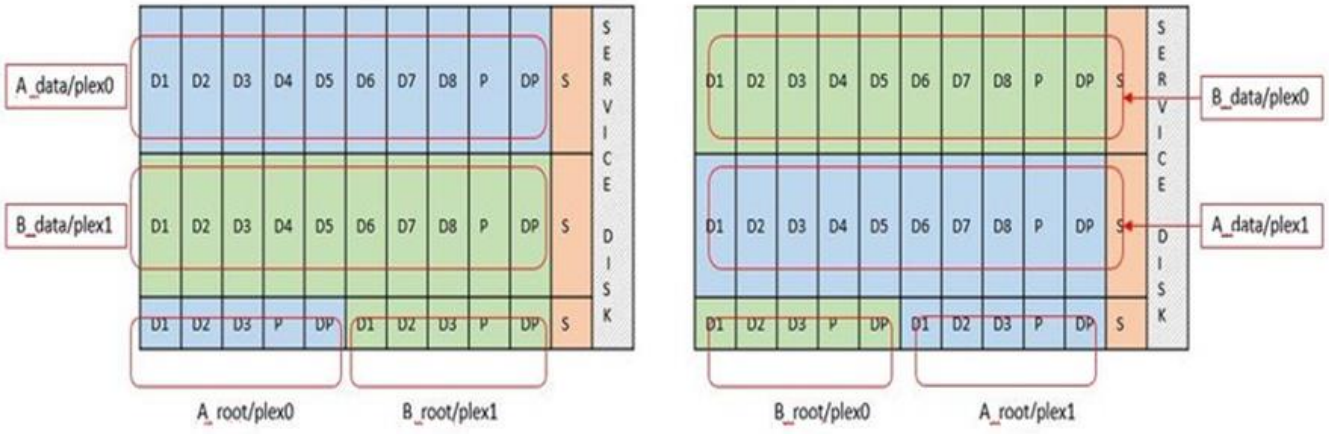
Chaque disque de données est divisé en trois parties : une petite partition racine (stripe) et deux partitions de taille égale pour créer deux disques de données visibles dans la machine virtuelle ONTAP Select. Les partitions utilisent le schéma Root Data Data (RD2), comme illustré dans les figures suivantes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

`P` désigne un disque de parité, `DP` désigne un disque de parité double, et `S` désigne un disque de secours.

Partitionnement de disque RDD pour les clusters à nœud unique



Partitionnement de disque RDD pour les clusters multi-nœuds (paires haute disponibilité)



Le logiciel RAID d'ONTAP prend en charge les types de RAID suivants : RAID 4, RAID-DP et RAID-TEC. Ce sont les mêmes constructions RAID utilisées par les plateformes FAS et AFF. Pour le provisionnement de la racine, ONTAP Select prend uniquement en charge RAID 4 et RAID-DP. Lorsque vous utilisez RAID-TEC pour l'agrégat de données, la protection globale est de type RAID-DP. ONTAP Select HA utilise une architecture sans partage qui réplique la configuration de chaque nœud sur l'autre nœud. Cela signifie que chaque nœud doit stocker sa partition racine et une copie de la partition racine de son homologue. Un disque de données possède une seule partition racine. Cela signifie que le nombre minimal de disques de données varie selon que le nœud ONTAP Select fait partie d'une paire haute disponibilité.

Pour les clusters à nœud unique, toutes les partitions de données servent à stocker les données locales (actives). Pour les nœuds faisant partie d'une paire haute disponibilité, une partition de données est utilisée pour stocker les données locales (actives) de ce nœud et la seconde partition de données est utilisée pour dupliquer les données actives du homologue HA.

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

Les hyperviseurs ESXi et KVM ne prennent pas en charge les disques NVMe en tant que Raw Device Maps (RDM). Pour permettre à ONTAP Select de prendre le contrôle direct des disques NVMe, vous devez configurer ces disques comme périphériques passthrough au sein d'ESXi ou de KVM. Lors de la configuration d'un périphérique NVMe en tant que périphérique passthrough, cela nécessite la prise en charge du BIOS du

serveur et il se peut que vous deviez redémarrer l'hôte. De plus, il existe des limites quant au nombre de périphériques passthrough pouvant être attribués par hôte, qui peuvent varier selon la plateforme. Cependant, ONTAP Deploy limite ce nombre à 14 périphériques NVMe par nœud ONTAP Select. Cela signifie que la configuration NVMe offre une densité d'IOPS (IOPS/To) très élevée, au détriment de la capacité totale. Sinon, si vous souhaitez une configuration haute performance avec une capacité de stockage plus importante, la configuration recommandée est une grande taille de machine virtuelle ONTAP Select, une carte INTEL Optane pour le disque système et un nombre nominal de disques SSD pour le stockage des données.

Pour tirer pleinement parti des performances NVMe, envisagez la taille de machine virtuelle ONTAP Select large.

Il existe une différence supplémentaire entre les périphériques passthrough et les RDM. Les RDM peuvent être associés à une machine virtuelle en cours d'exécution. Les périphériques passthrough nécessitent un redémarrage de la machine virtuelle. Cela signifie que toute procédure de remplacement de disque NVMe ou d'expansion de capacité (ajout de disque) nécessitera un redémarrage de la machine virtuelle ONTAP Select. L'opération de remplacement de disque et d'expansion de capacité (ajout de disque) est pilotée par un workflow dans ONTAP Deploy. ONTAP Deploy gère le redémarrage d'ONTAP Select pour les clusters à nœud unique et le basculement/retour arrière pour les paires haute disponibilité. Cependant, il est important de noter la différence entre le travail avec des disques de données SSD (aucun redémarrage ou basculement d'ONTAP Select n'est requis) et le travail avec des disques de données NVMe (un redémarrage ou un basculement d'ONTAP Select est requis).

Provisionnement de disques physiques et virtuels

Pour offrir une expérience utilisateur plus fluide, ONTAP Deploy provisionne automatiquement les disques système (virtuels) à partir du datastore spécifié (disque système physique) et les attache à la VM ONTAP Select. Cette opération se produit automatiquement lors de la configuration initiale afin que la VM 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 haute disponibilité, les partitions de données sont automatiquement affectées à un pool de stockage local et à un pool de stockage miroir. Cette affectation se produit automatiquement lors des opérations de création de cluster et d'ajout de stockage.

Étant donné que les disques de données sur la VM ONTAP Select sont associés aux disques physiques sous-jacents, la création de configurations comportant 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 nœud dispose de suffisamment de disques, il utilise RAID-DP, sinon il crée un agrégat racine RAID-4.

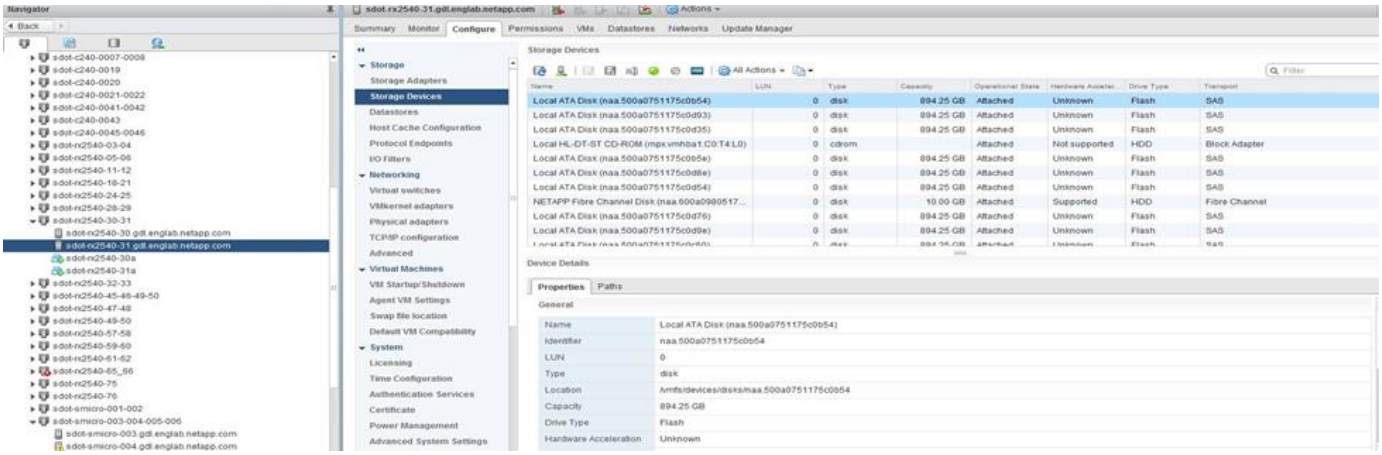
Lors de l'ajout de capacité à une machine virtuelle ONTAP Select à l'aide d'un RAID logiciel, l'administrateur doit tenir compte de la taille physique du disque et du nombre de disques requis. Pour plus de détails, consultez "Augmenter la capacité de stockage".

Comme pour les systèmes FAS et AFF, vous ne pouvez ajouter à un groupe RAID existant que des disques de capacité égale ou supérieure. Les disques de plus grande capacité sont dimensionnés automatiquement. Si vous créez de nouveaux groupes RAID, la taille du nouveau groupe RAID doit correspondre à celle du groupe RAID existant afin de garantir que les performances globales de l'agrégat ne se détériorent pas.

Associez un disque ONTAP Select au disque ESXi ou KVM correspondant

Les disques ONTAP Select sont généralement étiquetés NET x.y. 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 l'interface ESXi ou KVM, vous pouvez entrer la commande suivante pour faire clignoter la LED d'un disque physique donné (identifié par son naa.unique-id).

ESXi

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

KVM

```
cat /sys/block/<block_device_id>/device/wwid
```

Défaillances de plusieurs disques lors de l'utilisation d'un RAID logiciel

Il est possible qu'un système rencontre une situation où plusieurs disques tombent en panne simultanément. Le comportement du système dépend de la protection RAID de l'agrégat et du nombre de disques défaillants.

Un agrégat RAID4 peut survivre à une défaillance de disques, un agrégat RAID-DP peut survivre à deux

défaillances de disques, et un agrégat RAID-TEC peut survivre à trois défaillances de disques.

Si le nombre de disques défaillants est inférieur au nombre maximal de défaillances supporté par le type RAID, et si un disque de secours est disponible, la reconstruction démarre automatiquement. Si aucun disque de secours n'est disponible, l'agrégat sert les données dans un état dégradé jusqu'à ce que des disques de secours soient ajoutés.

Si le nombre de disques défaillants dépasse le nombre maximal de défaillances supporté par le type RAID, le plex local est marqué comme défaillant et l'état de l'agrégat est dégradé. Les données sont servies depuis le second plex situé sur le partenaire de la paire haute disponibilité. Cela signifie que toute requête d'E/S pour le nœud 1 est envoyée via le port d'interconnexion de cluster e0e (iSCSI) vers les disques physiquement situés sur le nœud 2. Si le second plex tombe également en panne, l'agrégat est marqué comme défaillant et les données sont indisponibles.

Pour que la mise en miroir des données reprenne correctement, un plex défaillant doit être supprimé puis recréé. Notez qu'une défaillance 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 impacter plusieurs agrégats, y compris 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.

Un plex défaillant est supprimé puis recréé dans l'exemple de sortie suivant :

```
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
          Position   Disk                      Type           Size
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..
```

Pour tester ou simuler une ou plusieurs pannes de disque, utilisez la commande `storage disk fail -disk NET-x.y -immediate`. Si un disque de rechange est disponible dans le système, l'agrégat commencera à se reconstruire. Vous pouvez vérifier l'état de la reconstruction à l'aide de la commande `storage aggregate show`. Vous pouvez supprimer le disque simulé défaillant à l'aide d'ONTAP Deploy. Notez qu'ONTAP a marqué le disque comme `Broken`. Le disque n'est pas réellement endommagé et peut être réintégré à l'aide d'ONTAP Deploy. Pour supprimer l'étiquette « Broken », saisissez les commandes suivantes dans la ligne de commandes ONTAP Select :

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

Le résultat de la dernière commande devrait être vide.

NVRAM virtualisée

NetApp FAS systems sont traditionnellement équipés d'une carte NVRAM PCI physique. Cette carte est une carte haute performance contenant une mémoire flash non volatile qui offre une amélioration significative des performances d'écriture. Elle permet à ONTAP d'accuser immédiatement réception des écritures entrantes auprès du client. Elle peut également planifier le déplacement des blocs de données modifiés vers des supports de stockage plus lents dans un processus appelé destaging.

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

ONTAP Select vSAN et configurations de baies externes

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

La configuration minimale requise est que l'hyperviseur utilisé (VMware ESXi ou KVM sur un hôte Linux compatible) prenne en charge la configuration sous-jacente. Si l'hyperviseur est ESXi, il doit figurer dans les listes de compatibilité matérielle (HCL) VMware correspondantes.

architecture vNAS

La nomenclature vNAS est utilisée pour toutes les configurations qui n'utilisent pas de DAS. Pour les clusters ONTAP Select multi-nœuds, cela inclut les architectures pour lesquelles les deux nœuds ONTAP Select d'une même paire haute disponibilité partagent un seul datastore (y compris les datastores vSAN). Les nœuds

peuvent également être installés sur des datastores distincts issus de la même baie externe partagée. Cela permet d'obtenir des efficacités de stockage côté baie afin de réduire l'encombrement global de l'ensemble de la paire haute disponibilité ONTAP Select. L'architecture des solutions ONTAP Select vNAS est très similaire à celle d'ONTAP Select sur DAS avec un contrôleur RAID local. C'est-à-dire que chaque nœud ONTAP Select continue d'avoir une copie des données de son partenaire de haute disponibilité. Les politiques d'efficacité du stockage ONTAP sont appliquées au niveau du nœud. Par conséquent, les efficacités de stockage côté baie sont préférables car elles peuvent potentiellement être appliquées aux ensembles de données des deux nœuds ONTAP Select.

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

Lorsqu'on utilise des baies externes distinctes pour chaque nœud ONTAP Select, il est très important que les deux baies offrent des caractéristiques de performance similaires à celles de 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 un contrôleur RAID. Dans les deux cas, ONTAP Select consomme de l'espace du datastore. Cet espace du datastore est découpé en VMDK, et ces VMDK forment les agrégats de données ONTAP traditionnels. ONTAP Deploy s'assure que les VMDK sont correctement dimensionnés et affectés au plex approprié (dans le cas de paires haute disponibilité) 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 différence la plus immédiate est que le vNAS ne nécessite pas de contrôleur RAID. Le vNAS part du principe que la baie externe sous-jacente assure la persistance des données et la résilience que fournirait un DAS avec un contrôleur RAID. La seconde, plus subtile, concerne les performances de la NVRAM.

NVRAM vNAS

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

Pour les configurations DAS avec un contrôleur RAID matériel, le cache du contrôleur RAID matériel fait office de cache NVRAM, car toutes les écritures dans le NVRAM VMDK 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 de démarrage est présent, ONTAP Select contourne la NVRAM et écrit directement la charge utile des données dans l'agrégat de données. La NVRAM est uniquement utilisée pour enregistrer l'adresse des blocs modifiés par l'opération WRITE. L'avantage de cette fonctionnalité est d'éviter une double écriture : une écriture dans la NVRAM et une seconde écriture lorsque la NVRAM est déstockée. Cette fonctionnalité est activée uniquement pour les VNAS car les écritures locales dans le cache du contrôleur RAID n'entraînent qu'une latence supplémentaire négligeable.

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

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

```
-logging off
```

Les performances d'écriture sont affectées si la fonction SIDL est désactivée. Il est possible de réactiver la fonction SIDL après que toutes les stratégies d'efficacité du stockage sur tous les volumes de cet agrégat ont été désactivées :

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

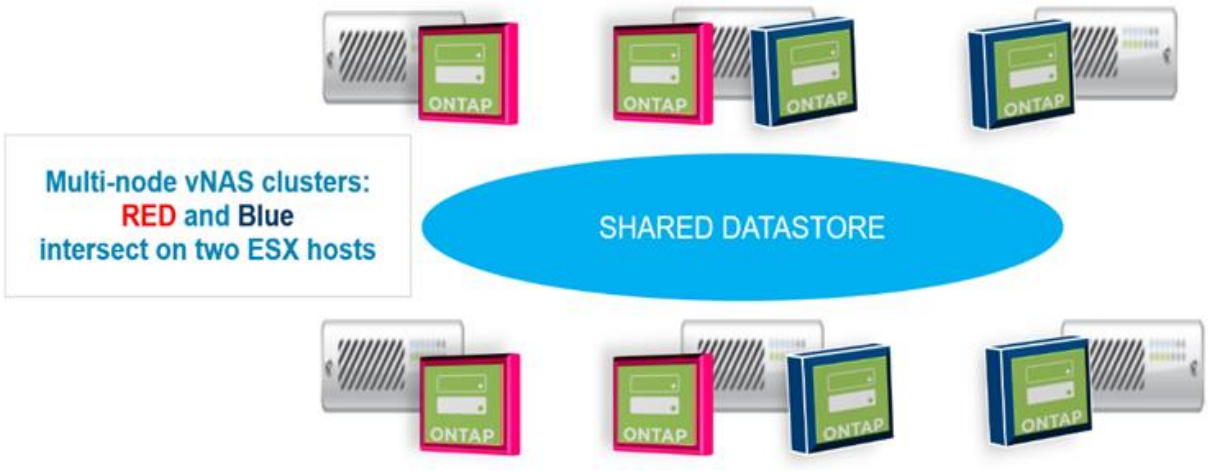
Colocaliser les nœuds ONTAP Select lors de l'utilisation de vNAS sur ESXi

ONTAP Select inclut la prise en charge des clusters ONTAP Select multi-nœuds sur stockage partagé. ONTAP Deploy permet la configuration de plusieurs nœuds ONTAP Select sur le même hôte ESXi, à condition que ces nœuds ne fassent pas partie du même cluster.

Cette configuration est uniquement valable pour les environnements VNAS (banques de données partagées). Plusieurs instances ONTAP Select par hôte ne sont pas prises en charge lors de l'utilisation d'un stockage DAS, car ces instances se disputent le même contrôleur RAID matériel.

ONTAP Deploy garantit que le déploiement initial du cluster VNAS multi-nœuds n'entraîne pas le placement de plusieurs instances ONTAP Select issues du même cluster sur le même hôte. La figure suivante illustre un exemple de déploiement correct de deux clusters à quatre nœuds qui se recoupent sur deux hôtes.

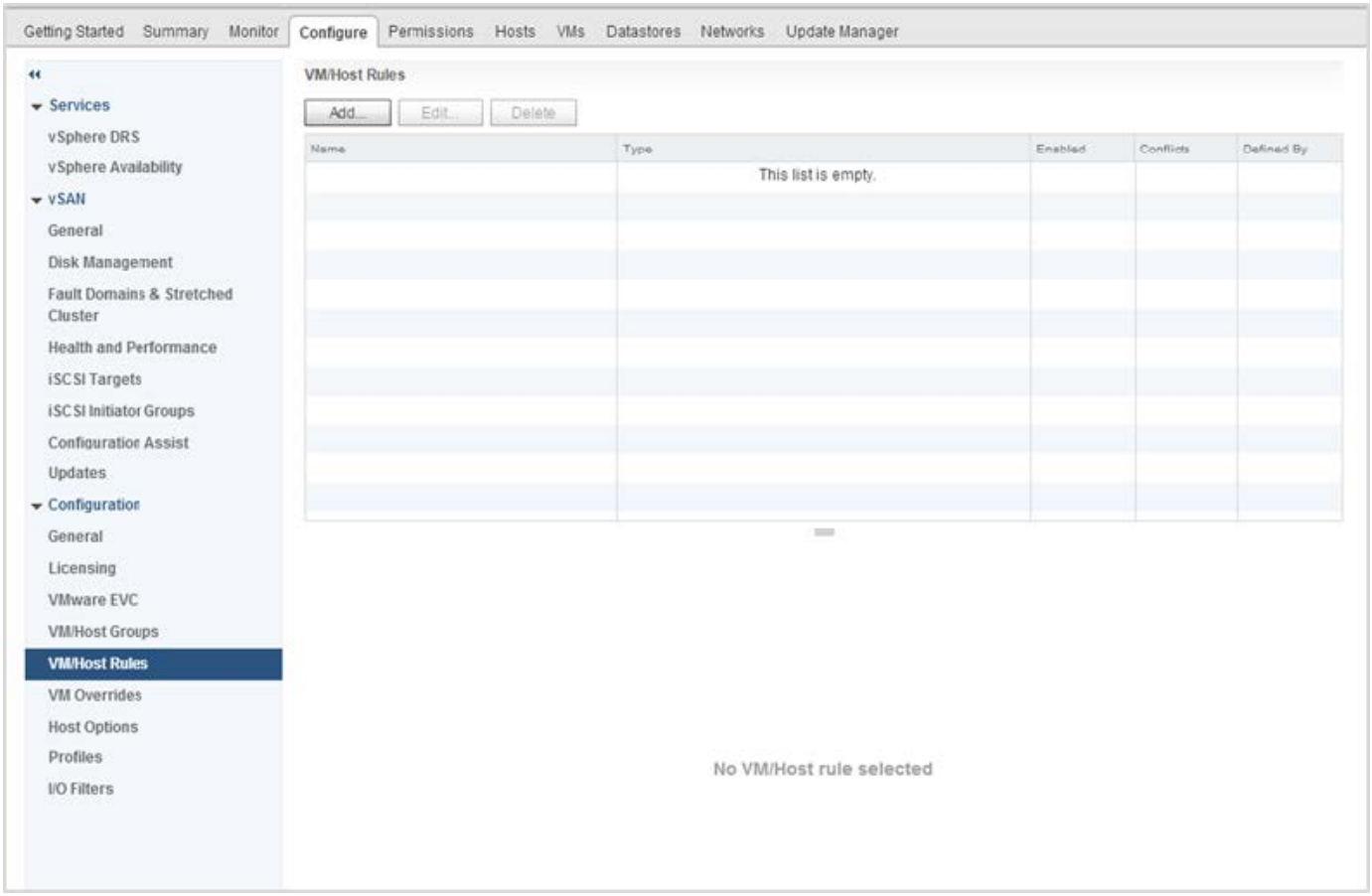
Déploiement initial de clusters VNAS multi-nœuds

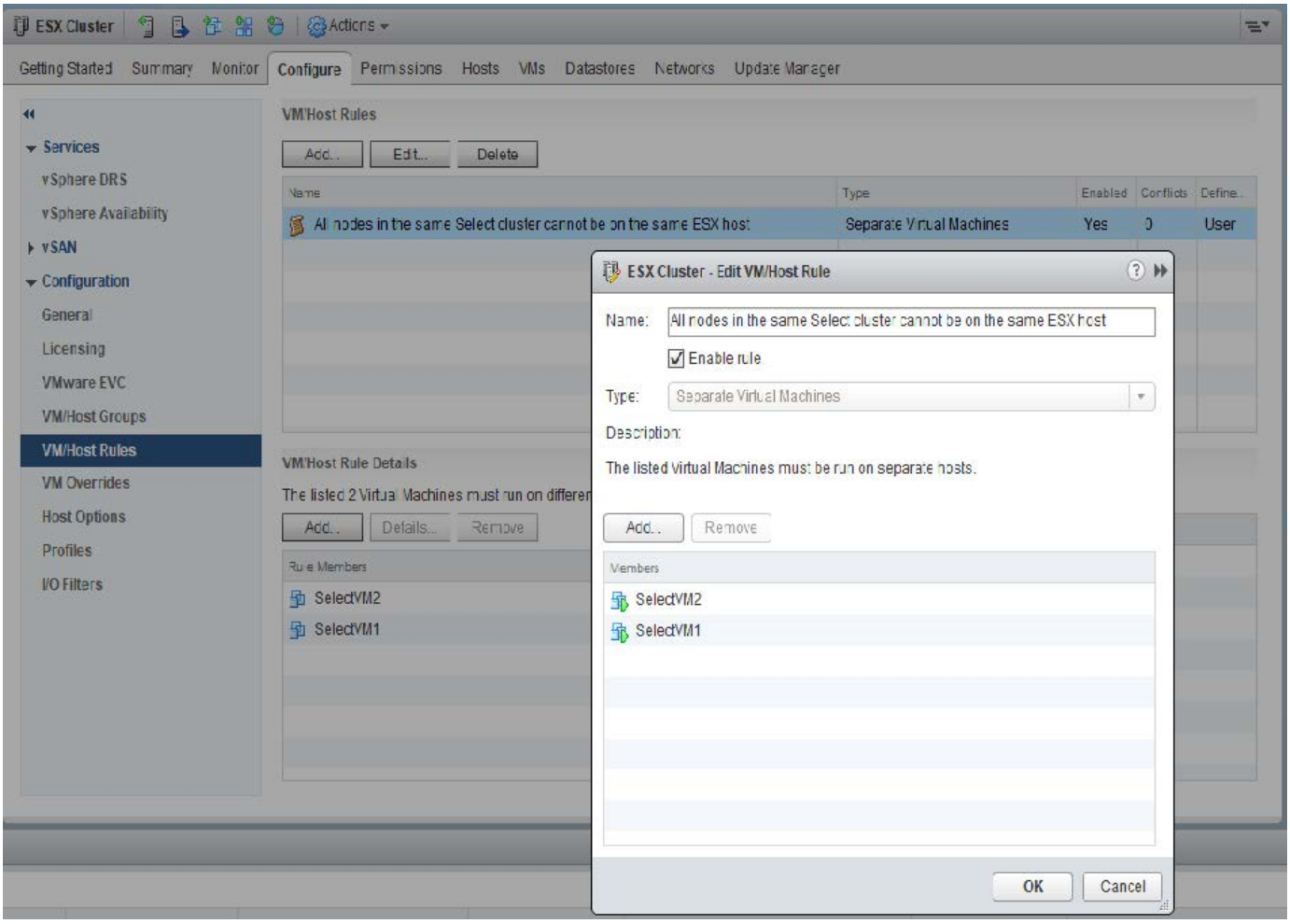


Après le déploiement, les nœuds ONTAP Select peuvent être migrés entre hôtes. Cela peut entraîner des configurations non optimales et non prises en charge, où deux nœuds ONTAP Select ou plus d'un même cluster partagent le même hôte sous-jacent. NetApp recommande la création manuelle de règles d'anti-affinité pour les machines virtuelles 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 haute disponibilité.

Les règles anti-affinité exigent que DRS soit activé sur le cluster ESXi.

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 plus d’une paire haute disponibilité, tous les nœuds du cluster doivent être inclus dans cette règle.





Deux ou plusieurs nœuds ONTAP Select du même cluster ONTAP Select peuvent potentiellement se trouver sur le même hôte ESXi pour l'une des raisons suivantes :

- DRS n'est pas présent en raison des limitations de la 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 VM initiée par l'administrateur est prioritaire.

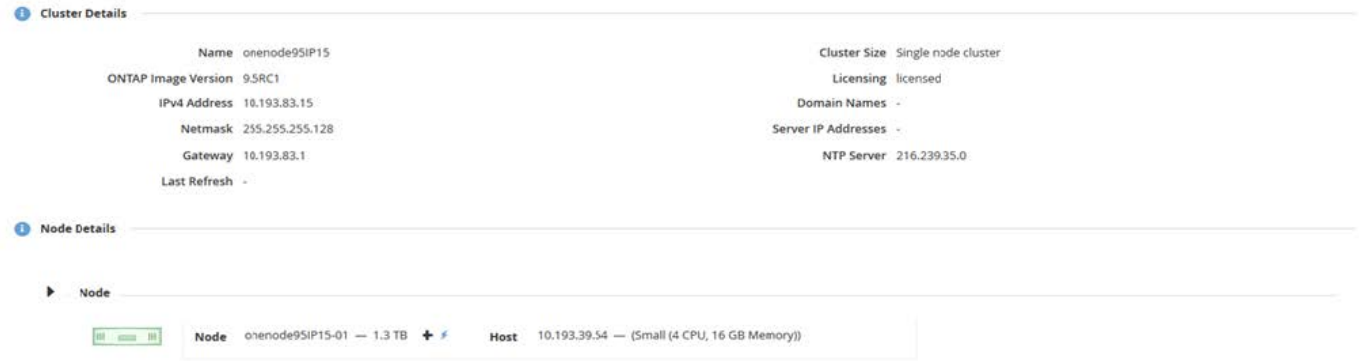
ONTAP Deploy ne surveille pas de manière proactive les emplacements des machines virtuelles ONTAP Select. Cependant, une opération d'actualisation du cluster reflète cette configuration non prise en charge dans les journaux ONTAP Deploy :



Considérations relatives à l'augmentation de la capacité de stockage d'ONTAP Select

ONTAP Deploy peut être utilisé pour ajouter et licencier du 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 le stockage géré, et la modification directe de la machine virtuelle ONTAP Select n'est pas prise en charge. La figure suivante illustre l'icône « + » qui lance l'assistant d'ajout de stockage.



Les points suivants sont essentiels à la réussite de l'extension de capacité. L'ajout de capacité nécessite que la licence existante couvre la totalité de l'espace (existant plus nouvel espace). Toute opération d'ajout de stockage entraînant un dépassement de la capacité sous licence du nœud échouera. Une nouvelle licence disposant d'une capacité suffisante doit être installée en premier.

Si la capacité supplémentaire est ajoutée à un agrégat ONTAP Select existant, alors le nouveau pool de stockage (datastore) devrait avoir un profil de performance similaire à celui du pool de stockage (datastore) existant. Notez qu'il n'est pas possible d'ajouter du stockage non SSD à un nœud ONTAP Select installé avec une personnalité de type AFF (flash activé). Le mélange de stockage DAS et de stockage externe n'est pas pris en charge.

Si vous ajoutez du stockage local à un système pour fournir des pools de stockage local supplémentaires (DAS), vous devez créer un groupe RAID et un ou plusieurs LUN supplémentaires. Comme pour les systèmes FAS, il est important de veiller 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 sur les performances du nouvel agrégat sont bien comprises.

Le nouvel espace peut être ajouté à ce même data store en tant qu'extent si la taille totale du data store ne dépasse pas la taille maximale de data store prise en charge. L'ajout d'un extent de data store au data store dans lequel ONTAP Select est déjà installé peut se faire dynamiquement et n'affecte pas les opérations du nœud ONTAP Select.

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

Dans une paire haute disponibilité, chaque nœud contient une copie conforme des données de son partenaire. L'ajout d'espace au nœud 1 nécessite l'ajout d'une quantité identique d'espace 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 intégralement protégées lors d'un événement de haute disponibilité.

Il convient de prendre en compte un autre facteur lié aux performances. Les données du nœud 1 sont répliquées de manière synchrone sur le nœud 2. Par conséquent, les performances du nouvel espace (magasin de données) sur le nœud 1 doivent correspondre à celles du nouvel espace (magasin de données) sur le nœud 2. Autrement dit, l'ajout d'espace sur les deux nœuds, mais avec des technologies de disques ou des tailles de groupes RAID différentes, peut entraîner des problèmes de performances. Ceci est dû à l'opération RAID SyncMirror utilisée pour maintenir une copie des données sur le nœud partenaire.

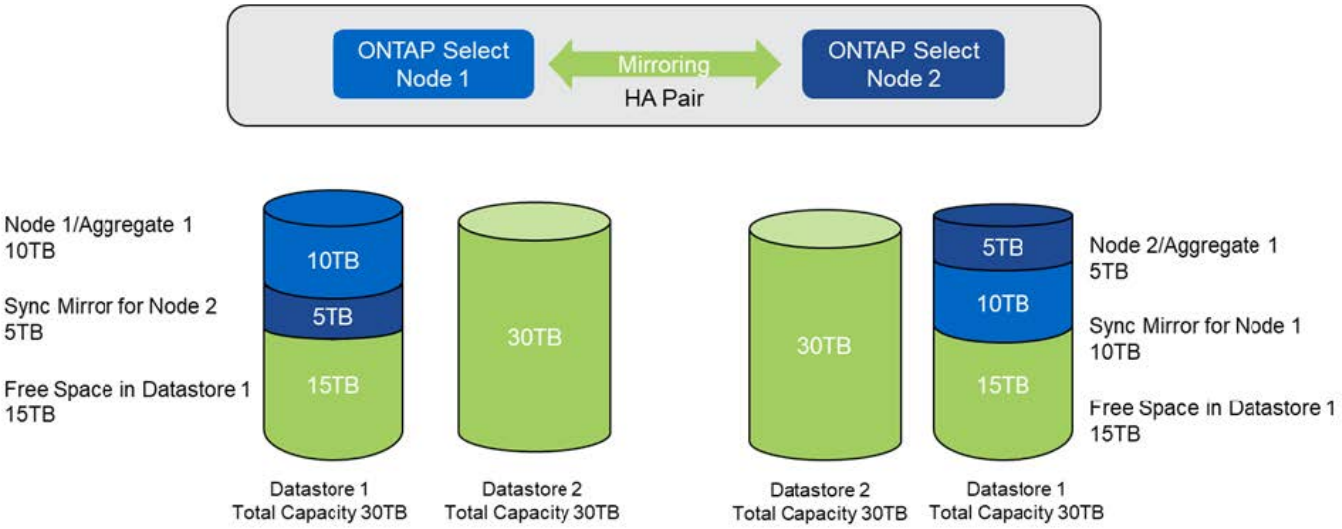
Pour augmenter la capacité accessible aux utilisateurs sur les deux nœuds d'une paire haute disponibilité,

deux opérations d'ajout de stockage doivent être effectuées, une pour chaque nœud. Chaque opération d'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.

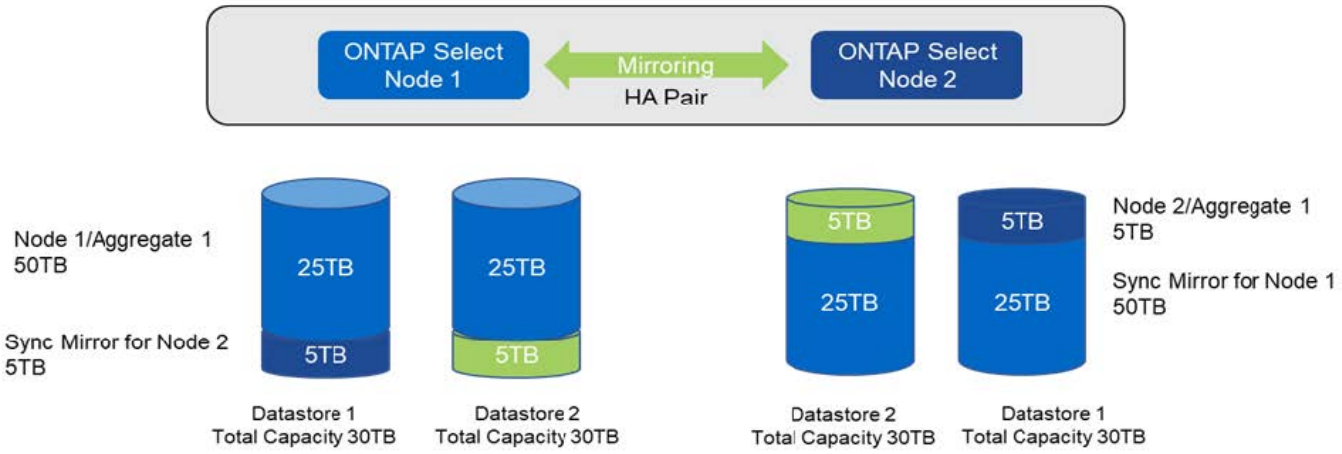
La figure suivante illustre les résultats d'une opération d'ajout de stockage pour le nœud 1. ONTAP Select utilise toujours la même quantité de stockage (15TB) sur chaque nœud. Cependant, le nœud 1 dispose de plus de stockage actif (10TB) que le nœud 2 (5TB). Les deux nœuds sont entièrement protégés,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

Répartition des capacités : espace alloué 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 de la banque de données 1 et une partie de la banque de données 2 (dans la limite de capacité). La première opération d'ajout de stockage utilise les 15 To d'espace libre restants dans la banque de données 1. La figure suivante illustre le résultat de la seconde opération d'ajout de stockage. À ce stade, le nœud 1 gère 50 To de données actives, tandis que le nœud 2 conserve ses 5 To initiaux.

Répartition des capacités : allocation et espace libre après deux opérations d'ajout de stockage supplémentaire pour le nœud 1



La taille maximale des VMDK utilisées lors des opérations d'ajout de capacité est de 16 To. La taille maximale des VMDK utilisées lors des opérations de création de cluster reste de 8 To. ONTAP Deploy crée des VMDK de taille adaptée selon votre configuration (cluster à un seul nœud ou à plusieurs nœuds) et 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 de ONTAP Select avec le RAID logiciel

L'assistant d'ajout de stockage permet également d'augmenter la capacité gérée pour les nœuds ONTAP Select utilisant le RAID logiciel. L'assistant affiche uniquement les disques SSD DAS disponibles pouvant être mappés comme RDM à la machine virtuelle ONTAP Select.

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

> Dans une paire haute disponibilité, l'ajout de stockage au nœud 1 nécessite qu'un nombre identique de disques soit également disponible sur la paire haute disponibilité du nœud (nœud 2). Les disques locaux et distants sont utilisés par une seule opération d'ajout de stockage sur le nœud 1. Autrement dit, les disques distants sont utilisés pour s'assurer que le nouveau stockage sur le 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 identique de disques doivent être disponibles sur les deux nœuds.

ONTAP Select partitionne les nouveaux disques en utilisant les mêmes partitions racine, de données et de données que les disques existants. Cette opération de partitionnement a lieu lors de la création d'un nouvel agrégat ou lors de l'extension d'un agrégat existant. La taille de la bande de la partition racine de chaque disque est définie pour correspondre à la taille de la partition racine existante sur les disques existants. Ainsi, la taille de chacune des deux partitions de données, de taille égale, se calcule en soustrayant la taille de la partition racine de la capacité totale du disque, puis en divisant le résultat par deux. La taille de la bande de la partition racine est variable et elle est calculée lors de la configuration initiale du cluster comme suit. L'espace total requis pour la partition racine (68 Go pour un cluster à nœud unique et 136 Go pour les paires haute disponibilité) est réparti entre le nombre initial de disques, moins les disques de secours et de parité. La taille de la bande de la partition racine reste constante sur tous les disques ajoutés au système.

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

Lors de l'ajout de stockage à un agrégat de données existant, certaines considérations supplémentaires s'imposent. Il est possible d'ajouter des disques à un groupe RAID existant, à condition que ce dernier n'ait pas déjà atteint sa limite maximale. Les bonnes pratiques traditionnelles FAS et AFF pour l'ajout de disques à des groupes RAID existants s'appliquent également ici, et la création d'un point chaud sur le nouveau disque est une préoccupation potentielle. De plus, seuls les disques dont la taille des partitions de données est égale ou supérieure peuvent être ajoutés à un groupe RAID existant. Comme expliqué ci-dessus, la taille de la partition de données n'est pas la même que la taille brute du disque. Si les partitions de données ajoutées sont plus grandes que les partitions existantes, les nouveaux disques sont dimensionnés de manière optimale. 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 propose des options d'efficacité du stockage similaires à celles présentes sur les baies FAS et AFF.

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

Une configuration de type AFF est automatiquement activée sur les nouvelles installations, à condition de disposer d'un stockage DAS avec disques SSD et d'une licence premium.

Avec une personnalité similaire à celle d'AFF, les fonctionnalités SE intégrées suivantes sont automatiquement activées lors de l'installation :

- Détection de motifs zéros en ligne
- Déduplication en ligne du volume
- Déduplication en arrière-plan du volume
- Compression adaptative en ligne
- Compaction des données en ligne
- Déduplication en ligne d'agrégat
- Déduplication en arrière-plan de l'agrégat

Pour vérifier qu'ONTAP Select a activé toutes les règles 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 Si                          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 d'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 à cocher **Activer les règles d'efficacité du stockage** lors de l'installation initiale du cluster avec ONTAP Deploy. L'activation d'une personnalité de type AFF après une mise à niveau d'ONTAP, lorsque les conditions préalables n'ont pas été remplies, nécessite la création manuelle d'un argument de démarrage et le redémarrage du nœud. Contactez le support technique pour plus de détails.

Configurations d'efficacité de stockage ONTAP Select

Le tableau suivant récapitule 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 logicielle.

Fonctionnalités ONTAP Select	DAS SSD (premium ou premium XL[1])	Disque dur DAS (toutes licences)	vNAS (toutes licences)
Détection de 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 sur une base par 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 en 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 à compression	Oui	Oui	Oui, activé par l'utilisateur sur une base par volume
Compaction 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 en ligne d'agrégat	Oui (par défaut)	S/O	Non pris en charge
Déduplication en 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 en arrière-plan de l'agrégat	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

machine virtuelle large est uniquement prise en charge pour les configurations DAS utilisant le RAID logiciel. Les configurations RAID matériel et vNAS ne sont pas prises en charge avec la machine virtuelle ONTAP Select large dans la version 9.6.

Remarques sur le comportement de mise à niveau des configurations SSD DAS

Après la mise à niveau vers ONTAP Select 9.6 ou une version ultérieure, attendez que la commande `system node upgrade-revert show` indique 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 une version ultérieure, un nouveau volume créé sur un agrégat existant ou sur un agrégat nouvellement créé se comporte de la même manière qu'un volume créé lors d'un déploiement neuf. Les volumes existants qui subissent la mise à niveau du code ONTAP Select conservent la plupart des mêmes règles d'efficacité de stockage qu'un volume nouvellement créé, à quelques variations près :

Scénario 1

Si aucune règle d'efficacité n'était activée sur un volume avant la mise à niveau, alors :

- Les volumes avec `space guarantee = volume` ne disposent pas de la compaction des données en ligne, de la déduplication en ligne agrégée ni de la déduplication en arrière-plan agrégée. Ces options peuvent être activées après la mise à niveau.
- Les volumes avec `space guarantee = none` ne disposent pas de la compression en arrière-plan activée. Cette option peut être activée après la mise à niveau.
- Après la mise à niveau, la règle d'efficacité du stockage sur les volumes existants est définie sur auto.

Scénario 2

Si certaines fonctionnalités d'optimisation du stockage sont déjà activées sur un volume avant la mise à niveau, alors :

- Les volumes avec `space guarantee = volume` ne constatent aucune différence après la mise à niveau.
- Volumes avec `space guarantee = none` la déduplication en arrière-plan de l'agrégat activée.
- Les volumes avec `storage policy inline-only` ont leur politique définie sur automatique.
- Les volumes dotés de règles 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 la déduplication en arrière-plan agrégée activée.

Réseautique

Concepts et caractéristiques de mise en réseau ONTAP Select

Commencez par vous familiariser avec les concepts généraux de mise en réseau applicables à l'environnement ONTAP Select. Explorez ensuite les caractéristiques et options spécifiques offertes par les clusters à nœud unique et à nœuds multiples.

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 relative au réseau physique comprend à la fois l'hôte de l'hyperviseur et l'environnement réseau commuté au sens large.

Options de carte réseau hôte

Chaque hôte hyperviseur ONTAP Select doit être configuré avec deux ou quatre ports physiques. La configuration exacte que vous choisissez dépend de plusieurs facteurs, notamment :

- Que le cluster contienne un ou plusieurs hôtes ONTAP Select
- Quel système d'exploitation hyperviseur est utilisé
- Configuration du commutateur virtuel
- Que LACP soit utilisé avec les liens ou non

Configuration physique du commutateur

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. La configuration exacte que vous choisissez dépend de plusieurs facteurs. Les principaux points à prendre en compte sont les suivants :

- Comment comptez-vous maintenir la séparation entre les réseaux internes et externes ?
- Maintiendrez-vous 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. Concrètement, le trafic peut circuler entre les hôtes du cluster ainsi qu'avec les clients de stockage et les autres machines en dehors du cluster. Les commutateurs virtuels gérés par les hyperviseurs contribuent à la prise en charge du réseau logique.

Réseau interne

Dans un déploiement en cluster multi-nœuds, les nœuds ONTAP Select individuels communiquent via un réseau « interne » isolé. Ce réseau n'est pas exposé ni accessible 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 à haute disponibilité (HA-IC)
  - Miroir de synchronisation RAID (RSM)
- Réseau de couche 2 unique basé sur un VLAN
- Les adresses IP statiques sont attribuées par ONTAP Select :
  - IPv4 uniquement
  - DHCP non utilisé
  - Adresse locale du lien
- La taille MTU est de 9000 octets par défaut et peut être ajustée dans une plage de 7500 à 9000 (inclus)

Réseau externe

Le réseau externe traite le trafic entre les nœuds d'un cluster ONTAP Select et les clients de stockage externe, ainsi que les autres machines. Le réseau externe fait partie de chaque 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)
- Prise en charge optionnelle des VLAN :
  - groupe de ports de données
  - Groupe de ports de gestion
- Adresses IP attribuées en fonction des choix de configuration de l'administrateur :
  - IPv4 ou IPv6
- La taille MTU est de 1500 octets par défaut (peut être ajustée)

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

Environnement réseau de machine virtuelle

L'hôte hyperviseur fournit plusieurs fonctionnalités réseau.

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

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

Commutateur virtuel
: Le logiciel de commutation virtuelle dans l'environnement de l'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, selon votre environnement.

Configurations réseau ONTAP Select à nœud unique et à nœuds multiples

ONTAP Select prend en charge les configurations réseau à nœud unique et à nœud multiple.

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, d'interconnexion haute disponibilité ou de mise en miroir.

Contrairement à la version multi-nœuds 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.

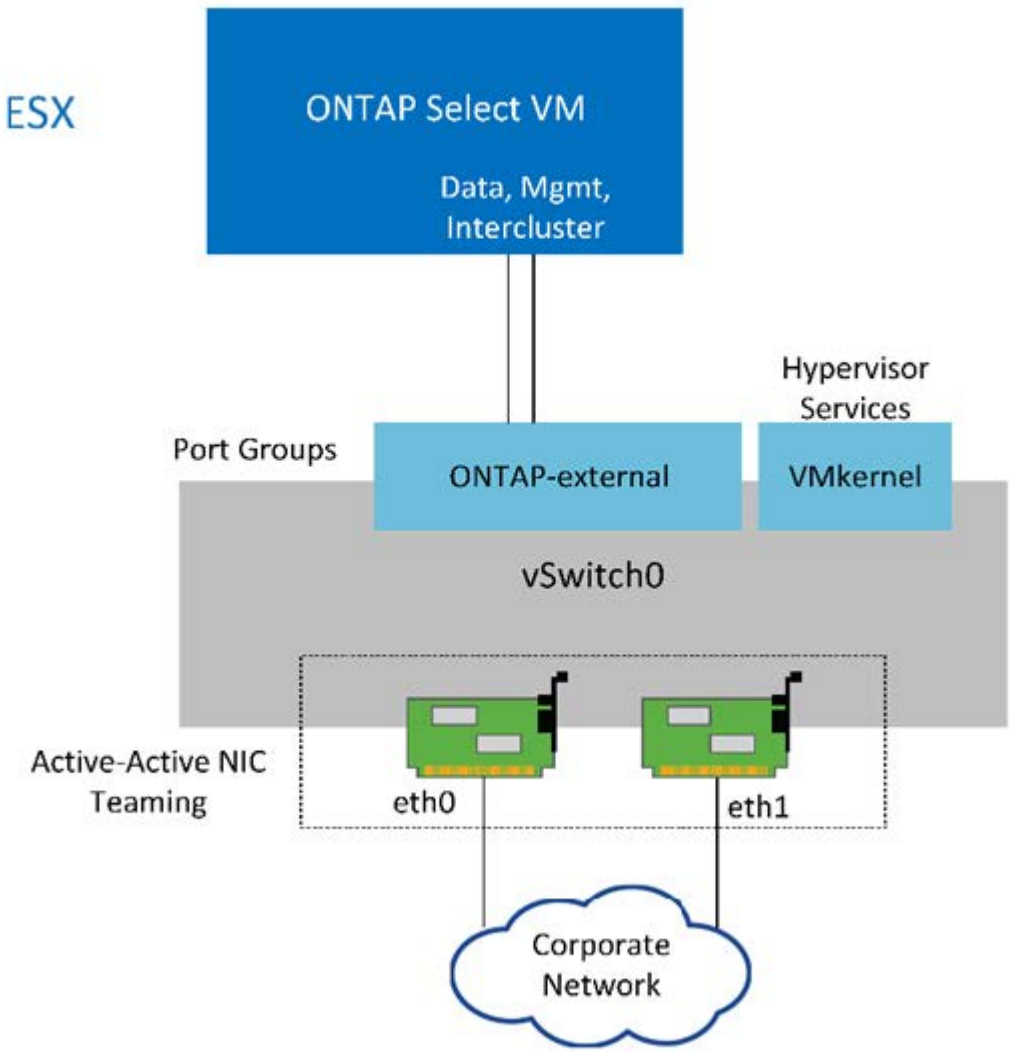
KVM

Vous pouvez déployer ONTAP Select sous forme de cluster à nœud unique. L'hôte hyperviseur inclut un commutateur virtuel qui permet d'accéder au réseau externe.

ESXi

La figure suivante illustre la relation entre ces ports et les adaptateurs physiques sous-jacents. La figure représente un nœud de cluster ONTAP Select sur l'hyperviseur ESXi.

Configuration réseau d'un cluster ONTAP Select à nœud unique



Même si deux adaptateurs suffisent pour un cluster à nœud unique, l'agrégation de cartes réseau reste nécessaire.

Affectation de LIF

Comme expliqué dans la section relative à l'attribution des LIF multi-nœuds de ce document, ONTAP utilise des espaces IP pour séparer le trafic réseau du cluster du trafic de données et de gestion. La version mono-nœud de cette plateforme ne comporte pas de réseau de cluster. Par conséquent, aucun port n'est présent dans l'espace IP du cluster.

Les LIF de gestion de cluster et de nœud sont créées automatiquement lors de la configuration du cluster ONTAP Select. Vous pouvez créer les LIF restantes après le déploiement.

LIF de gestion et de données (e0a, e0b et e0c)

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

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

Configuration réseau multi-nœuds

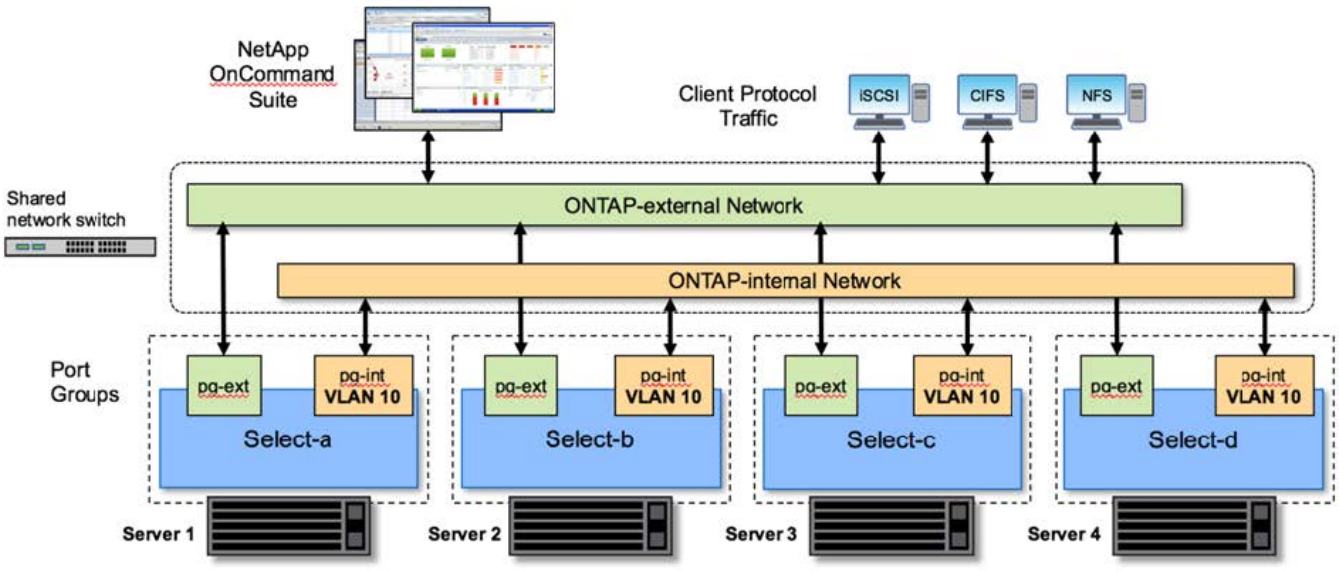
La configuration réseau multi-nœuds ONTAP Select se compose de deux réseaux.

Il s'agit d'un réseau interne, chargé de fournir les services de cluster et de réplication interne, et d'un réseau externe, chargé de fournir les services d'accès aux données et de gestion. L'isolation de bout en bout du trafic circulant au sein de ces deux réseaux est extrêmement importante pour construire un environnement adapté à la résilience du cluster.

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

Chaque instance d'ONTAP Select réside sur un serveur physique distinct. Le trafic interne et externe est isolé grâce à des groupes de ports réseau séparés, qui sont attribués à chaque interface réseau virtuelle et permettent aux nœuds du cluster de partager la même infrastructure de commutation physique.

Aperçu d'une configuration réseau de cluster multi-nœuds ONTAP Select



Chaque machine virtuelle ONTAP Select contient sept adaptateurs réseau virtuels présentés à ONTAP comme un ensemble de sept ports réseau, de e0a à e0g. Bien qu'ONTAP traite ces adaptateurs comme des cartes

réseau physiques, ils sont en réalité virtuels et correspondent à un ensemble d'interfaces physiques via une couche réseau virtualisée. Par conséquent, chaque serveur hôte ne nécessite pas six ports réseau physiques.

L'ajout de cartes réseau virtuelles à la machine virtuelle ONTAP Select n'est pas pris en charge.

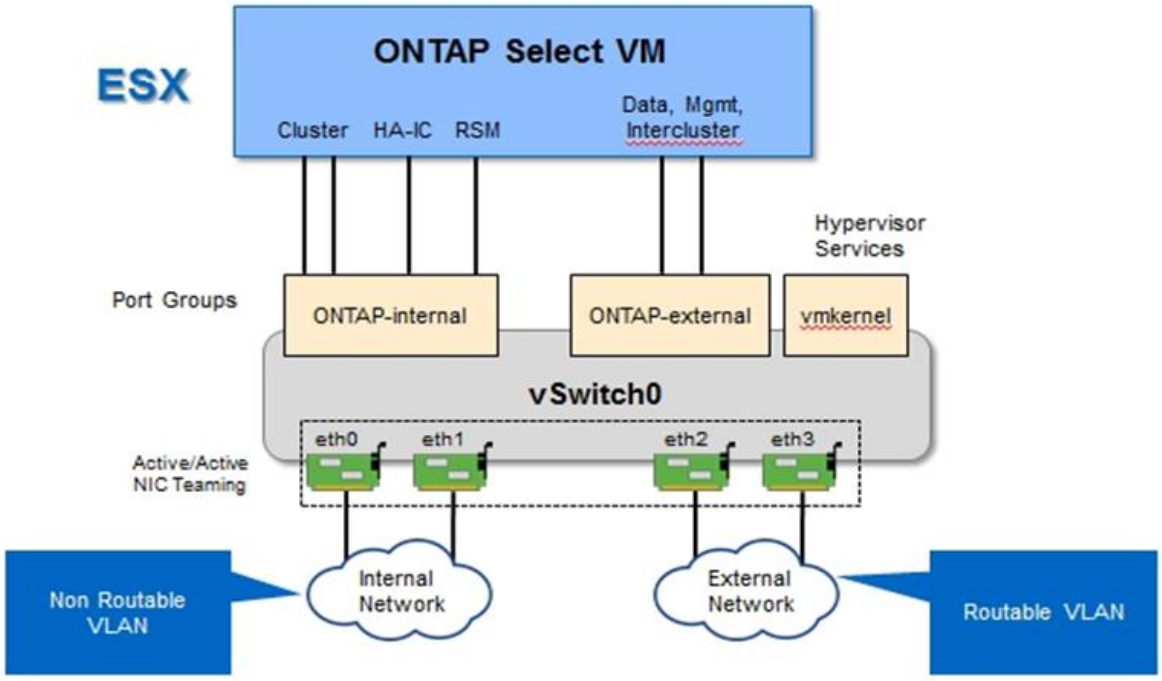
Ces ports sont préconfigurés pour fournir les services suivants :

- e0a, e0b et e0g. LIF de gestion et de données
- e0c, e0d. LIF de réseau de cluster
- e0e. RSM
- e0f. Interconnexion haute disponibilité

Les ports e0a, e0b et e0g se trouvent sur le réseau externe. Bien que les ports e0c à e0f remplissent plusieurs fonctions différentes, ils constituent ensemble le réseau interne Select. Lors de la prise de décisions de conception du réseau, vous devez placer ces ports sur un seul réseau de couche 2. Il n'est pas nécessaire de répartir 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 ESXi.

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 un accès insuffisant aux ressources réseau, introduisant des latences dans le système. De plus, l'agrégation via le teaming de cartes réseau permet au nœud de cluster ONTAP Select de continuer à accéder au réseau si un seul adaptateur réseau tombe en panne.

Les groupes de ports du réseau externe et du réseau interne contiennent chacun les quatre cartes réseau de manière symétrique. Les ports actifs du groupe de ports du réseau externe

sont les ports de secours du réseau interne. Inversement, les ports actifs du groupe de ports du réseau interne sont les ports de secours du groupe de ports du réseau externe.

Affectation de LIF

Avec l'introduction des IPspaces, les rôles de ports ONTAP ont été dépréciés. Comme les baies FAS, les clusters ONTAP Select contiennent à la fois un IPspace par défaut et un IPspace de cluster. En plaçant les ports réseau e0a, e0b et e0g dans l'IPspace par défaut et les ports e0c et e0d dans l'IPspace de cluster, ces ports ont essentiellement été isolés de l'hébergement de LIFs qui ne leur appartiennent pas. Les ports restants au sein du cluster ONTAP Select sont utilisés via l'attribution automatique d'interfaces fournissant des services internes. Ils ne sont pas exposés via le shell ONTAP, comme c'est le cas pour les interfaces RSM et d'interconnexion haute disponibilité.

Toutes les interfaces logiques (LIF) ne sont pas visibles via l'interface de ligne de commande ONTAP. Les interfaces d'interconnexion haute disponibilité (HA interconnect) et RSM sont masquées dans ONTAP et utilisées en interne pour fournir leurs services respectifs.

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

LIF de gestion et de données (e0a, e0b et e0g)

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

- Trafic des protocoles 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 de cluster et de nœud sont créées automatiquement lors de la configuration du cluster ONTAP Select. Vous pouvez créer les LIF restantes après le déploiement.

LIF du réseau de cluster (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 d'ONTAP à l'aide d'adresses IP link local (169.254.x.x).

Vous ne pouvez pas attribuer d'adresse IP statique à ces interfaces et vous ne devez pas créer d'interfaces de cluster supplémentaires.

Le trafic réseau du cluster doit transiter par un réseau de couche 2 non routé et à faible latence. Compte tenu des exigences de débit et de latence du cluster, vous devez installer physiquement le cluster ONTAP Select à proximité (par exemple, multipack, centre de données unique). La création de configurations de cluster étendu à quatre, six, huit, dix ou douze nœuds en séparant les nœuds de la paire haute disponibilité via 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, consultez la section "Meilleures pratiques pour la paire 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 (MTU de 7500 à 9000). Pour un fonctionnement correct 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 sur le port réseau e0e. Cette fonctionnalité se produit automatiquement, en utilisant les interfaces réseau configurées par ONTAP lors de la mise en place 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 l'interface LIF hébergée ne sont visibles dans l'interface de ligne de commande ONTAP ou dans System Manager. Cette interface est configurée pour utiliser une adresse IP locale de liaison générée automatiquement, et il est impossible d'attribuer une autre adresse IP. Ce port réseau requiert l'utilisation de trames jumbo (7500 à 9000 MTU).

interconnexion haute disponibilité (e0f)

NetApp FAS utilise du matériel spécialisé pour transmettre des informations entre les paires HA dans un cluster ONTAP. Cependant, les environnements définis par logiciel 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 d'ONTAP relatives au transport d'interconnexion ont imposé que cette fonctionnalité soit émulée par logiciel. Par conséquent, au sein d'un cluster ONTAP Select, la fonctionnalité de l'interconnexion haute disponibilité (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, qui est responsable de deux fonctions principales :

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

Le trafic d'interconnexion haute disponibilité transite par ce port réseau en utilisant une seule interface réseau, en superposant des trames d'accès direct à la mémoire à distance (RDMA) dans des paquets Ethernet.

À l'instar du 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 depuis System Manager. Par conséquent, vous ne pouvez pas modifier l'adresse IP de cette interface et vous ne pouvez pas changer l'état du port. Ce port réseau requiert l'utilisation de trames jumbo (7500 à 9000 MTU).

Réseaux internes et externes ONTAP Select

Caractéristiques des réseaux internes et externes ONTAP Select.

Réseau interne ONTAP Select

Le réseau interne ONTAP Select, présent uniquement dans la variante multi-nœuds du produit, est responsable de fournir au cluster ONTAP Select la communication de cluster, l'interconnexion haute disponibilité et les services de réplication synchrone. Ce réseau comprend les ports et interfaces suivants :

- **e0c, e0d.** Hébergement des LIF réseau de cluster
- **e0e.** Hébergement de la LIF RSM
- **e0f.** Hébergement de la LIF d'interconnexion haute disponibilité

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 indispensable à la sécurité du cluster et permet de garantir que les interfaces système restent séparées 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 le trafic autre que celui du cluster Select, comme le trafic applicatif ou de gestion, n'est pas prise en charge. Aucun autre hôte ou machine virtuelle ne peut se trouver sur le VLAN interne ONTAP.

Les paquets réseau transitant par le réseau interne doivent emprunter un réseau de couche 2 dédié et étiqueté VLAN. Ceci peut être réalisé en effectuant 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 lorsque ce VLAN natif n'est utilisé pour aucun autre trafic (attribution d'un groupe de ports sans ID de VLAN, c'est-à-dire en mode EST)

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



Seuls les vSwitches standard et distribués ESXi sont pris en charge. Les autres commutateurs virtuels ou la connectivité directe entre hôtes ESXi ne sont pas pris en charge. Le réseau interne doit être entièrement ouvert ; les NAT ou 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 à l'aide d'objets réseau virtuels de couche 2 appelés groupes de ports. L'attribution correcte du vSwitch à ces groupes de ports est extrêmement importante, en particulier pour le réseau interne, qui est responsable de fournir les services de cluster, d'interconnexion haute disponibilité et de réplication miroir. Une bande passante réseau insuffisante pour ces ports réseau peut entraîner une dégradation des performances et même affecter la stabilité du nœud de cluster. Par conséquent, les clusters à quatre, six, huit, dix et douze nœuds exigent que le réseau interne ONTAP Select utilise une connectivité 10 Gb ; les cartes réseau 1 Gb ne sont pas prises en charge. Des compromis peuvent toutefois être faits pour le réseau externe, car limiter le flux de données entrantes vers un cluster ONTAP Select n'affecte pas sa capacité à fonctionner de manière fiable.

Un cluster à deux nœuds peut utiliser soit quatre ports 1 Gb pour le trafic interne, soit un seul port 10 Gb au lieu des deux ports 10 Gb requis par un cluster à quatre nœuds. Dans un environnement où les conditions empêchent l'installation de quatre cartes réseau 10 Gb sur le serveur, deux cartes réseau 10 Gb peuvent être utilisées pour le réseau interne et deux cartes réseau 1 Gb pour le réseau ONTAP externe.

Validation et dépannage du réseau interne

Le réseau interne d'un cluster multi-nœuds peut être validé à l'aide de la fonctionnalité de vérification de la connectivité réseau. Cette fonction peut être appelée depuis l'interface de ligne de commande de déploiement en exécutant la commande `network connectivity-check start`.

Exécutez la commande suivante pour afficher le résultat 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. L'outil ne doit pas être utilisé pour le dépannage des clusters mononœud (y compris les configurations vNAS), la connectivité entre ONTAP Deploy et ONTAP Select, ni les problèmes de connectivité côté client.

L'assistant de création de cluster (faisant partie de l'interface utilisateur ONTAP Deploy) inclut le vérificateur de réseau interne comme étape optionnelle disponible lors de la création de clusters multi-nœuds. Étant donné le rôle important que le réseau interne joue dans les clusters multi-nœuds, faire de cette étape une partie du flux de travail de création de cluster améliore le taux de réussite des opérations de création de cluster.

À partir d'ONTAP Deploy 2.10, la taille MTU utilisée par le réseau interne peut être définie entre 7 500 et 9 000. L'outil de vérification de la connectivité réseau peut également être utilisé pour tester une taille de MTU comprise entre 7 500 et 9 000. La valeur MTU par défaut est définie sur la valeur du commutateur réseau virtuel. Cette valeur par défaut doit être remplacée par une valeur plus petite si une superposition réseau telle que VXLAN est présente dans l'environnement.

Réseau externe ONTAP Select

Le réseau externe ONTAP Select est responsable de toutes les communications sortantes du cluster et, par conséquent, est présent dans les configurations à nœud unique et à nœuds multiples. Bien que ce réseau ne soit pas soumis aux mêmes exigences de débit strictes que le réseau interne, l'administrateur doit veiller à ne pas créer de goulots d'étranglement réseau entre le client et la machine virtuelle ONTAP, car des problèmes de performance pourraient être à tort considérés comme des problèmes ONTAP Select.

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

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

Référence rapide : réseau interne versus réseau externe

Description	Réseau interne	Réseau externe
Services réseau	RAID haute disponibilité/interconnexion SyncMirror (RSM)	Gestion des données Intercluster (SnapMirror et SnapVault)
Isolation du réseau	Obligatoire	Facultatif
Taille de 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 (NIC teaming)

Pour garantir que les réseaux internes et externes disposent à la fois de la bande passante et des caractéristiques de résilience nécessaires pour offrir des performances élevées et une tolérance aux pannes, il est recommandé d'utiliser l'agrégation de cartes réseau physiques. Les configurations de cluster à deux nœuds avec une seule liaison 10 Gb sont prises en charge. Cependant, la NetApp bonne pratique recommandée consiste à utiliser l'agrégation de cartes réseau (NIC teaming) à la fois 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. L'utilitaire utilise un identifiant unique d'organisation (OUI) spécifique à la plateforme, propre à NetApp, pour s'assurer qu'il n'y a pas de conflit avec les systèmes FAS. Une copie de cette adresse est ensuite stockée dans une base de données interne au sein de la machine virtuelle d'installation ONTAP Select (ONTAP Deploy), afin d'empêcher toute réattribution accidentelle lors de futurs déploiements de nœuds. L'administrateur ne doit à aucun moment modifier l'adresse MAC attribuée à un port réseau.

Configurations réseau ONTAP Select prises en charge

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

Les fournisseurs de serveurs savent que les besoins des clients sont variés et que le choix est essentiel. C'est pourquoi, lors de l'achat d'un serveur physique, de nombreuses options sont disponibles lors de la prise de décisions concernant la connectivité réseau. La plupart des systèmes standard sont livrés avec différents choix de cartes réseau (NIC) offrant des options mono-port et multiport avec diverses permutations de vitesse et de débit. Cela inclut la prise en charge des adaptateurs NIC 25Gb/s et 40Gb/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 en sélectionnant des cartes réseau (NIC) plus rapides permet d'obtenir un cluster plus performant et une meilleure expérience utilisateur globale. Quatre NIC 10 Gb ou deux NIC plus rapides (25/40 Gb/s) peuvent être utilisés pour obtenir une architecture réseau haute performance. Un certain nombre d'autres configurations sont également prises en charge. Pour les clusters à deux nœuds, 4 ports 1 Gb ou 1 port 10 Gb sont pris en charge. Pour les clusters à un seul nœud, 2 ports 1 Gb sont pris en charge.

Configurations réseau minimales et recommandées

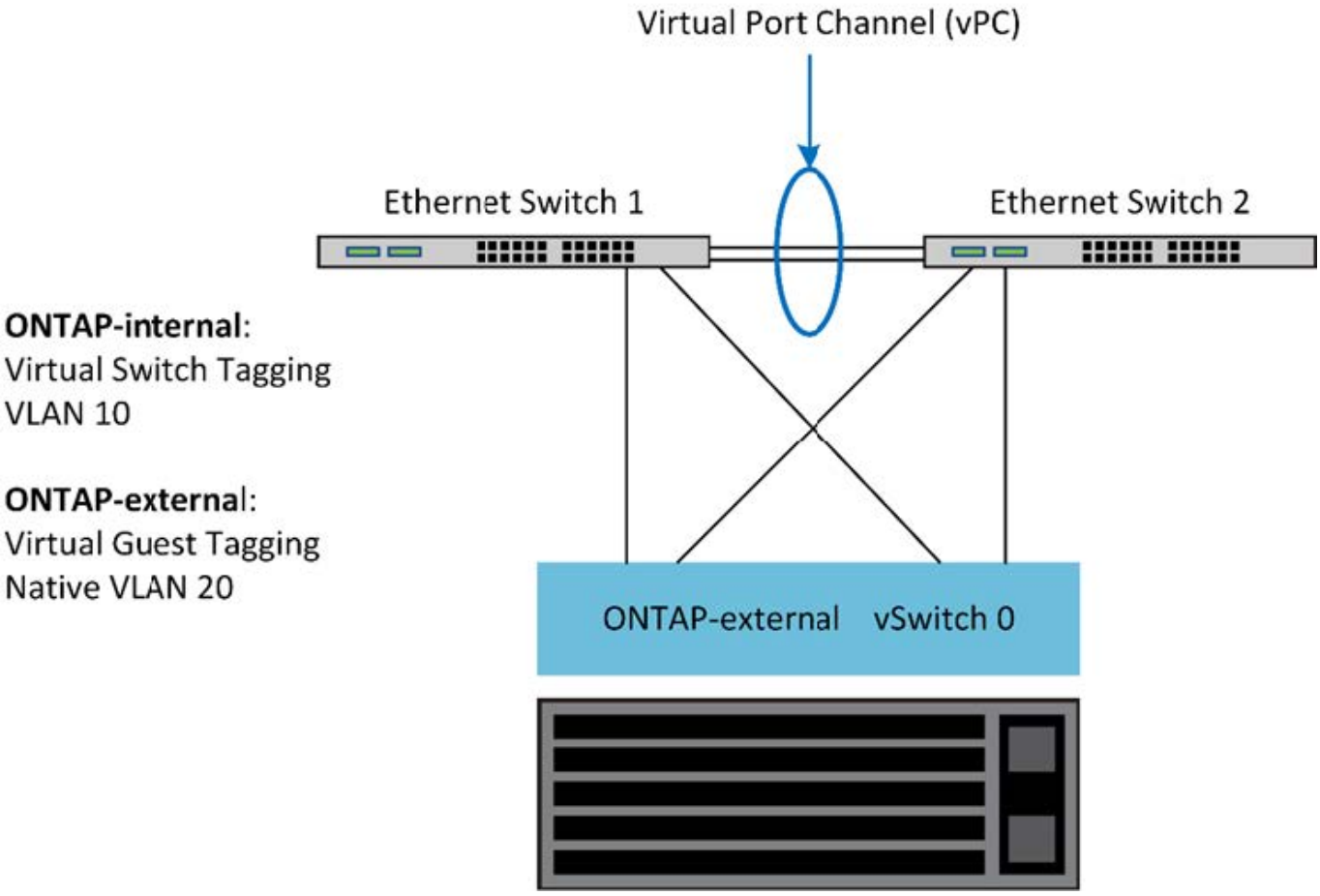
Plusieurs configurations Ethernet sont prises en charge en fonction de la taille du cluster.

Taille du cluster	Exigences minimales	Recommandation
Cluster à nœud unique	2 x 1GbE	2 x 10GbE
Cluster à deux nœuds ou MetroCluster SDS	4 x 1 GbE ou 1 x 10 GbE	2 x 10GbE
Cluster de quatre, six, huit, dix ou douze nœuds	2 x 10GbE	4 x 10GbE ou 2 x 25/40GbE

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 possible de convertir entre différentes configurations d'agrégation de cartes réseau requises pour chaque topologie.

Configuration réseau utilisant plusieurs commutateurs physiques

Lorsque le matériel disponible est suffisant, NetApp recommande d'utiliser la configuration multiswitch illustrée dans la figure suivante, en raison de la protection supplémentaire contre les pannes physiques des commutateurs.



Configuration ONTAP Select VMware vSphere vSwitch sur ESXi

Configuration ONTAP Select vSwitch et politiques d'équilibrage de charge pour les configurations à deux et quatre cartes réseau.

ONTAP Select prend en charge l'utilisation des configurations vSwitch standard et distribuées. Les vSwitches distribués prennent en charge les constructions d'agrégation de liens (LACP). L'agrégation de liens est une construction réseau courante utilisée pour agréger la bande passante de l'agrégat sur plusieurs adaptateurs physiques. LACP est une norme indépendante des fournisseurs. Il fournit un protocole ouvert pour les points de terminaison réseau qui regroupent des groupes de ports réseau physiques en un seul canal logique. ONTAP Select peut fonctionner avec des groupes de ports configurés comme groupe d'agrégation de liens (LAG). Cependant, NetApp recommande d'utiliser les ports physiques individuels comme ports de liaison montante (trunk) simples afin d'éviter la configuration LAG. Dans ces cas, les bonnes pratiques pour les vSwitches standard et distribués sont identiques.

Cette section décrit la configuration du vSwitch et les politiques d'équilibrage de charge à utiliser dans les configurations à deux et quatre cartes réseau.

Lors de la configuration des groupes de ports pour ONTAP Select, suivez ces bonnes pratiques ; la stratégie d'équilibrage de charge au niveau du groupe de ports est Route Based on Originating Virtual Port ID. VMware

recommande de configurer le protocole STP sur Portfast sur les ports du commutateur connectés aux hôtes ESXi.

Toutes les configurations vSwitch nécessitent au minimum deux cartes réseau physiques regroupées en une seule équipe NIC. ONTAP Select prend en charge une liaison unique de 10Gb pour les clusters à deux nœuds. Cependant, NetApp recommande d'utiliser l'agrégation de cartes réseau afin de garantir la redondance matérielle.

Sur un serveur vSphere, les équipes de cartes réseau (NIC teams) sont la structure d'agrégation utilisée pour regrouper plusieurs adaptateurs réseau physiques en un seul canal logique, permettant ainsi de répartir la charge réseau entre tous les ports membres. Il est important de se rappeler que les équipes de cartes réseau peuvent être créées sans le support du commutateur physique. Les stratégies d'équilibrage de charge et de basculement peuvent être appliquées directement à une équipe de cartes réseau, qui ne connaît pas la configuration du commutateur en amont. Dans ce cas, les stratégies ne s'appliquent qu'au trafic sortant.

Les canaux de ports 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 afin d'utiliser un groupe de ports pour le réseau externe et soit le même groupe de ports, soit un groupe de ports différent (en option) pour le trafic de gestion du cluster et du nœud. Pour les clusters à nœud unique, vous pouvez ajouter le nombre souhaité de ports physiques au groupe de ports externe en tant qu'adaptateurs actifs.

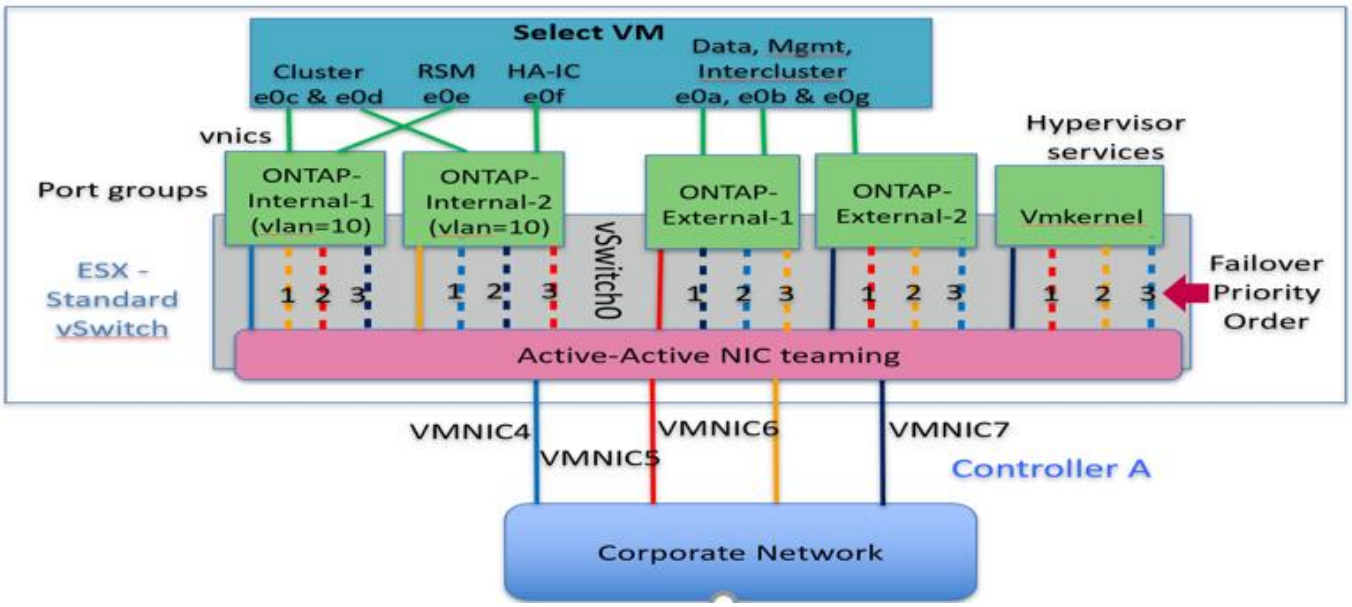
Pour les clusters multi-nœuds, ONTAP Deploy configure chaque machine virtuelle ONTAP Select pour utiliser un ou deux groupes de ports pour le réseau interne et, séparément, un ou deux groupes de ports pour le réseau externe. Le trafic de gestion du cluster et du nœud peut soit utiliser le même groupe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

ONTAP Select prend en charge un maximum de quatre VMNIC.

Standard ou distribué vSwitch et quatre ports physiques par nœud

Vous pouvez affecter quatre groupes de ports à chaque nœud dans un cluster multi-nœuds. Chaque groupe de ports dispose d'un port physique actif et de trois ports physiques de secours, comme dans la figure suivante.

vSwitch avec quatre ports physiques par nœud



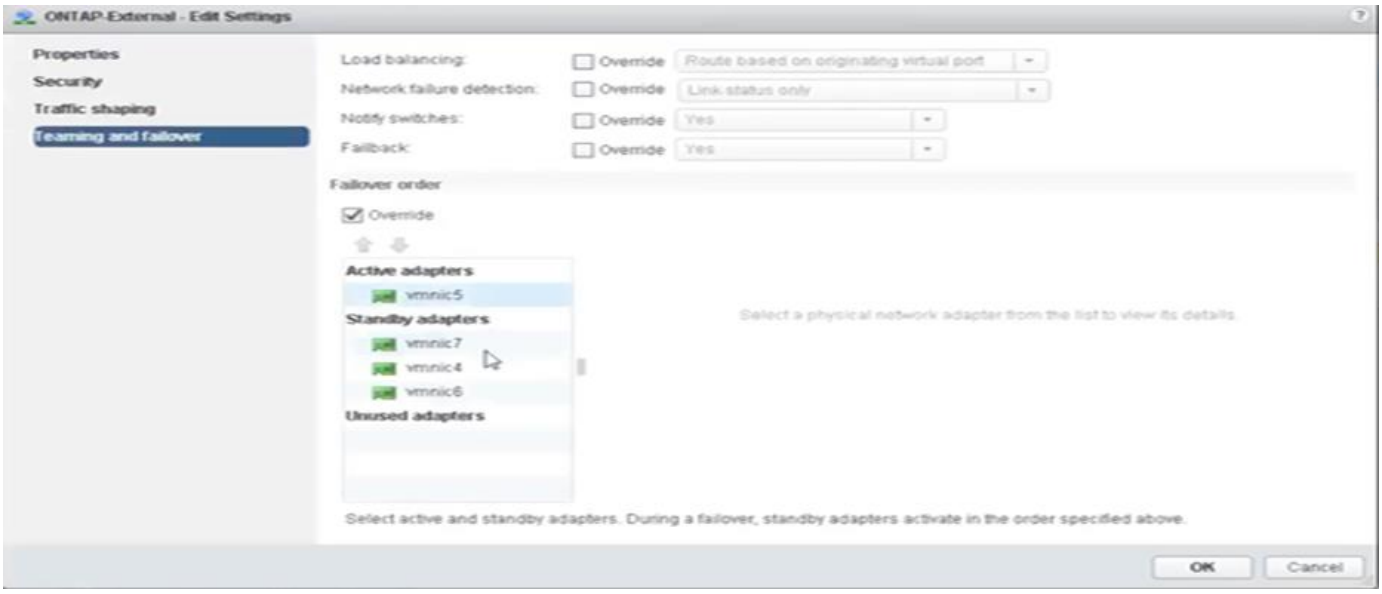
L'ordre des ports dans la liste de secours est important. Le tableau suivant présente un exemple de répartition physique des ports entre les quatre groupes de ports.

Configurations réseau minimales et recommandées

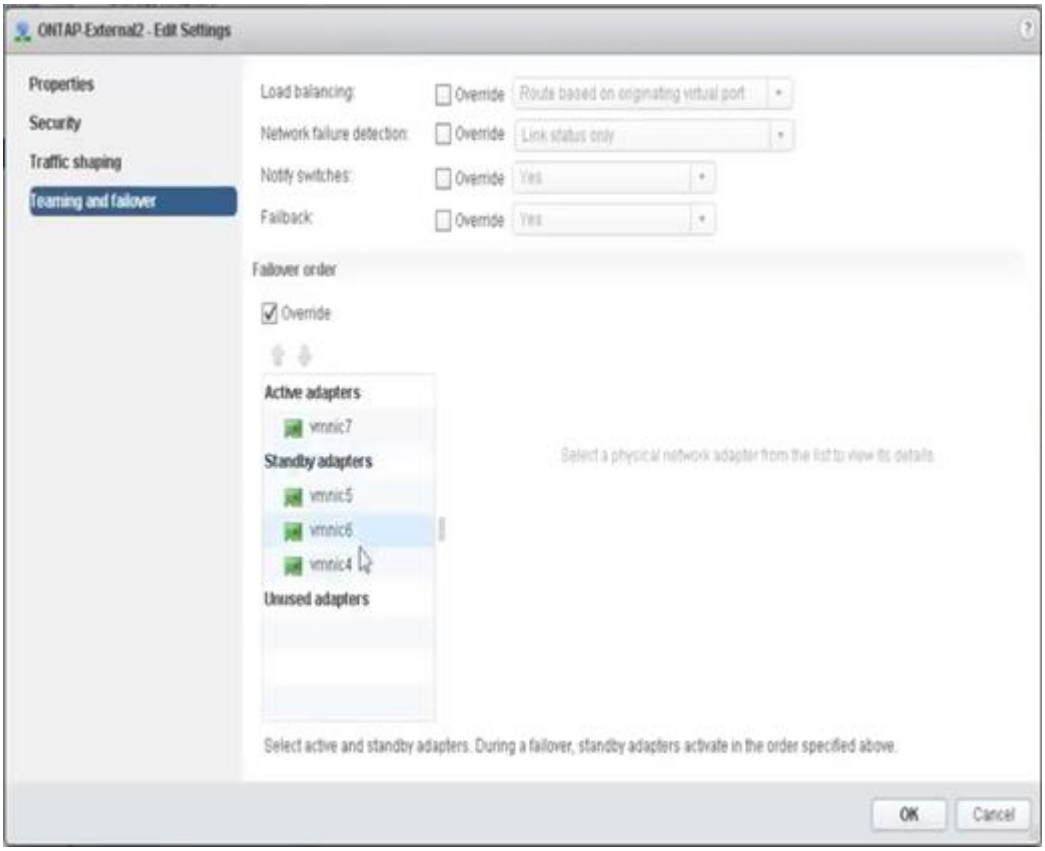
Groupe de ports	Externe 1	Externe 2	Interne 1	Interne 2
Actif	vmnic0	vmnic1	vmnic2	vmnic3
Veille 1	vmnic1	vmnic0	vmnic3	vmnic2
Veille 2	vmnic2	vmnic3	vmnic0	vmnic1
Veille 3	vmnic3	vmnic2	vmnic1	vmnic0

Les figures suivantes illustrent la configuration des groupes de ports réseau externes depuis l'interface utilisateur 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 (vmnic 0 à 3 ne sont pas utilisés dans cet exemple). L'ordre des adaptateurs de secours assure une bascule hiérarchique, les ports du réseau interne étant placés en dernier. L'ordre des ports internes dans la liste de secours est également inversé entre les deux groupes de ports externes.

Partie 1 : configurations des groupes de ports externes ONTAP Select



Partie 2 : configurations des groupes de ports externes ONTAP Select



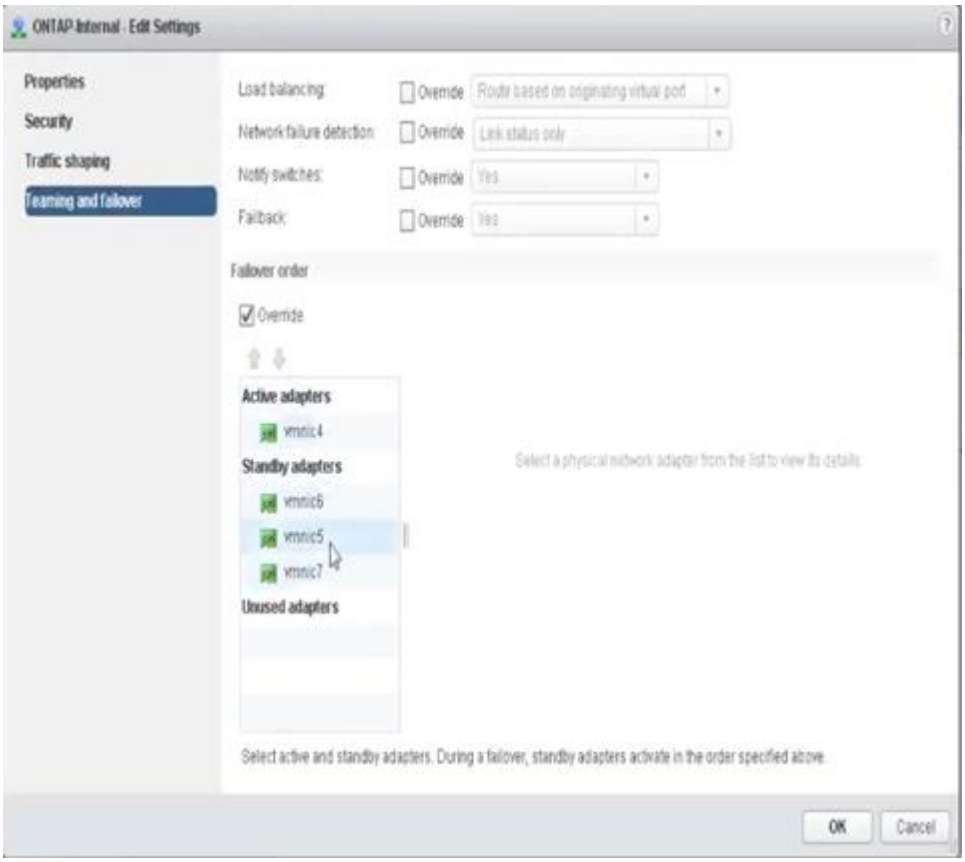
Pour plus de clarté, les affectations sont les suivantes :

ONTAP-External	ONTAP-External2
Adaptateurs actifs : vmnic5 Adaptateurs en veille : vmnic7, vmnic4, vmnic6	Adaptateurs actifs : vmnic7 Adaptateurs en veille : vmnic5, vmnic6, vmnic4

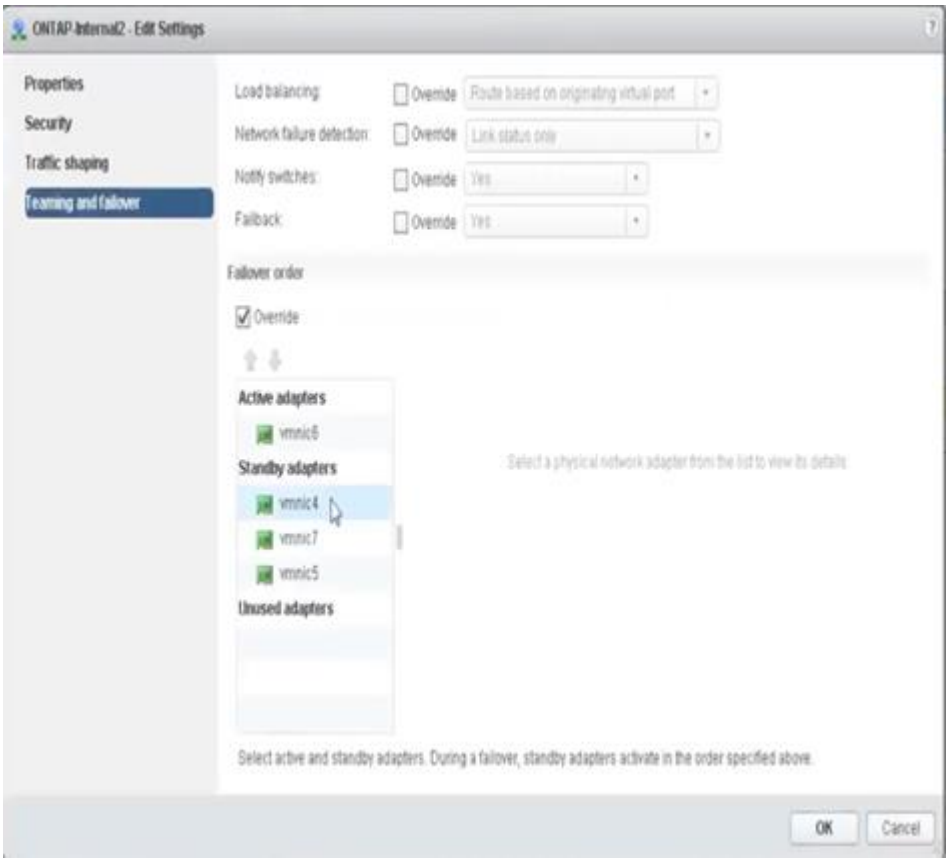
Les figures suivantes illustrent la configuration des groupes de ports du réseau interne (ONTAP-Internal et ONTAP-Internal2). Notez que les adaptateurs actifs proviennent de cartes réseau différentes. Dans cette

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 une bascule hiérarchique, les ports du réseau externe étant placés en dernier. L'ordre des ports externes dans la liste de secours est également inversé entre les deux groupes de ports internes.

Partie 1 : Configurations des groupes de ports internes ONTAP Select



Partie 2 : groupes de ports internes ONTAP Select



Pour plus de clarté, les affectations sont les suivantes :

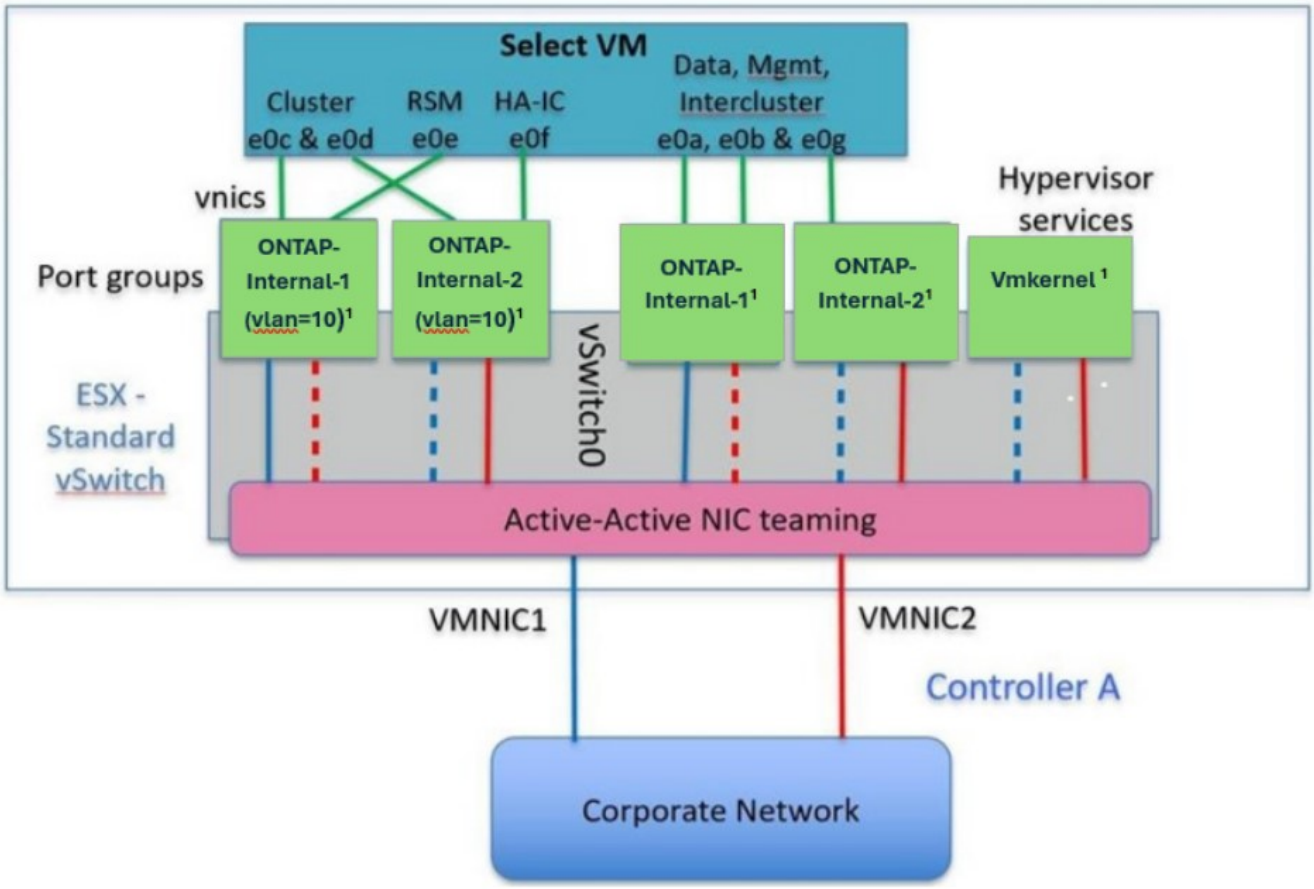
ONTAP-Interne	ONTAP-Internal2
Adaptateurs actifs : vmnic4 Adaptateurs en veille : vmnic6, vmnic5, vmnic7	Adaptateurs actifs : vmnic6 Adaptateurs en veille : vmnic4, vmnic7, vmnic5

Standard ou distribué vSwitch et deux ports physiques par nœud

Lorsqu'on utilise deux cartes réseau haut débit (25/40Gb), la configuration recommandée des groupes de ports est conceptuellement très similaire à celle avec quatre adaptateurs 10Gb. Vous devriez utiliser quatre groupes de ports même lorsque vous n'utilisez que deux adaptateurs physiques. Les attributions des groupes de ports sont les suivantes :

Groupe de ports	Externe 1 (e0a,e0b)	Interne 1 (e0c,e0e)	Interne 2 (e0d,e0f)	Externe 2 (e0g)
Actif	vmnic0	vmnic0	vmnic1	vmnic1
Veille	vmnic1	vmnic1	vmnic0	vmnic0

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

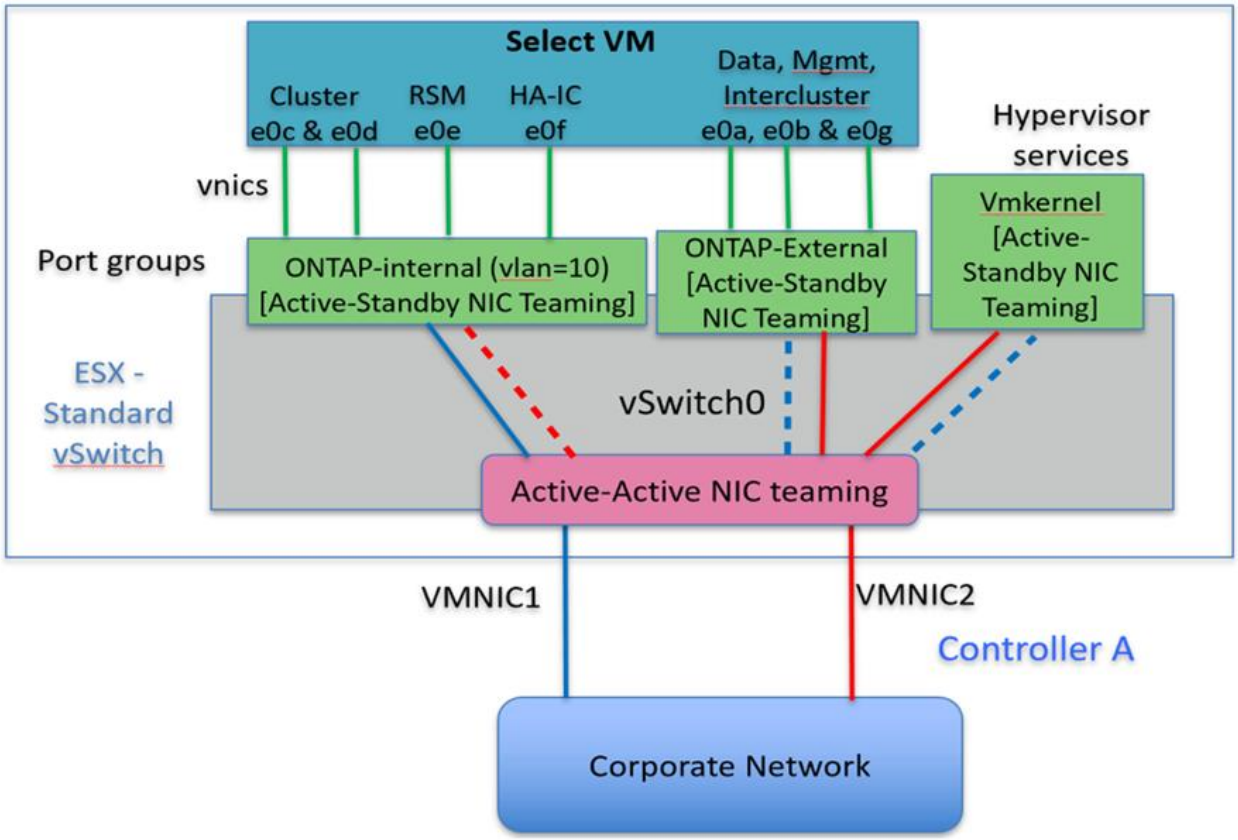


[1] The port groups attached to the virtual NICs are configured to use one NIC as active and the rest as standby.

Lors de l'utilisation de deux ports physiques (10Gb ou moins), chaque groupe de ports doit comporter un adaptateur actif et un adaptateur de secours configurés de manière opposée. Le réseau interne est uniquement présent pour les clusters ONTAP Select multi-nœuds. Pour les clusters mono-nœud, les deux adaptateurs peuvent être configurés comme actifs dans le groupe de ports externe.

L'exemple suivant montre la configuration d'un vSwitch et des deux groupes de ports responsables de la gestion des services de communication interne et externe pour un cluster ONTAP Select multinœud. Le réseau externe peut utiliser la VMNIC du réseau interne en cas de panne réseau, car les VMNICs du réseau interne font partie de ce groupe de ports et sont configurées en mode veille. L'inverse est vrai pour le réseau interne. L'alternance des VMNIC actives et de secours entre les deux groupes de ports est essentielle pour le basculement correct des machines virtuelles ONTAP Select lors des pannes réseau.

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

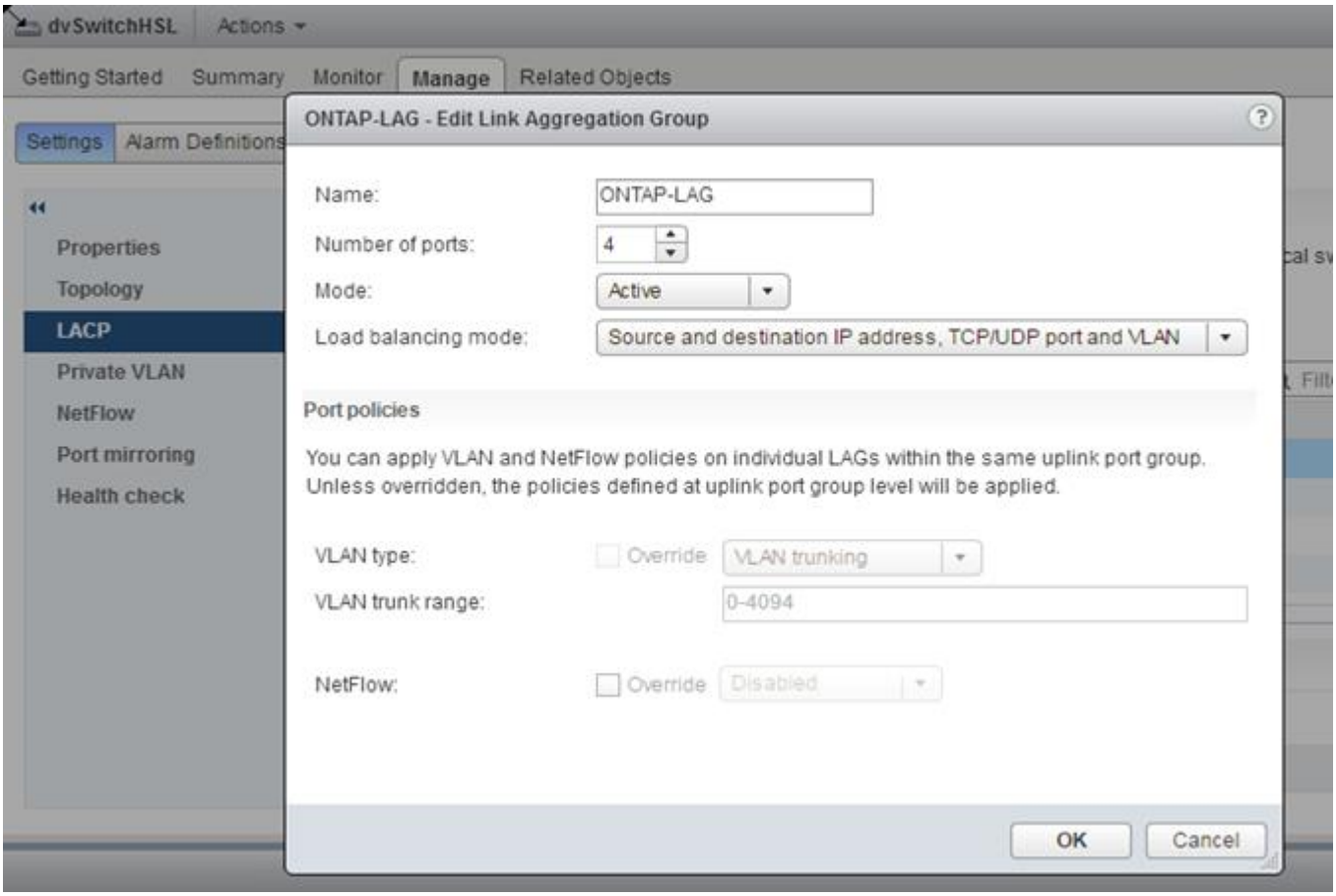


Distributed vSwitch avec LACP

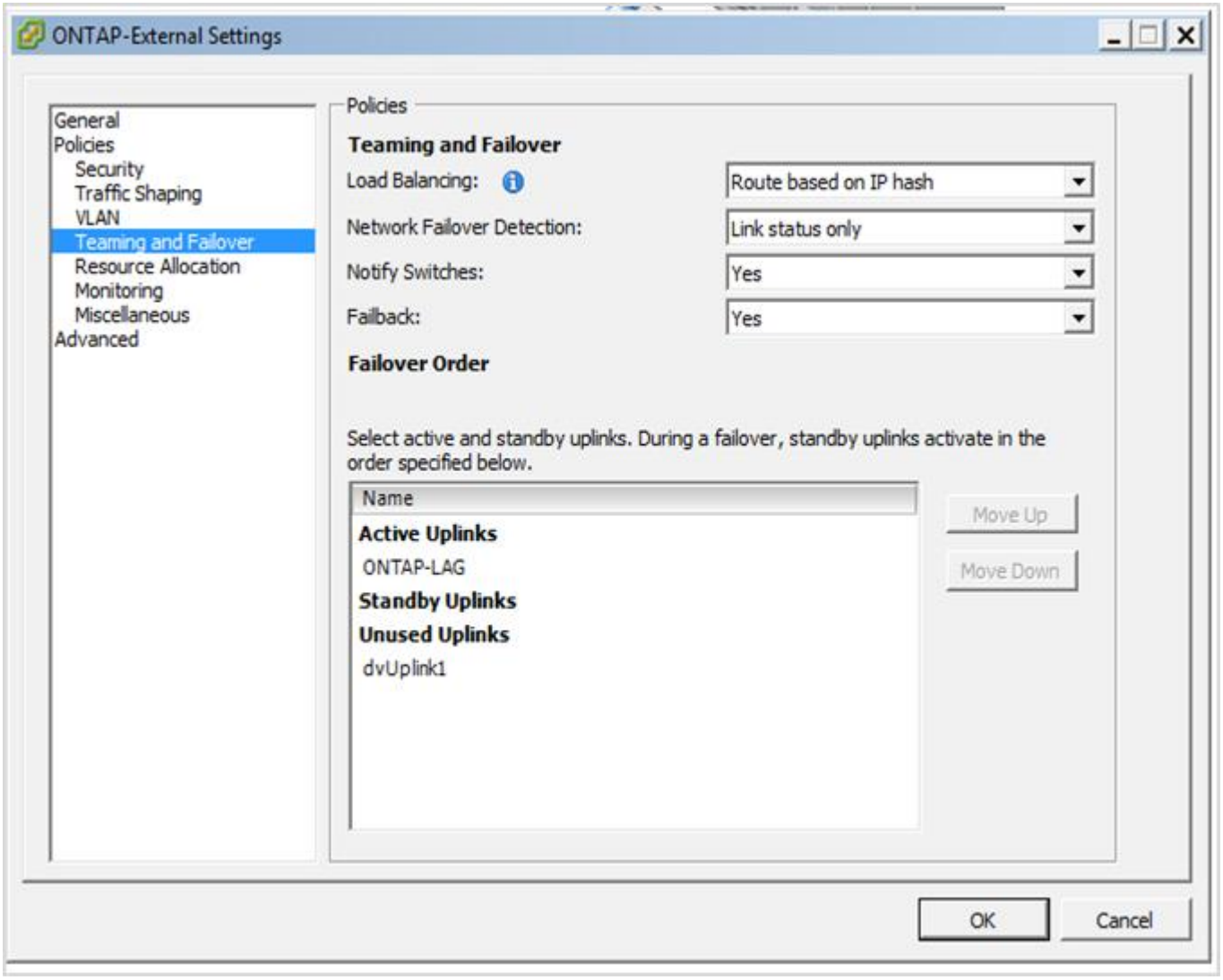
Lors de l'utilisation de vSwitches distribués dans votre configuration, LACP peut être utilisé (bien que cela ne soit pas une bonne pratique) afin de simplifier la configuration réseau. La seule configuration LACP prise en charge exige que toutes les VMNIC soient dans un seul LAG. Le commutateur physique de liaison montante doit prendre en charge une taille MTU comprise entre 7 500 et 9 000 sur tous les ports du canal. Les réseaux internes et externes ONTAP Select 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 illustrent la configuration distribuée de vSwitch utilisant LACP.

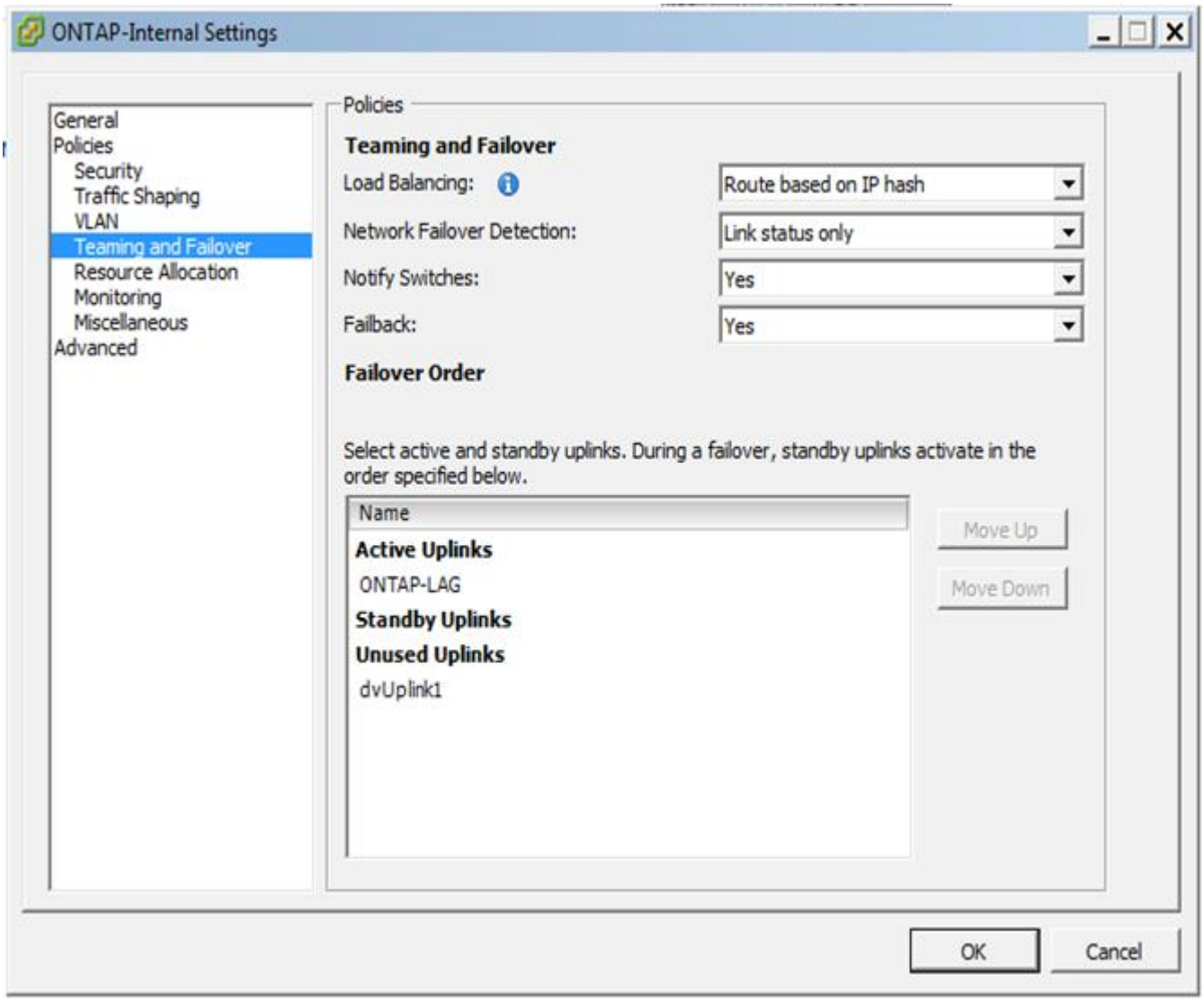
Propriétés LAG lors de l'utilisation de LACP



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



Pour utiliser LACP, vous devez configurer les ports du commutateur en amont en tant que canal de ports. Avant d'activer cette configuration sur le vSwitch distribué, assurez-vous qu'un canal de ports compatible LACP est correctement configuré.

Configuration du commutateur physique ONTAP Select

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

Il convient d'être particulièrement attentif 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 jusqu'à la couche réseau physique en amont grâce à l'isolation fournie par les VLAN de couche 2.

Les ports physiques du commutateur doivent être configurés en tant que trunkports. Le trafic externe ONTAP Select peut être séparé sur plusieurs réseaux de couche 2 de deux manières. Une méthode consiste à utiliser des ports virtuels ONTAP VLAN-tagged avec un seul groupe de ports. L'autre méthode consiste à affecter des groupes de ports distincts en mode VST au port de gestion e0a. Vous devez également affecter 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 séparé sur plusieurs réseaux de couche 2, les ports physiques de liaison montante du

commutateur doivent avoir ces VLAN dans leur liste de VLAN autorisés.

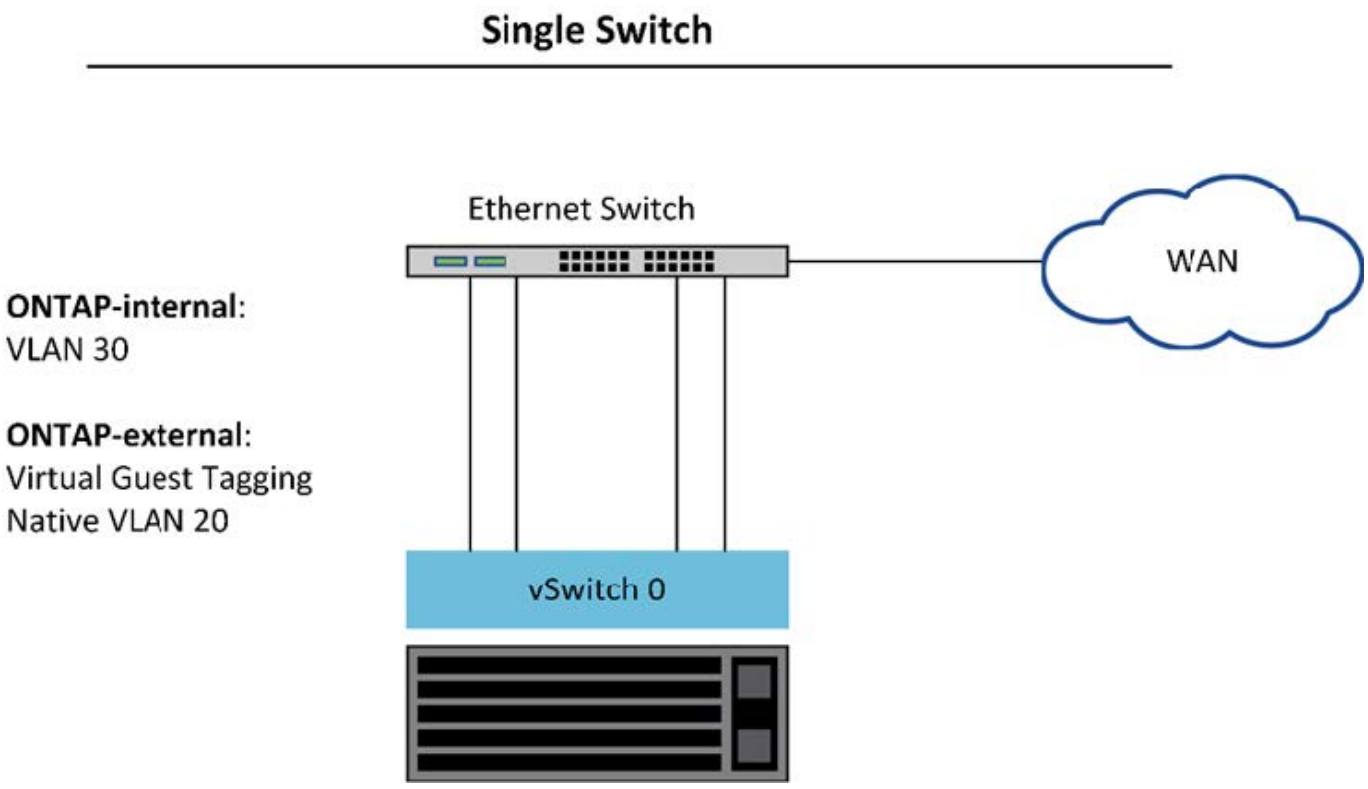
Le trafic réseau interne d'ONTAP Select s'effectue via des interfaces virtuelles définies avec des adresses IP locales. Ces adresses IP n'étant pas routables, le trafic interne entre les nœuds de cluster doit transiter par un seul réseau de couche 2. Les sauts de routage entre les nœuds de cluster ONTAP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

Commutateur physique partagé

La figure suivante illustre une configuration de commutateur possible utilisée par un nœud dans un cluster ONTAP Select multi-nœuds. Dans cet exemple, les cartes réseau physiques utilisées par les vSwitches hébergeant à la fois les groupes de ports réseau internes et externes sont connectées au même commutateur en amont. Le trafic de commutateur est maintenu isolé grâce à des domaines de diffusion contenus dans des VLAN distincts.

Pour le réseau interne ONTAP Select, le balisage 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 utilisant un commutateur physique partagé



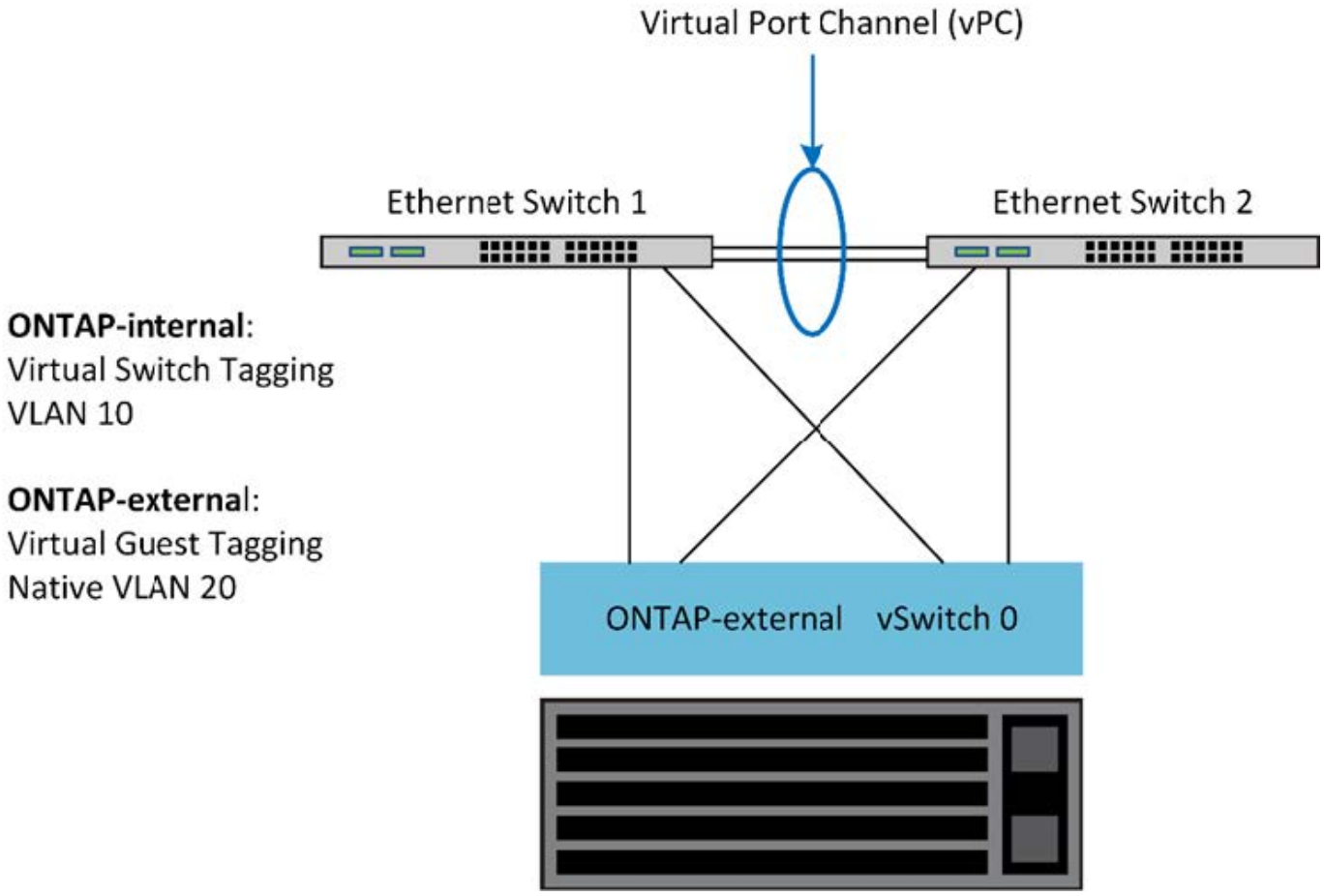
Dans cette configuration, le commutateur partagé devient un point de défaillance unique. Dans la mesure du possible, il convient d'utiliser plusieurs commutateurs afin d'éviter qu'une défaillance matérielle n'entraîne une interruption du réseau du cluster.

Plusieurs commutateurs physiques

Lorsque la redondance est nécessaire, il convient d'utiliser plusieurs commutateurs réseau physiques. La figure suivante illustre une configuration recommandée utilisée par un nœud dans un cluster ONTAP Select

multinœud. Les cartes réseau des groupes de ports internes et externes sont connectées à différents commutateurs physiques, protégeant l'utilisateur contre la défaillance d'un commutateur matériel. Un canal de ports virtuel est configuré entre les commutateurs pour éviter les problèmes de spanning tree.

Configuration réseau utilisant plusieurs commutateurs physiques



Séparation du trafic de données et de gestion ONTAP Select

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

Le trafic réseau externe d'ONTAP Select est défini comme 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 qui doit être hébergée sur un port réseau virtuel. Dans la configuration multinœud d'ONTAP Select, ces ports sont désignés comme e0a et e0b/e0g. Dans la configuration mononœud, ils sont désignés comme e0a et e0b/e0c, tandis que les ports restants sont réservés aux services internes du cluster.

NetApp recommande d'isoler le trafic de données et le trafic de gestion sur des réseaux de couche 2 distincts. Dans l'environnement ONTAP Select, cela se fait à l'aide de balises VLAN. Cela peut être réalisé en attribuant un groupe de ports balisé VLAN à l'adaptateur 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 à nœud unique) et e0b et e0g (clusters à nœuds multiples) pour le trafic de données.

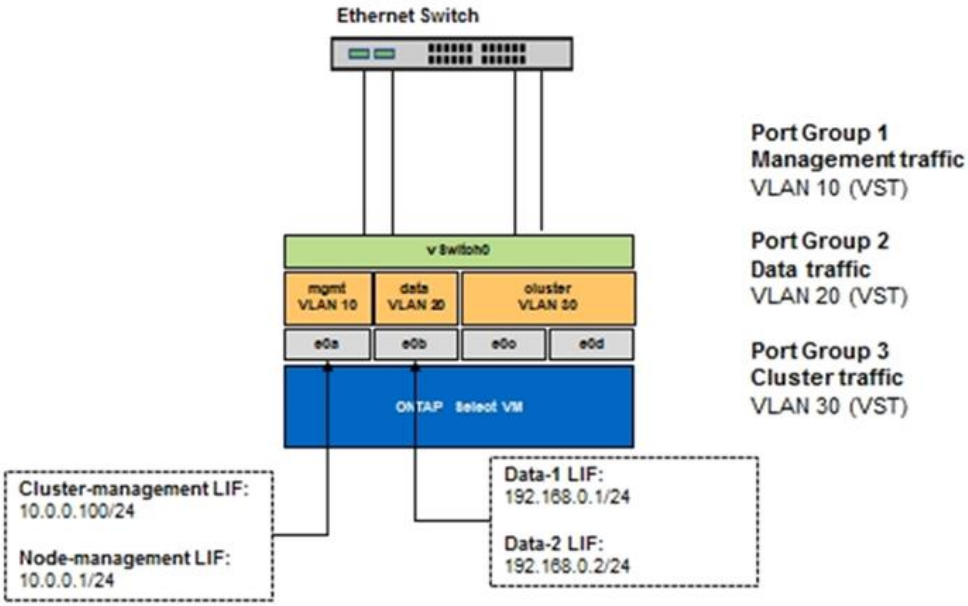
Si la solution VST décrite précédemment dans ce document ne suffit pas, il peut être nécessaire de regrouper les LIF de données et de gestion sur le même port virtuel. Pour ce faire, utilisez un processus appelé VGT, dans lequel le marquage VLAN est effectué par la VM.

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

Une limitation supplémentaire s'applique 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 avec le médiateur avant qu'ONTAP ne soit pleinement disponible. Par conséquent, seuls les marquages EST et VST sont pris en charge sur le groupe de ports associé 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 marquages 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 étiqueté au niveau de la couche vSwitch via le groupe de ports attribué. Dans cette configuration, les LIF de gestion du cluster et des nœuds sont affectées au port ONTAP e0a et étiquetées avec l'ID de VLAN 10 via le groupe de ports attribué. Les LIF de données sont affectées aux ports e0b et soit e0c soit e0g et se voient attribuer l'ID de VLAN 20 via un second groupe de ports. Les ports du cluster utilisent un troisième groupe de ports et appartiennent au VLAN 30.

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



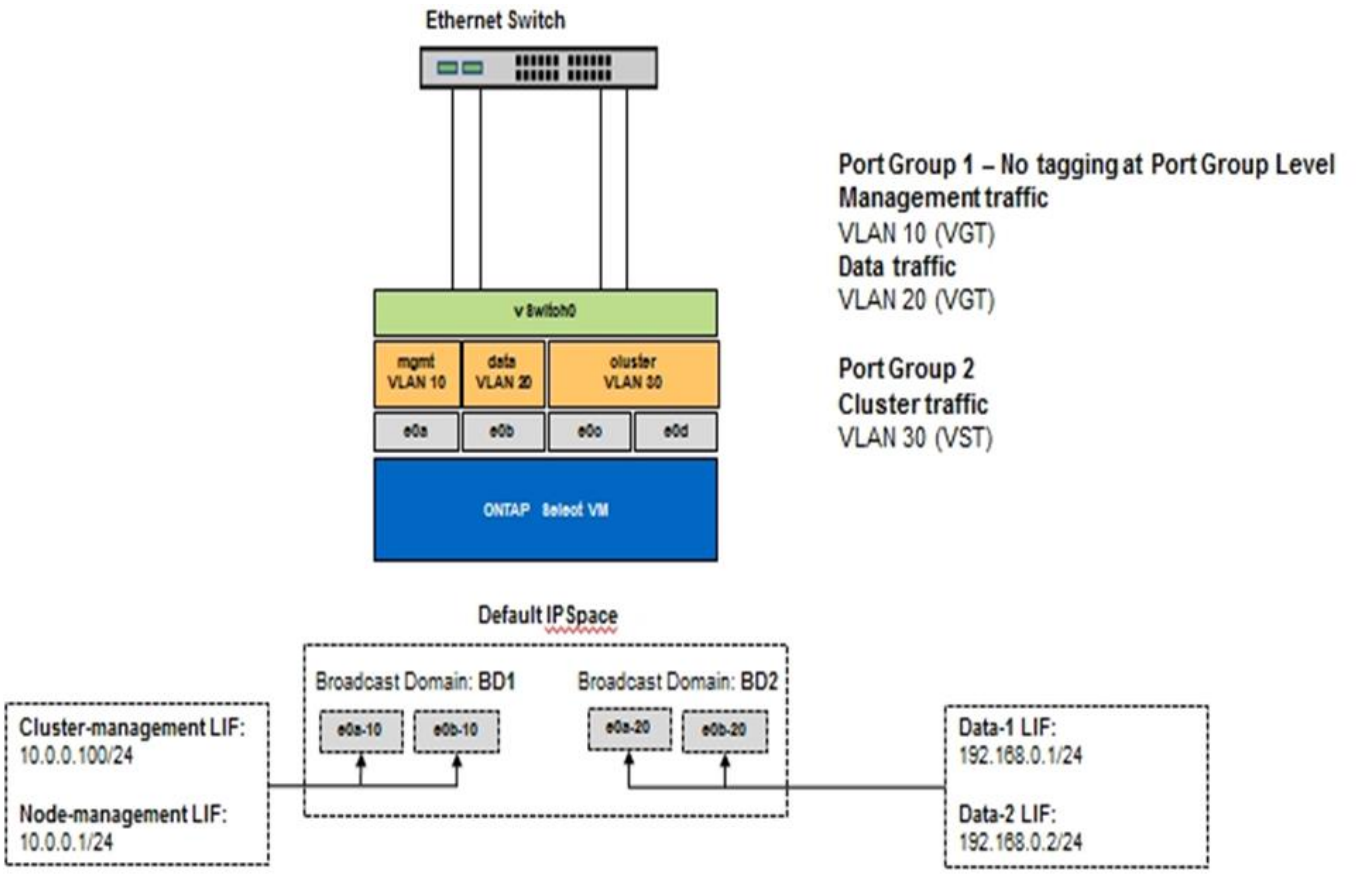
La figure suivante illustre le second scénario, VGT, dans lequel le trafic est étiqueté par la machine virtuelle ONTAP à l'aide de ports VLAN placés dans des domaines de diffusion distincts. Dans cet exemple, les ports virtuels e0a-10/e0b-10/(e0c ou e0g)-10 et e0a-20/e0b-20 sont placés au-dessus des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

Remarques :

- Ce type de configuration est particulièrement recommandé lors de l'utilisation de plusieurs IPspaces. Regroupez les ports VLAN dans des IPspaces personnalisés distincts si une isolation logique plus poussée et la multitenance sont souhaitées.

- Pour prendre en charge VGT, les cartes réseau de l'hôte ESXi/ESX doivent être connectées aux ports trunk du commutateur physique. Les groupes de ports connectés au commutateur virtuel doivent avoir leur ID de VLAN 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 à haute disponibilité

Configurations à 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 charges de travail applicatives des systèmes de stockage d'entreprise vers des solutions logicielles exécutées sur matériel ordinaire, les exigences et besoins en matière de résilience et de tolérance aux pannes restent inchangés. Une solution de paire haute disponibilité (HA) garantissant un objectif de point de récupération (RPO) zéro protège le client contre la perte de données due à une défaillance de n'importe quel composant de l'infrastructure.

Une grande partie du marché des SDS repose sur le principe du stockage sans partage, la réplication logicielle assurant la résilience des données grâce au stockage de plusieurs copies des données utilisateur sur différents silos de stockage. ONTAP Select exploite 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 s'effectue dans le cadre d'une paire haute disponibilité. Chaque paire haute disponibilité stocke deux copies des données utilisateur : l'une sur le stockage fourni par le nœud local, et l'autre sur le stockage fourni par le partenaire HA. Au sein d'un cluster ONTAP Select, la haute disponibilité et

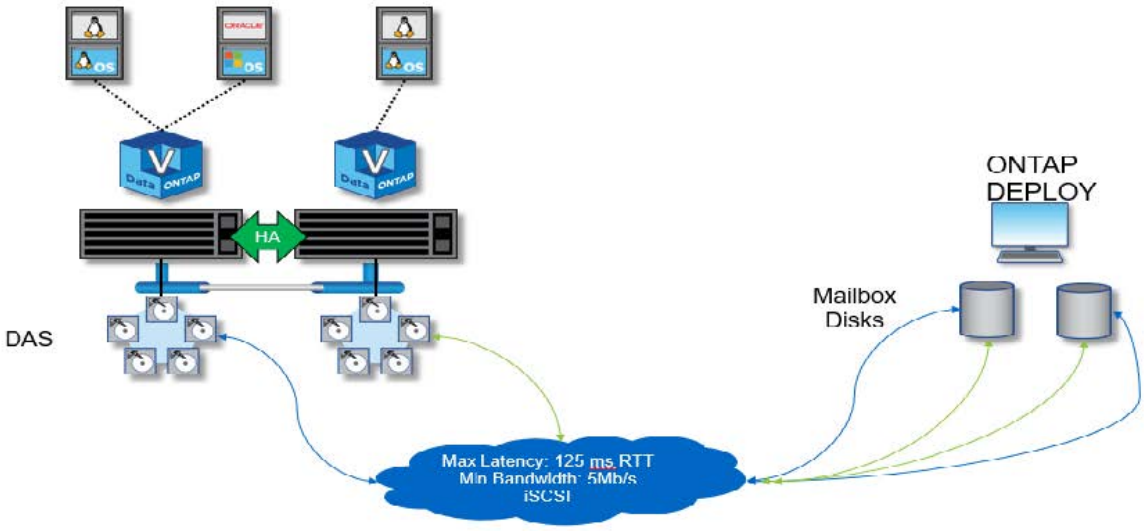
la réplication synchrone sont liées, et la fonctionnalité des deux ne peut pas être dissociée ni utilisée indépendamment. En conséquence, la fonctionnalité de réplication synchrone n'est disponible que dans l'offre multi-nœuds.

Dans un cluster ONTAP Select, la fonctionnalité de réplication synchrone est une fonction de l'implémentation de la haute disponibilité (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, huit, dix ou douze nœuds) et les clusters à deux nœuds. La principale caractéristique d'un cluster ONTAP Select à deux nœuds est l'utilisation d'un service médiateur externe pour résoudre les scénarios de split-brain. La machine virtuelle ONTAP Deploy fait office de médiateur par défaut pour toutes les paires haute disponibilité à deux nœuds qu'elle configure.

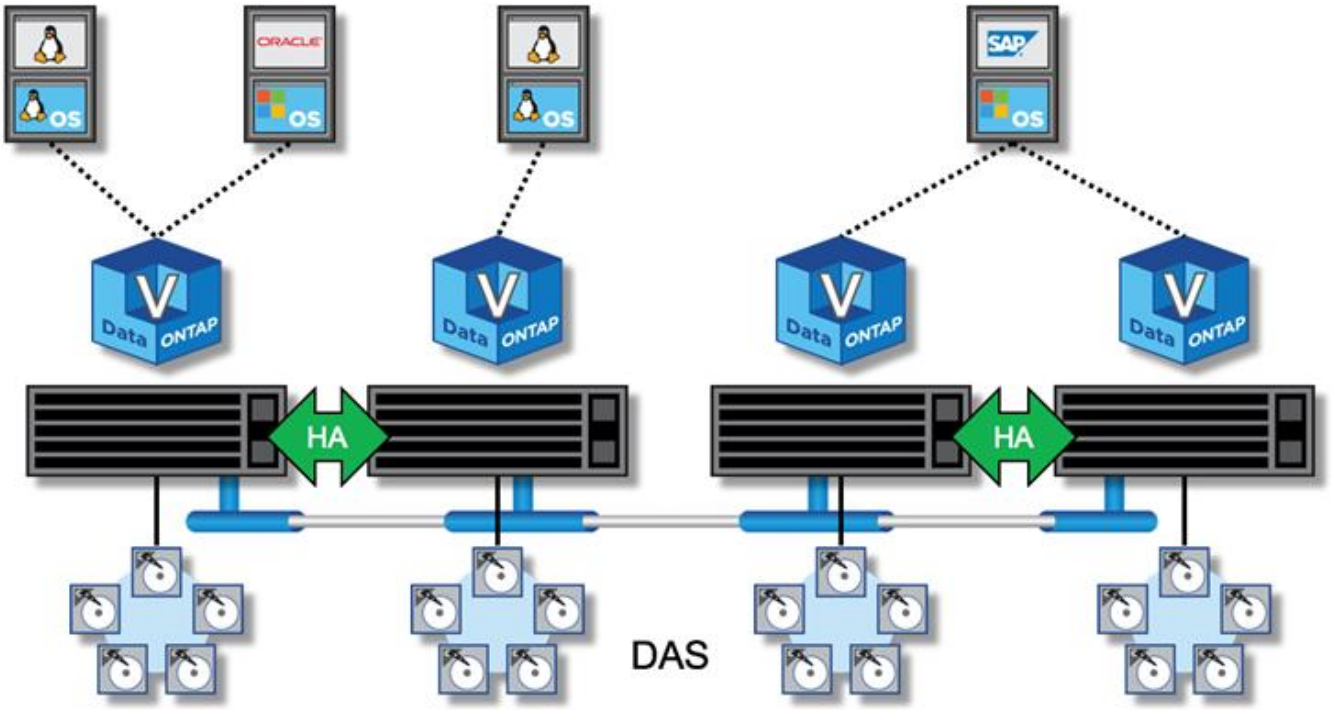
Les deux architectures sont représentées dans les figures suivantes.

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



Le cluster ONTAP Select à deux nœuds est composé d'une paire haute disponibilité et d'un médiateur. Au sein de la paire haute disponibilité, les agrégats de données de chaque nœud de cluster sont synchronement mis en miroir et, en cas de basculement, il n'y a aucune perte de données.

Cluster ONTAP Select à quatre nœuds utilisant un stockage local



- Le cluster ONTAP Select à quatre nœuds est composé de deux paires haute disponibilité. Les clusters à six, huit, dix et douze nœuds sont composés respectivement de trois, quatre, cinq et six paires haute disponibilité. Au sein de chaque paire haute disponibilité, les agrégats de données de chaque nœud de cluster sont synchronement mis en miroir et, en cas de basculement, il n'y a aucune perte de données.
- Une seule instance d'ONTAP Select peut être présente sur un serveur physique lors de l'utilisation du stockage DAS. ONTAP Select requiert un accès non partagé au contrôleur RAID local du système et est conçu pour gérer les disques connectés localement, ce qui serait impossible sans connexion physique au stockage.

HA à deux nœuds versus HA à plusieurs nœuds

Contrairement aux baies FAS, les nœuds ONTAP Select d'une paire haute disponibilité 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 la défaillance d'un nœud unique parce que le quorum du cluster peut être établi par les trois nœuds survivants ou plus. Le cluster à deux nœuds s'appuie sur le service de médiation hébergé par la VM 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 ONTAP Select à deux nœuds.

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

Si la machine virtuelle ONTAP Deploy agissant comme médiateur est temporairement ou potentiellement définitivement indisponible, une machine virtuelle ONTAP Deploy secondaire peut être utilisée pour rétablir le quorum du cluster à deux nœuds. Cela aboutit à une configuration dans laquelle la nouvelle machine virtuelle ONTAP Deploy ne peut pas gérer les nœuds ONTAP Select, mais elle 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 en utilisant le protocole iSCSI sur IPv4. L'adresse IP de gestion du nœud ONTAP Select est l'initiateur, et l'adresse IP 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 boîte aux lettres hébergés par ONTAP Deploy sont automatiquement créés et masqués vers les adresses IP de gestion appropriées des nœuds ONTAP Select lors de la création du cluster à deux nœuds. L'ensemble de la configuration est automatiquement effectué lors de l'installation, et aucune autre action administrative n'est requise. L'instance ONTAP Deploy qui crée le cluster est le médiateur par défaut pour ce cluster.

Une intervention administrative est nécessaire si l'emplacement du médiateur d'origine doit être modifié. Il est possible de rétablir le quorum du cluster même si la machine virtuelle ONTAP Deploy d'origine est perdue. Cependant, NetApp recommande de sauvegarder la base de données ONTAP Deploy après chaque instanciation d'un cluster à deux nœuds.

HA à deux nœuds versus HA étiré 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 potentiellement 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 inférieure à 300 m. En général, les deux nœuds disposent de liaisons montantes vers le même commutateur réseau ou le même ensemble de commutateurs réseau de liaison interswitch (ISL).

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

MetroCluster SDS est une fonctionnalité premium nécessitant une licence Premium ou une licence Premium XL. La licence Premium prend en charge la création de machines virtuelles petites et moyennes, ainsi que de supports HDD 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. MetroCluster SDS configurations doivent garantir une latence maximale de 10 ms entre les nœuds, latence du stockage partagé incluse. Autrement dit, la simple mesure de la latence entre les machines virtuelles Select est insuffisante, car la latence du stockage partagé n'est pas négligeable dans ces configurations.

ONTAP Select HA RSM et agrégats en miroir

Prévenez 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 au monde des serveurs ordinaires non partagés en utilisant la fonctionnalité RAID SyncMirror (RSM) présente dans ONTAP pour répliquer les blocs de données entre les nœuds de cluster, fournissant deux copies des données utilisateur réparties sur une paire haute disponibilité.

Un cluster à deux nœuds avec un médiateur peut s'étendre sur deux centres de données. Pour plus d'informations, consultez la section "Meilleures pratiques pour la paire haute disponibilité étendue à deux nœuds (MetroCluster SDS)".

Agrégats en miroir

Un cluster ONTAP Select est composé de deux à douze nœuds. Chaque paire haute disponibilité contient deux copies des données utilisateur, synchronisées de manière synchrone entre les nœuds via un réseau IP. Ce miroir est transparent pour l'utilisateur et constitue une propriété de l'agrégat de données, qui est automatiquement configurée lors du processus de création de l'agrégat de données.

Tous les agrégats dans un cluster ONTAP Select doivent être mis en miroir pour la disponibilité des données en cas de basculement d'un nœud et pour éviter un SPOF en cas de défaillance matérielle. Les agrégats dans un cluster ONTAP Select sont construits à partir de disques virtuels fournis par chaque nœud de la paire haute disponibilité 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 de la paire haute disponibilité du nœud actuel)

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

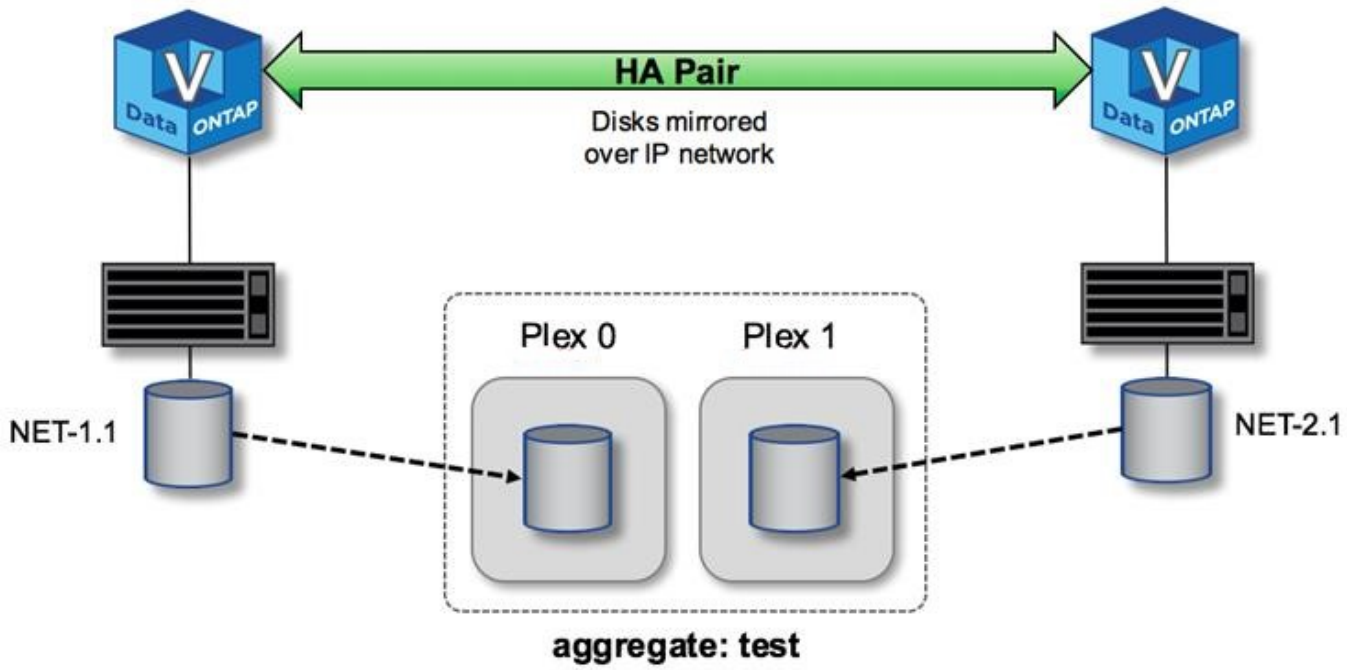
Cette approche diffère fondamentalement de la façon dont les clusters ONTAP standard fonctionnent. Cela s'applique à tous les disques racine et de données au sein du cluster ONTAP Select. L'agrégat contient à la fois des copies locales et des copies miroir des données. Par conséquent, un agrégat contenant N disques virtuels offre une capacité de stockage unique équivalente à N/2 disques, car la seconde copie des données réside sur ses propres disques uniques.

La figure suivante illustre une paire haute disponibilité au sein d'un cluster ONTAP Select à quatre nœuds. Ce cluster comprend un agrégat unique (test) utilisant 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 ONTAP Select (Plex 0), et un ensemble distant, fourni par le partenaire de basculement (Plex 1).

Plex 0 est le compartiment contenant tous les disques locaux. Plex 1 est le compartiment contenant les disques miroirs, c'est-à-dire les disques responsables du stockage d'une seconde copie répliquée des données utilisateur. Le nœud propriétaire de l'agrégat contribue des disques à Plex 0, et le partenaire de paire haute disponibilité de ce nœud contribue des disques à Plex 1.

La figure suivante illustre un agrégat en miroir composé de deux disques. Le contenu de cet agrégat est répliqué sur nos deux nœuds de cluster, le disque local NET-1.1 étant placé dans le compartiment Plex 0 et le disque distant NET-2.1 dans le compartiment Plex 1. Dans cet exemple, l'agrégat test appartient au nœud de cluster situé à gauche et utilise le disque local NET-1.1 ainsi que le disque miroir du partenaire de la paire haute disponibilité, NET-2.1.

*Agrégat en miroir ONTAP
Select*



Lors du déploiement d'un cluster ONTAP Select, tous les disques virtuels présents sur le système sont automatiquement affectés au plex approprié, sans intervention supplémentaire de l'utilisateur concernant l'affectation des disques. Ceci évite toute affectation accidentelle de disques à un plex incorrect et garantit une configuration optimale des disques en miroir.

Chemin d'écriture

La mise en miroir synchrone des blocs de données entre les nœuds du cluster et l'exigence d'absence de perte de données en cas de défaillance système ont un impact significatif sur le chemin emprunté par une écriture entrante lors de sa propagation à travers un cluster ONTAP Select. Ce processus comporte deux étapes :

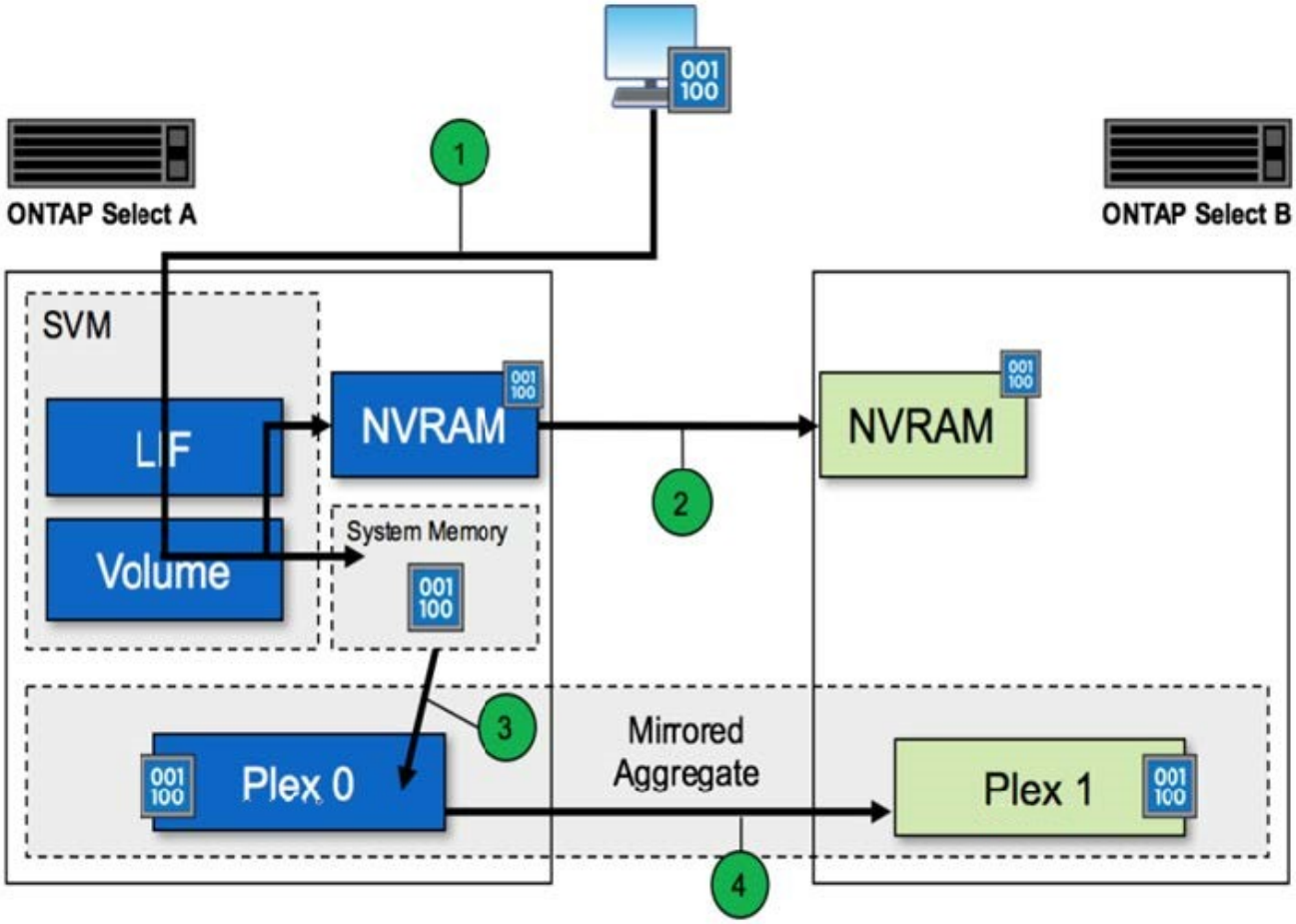
- Reconnaissance
- Vidage

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. Sur une configuration paire haute disponibilité, une étape supplémentaire a lieu, car ces écritures NVRAM sont immédiatement répliquées sur le partenaire HA du propriétaire du volume cible avant d'être confirmées. Ce processus garantit la cohérence du système de fichiers sur le nœud partenaire HA en cas de défaillance matérielle du nœud d'origine.

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

La figure suivante illustre le chemin d'écriture d'une requête d'écriture entrante vers un nœud ONTAP Select.

Workflow du chemin d'écriture ONTAP Select



L'accusé de réception d'une écriture entrante comprend les étapes suivantes :

- Les écritures entrent dans le système via une interface logique appartenant au nœud ONTAP Select A.
- Les écritures sont enregistrées dans la NVRAM du nœud A et répliquées sur le partenaire HA, le nœud B.
- Une fois la requête d'E/S présente sur les deux nœuds HA, la requête est alors confirmée au client.

Le transfert des données de NVRAM vers l'agrégat de données (ONTAP CP) dans ONTAP Select comprend les étapes suivantes :

- Les écritures sont transférées de la NVRAM virtuelle vers l'agrégat de données virtuel.
- Le moteur Mirror réplique de manière synchrone les blocs sur les deux plexes.

Améliorations de la protection des données avec ONTAP Select HA

Les fonctions de pulsation de disque à haute disponibilité (HA), de boîte aux lettres HA, de pulsation HA, de basculement HA et de restauration contribuent à améliorer la protection des données.

battement cardiaque du disque

Bien que l'architecture ONTAP Select HA exploite de nombreux chemins de code utilisés par les baies FAS traditionnelles, certaines exceptions existent. L'une de ces exceptions concerne l'implémentation du mécanisme de pulsation basé sur disque, une méthode de communication non basée sur le réseau utilisée par

les nœuds de cluster pour éviter que l'isolation réseau ne provoque un comportement de split-brain. Un scénario de split-brain résulte d'un partitionnement du cluster, généralement dû à des pannes réseau, où chaque côté croit que l'autre est hors service et tente de prendre le contrôle des ressources du cluster.

Les implémentations de haute disponibilité (HA) de niveau entreprise doivent gérer ce type de scénario avec élégance. ONTAP le fait grâce à une méthode personnalisée de pulsation basée sur le disque. C'est le rôle de la boîte aux lettres HA, un emplacement sur le stockage physique utilisé par les nœuds de cluster pour transmettre les messages de pulsation. Cela aide le cluster à déterminer la connectivité et donc à 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 de la manière suivante :

- réservations persistantes SCSI
- Métadonnées HA persistantes
- L'état HA est transmis via interconnexion haute disponibilité

Cependant, dans l'architecture sans partage d'un cluster ONTAP Select, un nœud ne peut accéder qu'à son propre stockage local et non à celui de son partenaire de haute disponibilité. Par conséquent, lorsque le partitionnement réseau isole chaque côté d'une paire haute disponibilité, les méthodes précédentes de détermination du quorum du cluster et du comportement en cas de basculement deviennent inapplicables.

Bien que la méthode actuelle de détection et d'évitement du split-brain soit inutilisable, une méthode de médiation reste nécessaire, adaptée aux contraintes d'un environnement sans partage de ressources. ONTAP Select étend l'infrastructure de messagerie existante, lui permettant de servir de méthode de médiation en cas de partitionnement du réseau. L'absence de stockage partagé implique que la médiation s'effectue via l'accès aux disques de messagerie sur NAS. Ces disques sont répartis sur l'ensemble du cluster, y compris sur le médiateur dans un cluster à deux nœuds, grâce au protocole iSCSI. Ainsi, des décisions intelligentes de basculement peuvent être prises par un nœud de cluster en fonction de l'accès à ces disques. Si un nœud peut accéder aux disques de messagerie d'autres nœuds en dehors de son partenaire de haute disponibilité, il est probablement opérationnel et sain.



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

Publication de la boîte aux lettres HA

L'architecture de boîtes aux lettres à haute disponibilité utilise un modèle de publication de messages. À intervalles réguliers, les nœuds de cluster publient des messages sur tous les autres disques de boîtes aux lettres du cluster, y compris le médiateur, indiquant que le nœud est en ligne et fonctionne. Dans un cluster sain, à tout moment, un disque de boîte aux lettres sur un nœud de cluster contient des messages publiés par tous les autres nœuds de cluster.

Un disque virtuel est attaché à chaque nœud de cluster Select et est utilisé spécifiquement pour l'accès partagé à la boîte aux lettres. 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 défaillance de nœud ou de partitionnement du réseau. Ce disque de boîte aux lettres contient des partitions pour chaque nœud de cluster et est monté sur un réseau iSCSI par les autres nœuds de cluster Select. Périodiquement, ces nœuds publient des états de santé sur la partition appropriée du disque de boîte aux lettres. L'utilisation de disques de boîte aux lettres accessibles via le réseau et répartis dans le cluster permet d'inférer l'état de santé des nœuds à l'aide d'une matrice de connectivité. Par exemple, les nœuds de cluster A et B peuvent publier sur la boîte aux

lettres du nœud de cluster D, mais pas sur celle du nœud C. De plus, le nœud de cluster D ne peut pas publier sur la boîte aux lettres du nœud C, il est donc probable que le nœud C soit soit hors service, soit isolé du réseau et doive être pris en charge.

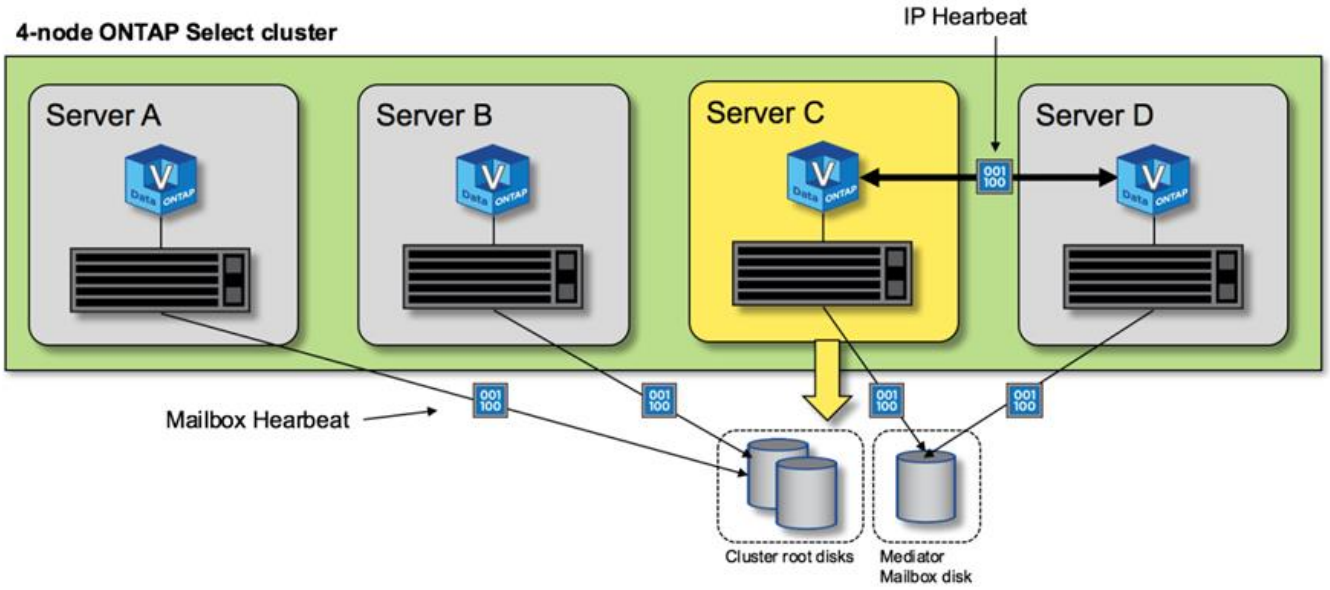
Battement de cœur HA

Comme avec les plateformes NetApp FAS, ONTAP Select envoie périodiquement des messages de pulsation HA via l'interconnexion haute disponibilité. Au sein du cluster ONTAP Select, cela s'effectue via une connexion réseau TCP/IP établie entre les partenaires HA. De plus, des messages de pulsation basés sur disque sont transmis à tous les disques de boîte aux lettres HA, y compris les disques de boîte aux lettres médiateur. Ces messages sont transmis toutes les quelques secondes et relus périodiquement. La fréquence à laquelle ceux-ci sont envoyés et reçus permet au cluster ONTAP Select de détecter les événements de défaillance HA en environ 15 secondes, soit la même fenêtre disponible sur les plateformes FAS. Lorsque les messages de pulsation ne sont plus lus, un événement de basculement est déclenché.

La figure suivante illustre le processus d'envoi et de réception de messages de pulsation sur l'interconnexion haute disponibilité et les disques médiateurs du point de vue d'un seul nœud de cluster ONTAP Select, le nœud C.

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

Battements cardiaques HA dans un cluster à quatre nœuds : état stable



Basculement et restauration HA

Lors d'un basculement, le nœud survivant prend en charge le traitement des données de son nœud partenaire en utilisant la copie locale des données de son partenaire de l'interconnexion haute disponibilité. Les E/S client restent ininterrompues, mais les modifications apportées à ces données doivent être répliquées avant que la reprise ne puisse avoir lieu. Notez qu'ONTAP Select ne prend pas en charge une reprise forcée, car cela entraînerait la perte des modifications stockées sur le nœud survivant.

L'opération de synchronisation de retour est automatiquement déclenchée lorsque le nœud redémarré rejoint le cluster. Le temps nécessaire pour la synchronisation de retour dépend de plusieurs facteurs. Ces facteurs incluent le nombre de modifications devant être répliquées, la latence réseau entre les nœuds et la vitesse des

sous-systèmes de disque sur chaque nœud. Il est possible que le temps nécessaire pour la synchronisation de retour dépasse la fenêtre de restauration automatique de 10 minutes. Dans ce cas, une restauration manuelle après la synchronisation de retour est requise. La progression de la synchronisation de retour peut être surveillée à l'aide de la commande suivante :

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

Découvrez les performances d'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 influençant les performances d'une instance ONTAP Select donnée. Voici quelques-uns des facteurs qui affectent les performances d'une instance ONTAP Select :

- **Fréquence fondamentale**. En général, une fréquence plus élevée est préférable.
- **Configuration mono-socket versus multi-socket**. ONTAP Select n'utilise pas les fonctionnalités multi-socket, mais la surcharge de l'hyperviseur liée à la prise en charge des configurations multi-socket explique une certaine différence dans les performances totales.
- **Configuration de la carte RAID et du pilote d'hyperviseur associé**. Le pilote par défaut fourni par l'hyperviseur peut nécessiter un remplacement par le pilote du fabricant du matériel.
- **Type de disque et nombre de disques dans le(s) groupe(s) RAID**.
- **Version de l'hyperviseur et niveau de correctif**.

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 DÉLICTUELLE (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.