Plongée en profondeur

ONTAP Select

NetApp
April 01, 2025

This PDF was generated from https://docs.netapp.com/fr-fr/ontap-select-9121/concept_stor_concepts_chars.html on April 01, 2025. Always check docs.netapp.com for the latest.

Sommaire

Plongée en profondeur . . . 1
- Stockage . . . 1
  - Stockage : concepts et caractéristiques générales . . . 1
  - Services RAID matériels pour le stockage ONTAP Select local . . . 6
  - Services RAID logiciels pour le stockage local . . . 13
  - Configurations de baies externes et VSAN . . . 22
  - Augmenter la capacité de stockage . . . 26
  - Prise en charge de l’efficacité du stockage . . . 30
- Mise en réseau . . . 33
  - Mise en réseau : concepts et caractéristiques généraux . . . 33
  - Configurations réseau à un ou plusieurs nœuds . . . 35
  - Réseau interne et externe ONTAP Select . . . 40
  - Configurations réseau prises en charge . . . 43
  - Configuration du vSwitch VMware vSphere . . . 44
  - Configuration physique des commutateurs . . . 53
  - Séparation du trafic de données et de gestion . . . 55
- Architecture haute disponibilité . . . 57
  - Configurations haute disponibilité . . . 57
  - RSM HAUTE DISPONIBILITÉ et agrégats en miroir . . . 60
  - HA de détails supplémentaires . . . 63
- Performance . . . 66
  - Performance . . . 66
  - Performances : stockage SSD DAS haute disponibilité premium . . . 66

Plongée en profondeur

Stockage

Stockage : concepts et caractéristiques générales

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

Phases de la configuration du stockage

Le stockage hôte ONTAP Select comprend les principales phases de configuration suivantes :

- Conditions préalables au déploiement
  - Assurez-vous que chaque hôte d'hyperviseur est configuré et prêt pour un déploiement ONTAP Select.
  - La configuration implique les disques physiques, les contrôleurs et groupes RAID, les LUN, ainsi que la préparation réseau associée.
  - Cette configuration est effectuée en dehors de 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 de ONTAP Select.
- Configuration à l'aide de l'utilitaire d'administration de ONTAP Select Deploy
  - Vous pouvez utiliser l'utilitaire d'administration Deploy pour configurer les constructions de stockage logique de base.
  - Cette opération est réalisée de manière explicite via des commandes CLI ou automatiquement par l'utilitaire dans le cadre d'un déploiement.
- Configuration après déploiement
  - Une fois ONTAP Select le déploiement terminé, vous pouvez configurer le cluster à l'aide de l'interface de ligne de commande de ONTAP ou de System Manager.
  - Cette configuration est effectuée en dehors du déploiement ONTAP Select.

Gestion par rapport au stockage non géré

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

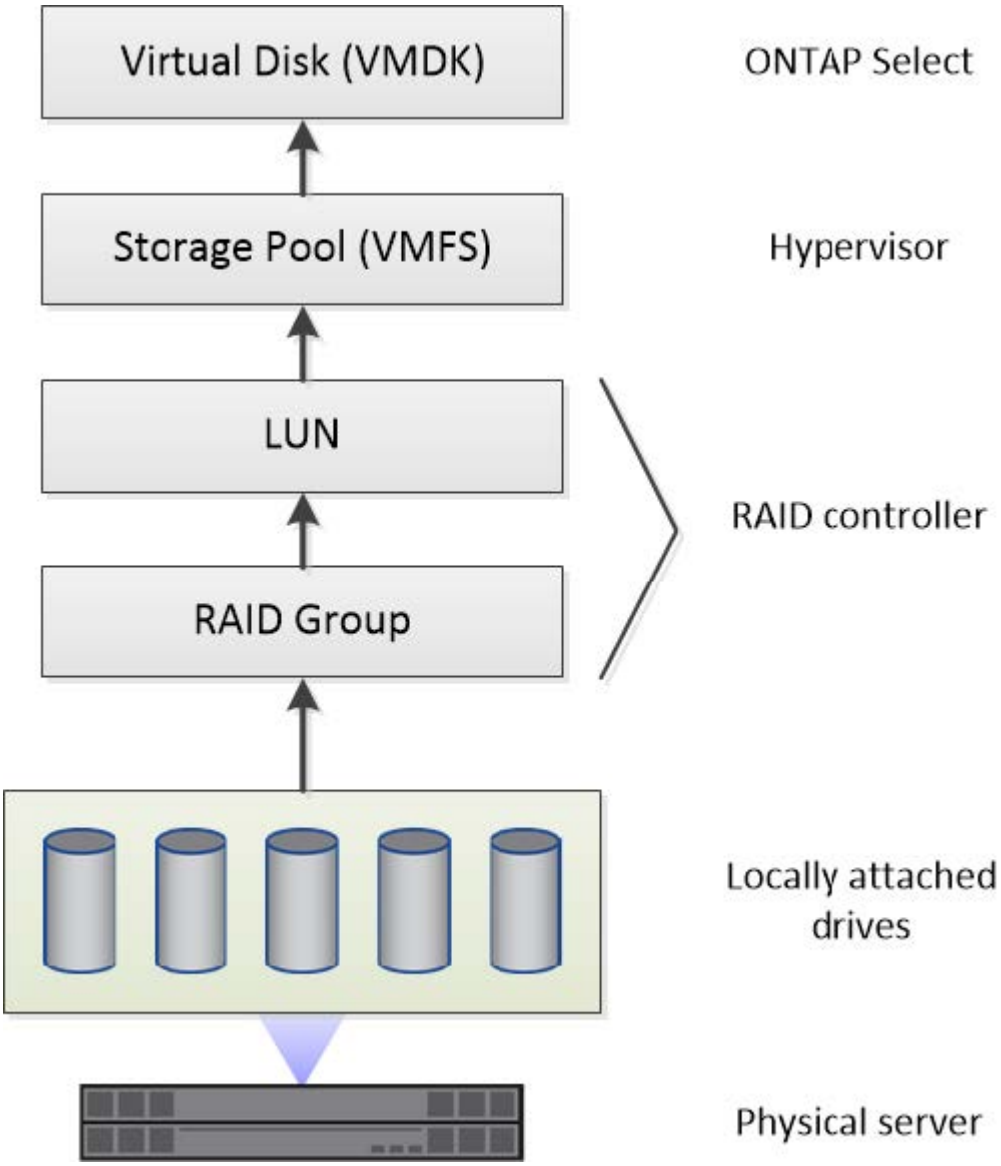
Stockage physique homogène

Tous les disques physiques composant le stockage géré par ONTAP Select doivent être homogènes. En d'autres termes, tout le matériel doit être identique concernant 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 qui peuvent être utilisés par ONTAP Select. Ces composants de stockage sont disposés 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. Avant de préparer un déploiement ONTAP Select, vous devez connaître ces concepts. Ces concepts sont organisés selon la catégorie : groupes RAID et LUN, pools de stockage et disques virtuels.

Regroupement de disques physiques en groupes RAID et LUN

Un ou plusieurs disques physiques peuvent être connectés localement au serveur hôte et disponibles à ONTAP Select. Les disques physiques sont affectés aux groupes RAID, qui sont ensuite présentés au système d'exploitation hôte de l'hyperviseur sous forme d'une ou de plusieurs LUN. Chaque LUN est présentée au système d'exploitation hôte de l'hyperviseur en tant que disque dur physique.

Lors de la configuration d'un hôte ONTAP Select, il faut prendre en compte les 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. Vous devez créer un seul groupe RAID pour ONTAP Select. Cependant, dans certaines situations, vous pouvez envisager de créer plusieurs groupes RAID. Reportez-vous à la section "Résumé des meilleures pratiques".

Considérations relatives au pool de stockage

Dans le cadre de votre préparation au déploiement de ONTAP Select, vous devez connaître plusieurs problèmes liés aux pools de stockage.

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

Pools de stockage et LUN

Chaque LUN est considérée comme un disque local sur l'hôte de l'hyperviseur et peut faire partie d'un pool de stockage unique. Chaque pool de stockage est formaté avec un système de fichiers que l'hyperviseur hôte 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 transféré à l'utilitaire d'administration ONTAP Select Deploy.

Gérer les disques virtuels

Dans le cadre de la préparation au déploiement de ONTAP Select, vous devez connaître plusieurs problèmes liés aux disques virtuels.

Disques virtuels et systèmes de fichiers

La machine virtuelle ONTAP Select est allouée à plusieurs disques virtuels. Chaque disque virtuel est en fait 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 connaître les éléments suivants concernant les disques virtuels :

- Le pool de stockage doit être disponible avant la création des disques virtuels.
- Les disques virtuels ne peuvent pas être créés avant la création de la machine virtuelle.
- Vous devez vous appuyer sur l'utilitaire d'administration ONTAP Select Deploy pour créer tous les disques virtuels (l'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

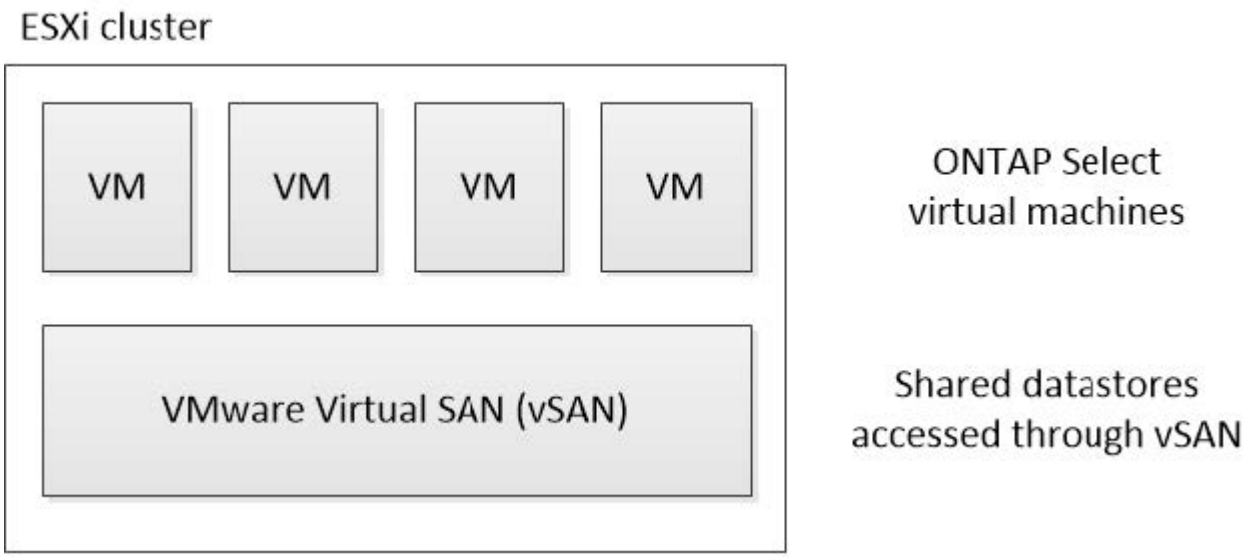
La solution ONTAP Select vNAS permet à ONTAP Select d'utiliser des datastores résidant sur une infrastructure de stockage externe à l'hôte de l'hyperviseur. Les datastores sont accessibles via le réseau à l'aide de VMware VSAN ou directement sur une baie de stockage externe.

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

- VSAN (SAN virtuel)
- VMFS
- NFS

Datastores VSAN

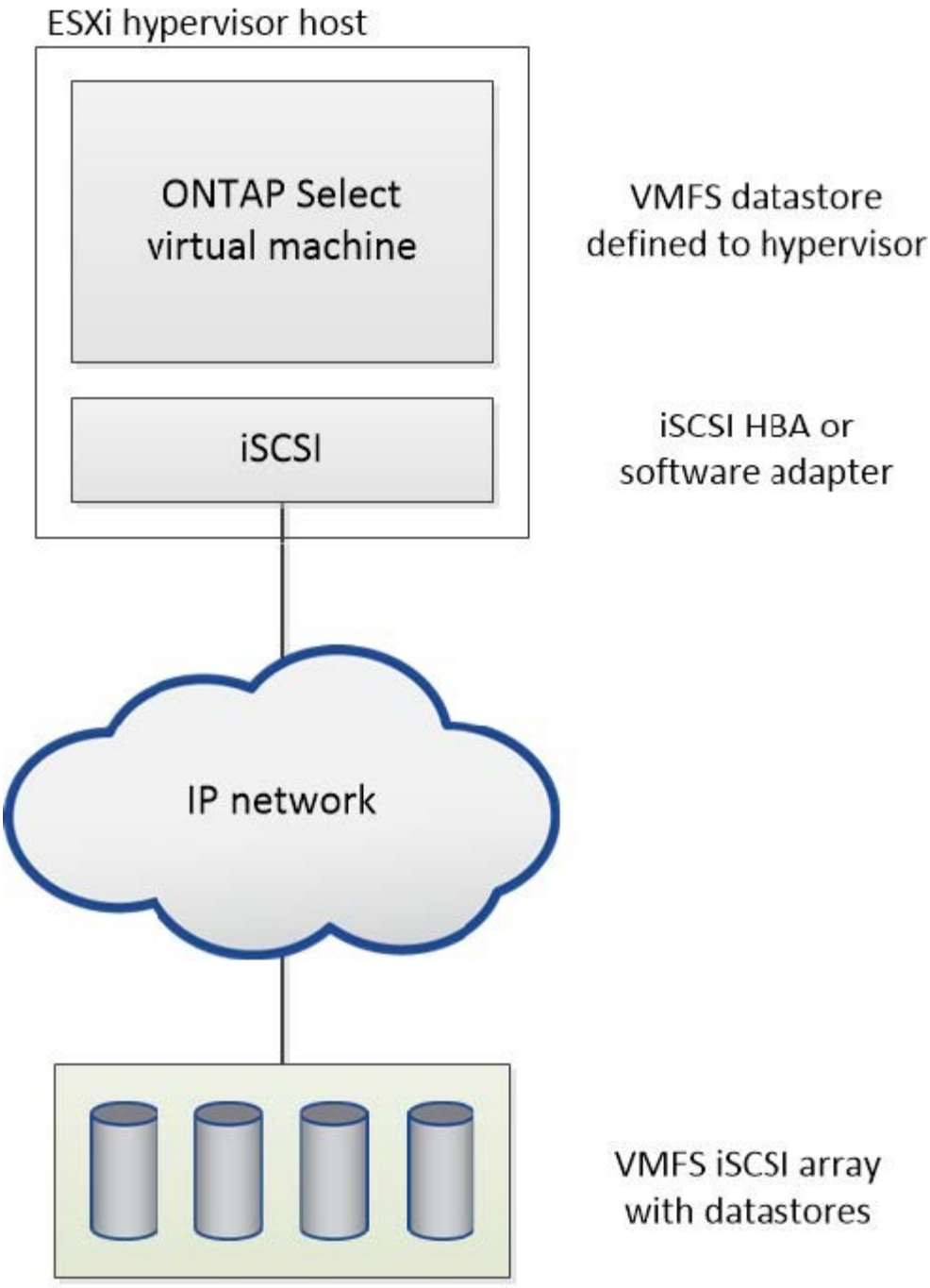
Chaque hôte ESXi peut avoir un ou plusieurs datastores VMFS locaux. En principe, ces datastores ne sont accessibles qu'à l'hôte local. Cependant, VMware VSAN autorise chacun des hôtes d'un cluster ESXi à partager tous les datastores du cluster comme s'ils étaient locaux. La figure suivante montre comment VSAN crée un pool de datastores partagés entre les hôtes du cluster ESXi.



Datastore VMFS sur une baie de stockage externe

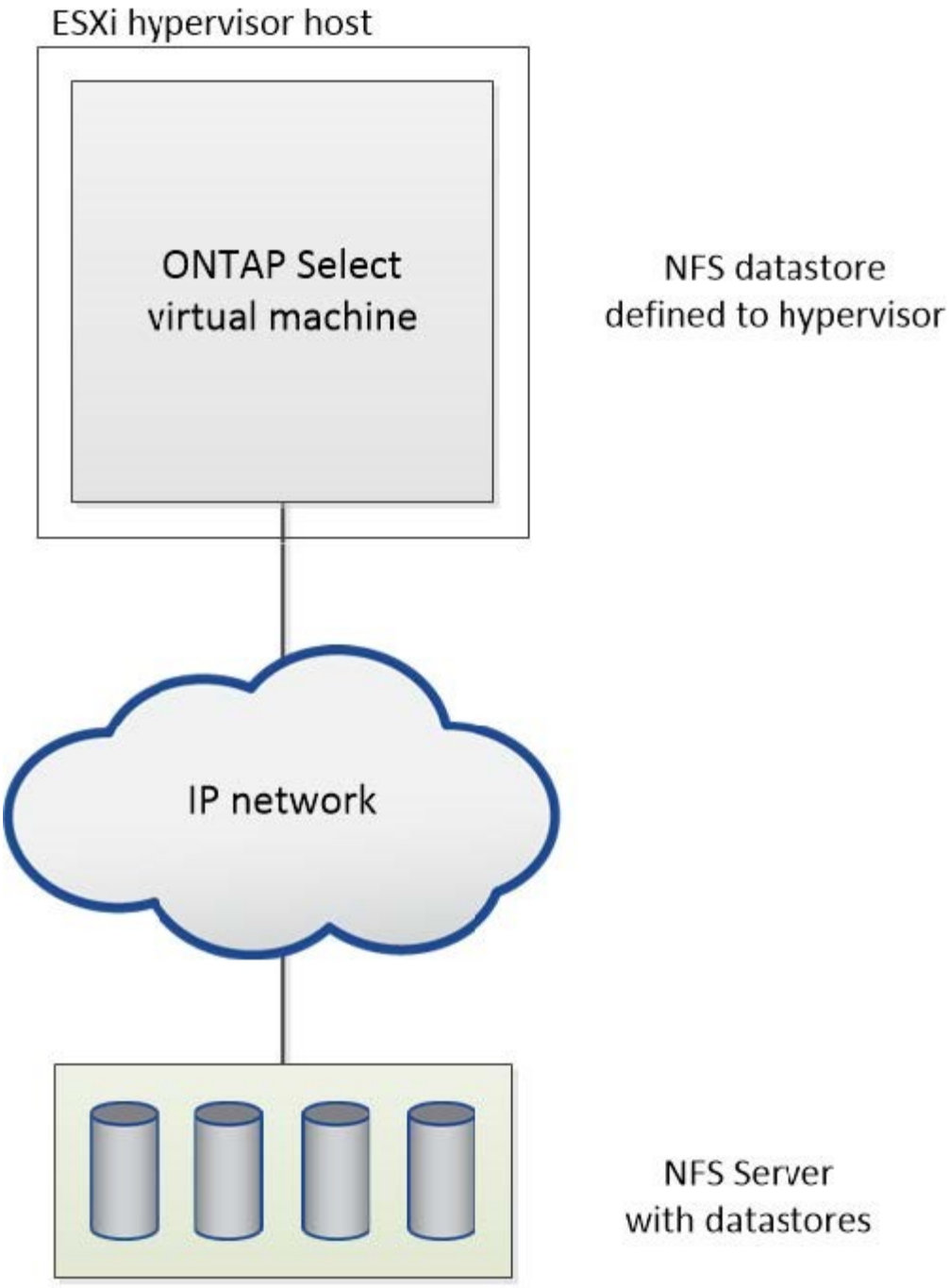
Vous pouvez créer un datastore VMFS résidant sur une baie de stockage externe. Le stockage est accessible via l'un des nombreux protocoles réseau différents. La figure suivante illustre un datastore VMFS sur une baie de stockage externe accessible à l'aide du protocole iSCSI.

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



Datastore NFS sur baie de stockage externe

Vous pouvez créer un datastore NFS résidant 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 système de stockage externe accessible via l'appliance de serveur NFS.



Services RAID matériels pour le stockage ONTAP Select local

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

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

Configuration du contrôleur RAID pour le stockage local

Tous les disques connectés localement qui fournissent à ONTAP Select une capacité de stockage secondaire doivent se trouver derrière un contrôleur RAID. La plupart des serveurs génériques sont dotés de plusieurs options de contrôleurs RAID pour différents prix, chacun avec différents niveaux de fonctionnalités. L'objectif est de prendre en charge autant de ces options que possible, en fournissant à ce dernier certaines exigences minimales placées sur le contrôleur.

> Vous ne pouvez pas détacher de machines virtuelles ONTAP Select qui utilisent la configuration RAID matérielle. Le détachement de disques n'est pris en charge que pour les machines virtuelles ONTAP Select qui utilisent la configuration RAID logicielle. Voir "Remplacez un disque défectueux dans une configuration RAID logicielle ONTAP Select" pour plus d'informations.

Le contrôleur RAID qui gère les disques ONTAP Select doit respecter les conditions suivantes :

- Le contrôleur RAID matériel doit disposer d'une unité de sauvegarde sur batterie (BBU) ou d'un cache d'écriture protégé par Flash (FBWC) et prendre en charge un débit de 12 Gbit/s.
- Le contrôleur RAID doit prendre en charge un mode qui résiste à au moins une ou deux pannes disques (RAID 5 et RAID 6).
- Le cache du lecteur doit être défini sur Désactivé.
- La stratégie d'écriture doit être configurée pour le mode écriture différée avec un remplacement pour pouvoir écrire en cas d'échec de BBU ou de mémoire flash.
- La stratégie d'E/S pour les lectures doit être définie sur mise en cache.

Tous les disques connectés en local qui fournissent à ONTAP Select le stockage de sauvegarde doivent être placés dans des groupes RAID exécutant RAID 5 ou RAID 6. Pour les disques SAS et SSD, grâce à des groupes RAID de 24 disques maximum, ONTAP tire parti de l'extension des demandes de lecture entrantes sur un plus grand nombre de disques. Les performances sont ainsi considérablement améliorées. Avec les configurations SAS/SSD, des tests de performances ont été réalisés sur des configurations à LUN unique plutôt que sur des configurations à LUN multiples. Aucune différence importante n'a été trouvée. Par souci de simplicité, NetApp recommande de créer le moins de LUN nécessaires pour prendre en charge vos besoins de configuration.

Les disques NL-SAS et SATA requièrent un ensemble de meilleures pratiques. Pour des raisons de performances, le nombre minimal de disques est toujours huit, mais la taille du groupe RAID ne doit pas être supérieure à 12 disques. NetApp recommande également d'utiliser un disque de secours par groupe RAID. Cependant, il est possible d'utiliser des disques de secours globaux pour tous les groupes RAID. Par exemple, vous pouvez utiliser deux unités de rechange pour trois groupes RAID, chaque groupe RAID composé de huit à 12 disques.

> L'extension et la taille maximales des datastores pour les versions ESX plus anciennes sont de 64 To, ce qui peut affecter le nombre de LUN nécessaires pour prendre en charge la capacité brute totale fournie par ces disques 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 pris par les demandes d'écriture. Ces trois modes sont les suivants :

- Écriture. Toutes les demandes d'E/S entrantes sont écrites dans le cache du contrôleur RAID, puis immédiatement transférées sur le disque avant d'accuser réception de la demande vers l'hôte.

- Poignet. Toutes les demandes d'E/S entrantes sont écrites directement sur le disque, contournement du cache du contrôleur RAID.
- Retour. Toutes les demandes d'E/S entrantes sont écrites directement sur le cache du contrôleur et immédiatement réceptionnées sur l'hôte. Les blocs de données sont transmis de manière asynchrone au disque à l'aide du contrôleur.

Le mode Writeback offre le chemin de données le plus court, avec accusé de réception des E/S immédiatement après que les blocs entrent dans le cache. Ce mode offre la plus faible latence et un débit maximal pour les charges de travail mixtes de lecture/écriture. Toutefois, sans la présence d'une technologie BBU ou Flash non volatile, les utilisateurs risquent de perdre des données en cas de panne de courant lors du fonctionnement dans ce mode.

ONTAP Select requiert la présence d'une unité de sauvegarde sur batterie ou Flash. Ainsi, nous pouvons être certains que les blocs en cache sont transmis au disque en cas de ce type de défaillance. C'est pourquoi il est obligatoire de configurer le contrôleur RAID en mode écriture différée.

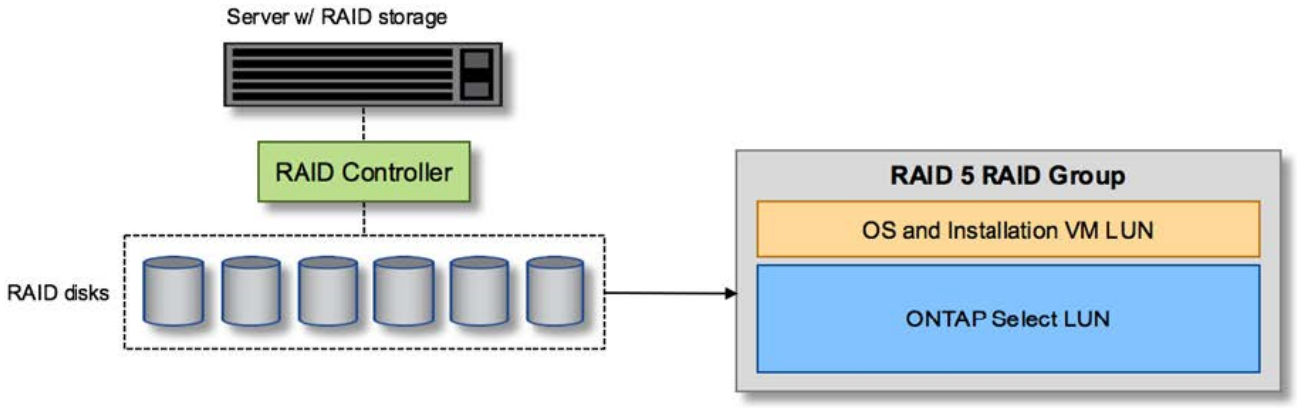
Disques locaux partagés entre le système ONTAP Select et le système d'exploitation

La configuration de serveur la plus courante est celle dans laquelle toutes les piles de disques connectées localement sont situées derrière un seul contrôleur RAID. Vous devez provisionner au moins deux LUN : un pour l'hyperviseur et un pour la machine virtuelle ONTAP Select.

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

La figure suivante montre ce style 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 de LUN de serveur avec uniquement des piles de disques gérées par RAID



Le provisionnement des LUN de système d'exploitation à partir du même groupe RAID que ONTAP Select permet à l'OS de l'hyperviseur (et à toute machine virtuelle client qui est également provisionnée à partir de ce stockage) de bénéficier de la protection RAID. Cette configuration évite qu'une panne de disque unique n'entraînent l'arrêt du système dans son ensemble.

Disques locaux séparés entre ONTAP Select et OS

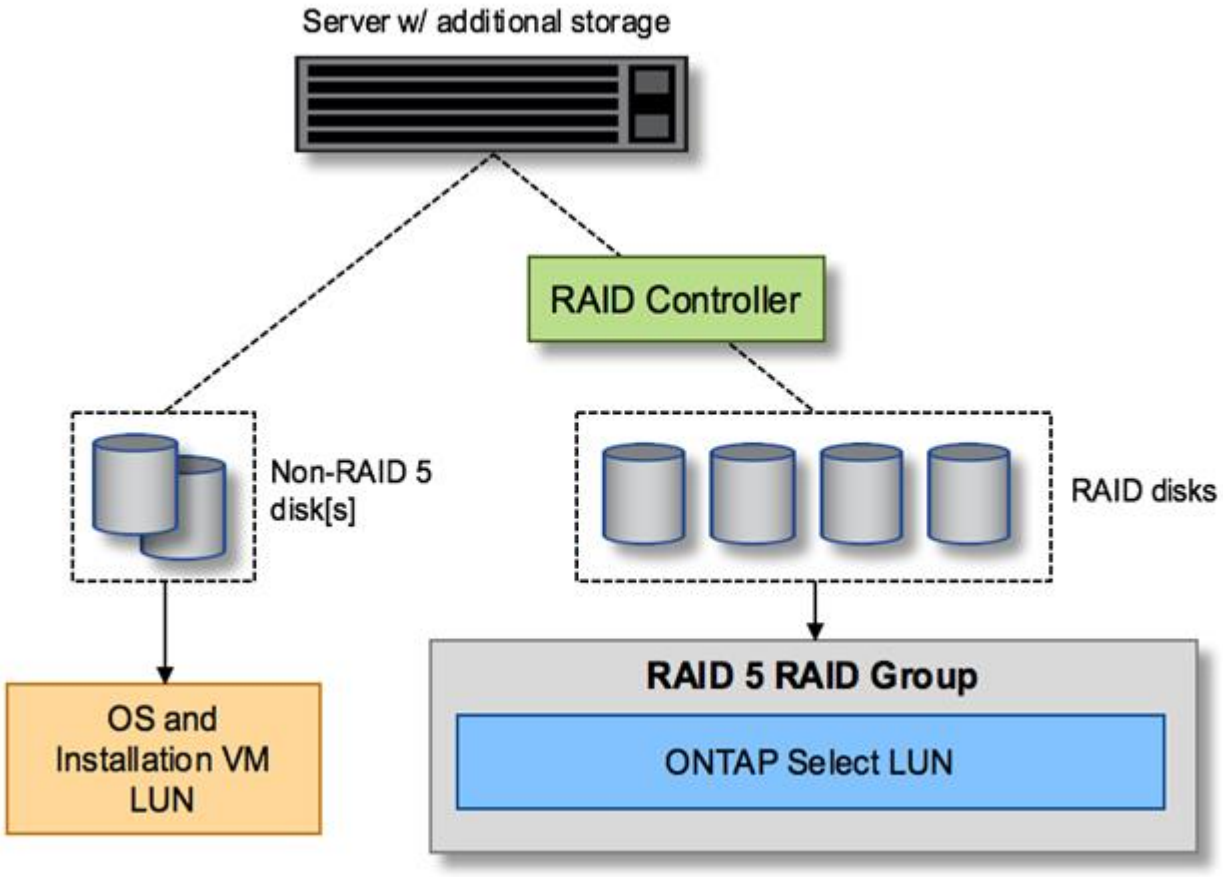
L'autre configuration possible proposée par les fournisseurs de serveurs implique de configurer le système avec plusieurs contrôleurs RAID ou de disque. Dans cette configuration, un jeu de disques est géré par un

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

Dans ce style de configuration, l'ensemble des piles de disques derrière le contrôleur RAID qui peut fournir des services RAID 5/6 doit être utilisé exclusivement par la machine virtuelle ONTAP Select. Selon la capacité de stockage totale que vous gérez, vous devez configurer les piles de disques en un ou plusieurs groupes RAID et une ou plusieurs LUN. Ces LUN seraient ensuite utilisées pour créer un ou plusieurs datastores, tous protégés par le contrôleur RAID.

Le premier jeu de disques est réservé pour le système d'exploitation de l'hyperviseur et toute machine virtuelle client qui n'utilise pas le stockage ONTAP, comme illustré dans la figure ci-dessous.

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



Plusieurs LUN

Dans deux cas, vous devez modifier les configurations single-RAID group/single-LUN. Si vous utilisez des disques NL-SAS ou SATA, la taille du groupe RAID ne doit pas dépasser 12 disques. En outre, une LUN peut être plus grande que les limites de stockage de l'hyperviseur sous-jacent : taille maximale de l'extension du système de fichiers individuels ou taille maximale du pool de stockage total. Ensuite, le stockage physique sous-jacent doit être réparti en plusieurs LUN pour permettre la création du système de fichiers.

Limites du système de fichiers de machines virtuelles VMware vSphere

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

Si un serveur possède plus de 64 To de stockage connecté, plusieurs LUN peuvent avoir besoin d'être provisionnés, chacun d'être plus petits que 64 To. La création de plusieurs groupes RAID pour améliorer la durée de reconstruction RAID sur les disques SATA/NL-SAS entraîne également le provisionnement de plusieurs LUN.

Lorsque plusieurs LUN sont requises, un élément majeur est l'élément à prendre en compte afin de s'assurer que ces LUN présentent des performances similaires et cohérentes. Ceci est particulièrement important si l'ensemble des LUN doivent être utilisés dans un seul agrégat ONTAP. Si un sous-ensemble d'un ou plusieurs LUN présente un profil de performances différent, nous vous recommandons vivement d'isoler ces LUN dans un agrégat ONTAP distinct.

Plusieurs extensions du système de fichiers peuvent être utilisées pour créer un datastore unique jusqu'à la taille maximale du datastore. Pour limiter la capacité qui nécessite une licence ONTAP Select, veillez à spécifier un plafond de capacité lors de l'installation du cluster. Cette fonctionnalité permet à ONTAP Select d'utiliser (et donc de nécessiter une licence) uniquement un sous-ensemble de l'espace d'un datastore.

Vous pouvez également commencer par créer un datastore unique sur une LUN. Lorsque de l'espace supplémentaire requiert une licence de capacité ONTAP Select supérieure, cet espace peut être ajouté au même datastore dans une certaine mesure, dans la limite de 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 la capacité sont pris en charge et peuvent être obtenus à l'aide de la fonctionnalité d'ajout de stockage de ONTAP. Chaque nœud ONTAP Select peut être configuré pour prendre en charge jusqu'à 400 To de stockage. Le provisionnement des capacités depuis plusieurs datastores requiert un processus en deux étapes.

La création initiale du cluster peut être utilisée pour créer un cluster ONTAP Select qui consomme une partie ou l'intégralité de l'espace dans le datastore initial. Une deuxième étape consiste à effectuer une ou plusieurs opérations d'ajout de capacité en utilisant des datastores supplémentaires jusqu'à ce que la capacité totale souhaitée soit atteinte. Cette fonctionnalité est détaillée dans la section "Augmenter la capacité de stockage".

La surcharge VMFS n'est pas nulle (voir "VMware KB 1001618"), et la tentative d'utilisation de l'espace entier signalé comme libre par un datastore a entraîné des erreurs fallacieuses au cours des opérations de création du cluster.

Un tampon de 2 % reste inutilisé dans chaque datastore. Cet espace ne nécessite pas de licence de capacité, car ONTAP Select n'est pas utilisé. ONTAP Deploy calcule automatiquement le nombre exact de gigaoctets pour le tampon, tant que la capacité maximale n'est pas spécifiée. Si un bouchon de capacité est spécifié, cette taille est appliquée en premier. Si la taille du bouchon de capacité est comprise dans la taille du tampon, la création du cluster échoue et un message d'erreur indiquant le paramètre de taille maximale approprié qui peut 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 en tant que cible des opérations de stockage vMotion d'une machine virtuelle ONTAP Deploy ou ONTAP Select.

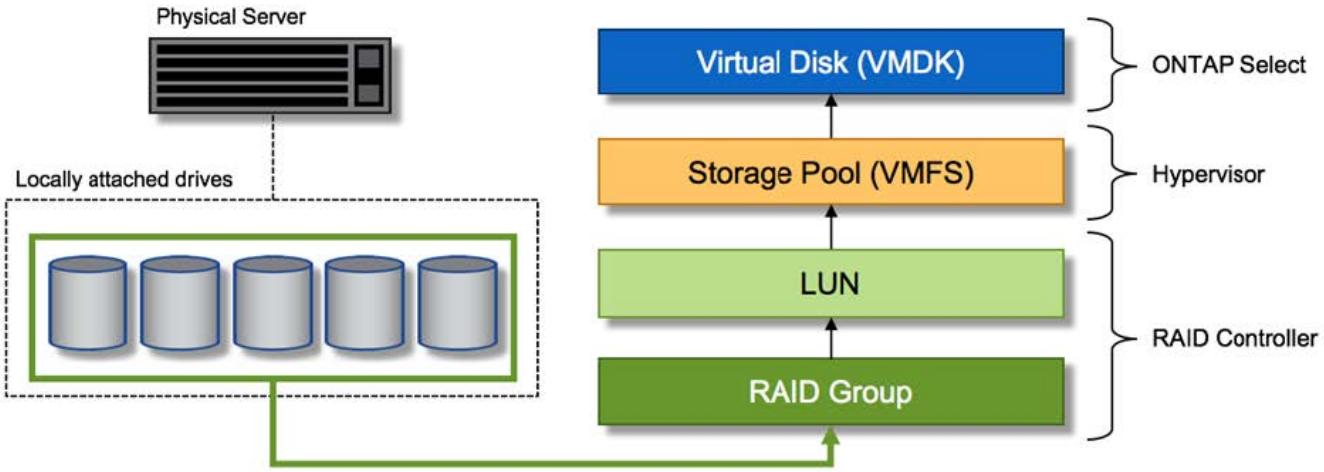
VMware ne prend pas en charge les mises à niveau sur place de VMFS 5 vers VMFS 6. Storage vMotion est donc le seul mécanisme qui permet aux machines virtuelles de passer d'un datastore VMFS 5 à un datastore VMFS 6. Cependant, la prise en charge de Storage vMotion avec ONTAP Select et ONTAP a été étendue pour couvrir d'autres scénarios, en plus de l'objectif spécifique de la transition de VMFS 5 vers VMFS 6.

Disques virtuels ONTAP Select

Dans le cœur, ONTAP Select présente à ONTAP un ensemble de disques virtuels provisionnés à partir d'un ou plusieurs pools de stockage. ONTAP est présenté avec un ensemble de disques virtuels qu'il traite comme physiques, et la partie restante de la pile de stockage est extraite par l'hyperviseur. La figure suivante montre cette relation plus en détail, mettant en évidence la relation entre le contrôleur RAID physique, l'hyperviseur et la machine virtuelle ONTAP Select.

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

Mappage disque virtuel sur disque physique



Provisionnement de disque virtuel

Pour optimiser l'expérience utilisateur, l'outil de gestion ONTAP Select déploie, ONTAP provisionne automatiquement les disques virtuels depuis le pool de stockage associé et les connecte à la machine virtuelle ONTAP Select. Cette opération a lieu automatiquement lors de la configuration initiale et 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 miroir.

ONTAP Select divise le stockage NAS sous-jacent en disques virtuels de taille équivalente, chacun ne dépassant pas 16 To. Si le nœud ONTAP Select fait partie d'une paire haute disponibilité, un minimum de deux disques virtuels sont créés sur chaque nœud de cluster et attribués au plex local et miroir à utiliser dans un agrégat en miroir.

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

L'ajout de capacité à une machine virtuelle ONTAP Select entraîne probablement des VMDK de différentes tailles. Pour plus de détails, reportez-vous à la section "Augmenter la capacité de stockage". Contrairement aux systèmes FAS, des VMDK de différentes tailles peuvent exister dans le même agrégat. ONTAP Select utilise une bande RAID 0 sur ces VMDK, ce qui permet d'exploiter pleinement l'ensemble de l'espace de chaque VMDK, quelle que soit sa taille.

NVRAM virtualisée

En général, les systèmes FAS de NetApp sont équipés d'une carte PCI NVRAM physique, une carte hautes performances contenant une mémoire Flash non volatile. Cette carte améliore considérablement les performances d'écriture en permettant à ONTAP d'accuser réception immédiate des écritures entrantes sur le client. Il peut également planifier le déplacement des blocs de données modifiés vers le support de stockage plus lent, dans le cadre d'un processus appelé déchargement.

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

La mémoire NVRAM est placée sur son propre VMDK. Le fractionnement de la mémoire NVRAM dans son propre VMDK permet à la machine virtuelle ONTAP Select d'utiliser le pilote vNVMe pour communiquer avec son VMDK NVRAM. La machine virtuelle ONTAP Select utilise également la version 13 du matériel compatible avec ESX 6.5 et les versions ultérieures.

Le chemin d'accès aux données s'est expliqué : contrôleur NVRAM et RAID

L'interaction entre la partition système NVRAM virtualisée et le contrôleur RAID peut être mise en évidence en parcourant le chemin d'accès aux données pris par une demande d'écriture lors de son entrée dans le système.

Les demandes d'écriture entrantes sur la machine virtuelle ONTAP Select sont ciblées sur la partition NVRAM de la machine virtuelle. Au niveau de la couche de virtualisation, cette partition existe sur un disque système ONTAP Select, un VMDK attaché à la machine virtuelle d'ONTAP Select. Au niveau de la couche physique, ces requêtes sont mises en cache dans le contrôleur RAID local, comme toutes les modifications de bloc qui sont destinées aux piles de disques sous-jacentes. À partir de là, l'écriture est acquittée vers l'hôte.

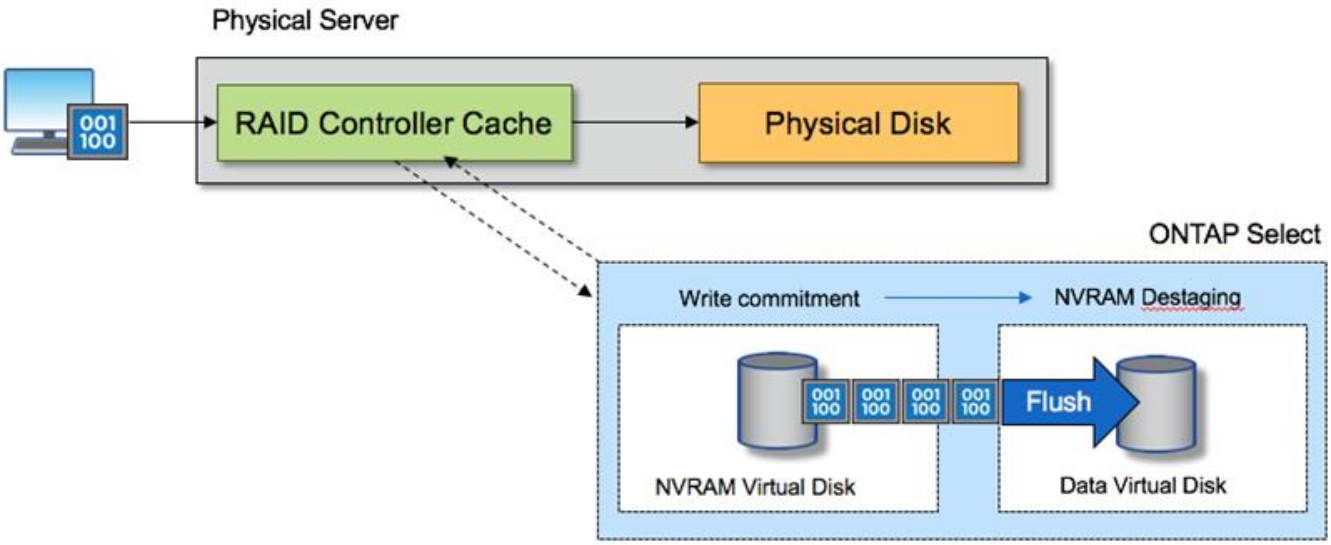
À ce stade, physiquement, le bloc réside dans le cache du contrôleur RAID, en attente d'être transféré vers le disque. Logiquement, le bloc réside dans la NVRAM qui attend le transfert vers les disques de données utilisateur appropriés.

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 régulièrement transférées sur le support de stockage physique. Cette opération ne doit pas être confondue avec le rinçage périodique du contenu NVRAM sur les disques de données ONTAP. Ces deux événements ne sont pas liés et se produisent à des moments et des fréquences différents.

La figure suivante montre le chemin d'E/S qu'une écriture entrante prend. Elle souligne la différence entre la couche physique (représentée par le cache du contrôleur RAID et les disques) et la couche virtuelle (représentée par la mémoire NVRAM de la machine virtuelle et les disques virtuels de données).

Bien que les blocs modifiés sur le VMDK NVRAM soient mis en cache dans le cache de contrôleur RAID local, le cache ne connaît pas la construction de la machine virtuelle ou ses disques virtuels. Il stocke tous les blocs modifiés sur le système, pour lesquels la mémoire NVRAM n'est qu'une partie. Cela inclut les demandes d'écriture liées à l'hyperviseur, si elles sont provisionnées à partir des mêmes piles de disques de support.

Écritures entrantes sur 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 ESX de 6.5 ou ultérieure. Ce changement est particulièrement important pour les installations ONTAP Select avec RAID logiciel, qui ne bénéficient pas du cache du contrôleur RAID.

Services RAID logiciels pour le stockage local

Le RAID logiciel est une couche d'abstraction RAID implémentée dans la pile logicielle ONTAP. Elle fournit les mêmes fonctionnalités que la couche RAID sur une plateforme ONTAP traditionnelle telle que FAS. La couche RAID effectue des calculs de parité des disques et protège contre les pannes de disques individuelles au sein d'un nœud ONTAP Select.

Indépendamment des configurations RAID matérielles, ONTAP Select propose également une option RAID logicielle. Un contrôleur RAID matériel peut ne pas être disponible ou n'être pas souhaitable dans certains environnements, notamment lors du déploiement de ONTAP Select sur un matériel générique petit format. Software RAID étend les options de déploiement disponibles pour inclure de tels environnements. Pour activer le RAID logiciel dans votre environnement, voici quelques points à retenir :

- Elle est disponible avec une licence Premium ou Premium XL.
- Elle prend uniquement en charge les disques SSD ou NVMe (requiert une licence Premium XL) pour les disques racine ONTAP et de données.
- Il nécessite un disque système distinct pour la partition de démarrage de la machine virtuelle ONTAP Select.

- Choisissez un disque distinct, soit un disque SSD, soit un lecteur NVMe, pour créer un datastore pour les disques système (NVRAM, carte Boot/CF, coredump et Mediator dans une configuration à plusieurs nœuds).

Notes

- Les termes « disque de service » et « disque système » sont utilisés de manière interchangeable.
  - Les disques de service sont les VMDK utilisés au sein de la machine virtuelle ONTAP Select pour assurer la maintenance des différents éléments, tels que la mise en cluster, le démarrage, etc.
  - Les disques de service sont situés physiquement sur un seul disque physique (appelé collectivement disque physique de service/système), vu depuis l'hôte. Ce disque physique doit contenir un datastore DAS. ONTAP Deploy crée ces disques de service pour la machine virtuelle d'ONTAP Select pendant le déploiement du cluster.
- Il est impossible de séparer davantage les disques du système ONTAP Select entre plusieurs datastores ou entre plusieurs disques physiques.
- Le RAID matériel n'est pas obsolète.

Configuration RAID logicielle pour le stockage local

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

- Le contrôleur RAID matériel doit être désactivé de sorte que les disques puissent être présentés directement au système (JBOD). Cette modification peut généralement être effectuée dans le BIOS du contrôleur RAID
- Ou le contrôleur RAID matériel doit être en mode SAS HBA. Par exemple, certaines configurations du BIOS autorisent un mode "AHCI" en plus de RAID, qui peut être choisi pour activer le mode JBOD. Cela active un passthrough, de sorte que les disques physiques puissent être vus comme étant sur l'hôte.

En fonction du nombre maximum de disques pris en charge par le contrôleur, un contrôleur supplémentaire peut être nécessaire. Avec le mode HBA SAS, vérifiez que le contrôleur d'E/S (SAS HBA) est pris en charge avec une vitesse de 6 Gbit/s. Toutefois, NetApp recommande un débit de 12 Gbit/s.

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

Les administrateurs doivent garder le contrôle des disques utilisés par la machine virtuelle ONTAP Select et éviter toute utilisation involontaire de ces disques sur l'hôte.

Disques virtuels et physiques ONTAP Select

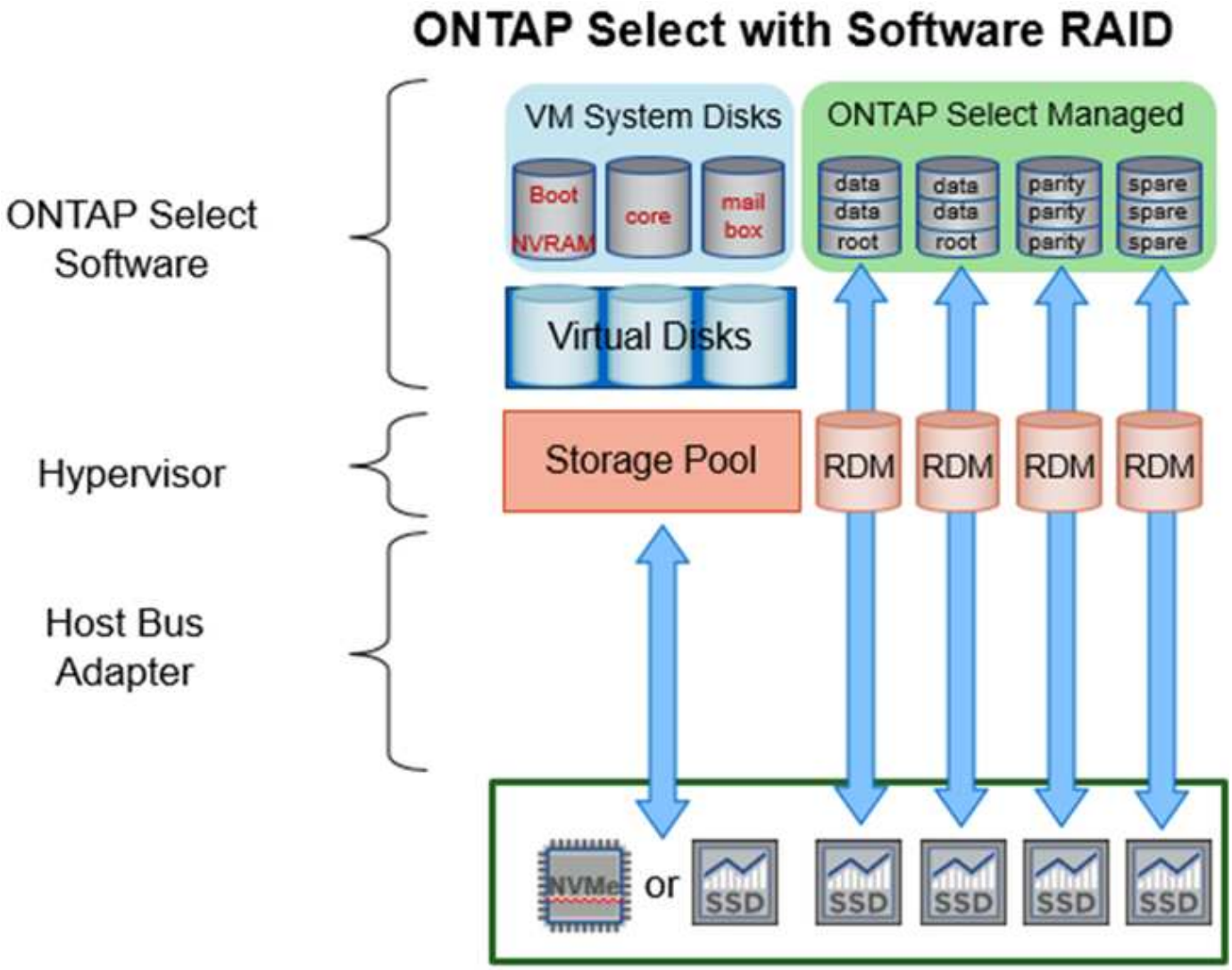
Pour les configurations avec contrôleurs RAID matériels, la redondance de disque physique est fournie par le contrôleur RAID. La solution ONTAP Select est présentée avec un ou plusieurs VMDK à partir desquels l'administrateur ONTAP peut configurer les agrégats de données. Ces VMDK sont répartis dans un format RAID 0, car le logiciel ONTAP RAID est redondant, inefficace et inefficace du fait de la résilience fournie au niveau matériel. En outre, les VMDK utilisés pour les disques système sont dans le même datastore que les VMDK utilisés pour stocker les données des utilisateurs.

Avec le RAID logiciel, ONTAP Deploy présente à ONTAP Select un jeu de disques virtuels (VMDK) et des disques physiques des mappages de périphériques Raw Device [RDM] pour les SSD et les périphériques

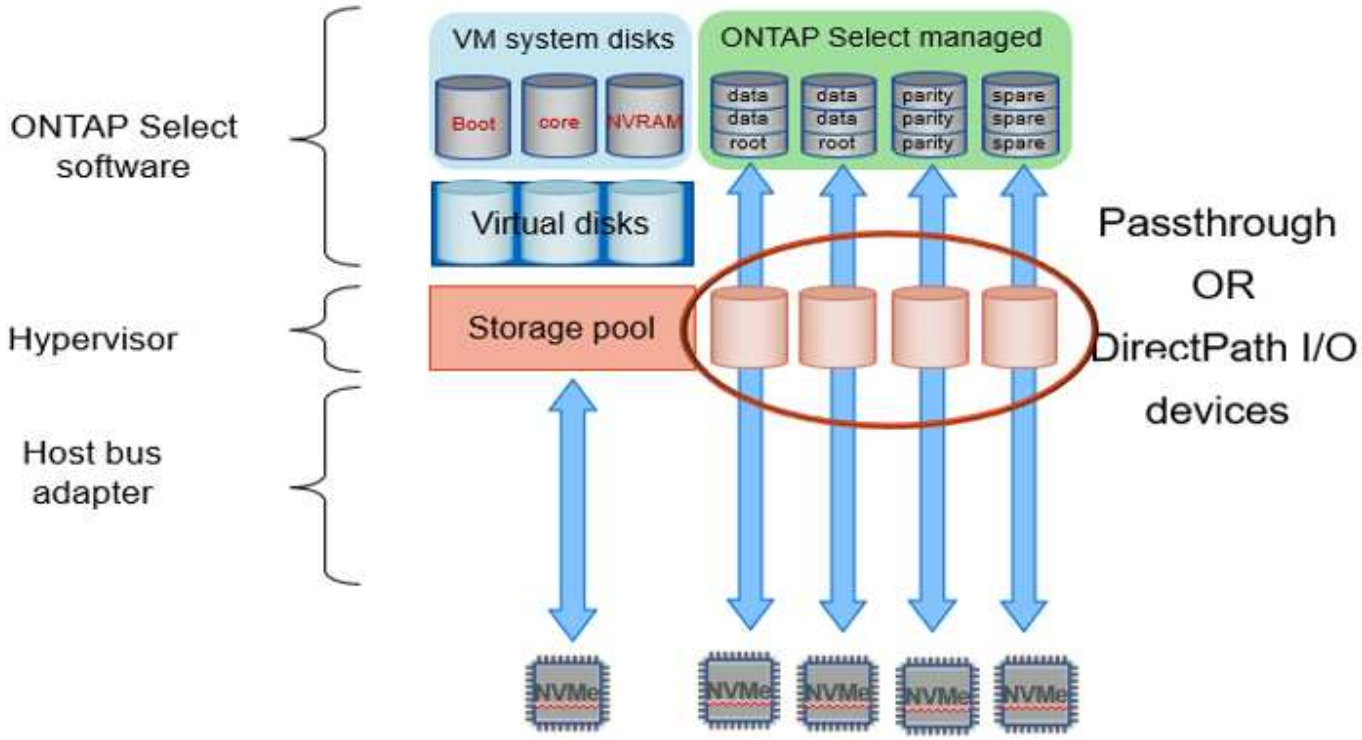
d’E/S de type passerelle ou DirectPath pour les NVMes.

Les figures suivantes montrent cette relation plus en détail, soulignant la différence entre les disques virtualisés utilisés pour les serveurs virtuels internes de ONTAP Select et les disques physiques utilisés pour stocker les données de l’utilisateur.

Logiciel ONTAP Select RAID : 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 requiert un support rapide et durable. Par conséquent, seuls les datastores NVMe et de type SSD sont pris en charge.



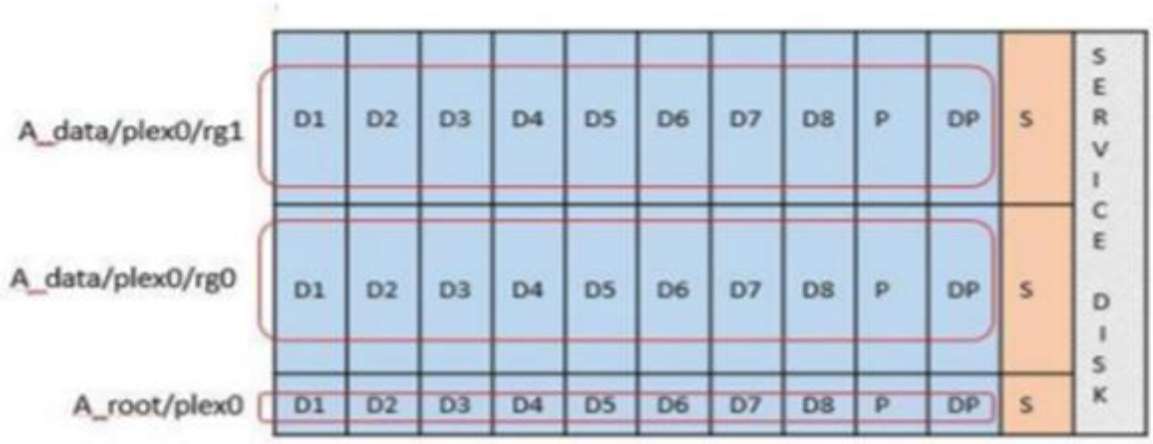
Les disques système (VMDK) résident dans le même datastore et sur le même disque physique. Le disque NVRAM virtuel requiert un support rapide et durable. Par conséquent, seuls les datastores NVMe et de type SSD sont pris en charge. Lorsque vous utilisez des disques NVMe pour les données, le disque du système doit également être un périphérique NVMe pour des raisons de performance. Les cartes INTEL Optane sont un bon candidat pour le disque système dans une configuration 100 % NVMe.

Avec la version actuelle, il est impossible de séparer davantage les disques du système ONTAP Select entre plusieurs datastores ou plusieurs disques physiques.

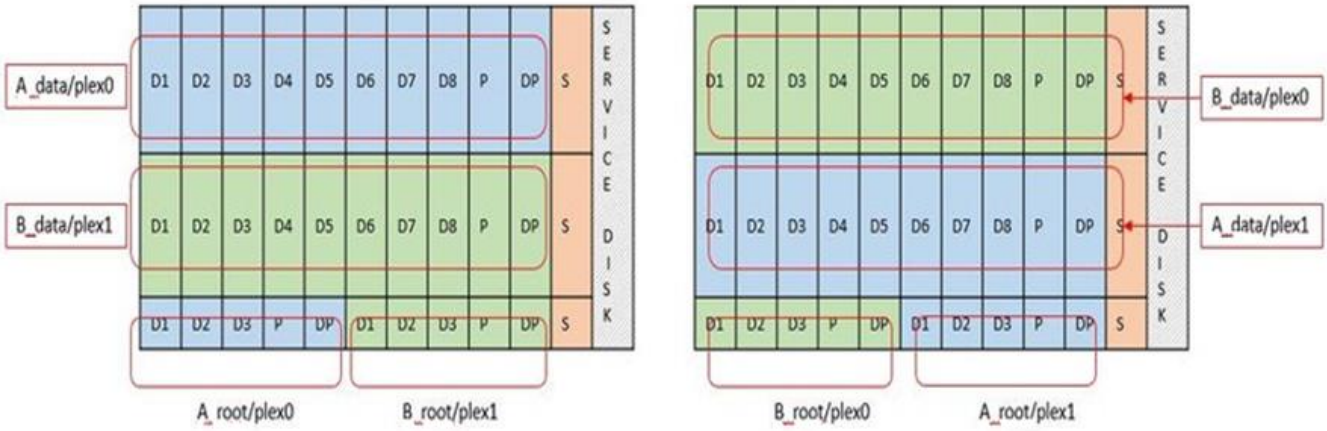
Chaque disque de données est divisé en trois parties : une petite partition racine (bande) et deux partitions de taille équivalente pour créer deux disques de données vus au sein de la machine virtuelle ONTAP Select. Les partitions utilisent le schéma RD2 (données racines) comme indiqué dans les figures suivantes pour un cluster à un seul nœud et pour un nœud dans une paire HA.

`P` signale un disque de parité. `DP` signale un disque double parité et `S` signale un disque de rechange.

Partitionnement de disque RDD pour les clusters à un seul nœud



Partitionnement de disque RDD pour les clusters multinœuds (paires HA)



Le logiciel RAID de ONTAP prend en charge les types RAID suivants : RAID 4, RAID-DP et RAID-TEC. Ce sont les mêmes constructions RAID que les plates-formes FAS et AFF. Pour le provisionnement racine, ONTAP Select ne prend en charge que RAID 4 et RAID-DP. Lorsque vous utilisez RAID-TEC pour l'agrégat de données, la protection globale est RAID-DP. ONTAP Select HA utilise une architecture sans partage qui réplique la configuration de chaque nœud sur l'autre nœud. Cela signifie que chaque nœud doit stocker sa partition racine et une copie de la partition racine de son homologue. Étant donné qu'un disque de données dispose d'une seule partition root, le nombre minimal de disques de données varie selon que le nœud ONTAP Select fait partie d'une paire HA ou non.

Pour les clusters à un seul nœud, toutes les partitions de données sont utilisées pour stocker des 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) pour ce nœud et la seconde partition de données est utilisée pour mettre en miroir les données actives depuis le homologue haute disponibilité.

Comparaison entre les périphériques Passthrough (DirectPath IO) et Cartes de périphériques brutes (RDM)

VMware ESX ne prend actuellement pas en charge les disques NVMe en tant que Raw Device Maps. Pour que ONTAP Select puisse prendre le contrôle direct des disques NVMe, les disques NVMe doivent être configurés dans ESX en tant que périphériques d'authentification. Veuillez noter que la configuration d'un périphérique NVMe en tant que périphérique d'authentification requiert la prise en charge du BIOS du serveur

et qu'il s'agit d'un processus perturbateur nécessitant un redémarrage de l'hôte ESX. De plus, le nombre maximal de périphériques d'authentification par hôte ESX est de 16. Cependant, le déploiement ONTAP limite ceci à 14. Cette limite de 14 dispositifs NVMe par nœud ONTAP Select implique qu'une configuration 100 % NVMe offre une densité très élevée d'IOPS (IOPS/To), au détriment de la capacité totale. De la même façon, si vous souhaitez une configuration haute performance avec une capacité de stockage supérieure, il est recommandé d'utiliser 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 exploiter pleinement les performances de NVMe, pensez à la taille importante des machines virtuelles ONTAP Select.

Il y a une différence supplémentaire entre les dispositifs de passants et les RDM. RDM peut être mappé à une machine virtuelle en cours d'exécution. Les terminaux passthrough requièrent un redémarrage VM. Cela signifie que toute procédure de remplacement de disque NVMe ou d'extension de capacité (ajout de disque) nécessite un redémarrage de machine virtuelle ONTAP Select. Le remplacement de disques et l'extension de capacité (ajout de disques) sont pilotés par un workflow du déploiement ONTAP. ONTAP Deploy gère le redémarrage de ONTAP Select pour les clusters à un seul nœud et le basculement/retour arrière pour les paires haute disponibilité. Toutefois, il est important de noter la différence entre l'utilisation des disques de données SSD (aucun redémarrage ONTAP Select/basculement n'est nécessaire) et la compatibilité avec les disques de données NVMe (le redémarrage ou le basculement ONTAP Select est requis).

Provisionnement des disques physiques et virtuels

Pour optimiser l'expérience utilisateur, le déploiement de ONTAP provisionne automatiquement les disques (virtuels) du datastore spécifié (disque du système physique) et les connecte à la machine virtuelle ONTAP Select. Cette opération a lieu automatiquement lors de la configuration initiale afin que la machine virtuelle ONTAP Select puisse démarrer. Les RDM sont partitionnés et l'agrégat racine est automatiquement créé. Si le nœud ONTAP Select fait partie d'une paire haute disponibilité, les partitions de données sont automatiquement attribuées à un pool de stockage local et à un pool de stockage en miroir. Cette affectation a lieu automatiquement lors des opérations de création de clusters et d'ajout de stockage.

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

Le type de groupe RAID de l'agrégat racine dépend du nombre de disques disponibles. Le déploiement de ONTAP sélectionne le type de groupe RAID approprié. S'il dispose de suffisamment de disques alloués au nœud, il utilise RAID-DP, sinon il crée un agrégat racine RAID-4.

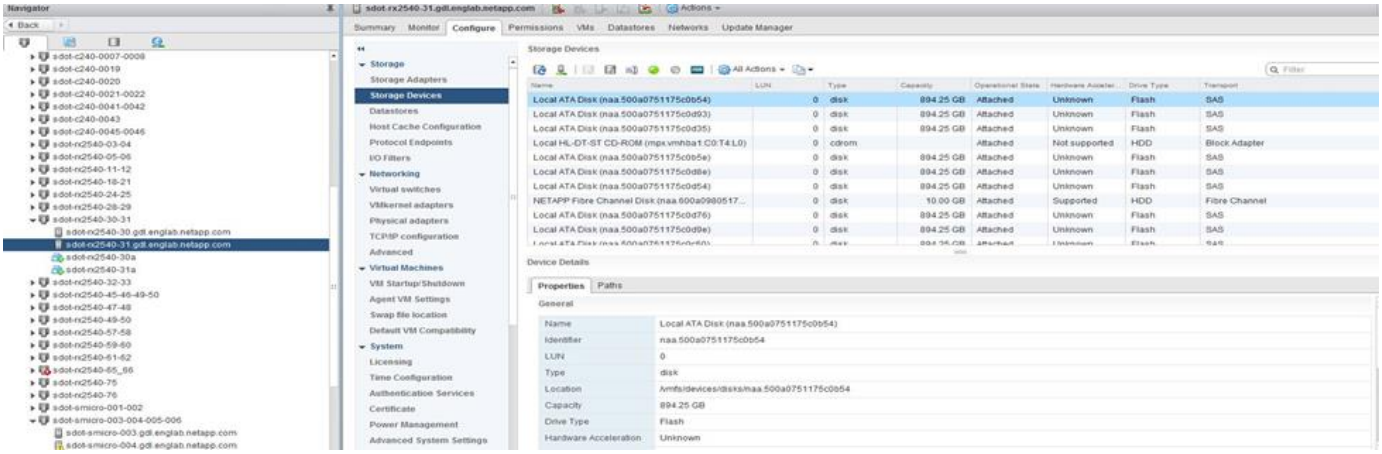
Lorsque vous ajoutez de la capacité à une machine virtuelle ONTAP Select via la technologie RAID logicielle, l'administrateur doit tenir compte de la taille du disque physique et du nombre de disques requis. Pour plus de détails, reportez-vous à la section "Augmenter la capacité de stockage".

À l'instar des systèmes FAS et AFF, seuls les disques de capacité égale ou supérieure peuvent être ajoutés à un groupe RAID existant. Les disques de grande capacité sont de la taille appropriée. Si vous créez de nouveaux groupes RAID, la taille du groupe RAID doit correspondre à celle du groupe RAID existant pour garantir que les performances globales de l'agrégat ne se détériorent pas.

Faites correspondre un disque ONTAP Select au disque ESX 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 le shell ESXi, vous pouvez entrer la commande suivante pour faire clignoter le voyant d'un disque physique donné (identifié par son naa.unique-ID).

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

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

Il est possible qu'un système rencontre une situation dans laquelle plusieurs disques sont en panne en même temps. 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 panne de disque, et un agrégat RAID-DP peut survivre à deux pannes de disque 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 pris en charge par ce type RAID et si un disque de spare est disponible, le processus de reconstruction démarre automatiquement. Si des disques de spare ne sont pas disponibles, l'agrégat transmet des données en état dégradé jusqu'à l'ajout de disques de spare.

Si le nombre de disques défaillants est supérieur au nombre maximal de défaillances pris en charge par le

type RAID, le plex local est marqué comme défectueux et l'état de l'agrégat est dégradé. Les données sont servies par le second plex résidant sur le partenaire de haute disponibilité. Cela signifie que toutes les demandes d'E/S du nœud 1 sont envoyées via le port d'interconnexion de cluster e0e (iSCSI) aux disques physiquement situés sur le nœud 2. Si le second plex tombe également en panne, l'agrégat est marqué comme étant en panne et les données sont indisponibles.

Un plex défaillant doit être supprimé et recréé pour la mise en miroir correcte des données à reprendre. Notez qu'une défaillance de plusieurs disques entraînant la dégradation de l'agrégat de données entraîne également la dégradation de l'agrégat racine. ONTAP Select utilise le schéma de partitionnement données-racines (RDD) pour séparer chaque disque physique en une partition racine et deux partitions de données. Par conséquent, perdre un ou plusieurs disques peut avoir un impact sur plusieurs agrégats, notamment la racine locale ou la copie de l'agrégat racine distant, ainsi que l'agrégat de données locales et la copie de l'agrégat de données distant.

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

C3111E67::> storage aggregate mirror -aggregate aggr1
Info: Disks would be added to aggregate "aggr1" on node "sti-rx2540-335a"
in the following manner:
      Second Plex
        RAID Group rg0, 5 disks (advanced_zoned checksum, raid_dp)
                                                                  Usable
Physical
          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 défaillances de lecteur, utiliser le `storage disk fail -disk NET-x.y -immediate` commande. Si un disque de secours se trouve dans le système, l'agrégat commence à reconstruire. Vous pouvez vérifier le statut de la reconstruction à l'aide de la commande `storage aggregate show`. Vous pouvez supprimer le disque défectueux simulé à l'aide de ONTAP Deploy. Notez que ONTAP a marqué le lecteur comme `Broken`. Le disque n'est pas réellement cassé et peut être ajouté à l'aide du logiciel ONTAP Deploy. Pour effacer l'étiquette interrompue, entrez les commandes suivantes dans l'interface de ligne de commande d'ONTAP Select :

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

La sortie de la dernière commande doit être vide.

NVRAM virtualisée

En général, les systèmes FAS de NetApp sont équipés d'une carte PCI NVRAM physique. Cette carte hautes performances contient une mémoire Flash non volatile qui permet de booster considérablement les performances en écriture. En effet, ONTAP permet à ce dernier de valider immédiatement les écritures entrantes sur le client. Il peut également planifier le déplacement des blocs de données modifiés vers le support de stockage plus lent, dans le cadre d'un processus appelé déchargement.

Les systèmes de produits de base ne sont généralement pas équipés de ce type d'équipement. La fonctionnalité de la carte NVRAM a donc été virtualisée et placée dans une partition sur le disque de démarrage du système ONTAP Select. C'est pour cette raison que le placement du disque virtuel système de l'instance est extrêmement important.

Configurations de baies externes et VSAN

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

La configuration sous-jacente est au minimum prise en charge par VMware et doit être répertoriée sur les HCLs VMware correspondants.

Architecture vNAS

La nomenclature vNAS est utilisée pour toutes les configurations qui n'utilisent pas DAS. Pour les clusters ONTAP Select multinœuds, ces architectures intègrent des architectures selon lesquelles les deux nœuds ONTAP Select de la même paire haute disponibilité partagent un seul datastore (y compris les datastores VSAN). Les nœuds peuvent également être installés sur des datastores séparés à partir de la même baie externe partagée. L'efficacité du stockage côté baie peut ainsi réduire l'empreinte globale de la paire haute disponibilité ONTAP Select entière. 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 HA. Les politiques d'efficacité du stockage ONTAP ont été définies au niveau des nœuds. Par conséquent, les fonctionnalités d'efficacité du stockage côté baie sont préférables car elles peuvent être appliquées à l'ensemble des jeux de données à partir des deux nœuds

ONTAP Select.

Il est également possible que chaque nœud ONTAP Select d'une paire HA utilise une baie externe distincte. Ce choix est commun avec ONTAP Select MetroCluster SDS et un stockage externe.

Lors de l'utilisation de baies externes distinctes pour chaque nœud ONTAP Select, il est très important que les deux baies présentent des caractéristiques de performances similaires à la machine virtuelle d'ONTAP Select.

Architectures vNAS par rapport à DAS local avec contrôleurs RAID matériels

L'architecture vNAS est logiquement la plus similaire à l'architecture d'un serveur avec DAS et un contrôleur RAID. Dans les deux cas, ONTAP Select utilise un espace de datastore. L'espace du datastore est divisé en VMDK et ces VMDK constituent les agrégats de données ONTAP traditionnels. ONTAP Deploy garantit que les VMDK sont correctement dimensionnés et attribués au plex approprié (dans le cas de paires haute disponibilité) lors des opérations de création et d'ajout de stockage en cluster.

Il existe deux différences majeures entre vNAS et DAS avec un contrôleur RAID. La différence la plus immédiate est que vNAS ne nécessite pas de contrôleur RAID. VNAS assure que la baie externe sous-jacente fournit la persistance et la résilience des données qu'un DAS avec une configuration de contrôleur RAID fournit. La deuxième différence, plus subtile, est quant aux performances de la NVRAM.

NVRAM vNAS

La NVRAM de ONTAP Select est un VMDK. En d'autres termes, ONTAP Select émule un espace adressable par octet (NVRAM traditionnelle) sur un périphérique adressable par bloc (VMDK). Cependant, la performance de la mémoire NVRAM est absolument cruciale pour les 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 de facto, car toutes les écritures sur le VMDK NVRAM sont d'abord hébergées dans le cache du contrôleur RAID.

Pour les architectures VNAS, ONTAP Deploy configure automatiquement les nœuds ONTAP Select avec un argument de démarrage appelé consignation de données à instance unique (SIDL). Lorsqu'il s'agit d'un argument de démarrage, ONTAP Select ignore la mémoire NVRAM et écrit les données directement dans l'agrégat de données. La mémoire NVRAM n'est utilisée que pour enregistrer l'adresse des blocs modifiés par l'opération D'ÉCRITURE. Cette fonctionnalité permet d'éviter une double écriture : une écriture sur la mémoire NVRAM et une seconde écriture lorsque la mémoire NVRAM est déchargée. Cette fonctionnalité n'est activée que sur vNAS, car les écritures locales sur le cache du contrôleur RAID présentent une latence supplémentaire négligeable.

La fonction SIDL n'est pas compatible avec toutes les fonctionnalités d'efficacité du stockage ONTAP Select. La fonction 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
```

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

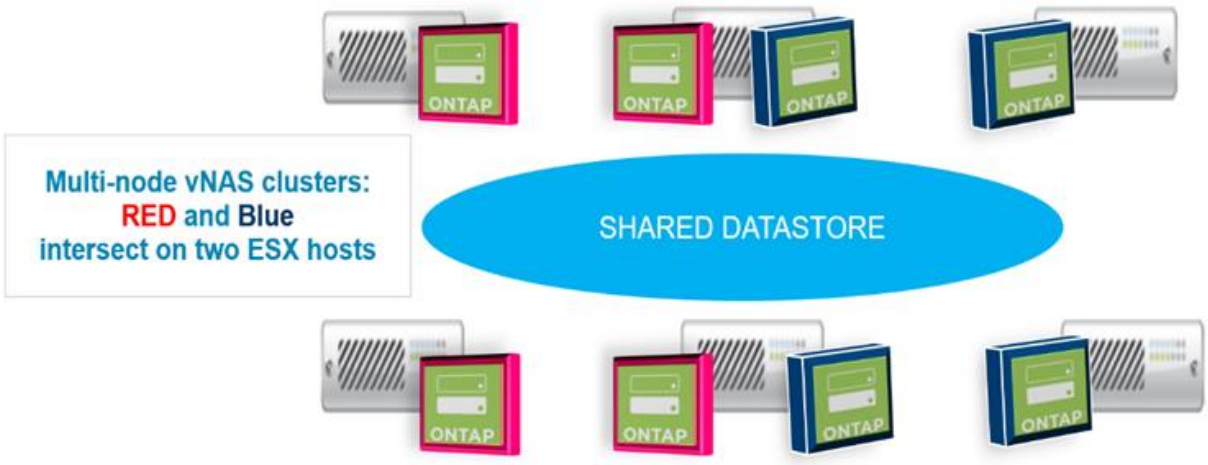
```
volume efficiency stop -all true -vserver * -volume * (all volumes in the
affected aggregate)
```

Colocaliser les nœuds ONTAP Select avec vNAS

ONTAP Select prend également en charge les clusters ONTAP Select multinœuds sur le stockage partagé. ONTAP Deploy permet la configuration de plusieurs nœuds ONTAP Select sur le même hôte ESX tant que ces nœuds ne font pas partie du même cluster. Notez que cette configuration n'est valide que pour les environnements VNAS (datastores partagés). En cas d'utilisation du stockage DAS, plusieurs instances ONTAP Select par hôte ne sont pas prises en charge, car ces instances se disputent le même contrôleur RAID matériel.

ONTAP Deploy s'assure que le déploiement initial du cluster NAS multinœud ne place pas plusieurs instances ONTAP Select à partir du même cluster sur le même hôte. La figure suivante montre un exemple de déploiement correct de deux clusters à quatre nœuds qui se croisent sur deux hôtes.

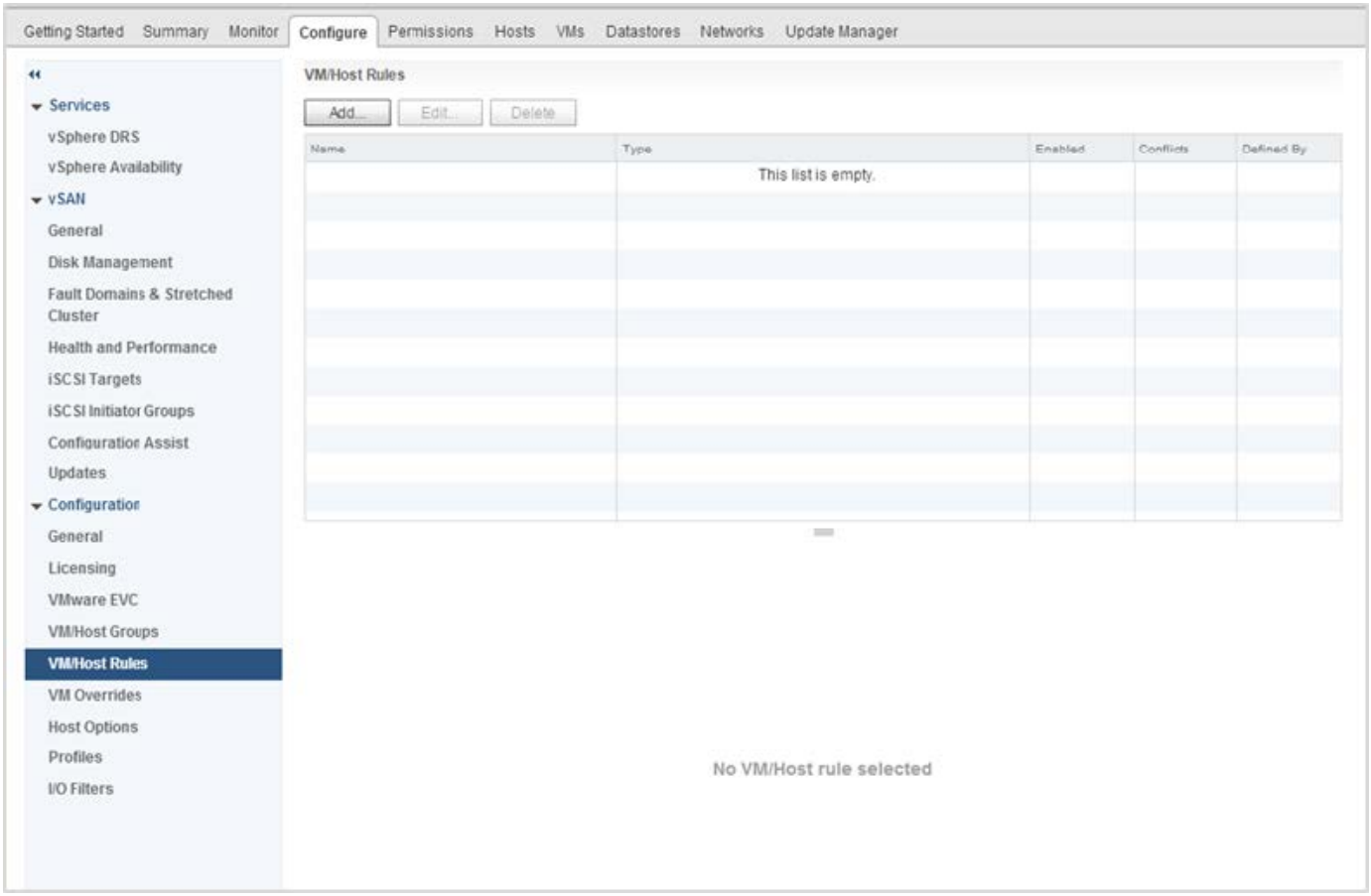
Déploiement initial des clusters VNAS multinœuds

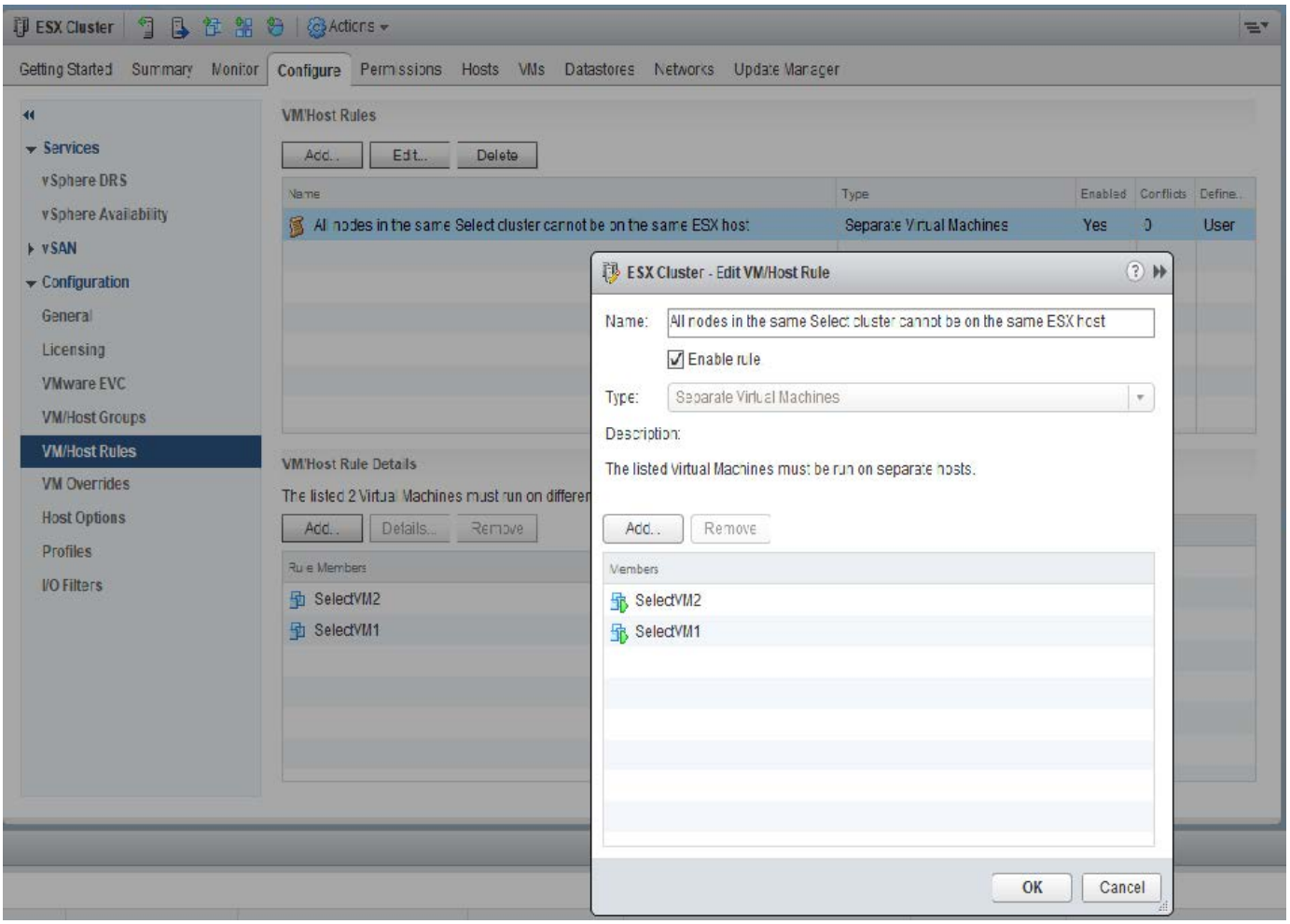


Une fois le déploiement terminé, les nœuds ONTAP Select peuvent être migrés entre des hôtes. Cela pourrait entraîner des configurations non optimales et non prises en charge pour lesquelles deux nœuds ONTAP Select ou plus du même cluster partagent le même hôte sous-jacent. NetApp recommande la création manuelle de règles d'anti-affinité des VM afin que VMware maintienne automatiquement la séparation physique entre les nœuds du même cluster, et pas seulement les nœuds de la même paire haute disponibilité.

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

Consultez l'exemple suivant sur la manière de créer une règle d'anti-affinité pour les machines virtuelles ONTAP Select. Si le cluster ONTAP Select contient plusieurs paires haute disponibilité, tous les nœuds du cluster doivent être inclus dans cette règle.


Getting Started
Summary
Monitor
Configure
Permissions
Hosts
VMs
Datastores
Networks
Update Manager
Services
vSphere DRS
vSphere Availability
vSAN
General
Disk Management
Fault Domains & Stretched Cluster
Health and Performance
iSCSI Targets
iSCSI Initiator Groups
Configuration Assist
Updates
Configuration
General
Licensing
VMware EVC
VM/Host Groups
VM/Host Rules
VM Overrides
Host Options
Profiles
I/O Filters
VM/Host Rules
Add...
Edit...
Delete
Name
Type
Enabled
Conflicts
Defined By
This list is empty.
No VM/Host rule selected



Deux nœuds ONTAP Select ou plus depuis le même cluster ONTAP Select peuvent être trouvés sur le même hôte ESX, pour l'une des raisons suivantes :

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

Notez que ONTAP Deploy ne surveille pas de manière proactive l'emplacement des machines virtuelles ONTAP Select. Toutefois, une opération de mise à jour du cluster reflète cette configuration non prise en charge dans les journaux de déploiement ONTAP :

Augmenter la capacité de stockage

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

La fonctionnalité d'ajout de stockage de ONTAP Deploy est la seule façon d'augmenter la gestion du stockage, et la modification directe de la machine virtuelle ONTAP Select n'est pas prise en charge. La figure suivante montre l'icône "+" qui lance l'assistant d'ajout de stockage.

Les considérations suivantes sont importantes pour la réussite de l'opération d'extension de capacité. L'ajout de capacité requiert la licence existante pour couvrir la quantité totale d'espace (existant plus nouveau). Une opération d'ajout de stockage entraînant une défaillance du nœud dépassant sa capacité sous licence. Une nouvelle licence ayant une capacité suffisante doit être installée en premier.

Si la capacité supplémentaire est ajoutée à un agrégat ONTAP Select existant, le nouveau pool de stockage (datastore) doit présenter un profil de performances similaire à celui du pool de stockage existant. Notez qu'il n'est pas possible d'ajouter un stockage non-SSD à un nœud ONTAP Select équipé d'une personnalisation semblable à l'AFF (Flash activé). Le mélange de DAS et de stockage externe n'est pas également pris en charge.

Si du stockage connecté localement est ajouté à un système afin de fournir des pools de stockage locaux (DAS) supplémentaires, vous devez créer un groupe RAID et une LUN (ou LUN) supplémentaires. Tout comme pour les systèmes FAS, il convient de s'assurer que les performances du nouveau groupe RAID sont similaires à celles du groupe RAID d'origine si vous ajoutez de l'espace au même agrégat. Si vous créez un nouvel agrégat, la nouvelle disposition des groupes RAID peut être différente si l'on comprend parfaitement les implications de performances du nouvel agrégat.

Le nouvel espace peut être ajouté au même datastore, dans la mesure où la taille totale du datastore ne dépasse pas la taille maximale du datastore prise en charge par ESX. L'ajout d'une extension au datastore dans lequel ONTAP Select est déjà installé peut s'effectuer de façon dynamique et n'affecte pas les opérations du nœud ONTAP Select.

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

Dans une paire haute disponibilité, chaque nœud contient une copie en miroir des données de son partenaire. L'ajout d'espace au nœud 1 requiert qu'une quantité d'espace identique soit ajoutée à son partenaire, le nœud 2, de sorte que toutes les données du nœud 1 soient répliquées vers le nœud 2. En d'autres termes, l'espace ajouté au nœud 2 dans le cadre de l'opération d'ajout de capacité pour le nœud 1 n'est pas visible ou accessible sur le nœud 2. L'espace est ajouté au nœud 2 afin que les données du nœud 1 soient entièrement protégées lors d'un événement HA.

Il y a une considération supplémentaire en ce qui concerne la performance. Les données du nœud 1 sont répliquées de manière synchrone sur le nœud 2. Par conséquent, les performances du nouvel espace (datastore) sur le nœud 1 doivent correspondre aux performances du nouvel espace (datastore) sur le nœud 2. En d'autres termes, l'ajout d'espace sur les deux nœuds, mais l'utilisation de technologies de disque différentes ou de tailles de groupe RAID différentes, peut entraîner des problèmes de performances. Cela est dû à l'opération RAID SyncMirror utilisée pour conserver une copie des données sur le nœud partenaire.

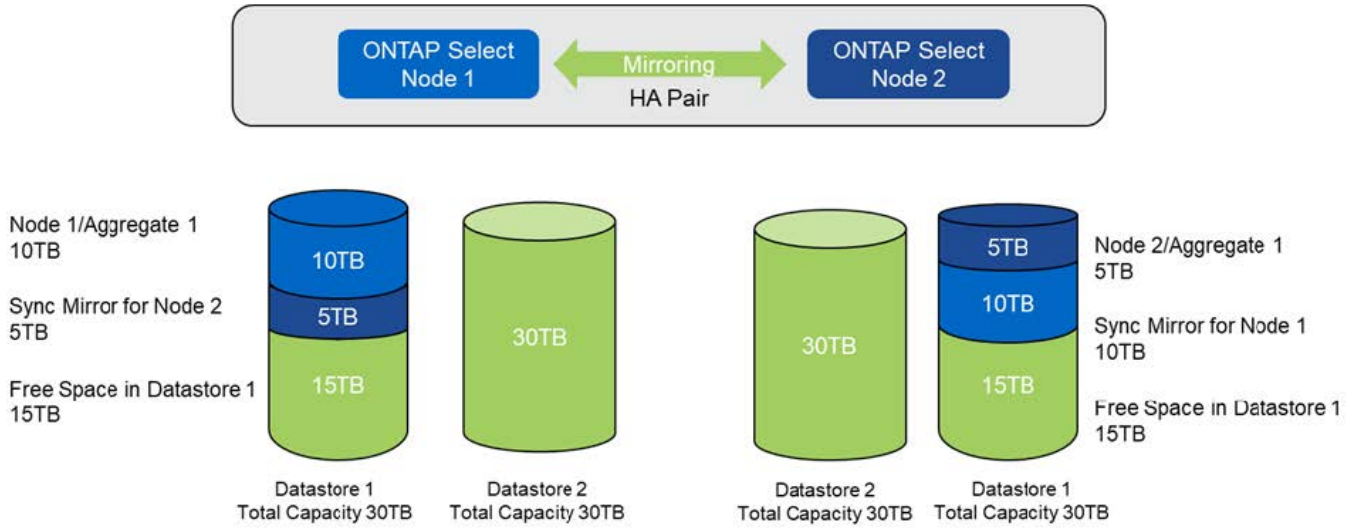
Pour augmenter la capacité accessible par l'utilisateur sur les deux nœuds d'une paire haute disponibilité, deux opérations d'ajout de stockage doivent être effectuées, une pour chaque nœud. Chaque opération d'ajout

de stockage requiert 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 et à l'espace requis sur le nœud 2.

La configuration initiale est effectuée avec deux nœuds. Chaque nœud dispose de deux datastores avec 30 To d'espace dans chaque datastore. ONTAP Deploy crée un cluster à deux nœuds dont chaque nœud consomme 10 To d'espace dans le datastore 1. ONTAP Deploy configure chaque nœud avec 5 To d'espace actif par nœud.

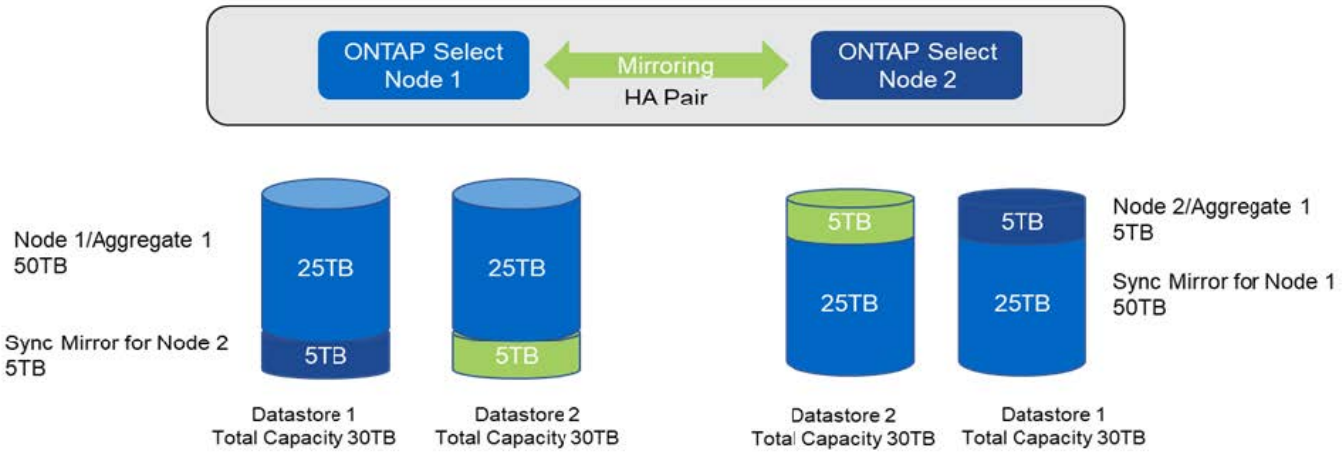
La figure suivante présente les résultats d'une opération d'ajout de stockage pour le nœud 1. ONTAP Select utilise toujours la même quantité de stockage (15 To) sur chaque nœud. Cependant, le nœud 1 dispose d'un stockage plus actif (10 To) que le nœud 2 (5 To). Les deux nœuds sont entièrement protégés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

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



Deux opérations d'ajout de stockage supplémentaires sur le nœud 1 utilisent le reste du datastore 1 et une partie du datastore 2 (avec le cache de capacité). La première opération d'ajout de stockage utilise 15 To d'espace libre dans le datastore 1. La figure suivante présente 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 en contient les 5 To d'origine.

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



La taille maximale de VMDK utilisée lors des opérations d'ajout de capacité est de 16 To. La taille VMDK maximale utilisée lors des opérations de création de clusters continue d'être de 8 To. Le déploiement de ONTAP crée des VMDK correctement dimensionnés en fonction de votre configuration (cluster à un ou plusieurs nœuds) et de la capacité ajoutée. Cependant, la taille maximale de chaque VMDK ne doit pas dépasser 8 To lors des opérations de création du cluster et 16 To lors des opérations d'ajout de stockage.

Augmentez la capacité du système ONTAP Select avec le RAID logiciel

De la même manière, l'assistant d'ajout de stockage permet d'augmenter la capacité de gestion des nœuds ONTAP Select grâce au RAID logiciel. L'assistant ne présente que les disques SSD DAS disponibles et peut être mappé en tant que RDM à la machine virtuelle ONTAP Select.

Bien qu'il soit possible d'augmenter la licence de capacité d'un seul To, lorsque l'on travaille avec le RAID logiciel, il n'est pas possible d'augmenter physiquement la capacité de un seul To. Tout comme l'ajout de disques à une baie FAS ou AFF, certains facteurs déterminent la quantité minimale de stockage que vous pouvez ajouter en une seule opération.

Notez que dans une paire haute disponibilité, l'ajout de stockage au nœud 1 requiert 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 opération d'ajout de stockage sur le nœud 1. C'est-à-dire que les disques distants sont utilisés pour s'assurer que le nouveau stockage du nœud 1 est répliqué et protégé sur le nœud 2. Pour ajouter du stockage utilisable localement sur le nœud 2, une opération d'ajout de stockage distincte et un nombre de disques distinct et égal doivent être disponibles sur les deux nœuds.

ONTAP Select partitionne les nouveaux disques dans les mêmes partitions racine, de données et de données que les disques existants. L'opération de partitionnement se déroule pendant la création d'un nouvel agrégat ou pendant l'extension sur un agrégat existant. La taille de la bande de partition racine sur chaque disque est définie pour correspondre à la taille de partition racine existante sur les disques existants. Par conséquent, chacune des deux tailles de partition de données identiques peut être calculée comme la capacité totale du disque moins la taille de la partition racine divisée par deux. La taille de bande de la partition racine est variable et est calculée au cours de la configuration initiale du cluster comme suit. L'espace racine total requis (68 Go pour un cluster à un seul nœud et 136 Go pour les paires HA) est divisé en différents disques moins de disques de secours et de parité. La taille de bande de la partition racine est maintenue constante sur tous les lecteurs ajoutés au système.

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

Si vous ajoutez du stockage à un agrégat existant, certaines considérations supplémentaires sont nécessaires. Il est possible d'ajouter des disques à un groupe RAID existant, en supposant que le groupe RAID n'est pas à

la limite maximale déjà. Les meilleures pratiques traditionnelles FAS et AFF pour l'ajout de piles de disques aux groupes RAID existants s'appliquent également ici, et la création d'un point fort sur la nouvelle pile de disques peut être un problème. En outre, seuls les disques de taille égale ou supérieure des partitions de données 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 lecteur. Si les partitions de données ajoutées sont supérieures aux partitions existantes, les nouveaux disques sont de bonne taille. En d'autres termes, une partie de la capacité de chaque nouveau disque reste inutilisé.

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

Prise en charge de l'efficacité du stockage

ONTAP Select fournit des options d'efficacité du stockage similaires aux options d'efficacité du stockage présentes sur les baies FAS et AFF.

Les déploiements NAS virtuels ONTAP Select (vNAS) avec VSAN 100 % Flash ou les baies Flash génériques doivent suivre les bonnes pratiques pour ONTAP Select avec stockage DAS non SSD.

La personnalisation du système AFF est automatiquement activée sur les nouvelles installations, à condition que vous disposiez d'un stockage DAS avec disques SSD et d'une licence Premium.

Les fonctionnalités de se à la volée suivantes sont automatiquement activées lors de l'installation.

- Détection en ligne de zéro motif
- Déduplication à la volée des volumes
- Déduplication en arrière-plan des volumes
- Compression à la volée évolutive
- Compaction des données à la volée
- Déduplication à la volée dans l'agrégat
- Déduplication en arrière-plan de l'agrégat

Pour vérifier que ONTAP Select a activé toutes les règles d'efficacité du stockage par défaut, exécutez la commande suivante sur un nouveau volume :

```
<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 ONTAP Select à partir de la version 9.6 et ultérieures, vous devez installer ONTAP Select sur un stockage SSD DAS avec une licence Premium. De plus, vous devez cocher la case **Activer l'efficacité du stockage** lors de l'installation initiale du cluster avec ONTAP Deploy. Pour activer la mise à niveau post-ONTAP de type AFF lorsque les conditions précédentes ne sont pas respectées, il est nécessaire de créer manuellement un argument de démarrage et de redémarrer un nœud. Contactez le support technique pour plus de détails.

Configurations d'efficacité du stockage ONTAP Select

Le tableau suivant récapitule les différentes options d'efficacité du stockage disponibles, activées par défaut ou non par défaut mais recommandées, en fonction du type de support et des licences logicielles.

Fonctionnalités ONTAP Select	SSD DAS (Premium ou Premium XL[1])	Disque dur DAS (toutes licences)	VNAS (toutes licences)
Détection nulle en ligne	Oui (par défaut)	Oui, activé par l'utilisateur et par volume	Oui, activé par l'utilisateur et par volume
Déduplication à la volée des volumes	Oui (par défaut)	Non disponible	Non pris en charge
Compression 32 K en ligne (compression secondaire)	Oui, activée par l'utilisateur sur la base d'un volume.	Oui, activé par l'utilisateur et par volume	Non pris en charge
Compression à la volée 8 Ko (compression adaptative)	Oui (par défaut)	Oui, activée par l'utilisateur sur la base d'un volume	Non pris en charge
Compression en arrière-plan	Non pris en charge	Oui, activée par l'utilisateur sur la base d'un volume	Oui, activé par l'utilisateur et par volume

Fonctionnalités ONTAP Select	SSD DAS (Premium ou Premium XL[1])	Disque dur DAS (toutes licences)	VNAS (toutes licences)
Lecteur de compression	Oui.	Oui.	Oui, activé par l'utilisateur et par volume
Compaction des données à la volée	Oui (par défaut)	Oui, activée par l'utilisateur sur la base d'un volume	Non pris en charge
Scanner de compactage	Oui.	Oui.	Non pris en charge
Déduplication à la volée dans l'agrégat	Oui (par défaut)	S/O	Non pris en charge
Déduplication en arrière-plan des volumes	Oui (par défaut)	Oui, activée par l'utilisateur sur la base d'un volume	Oui, activé par l'utilisateur et 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 (grande). Toutefois, la machine virtuelle grande taille n'est prise en charge que pour les configurations DAS utilisant RAID logiciel. Les configurations RAID et vNAS matérielles ne sont pas prises en charge avec la machine virtuelle ONTAP Select grande dans la version 9.6.

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

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

Sur un système mis à niveau vers ONTAP Select 9.6 ou version ultérieure, un nouveau volume créé dans un agrégat existant ou un agrégat nouvellement créé présente le même comportement qu'un volume créé dans un nouveau déploiement. Les volumes existants soumis à la mise à niveau du code ONTAP Select possèdent la plupart des mêmes règles d'efficacité du stockage que ceux d'un nouveau volume, avec quelques variantes :

Scénario 1

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

- Des volumes avec `space guarantee = volume` ne mettez pas la compaction des données à la volée, la déduplication à la volée dans les agrégats ou la déduplication en arrière-plan dans l'agrégat. Ces options peuvent être activées après la mise à niveau.
- Des volumes avec `space guarantee = none` la compression en arrière-plan n'est pas activée. Cette option peut être activée après la mise à niveau.
- La règle d'efficacité du stockage sur les volumes existants est définie sur automatique après la mise à niveau.

Scénario 2

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

- Des volumes avec `space guarantee = volume` ne constatez aucune différence après la mise à niveau.

- Des volumes avec `space guarantee = none` activer la déduplication en arrière-plan de l'agrégat.
- Des volumes avec `storage policy inline-only` définir leur stratégie sur auto.
- Les volumes dotés de règles d'efficacité du stockage définies par l'utilisateur n'ont pas de modification de règle, à l'exception des volumes avec `space guarantee = none`. Ces volumes ont activé la déduplication en arrière-plan de l'agrégat.

Mise en réseau

Mise en réseau : concepts et caractéristiques généraux

Commencez par vous familiariser avec les concepts généraux de mise en réseau qui s'appliquent à l'environnement ONTAP Select. Explorez ensuite les caractéristiques et les options spécifiques disponibles avec les clusters à un ou plusieurs nœuds.

Mise en réseau physique

Le réseau physique prend en charge le déploiement de clusters ONTAP Select principalement en fournissant l'infrastructure de commutation de couche 2 sous-jacente. La configuration associée au réseau physique inclut à la fois l'hôte de l'hyperviseur et l'environnement réseau commuté plus vaste.

Options de carte réseau de l'hôte

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

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

Configuration physique des commutateurs

Vous devez vous assurer que la configuration des switchs physiques prend en charge le déploiement ONTAP Select. Ces commutateurs physiques sont intégrés aux commutateurs virtuels basés sur l'hyperviseur. La configuration exacte choisie dépend de plusieurs facteurs. Plusieurs points sont à prendre en compte :

- Comment allez-vous maintenir une séparation entre les réseaux internes et externes ?
- Allez-vous conserver une séparation entre les réseaux de gestion et de données ?
- Comment seront configurés les VLAN de couche 2 ?

Mise en réseau logique

ONTAP Select utilise deux réseaux logiques différents, séparant le trafic par type. Plus précisément, le trafic peut circuler entre les hôtes dans le cluster, ainsi qu'vers les clients de stockage et d'autres machines à l'extérieur du cluster. Les commutateurs virtuels gérés par les hyperviseurs permettent la prise en charge du réseau logique.

Réseau interne

Dans le cas d'un déploiement de cluster à plusieurs nœuds, les nœuds ONTAP Select individuels communiquent à l'aide d'un réseau « interne » isolé. Ce réseau n'est pas exposé ni disponible en dehors des nœuds du cluster ONTAP Select.

Le réseau interne est uniquement présent avec un cluster à plusieurs nœuds.

Le réseau interne présente les caractéristiques suivantes :

- Permet de traiter le trafic intra-cluster ONTAP, notamment :
  - Cluster
  - Interconnexion haute disponibilité (HA-IC)
  - Miroir synchrone RAID (RSM)
- Réseau à couche unique - 2 basé sur un VLAN
- Les adresses IP statiques sont attribuées par ONTAP Select :
  - IPv4 uniquement
  - DHCP non utilisé
  - Adresse lien-local
- La taille de MTU est de 9000 octets par défaut et peut être ajustée dans une plage de 7500-9000 (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 ordinateurs. Le réseau externe fait partie intégrante de chaque déploiement de cluster et présente les caractéristiques suivantes :

- Permet de traiter le trafic ONTAP, notamment :
  - Données (NFS, CIFS, iSCSI)
  - Gestion (cluster et nœud ; option de SVM)
  - Intercluster (en option)
- Prise en charge facultative 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 de MTU est de 1 1500 octets par défaut (peut être ajustée)

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

Environnement réseau de machines virtuelles

L'hôte de l'hyperviseur fournit plusieurs fonctionnalités réseau.

ONTAP Select se base sur les capacités suivantes exposées à la machine virtuelle :

Ports de machines virtuelles

Plusieurs ports sont disponibles pour l'utilisation par ONTAP Select. Ils sont affectés et utilisés en fonction de
plusieurs facteurs, dont la taille du cluster.

Commutateur virtuel

Le logiciel du commutateur virtuel au sein de l'environnement de l'hyperviseur, que ce soit vSwitch (VMware) ou
Open vSwitch (KVM) relie les ports exposés par la machine virtuelle à l'Ethernet physique
Ports de carte réseau. Vous devez configurer un vSwitch pour chaque hôte ONTAP Select, en fonction de votre
environnement.

Configurations réseau à un ou plusieurs nœuds

ONTAP Select prend en charge les configurations réseau à un ou plusieurs nœuds.

Configuration réseau à un seul nœud

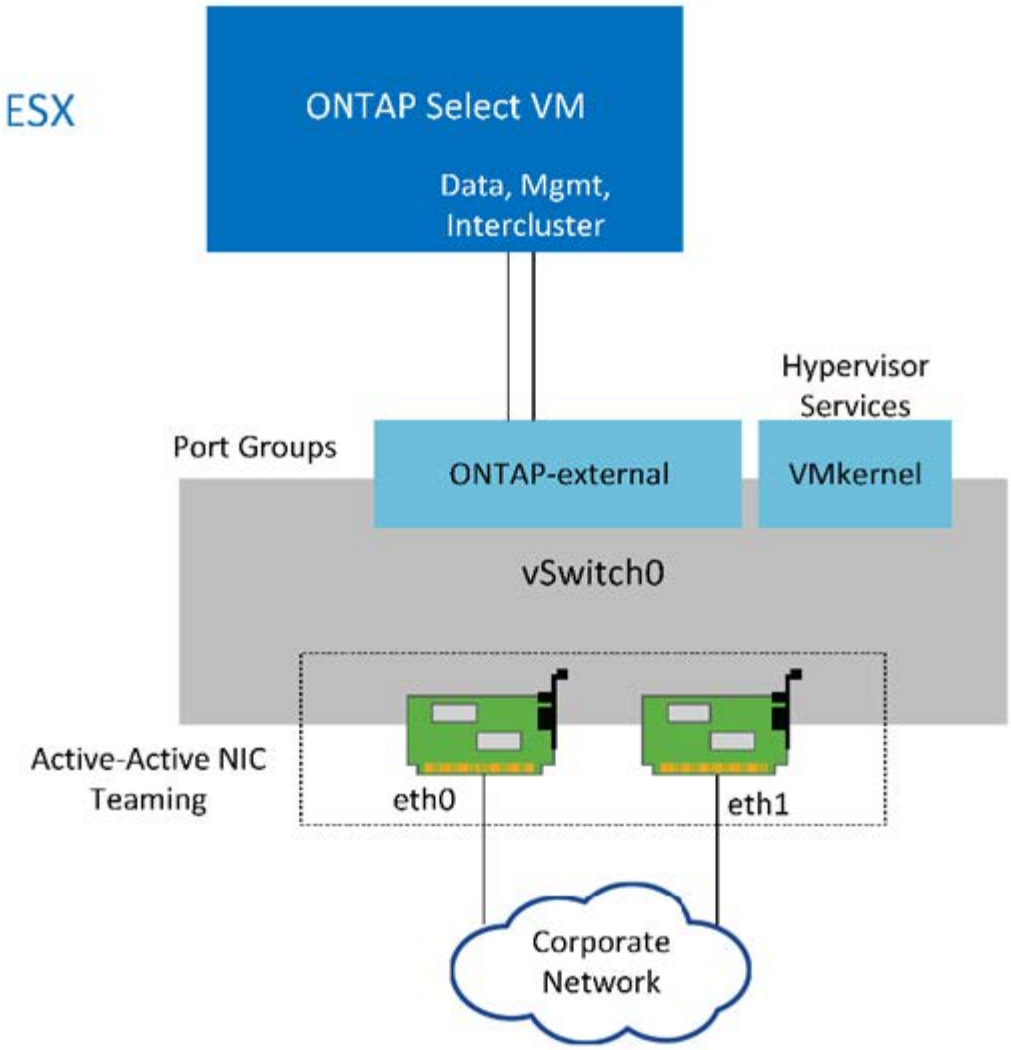
Les configurations ONTAP Select à un seul nœud ne nécessitent pas le réseau interne ONTAP, car il n'y a pas de trafic de cluster, de haute disponibilité ou de miroir.

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

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

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

Configuration réseau du cluster ONTAP Select à un seul nœud



Même si deux cartes sont suffisantes pour un cluster à un seul nœud, le regroupement de cartes réseau est toujours nécessaire.

Assignation de LIF

Comme expliqué dans la section affectation des LIF à plusieurs nœuds de ce document, les IPspaces sont utilisées par ONTAP Select pour séparer le trafic réseau du cluster du trafic de données et de gestion. La variante à un seul nœud de cette plate-forme ne contient pas de réseau de clusters. Par conséquent, aucun port n'est présent dans le cluster IPspace.

Les LIF de Cluster et node management sont automatiquement créées au cours du cluster ONTAP Select. Les LIF restantes peuvent être créées après le déploiement.

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

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

- Trafic de protocole SAN/NAS (CIFS, NFS et iSCSI)
- Trafic de gestion de cluster, de nœud et SVM

- Trafic intercluster (SnapMirror et SnapVault)

Configuration réseau multi-nœuds

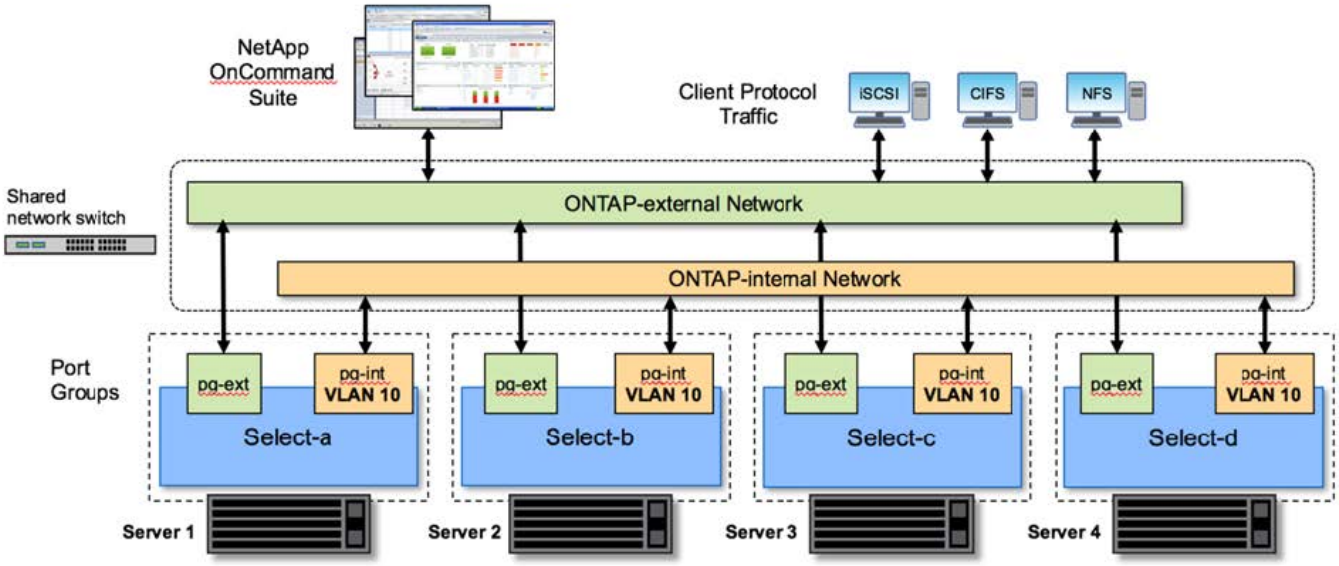
La configuration réseau multinœud ONTAP Select se compose de deux réseaux.

Il s'agit d'un réseau interne qui fournit des services de réplication interne et de cluster, ainsi qu'un réseau externe chargé de fournir des services d'accès et de gestion des données. L'isolation de bout en bout du trafic qui circule au sein de ces deux réseaux est extrêmement importante pour vous permettre de créer un environnement adapté à la résilience du cluster.

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

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

Présentation de la configuration réseau d'un cluster ONTAP Select multinœud



Chaque VM ONTAP Select contient sept adaptateurs réseau virtuels présentés à ONTAP comme un ensemble de sept ports réseau, e0a à e0g. Bien que ONTAP traite ces adaptateurs comme des cartes réseau physiques, ils sont en fait virtuels et associés à un ensemble d'interfaces physiques via une couche réseau virtualisée. Par conséquent, chaque serveur d'hébergement ne nécessite pas six ports réseau physiques.

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

Ces ports sont préconfigurés afin de fournir les services suivants :

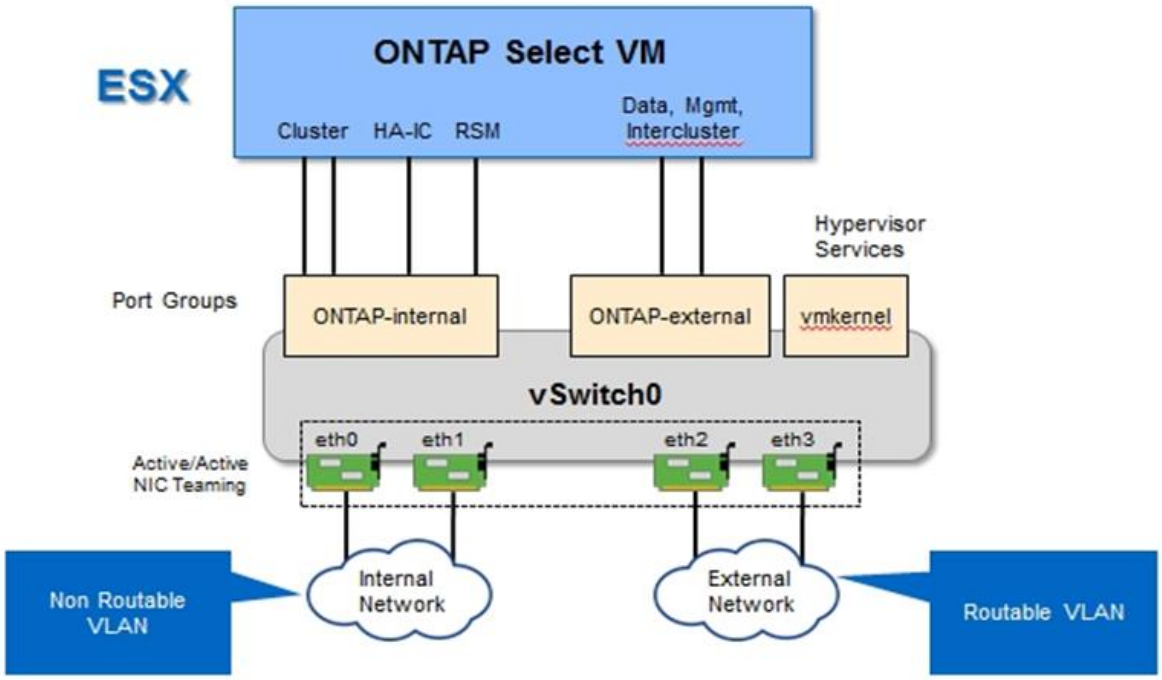
- e0a, e0b et e0g. Gestion et LIF de données
- e0c, e0d. LIF de cluster network

- e0e, RSM
- e0f. Interconnexion HAUTE DISPONIBILITÉ

Les ports e0a, e0b et e0g résident sur le réseau externe. Bien que les ports e0c à e0f exécutent plusieurs fonctions, ils composent collectivement le réseau Select interne. Lors de la prise de décisions de conception de réseau, ces ports doivent être placés sur un réseau de couche 2 unique. Il est inutile de séparer ces adaptateurs virtuels sur différents réseaux.

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

Configuration réseau d'un nœud unique faisant partie d'un cluster ONTAP Select multinœud



La segmentation du trafic interne et externe entre différentes cartes réseau physiques empêche l'introduction de latences dans le système en raison d'un accès insuffisant aux ressources réseau. En outre, l'agrégation via le regroupement de cartes réseau garantit que la défaillance d'une seule carte réseau n'empêche pas le nœud de cluster ONTAP Select d'accéder au réseau correspondant.

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

Assignation de LIF

Les rôles de ports ONTAP ont été obsolètes avec l'introduction des IPspaces. Tout comme les baies FAS, les clusters ONTAP Select contiennent à la fois un IPspace par défaut et un cluster IPspace. En plaçant les ports réseau e0a, e0b et e0g dans l'IPspace par défaut et les ports e0c et e0d dans le cluster IPspace, ces ports ont été essentiellement fermés de l'hébergement des LIF qui n'appartiennent pas. Les ports restants du cluster ONTAP Select sont consommés par le biais de l'assignation automatique d'interfaces fournissant les services internes. Ils ne sont pas exposés via le shell ONTAP, comme c'est le cas avec les interfaces d'interconnexion

RSM et haute disponibilité.

> Toutes les LIFs ne sont pas visibles via le shell de commandes de ONTAP. Les interfaces d'interconnexion haute disponibilité et RSM sont masquées de ONTAP et sont utilisées en interne pour fournir leurs services respectifs.

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

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

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

- Trafic de protocole SAN/NAS (CIFS, NFS et iSCSI)
- Trafic de gestion de cluster, de nœud et SVM
- Trafic intercluster (SnapMirror et SnapVault)

> Les LIF de Cluster et node management sont automatiquement créées au cours du cluster ONTAP Select. Les LIF restantes peuvent être créées après le déploiement.

LIF cluster network (e0c, e0d)

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

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

Le trafic du réseau en cluster doit circuler via un réseau de couche 2 à faible latence non routé. Compte tenu des exigences de débit et de latence du cluster, le cluster ONTAP Select doit se trouver physiquement à proximité (par exemple, plusieurs packs, un data Center unique). Il n'est pas pris en charge de construire des configurations de clusters étendus à quatre, six ou huit nœuds en séparant les nœuds HA sur un WAN ou sur des distances géographiques importantes. Une configuration à deux nœuds étirée avec un médiateur est prise en charge.

Pour plus de détails, reportez-vous à la section "Les bonnes pratiques de la HA étendue à deux nœuds (MetroCluster SDS)".

> Pour garantir un débit maximal pour le trafic réseau en cluster, ce port réseau est configuré de manière à utiliser des trames jumbo (7500 à 9000 MTU). Pour un fonctionnement correct du cluster, vérifiez que les trames Jumbo sont activées sur tous les commutateurs physiques et virtuels en amont et fournissent des services réseau internes aux nœuds de cluster ONTAP Select.

Trafic RAID SyncMirror (e0e)

La réplication synchrone des blocs entre les nœuds partenaires HA se fait à l'aide d'une interface réseau interne qui réside sur le port réseau e0e. Cette fonctionnalité se produit automatiquement, à l'aide d'interfaces réseau configurées par ONTAP lors de la configuration des clusters, et ne nécessite aucune configuration de la part de l'administrateur.

Le port e0e est réservé par ONTAP pour le trafic de réplication interne. Par conséquent, ni le port ni la LIF hébergée ne sont visibles dans l'interface de ligne de commandes de ONTAP ou dans System Manager. Cette interface est configurée pour utiliser une adresse IP locale de liaison générée automatiquement et la réaffectation d'une autre adresse IP n'est pas prise en charge. Ce port réseau nécessite l'utilisation de trames jumbo (7500 à 9000 MTU).

Interconnexion HAUTE DISPONIBILITÉ (e0f)

Les baies NetApp FAS utilisent du matériel spécialisé pour transmettre des informations entre les paires haute disponibilité d'un cluster ONTAP. Cependant, les environnements Software-defined n'ont pas tendance à disposer de ce type d'équipement (comme les appareils InfiniBand ou iWARP), ce qui signifie qu'il faut donc utiliser une autre solution. Bien que plusieurs possibilités aient été envisagées, les exigences ONTAP placées sur le transport d'interconnexion exigent que cette fonctionnalité soit émulée dans le logiciel. Par conséquent, dans un cluster ONTAP Select, la fonctionnalité de l'interconnexion haute disponibilité (généralement fournie par le matériel) a été conçue pour le système d'exploitation, utilisant l'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 assure deux fonctions principales :

- Mise en miroir du contenu de la mémoire NVRAM entre paires haute disponibilité
- Envoi/réception d'informations sur l'état HA et de messages de signal de détection du réseau entre les paires HA

Le trafic d'interconnexion HAUTE DISPONIBILITÉ transite par ce port réseau à l'aide d'une seule interface réseau en superposition de trames RDMA (Remote Direct Memory Access) dans les paquets Ethernet.

De la même manière que le port RSM (e0e), ni le port physique ni l'interface réseau hébergée ne sont visibles par les utilisateurs à partir de l'interface de ligne de commande ONTAP ou de System Manager. Par conséquent, l'adresse IP de cette interface ne peut pas être modifiée et l'état du port ne peut pas être modifié. Ce port réseau nécessite l'utilisation de trames jumbo (7500 à 9000 MTU).

Réseau interne et externe ONTAP Select

Caractéristiques des réseaux internes et externes ONTAP Select.

Réseau interne ONTAP Select

Le réseau ONTAP Select interne, qui n'est présent que dans la variante à plusieurs nœuds du produit, est chargé de fournir au cluster ONTAP Select des services de communication en cluster, d'interconnexion haute disponibilité et de réplication synchrone. Ce réseau comprend les ports et interfaces suivants :

- **E0c, e0d.** Hosting cluster network LIFs
- **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. Une isolation réseau est nécessaire pour assurer la sécurité des clusters et 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 le trafic Select cluster, tel que le trafic d'application ou de gestion, n'est pas prise en charge. Il ne peut y avoir aucun autre VM ou hôte sur le VLAN interne de ONTAP.

Les paquets réseau traversant le réseau interne doivent se trouver sur un réseau VLAN dédié de couche 2. Pour ce faire, vous pouvez effectuer l'une des tâches suivantes :

- Affectation d'un groupe de ports Tagged VLAN aux NIC virtuels internes (e0c à e0f) (mode VST)
- Utilisation du VLAN natif fourni par le commutateur en amont où le VLAN natif n'est pas utilisé pour un autre trafic (affecter un groupe de ports sans ID VLAN, c'est-à-dire le mode est)

Dans tous les cas, le balisage VLAN pour le trafic réseau interne est effectué en dehors du VM ONTAP Select.

Seuls les vswitches ESX standard et distribués sont pris en charge. Les autres commutateurs virtuels ou la connectivité directe entre les hôtes ESX 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.

Dans un cluster ONTAP Select, le trafic interne et le trafic externe sont séparés à l'aide d'objets réseau Virtual Layer-2 appelés port groups. Une affectation appropriée au vSwitch de ces port groups est extrêmement importante, en particulier pour le réseau interne, qui est responsable de la mise en œuvre de services de réplication de cluster, d'interconnexion haute disponibilité et de miroir. Une bande passante réseau insuffisante sur ces ports réseau peut entraîner une dégradation des performances, voire affecter la stabilité du nœud de cluster. Par conséquent, les clusters à quatre, six et huit nœuds nécessitent une connectivité ONTAP Select interne de 10 Go ; les cartes réseau 1 Gbit ne sont pas prises en charge. Toutefois, il est possible de faire des compromis sur le réseau externe, car limiter le flux de données entrantes vers un cluster ONTAP Select n'affecte pas sa capacité à fonctionner en toute fiabilité.

Un cluster à deux nœuds peut utiliser quatre ports 1 Gbit pour le trafic interne ou un seul port 10 Gbit à la place des deux ports 10 Gbit/s requis par le cluster à quatre nœuds. Dans un environnement dans lequel les conditions empêchent le serveur d'utiliser quatre cartes NIC de 10 Gbit/s, deux cartes NIC de 10 Gbit/s peuvent être utilisées pour le réseau interne et deux cartes réseau de 1 Gbit/s peuvent être utilisées pour le réseau ONTAP externe.

Validation et dépannage du réseau interne

Le réseau interne d'un cluster multinœud peut être validé à l'aide de la fonctionnalité de vérification de la connectivité réseau. Cette fonction peut être invoquée à partir de la CLI de déploiement exécutant le `network connectivity-check start` commande.

Lancer la commande suivante pour afficher les valeurs de sortie du test :

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

Cet outil n'est utile que pour le dépannage du réseau interne dans un cluster multinœud Select. Cet outil ne doit pas être utilisé pour résoudre les problèmes de clusters à un seul nœud (y compris les configurations vNAS), de connectivité ONTAP déployée sur ONTAP Select ou de connectivité côté client.

L'assistant de création de cluster (intégré à l'interface utilisateur graphique de déploiement ONTAP) inclut le vérificateur de réseau interne comme étape facultative disponible lors de la création de clusters multinœuds. Étant donné le rôle important que joue le réseau interne dans les clusters multinœuds, cette étape du flux de production cluster create améliore le taux de réussite des opérations cluster create.

Depuis ONTAP Deploy 2.10, la taille de MTU utilisée par le réseau interne peut être définie entre 7,500 et 9,000. Le vérificateur de connectivité réseau peut également être utilisé pour tester la taille de MTU 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 comme VXLAN est présente dans l'environnement.

Réseau externe ONTAP Select

Le réseau externe ONTAP Select est responsable de toutes les communications sortantes effectuées par le cluster. Il est donc présent sur les configurations à nœud unique et multinœud. Bien que ce réseau ne présente pas les exigences de débit strictement définies du réseau interne, l'administrateur doit veiller à ne pas créer de goulets d'étranglement entre le client et la machine virtuelle ONTAP, car des problèmes de performances pourraient se caractériser par un problème de ONTAP Select.

> D'une manière similaire au trafic interne, le trafic externe peut être balisé au niveau de la couche vSwitch (VST) et au niveau de la couche est (commutateur externe). De plus, le trafic externe peut être balisé par la machine virtuelle ONTAP Select elle-même dans un processus appelé VGT. Voir la section "Séparation du trafic de données et de gestion" pour plus d'informations.

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

Référence rapide de réseau interne et externe

Description	Réseau interne	Réseau externe
Services réseau	Cluster HA/IC SyncMirror RAID (RSM)	La gestion des données Intercluster (SnapMirror et SnapVault)
Isolement 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	Auto-générée	Défini par l'utilisateur
Prise en charge de DHCP	Non	Non

Regroupement de cartes réseau

Pour vous assurer que les réseaux internes et externes possèdent à la fois les caractéristiques de bande passante et de résilience requises pour assurer des performances et une tolérance aux pannes élevées, il est recommandé d'utiliser le regroupement des adaptateurs réseau physiques. Les configurations en cluster à 2 nœuds avec une seule liaison 10 Gb sont prises en charge. Cependant, la meilleure pratique recommandée par NetApp consiste à utiliser le regroupement des cartes réseau sur les réseaux internes et externes du cluster ONTAP Select.

Génération d'adresses MAC

Les adresses MAC attribuées à tous les ports réseau ONTAP Select sont générées automatiquement par l'utilitaire de déploiement inclus. La fonction de mise en service utilise un identifiant unique propre à la plateforme propre à l'entreprise (OUI) spécifique à NetApp pour s'assurer qu'il n'y a aucun conflit avec les systèmes FAS. Une copie de cette adresse est ensuite stockée dans une base de données interne au sein de

la VM d'installation de ONTAP Select (ONTAP Deploy), afin d'empêcher une réattribution accidentelle lors de futurs déploiements de nœuds. L'administrateur ne doit en aucun cas modifier l'adresse MAC attribuée à un port réseau.

Configurations réseau prises en charge

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

Les fournisseurs de serveurs savent que les clients ont des besoins différents et qu'ils ont un choix stratégique. Par conséquent, lors de l'achat d'un serveur physique, de nombreuses options sont disponibles pour les décisions de connectivité réseau. La plupart des systèmes de base sont fournis avec différentes options de carte réseau, qui proposent des options à un ou plusieurs ports et présentent différentes permutations de vitesse et de débit. Notamment la prise en charge des adaptateurs réseau 25 Gbit/s et 40 Gbit/s avec VMware ESX.

Les performances des machines virtuelles ONTAP Select étant directement liées aux caractéristiques du matériel sous-jacent, l'augmentation du débit des machines virtuelles en sélectionnant des cartes réseau haut débit permet d'obtenir un cluster plus performant et une meilleure expérience utilisateur globale. Quatre cartes NIC 10 Gbit/s ou deux cartes réseau haut débit (25/40 Gbit/s) peuvent être utilisées pour obtenir une disposition du réseau hautes performances. Un certain nombre d'autres configurations sont également prises en charge. Pour les clusters à deux nœuds, 4 ports 1 Gbit ou 1 port de 10 Go sont pris en charge. Dans le cas de clusters à un seul nœud, 2 ports de 1 Gbit sont pris en charge.

Configurations minimales et recommandées pour le réseau

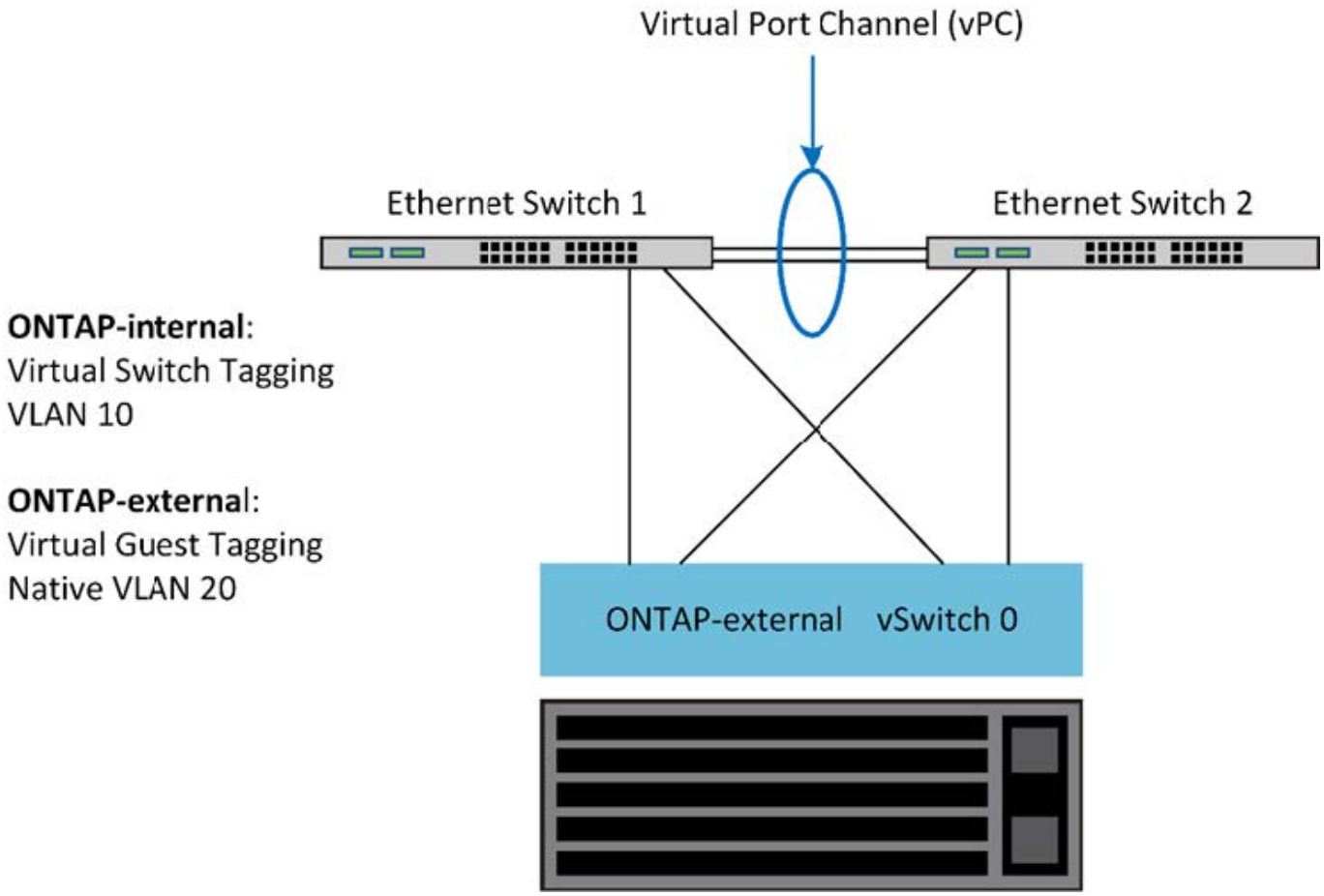
Il existe plusieurs configurations Ethernet prises en charge en fonction de la taille du cluster.

Taille du cluster	Configuration minimale requise	Recommandation
Cluster à un seul nœud	2 x 1 GbE	2 x 10 GbE
Cluster à deux nœuds ou MetroCluster SDS	4 x 1 GbE ou 1 x 10 GbE	2 x 10 GbE
cluster à 4 ou 6/8 nœuds	2 x 10 GbE	4 x 10 GbE ou 2 x 25 GbE

La conversion entre les topologies à liaison unique et à liaisons multiples sur un cluster en cours d'exécution n'est pas prise en charge car il est possible qu'il soit nécessaire de convertir les différentes configurations de regroupement de cartes réseau requises pour chaque topologie.

Configuration du réseau avec plusieurs commutateurs physiques

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



Configuration du vSwitch VMware vSphere

ONTAP Select vSwitch : stratégies de configuration et d'équilibrage de la charge pour les configurations à deux cartes réseau et à quatre cartes NIC.

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 commune utilisée pour agréger la bande passante sur plusieurs adaptateurs physiques. LACP est une norme indépendante du fournisseur qui fournit un protocole ouvert pour les terminaux réseau qui regroupe des groupes de ports réseau physiques en un canal logique unique. ONTAP Select peut fonctionner avec des groupes de ports configurés en tant que groupe d'agrégation de liens (LAG). Toutefois, NetApp recommande d'utiliser les ports physiques individuels comme ports de liaison ascendante simples pour éviter la configuration LAG. Dans ce cas, les meilleures pratiques pour les vswitches standard et distribués sont identiques.

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

Lors de la configuration des groupes de ports à utiliser par ONTAP Select, les meilleures pratiques suivantes doivent être respectées ; la règle d'équilibrage de charge au niveau des groupes de ports est route basée sur l'ID de port virtuel d'origine. VMware recommande de définir le protocole STP sur PortFast sur les ports de commutateur connectés aux hôtes ESXi.

Toutes les configurations vSwitch requièrent un minimum de deux cartes réseau physiques intégrées à une seule association NIC. ONTAP Select prend en charge une liaison 10 Gb unique pour les clusters à deux nœuds. Toutefois, il s'agit d'une meilleure pratique NetApp de garantir la redondance matérielle via

l'agrégation de cartes réseau.

Sur un serveur vSphere, les équipes NIC sont la construction d'agrégation utilisée pour regrouper plusieurs cartes réseau physiques en un seul canal logique, ce qui permet de partager la charge du réseau entre tous les ports membres. Il est important de se rappeler que les équipes NIC 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 NIC, qui ne connaît pas la configuration de commutateur en amont. Dans ce cas, les stratégies sont uniquement appliquées au trafic sortant.

Les canaux de port statiques ne sont pas pris en charge avec ONTAP Select. Les canaux compatibles LACP sont pris en charge par des vswitches distribués, mais l'utilisation de LAG LACP peut entraîner une distribution de charge non uniforme entre les membres LAG.

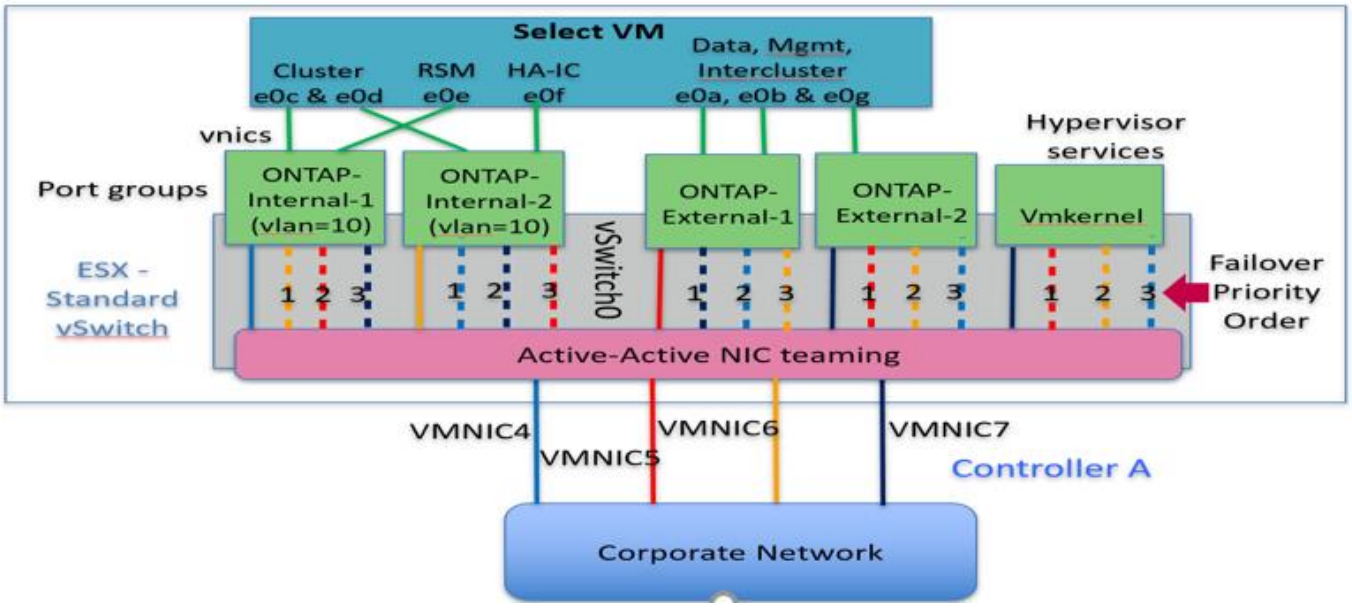
Pour les clusters à un seul nœud, ONTAP déploie la machine virtuelle ONTAP Select de manière à utiliser un groupe de ports pour le réseau externe et le même groupe de ports ou, éventuellement, un groupe de ports différent pour le trafic de gestion du cluster et des nœuds. Pour les clusters à un seul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

Pour les clusters à plusieurs nœuds, ONTAP Deploy configure chaque serveur virtuel ONTAP Select afin qu'il utilise un ou deux groupes de ports pour le réseau interne et, séparément, un ou deux groupes de ports pour le réseau externe. Le trafic de gestion du cluster et des nœuds peut utiliser le même groupe de ports que le trafic externe, ou un autre groupe de ports. Le trafic cluster et node management ne peut pas partager le même port group avec le trafic interne.

VSwitch standard ou distribué et quatre ports physiques par nœud

Quatre groupes de ports peuvent être attribués à chaque nœud d'un cluster multinœud. Chaque port group dispose d'un seul port physique actif et de trois ports physiques de secours, comme dans la figure suivante.

VSwitch avec quatre ports physiques par nœud



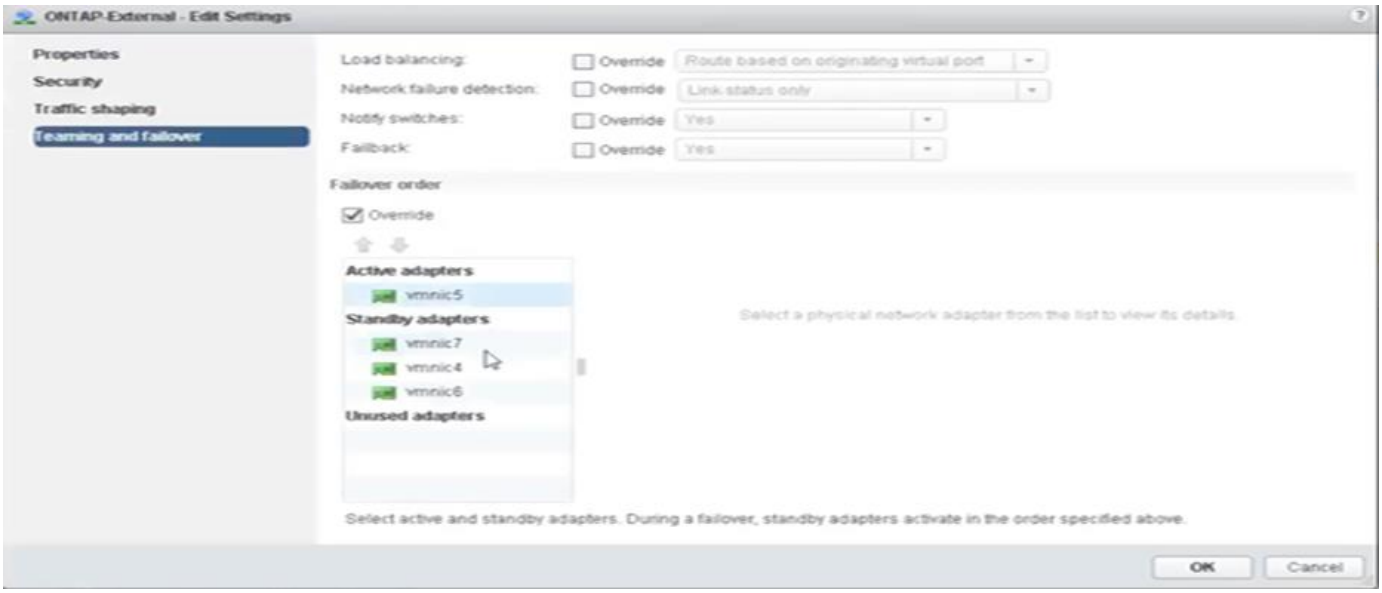
L'ordre des ports dans la liste de secours est important. Le tableau suivant fournit un exemple de distribution physique des ports entre les quatre groupes de ports.

Configurations minimales et recommandées pour le réseau

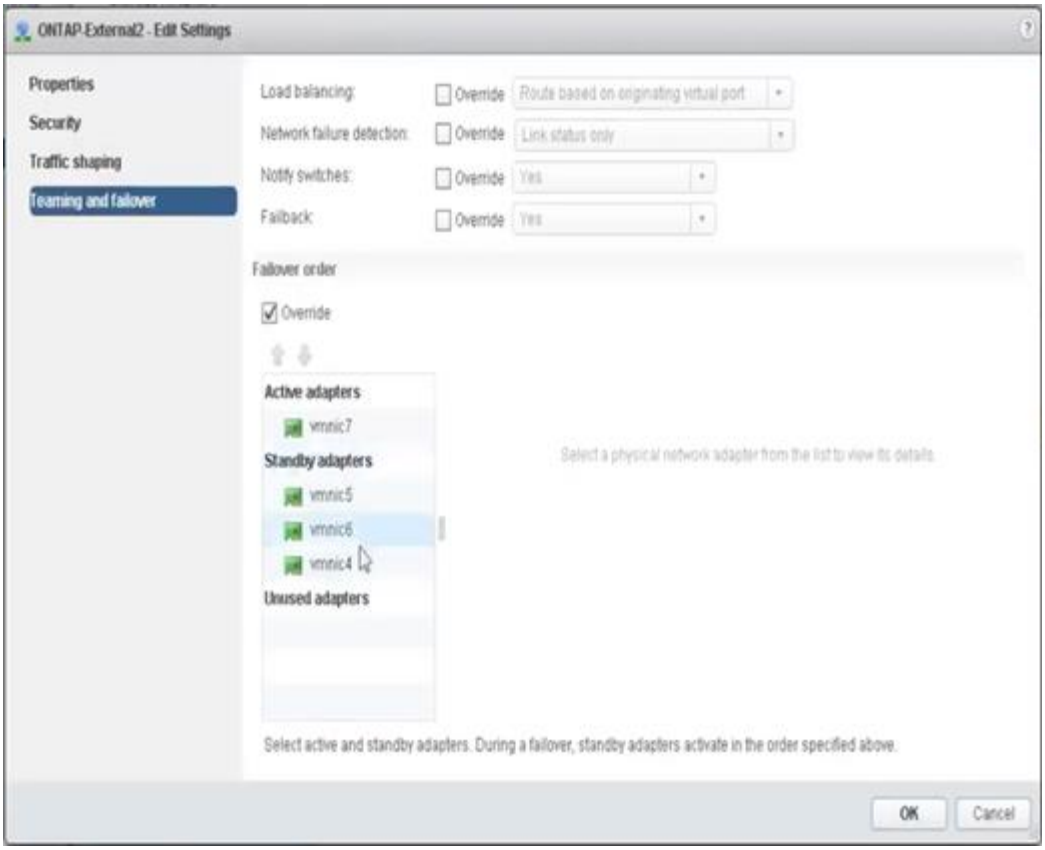
Groupes de ports	Externe 1	Externe 2	Interne 1	Interne 2
Actif	vmnic0	vmnic1	vmnic2	vmnic3
Veille 1	vmnic1	vmnic0	vmnic3	vmnic2
Veille 2	vmnic2	vmnic3	vmnic0	vmnic1
Veille 3	vmnic3	vmnic2	vmnic1	vmnic0

Les figures suivantes illustrent la configuration des groupes de ports réseau externes à partir de l'interface utilisateur graphique vCenter (ONTAP-External et ONTAP-External2). Notez que les cartes actives proviennent de différentes cartes réseau. Dans cette configuration, vmnic 4 et vmnic 5 sont des ports doubles sur la même carte réseau physique, tandis que vmnic 6 et vminc 7 sont de même des ports doubles sur une carte réseau séparée (les vnmacs 0 à 3 ne sont pas utilisés dans cet exemple). L'ordre des cartes de secours fournit un basculement hiérarchique avec les ports du réseau interne en dernier. L'ordre des ports internes dans la liste de secours est également intervertie entre les deux groupes de ports externes.

Partie 1 : configurations de groupes de ports externes ONTAP Select



Partie 2 : configurations de groupe de ports externes ONTAP Select

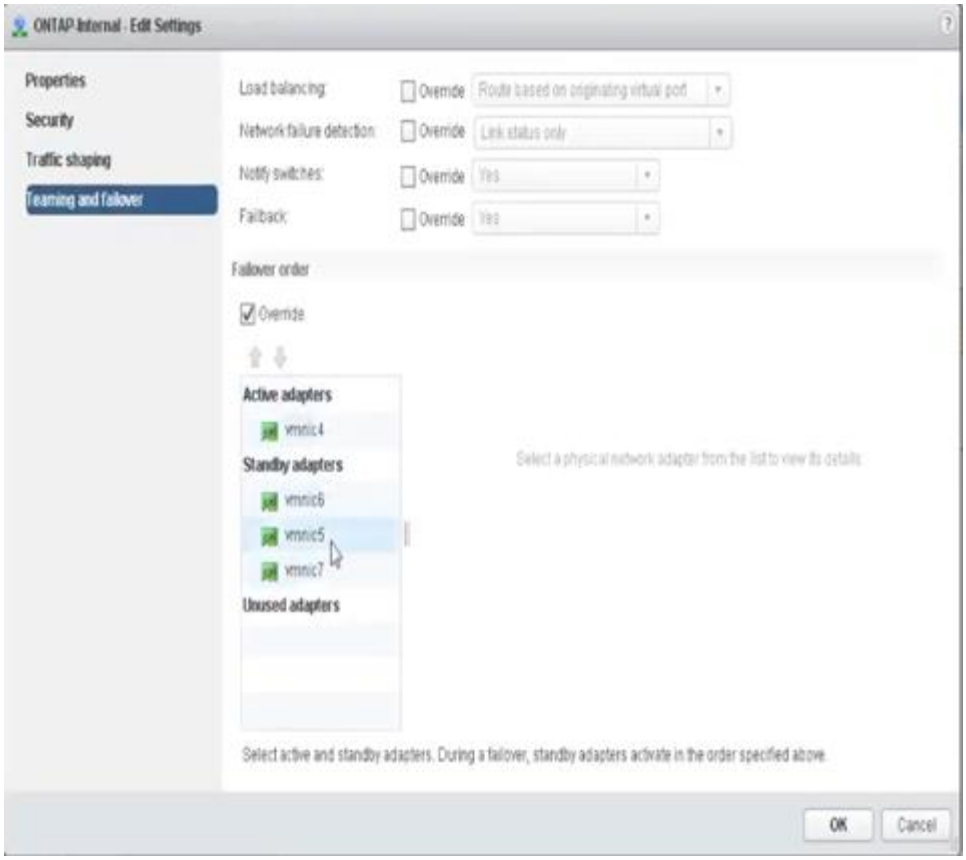


Pour la lisibilité, les affectations sont les suivantes :

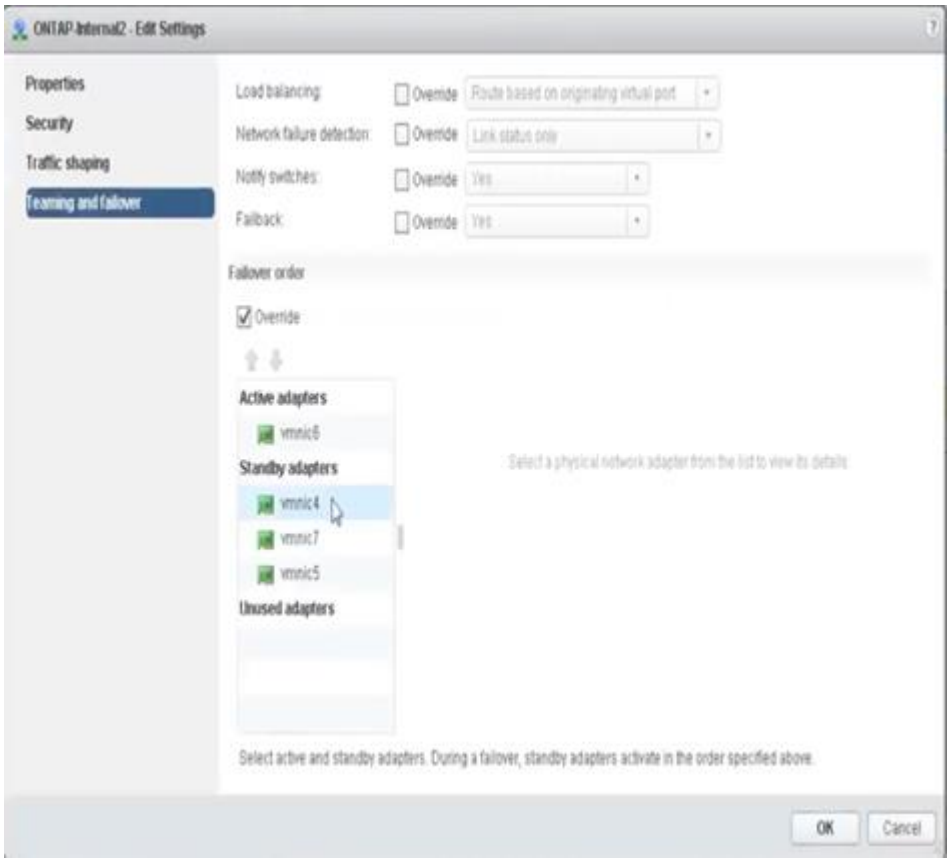
ONTAP-externe	ONTAP-External2
Adaptateurs actifs : vmnic5 Adaptateurs de secours : vmnic7, vmnic4, vmnic6	Adaptateurs actifs : vmnic7 Adaptateurs de secours : vmnic5, vmnic6, vmnic4

Les figures suivantes montrent les configurations des groupes de ports réseau internes (ONTAP-interne et ONTAP-Internal2). Notez que les cartes actives proviennent de différentes cartes réseau. Dans cette installation, vmnic 4 et vmnic 5 sont deux ports sur le même ASIC physique, tandis que vmnic 6 et vmnic 7 sont de même des ports doubles sur un ASIC séparé. L'ordre des cartes de secours fournit un basculement hiérarchique avec les ports du réseau externe en dernier. L'ordre des ports externes dans la liste de secours est également intervertie entre les deux groupes de ports internes.

Partie 1 : configurations de groupes de ports internes ONTAP Select



Partie 2 : groupes de ports internes ONTAP Select



Pour la lisibilité, les affectations sont les suivantes :

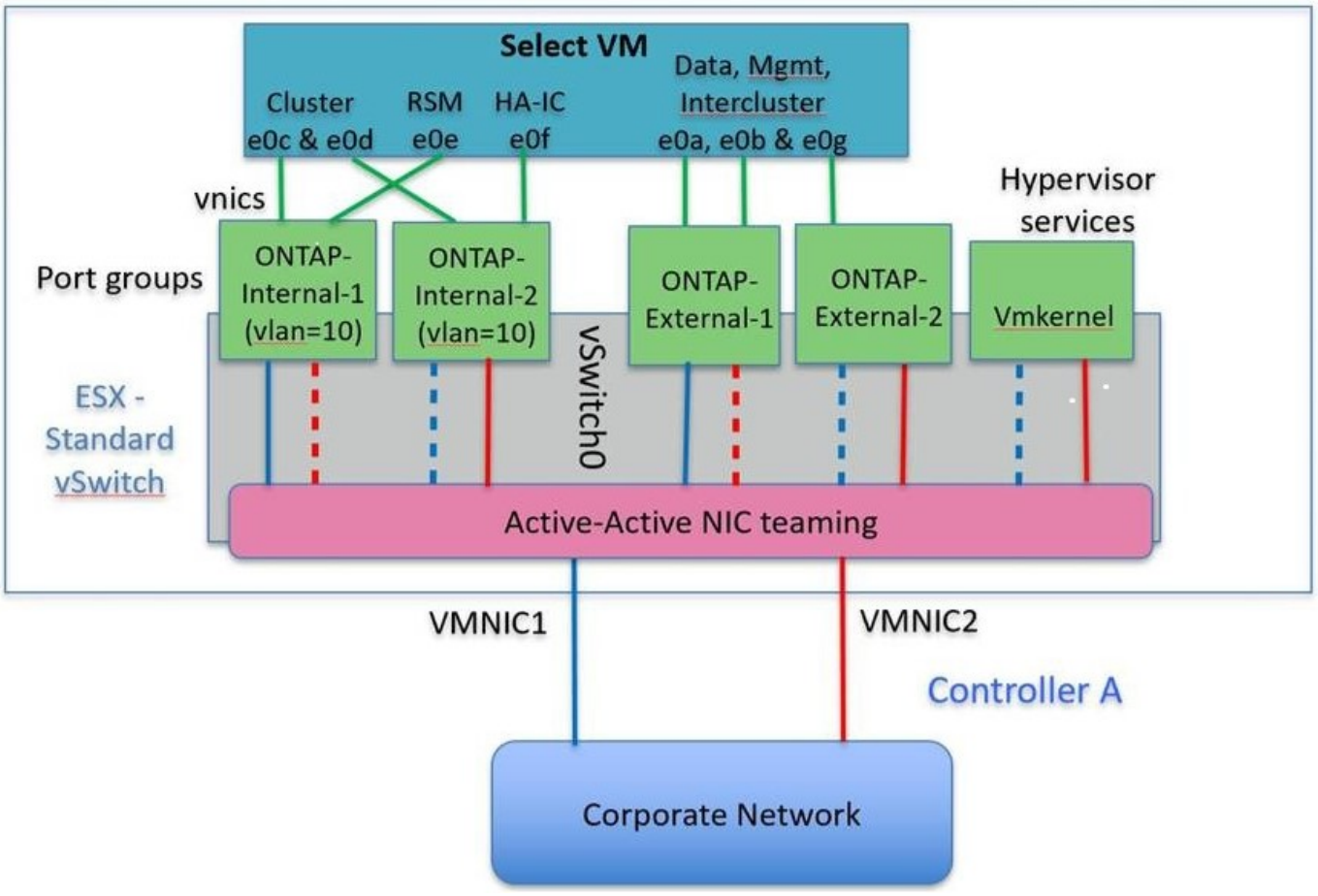
Interne ONTAP	Aéroport international ONTAP 2
Adaptateurs actifs : vmnic4 Adaptateurs de secours : vmnic6, vmnic5, vmnic7	Adaptateurs actifs : vmnic6 Adaptateurs de secours : vmnic4, vmnic7, vmnic5

VSwitch standard ou distribué et deux ports physiques par nœud

Dans le concept de deux cartes NIC à grande vitesse (25 Gb), la configuration de groupes de ports recommandée est très similaire à celle utilisée avec quatre adaptateurs 10 Gb. Quatre groupes de ports doivent être utilisés même avec deux adaptateurs physiques uniquement. Les affectations de groupes de ports sont les suivantes :

Groupes 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 Go) par nœud

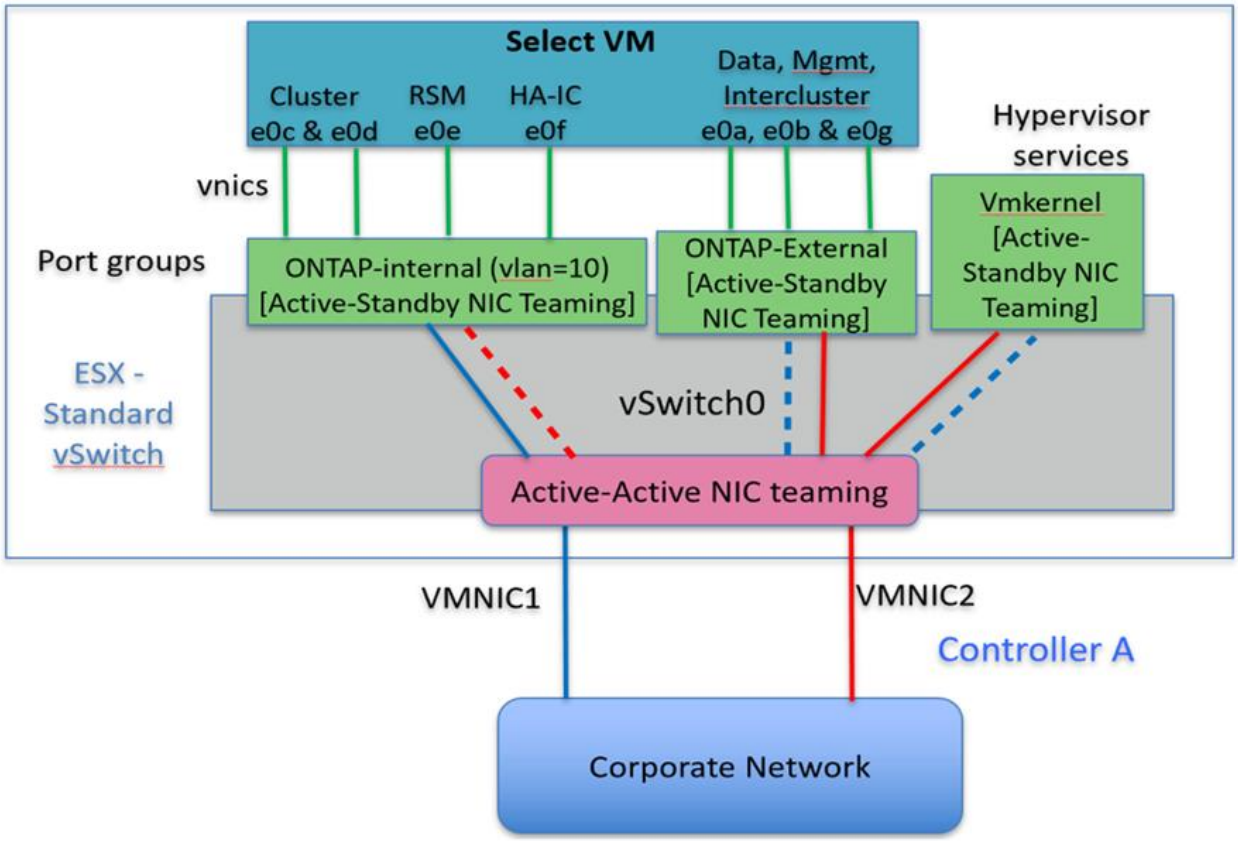


Lorsque vous utilisez deux ports physiques (10 Go ou moins), chaque port group doit disposer d'un adaptateur actif et d'un adaptateur de secours configuré l'un en face de l'autre. Le réseau interne n'est présent que pour les clusters ONTAP Select à plusieurs nœuds. Dans le cas des clusters à un seul nœud, les deux adaptateurs peuvent être configurés comme actifs dans le port group externe.

L'exemple suivant montre la configuration d'un vSwitch et des deux groupes de ports responsables de la

gestion des services de communication internes et externes pour un cluster ONTAP Select multinœud. Le réseau externe peut utiliser le VMNIC interne en cas de panne du réseau, car les vmnics du réseau interne font partie de ce groupe de ports et sont configurés en mode veille. Le contraire est le cas du réseau externe. L'alternance des vmniques actif et en veille entre les deux groupes de ports est cruciale pour assurer un basculement approprié des machines virtuelles ONTAP Select en cas de pannes réseau.

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

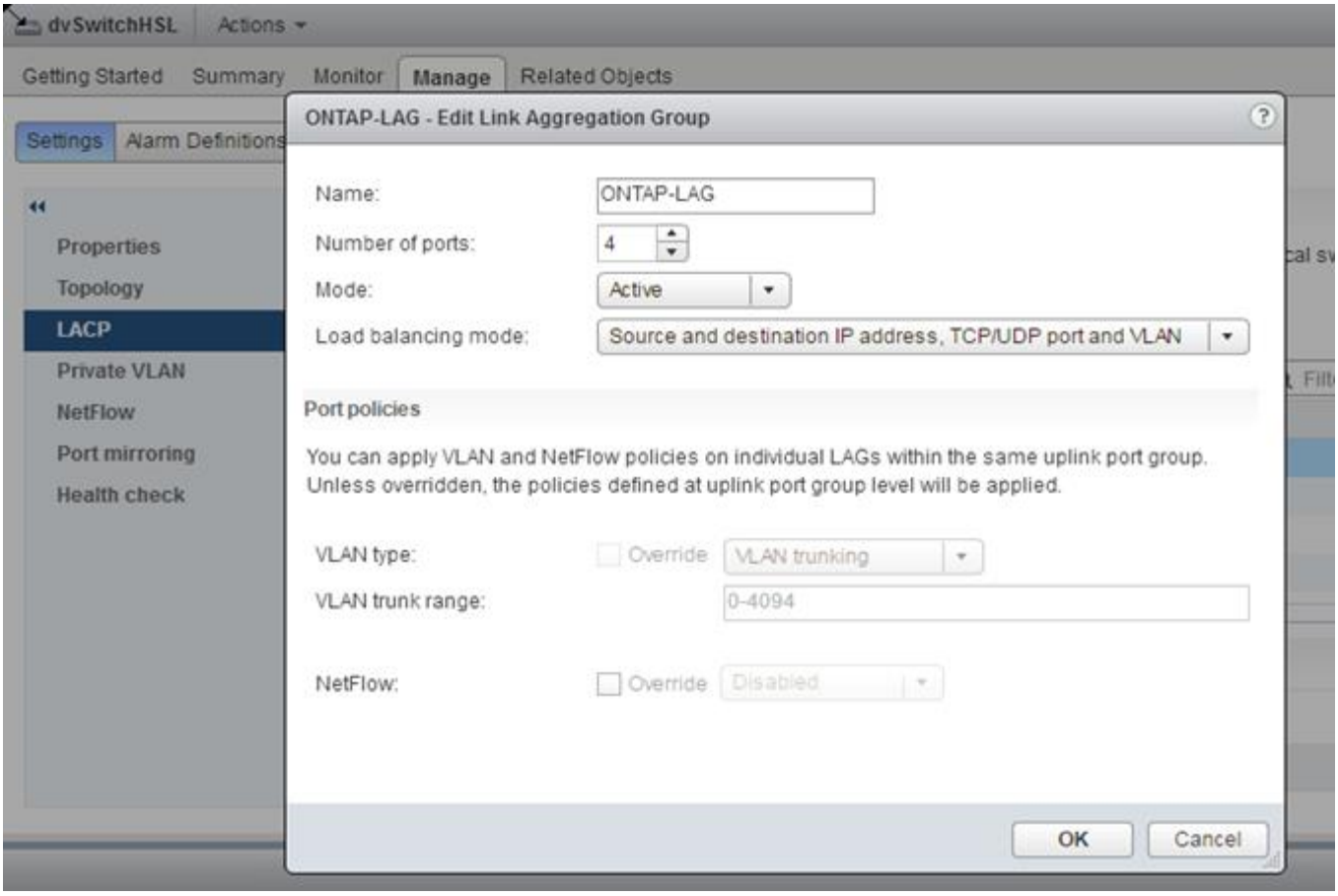


VSwitch distribué avec LACP

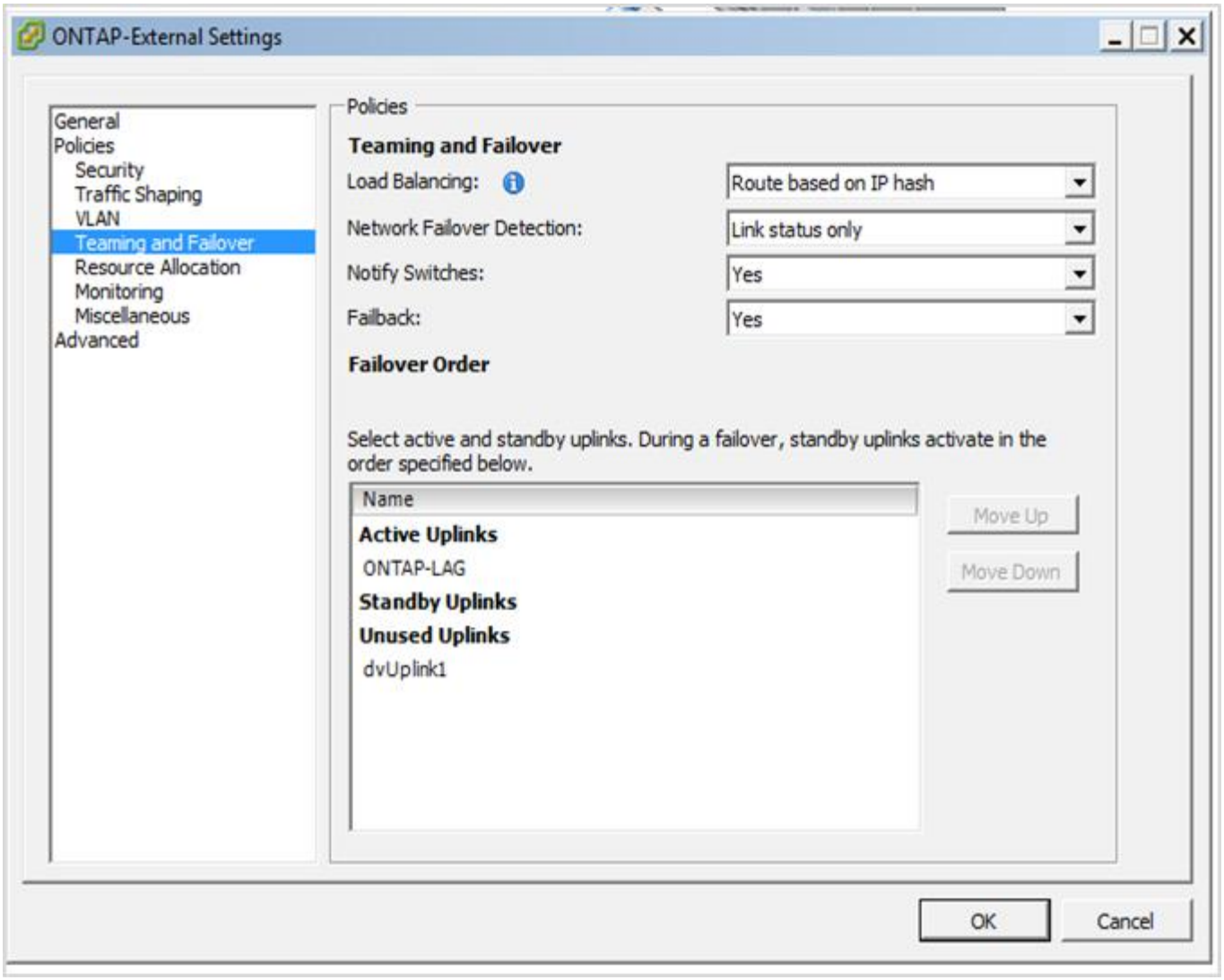
Lorsque vous utilisez des vswitches distribués dans votre configuration, vous pouvez utiliser LACP (bien qu'il ne s'agisse pas d'une meilleure pratique) pour simplifier la configuration du réseau. La seule configuration LACP prise en charge exige que tous les vmnics se trouvent dans un seul LAG. Le commutateur physique uplink doit prendre en charge une taille MTU comprise entre 7,500 et 9,000 sur tous les ports du canal. Les réseaux ONTAP Select internes et externes doivent être isolés au niveau du port group. Le réseau interne doit utiliser un VLAN non routable (isolé). Le réseau externe peut utiliser VST, est ou VGT.

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

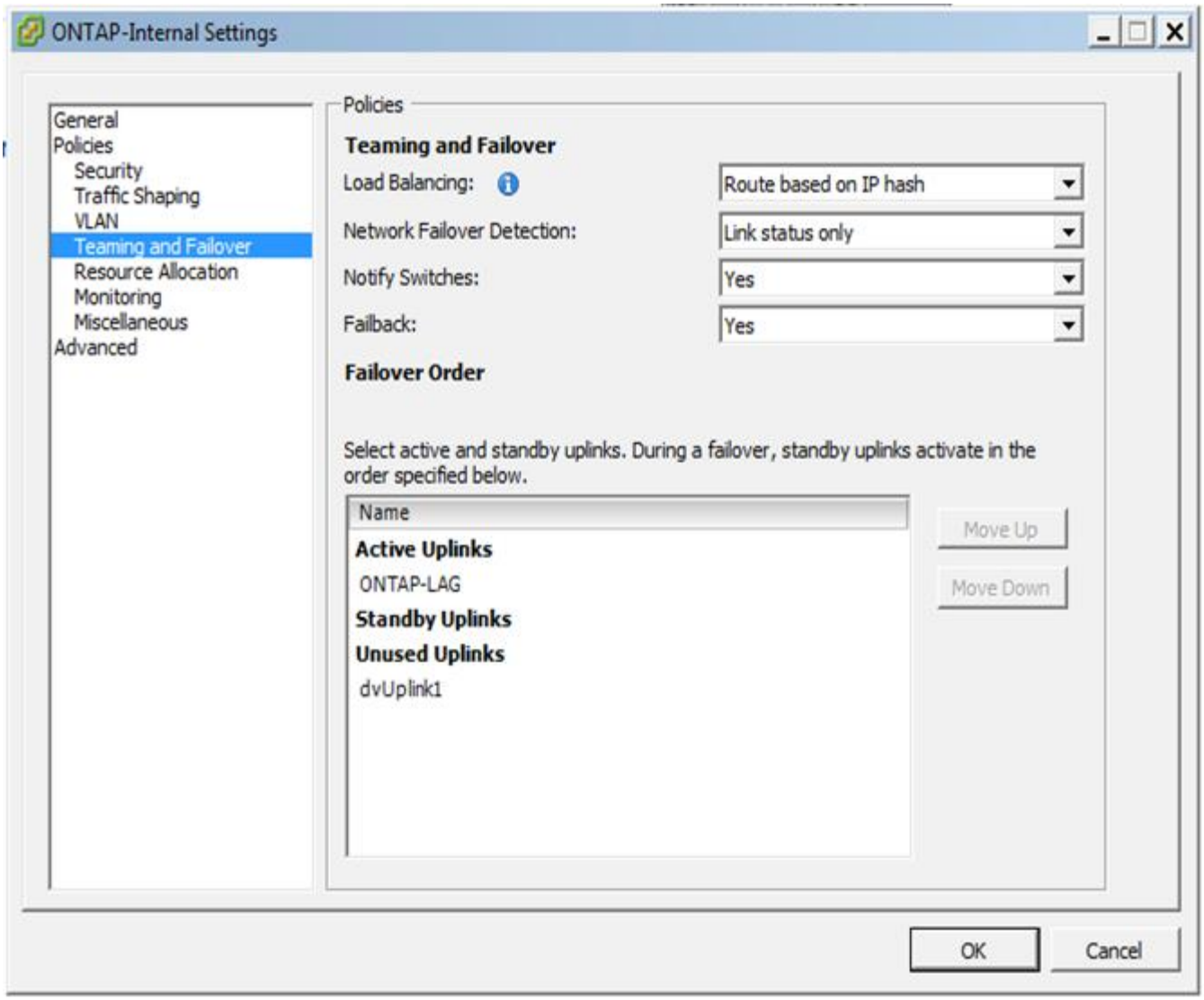
Propriétés LAG lors de l'utilisation de LACP



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



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



Le protocole LACP requiert la configuration des ports de switch amont comme un canal de port. Avant d'activer ceci sur le vSwitch distribué, assurez-vous qu'un canal de port compatible LACP est correctement configuré.

Configuration physique des commutateurs

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

Des considérations rigoureuses doivent être prises en compte lors des décisions de connectivité, de la couche du commutateur virtuel aux commutateurs physiques. La séparation du trafic interne sur le cluster des services de données externes doit s'étendre à la couche réseau physique en amont via l'isolation fournie par les VLAN de couche 2.

Les ports des switchs physiques doivent être configurés en tant que ports trunkports. Le trafic externe ONTAP Select peut être séparé sur plusieurs réseaux de couche 2 de deux manières. L'une des méthodes consiste à utiliser des ports virtuels balisés ONTAP VLAN avec un seul port group. L'autre méthode consiste à attribuer des groupes de ports distincts en mode VST au port de gestion e0a. Vous devez également attribuer des ports de données à e0b et e0c/e0g en fonction de la version du ONTAP Select et de la configuration à un ou plusieurs nœuds. Si le trafic externe est séparé sur plusieurs réseaux de couche 2, les ports de commutation

physique uplink doivent avoir ces VLAN dans sa liste VLAN autorisés.

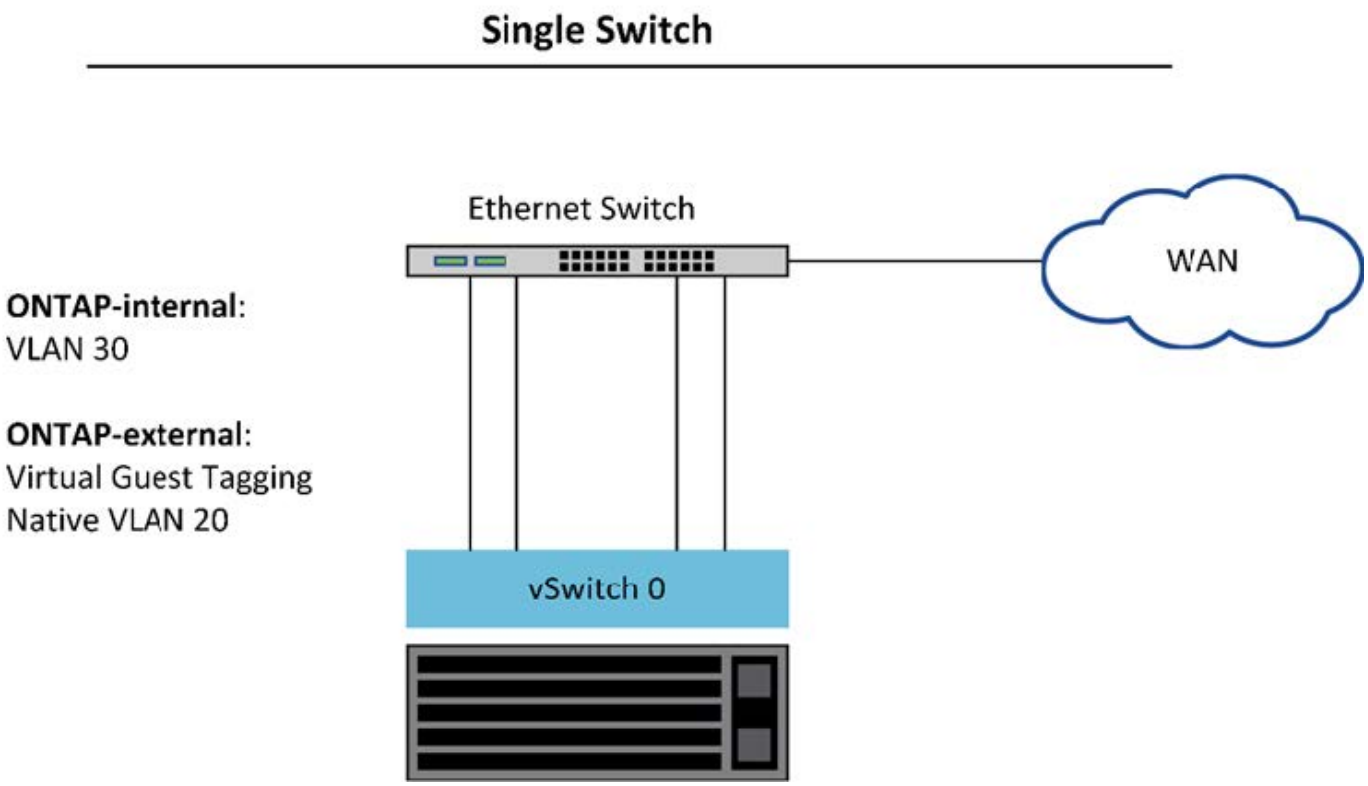
Le trafic réseau interne ONTAP Select se produit à l'aide d'interfaces virtuelles définies avec des adresses IP locales de liaison. Comme ces adresses IP ne sont pas routables, le trafic interne entre les nœuds du cluster doit circuler sur un réseau de couche 2 unique. Les sauts de route entre les nœuds du cluster ONTAP Select ne sont pas pris en charge.

Commutateur physique partagé

La figure suivante décrit une configuration de switch utilisée par un nœud d'un cluster ONTAP Select à plusieurs nœuds. Dans cet exemple, les cartes réseau physiques utilisées par les vswitches hébergeant les groupes de ports réseau interne et externe sont câblées sur le même commutateur en amont. Le trafic des switchs est maintenu isolé à l'aide des domaines de diffusion contenus dans des VLAN séparés.

Pour le réseau interne ONTAP Select, l'étiquetage s'effectue au niveau du port group. Alors que l'exemple suivant utilise VGT pour le réseau externe, VGT et VST sont pris en charge sur ce groupe de ports.

Configuration réseau utilisant un commutateur physique partagé



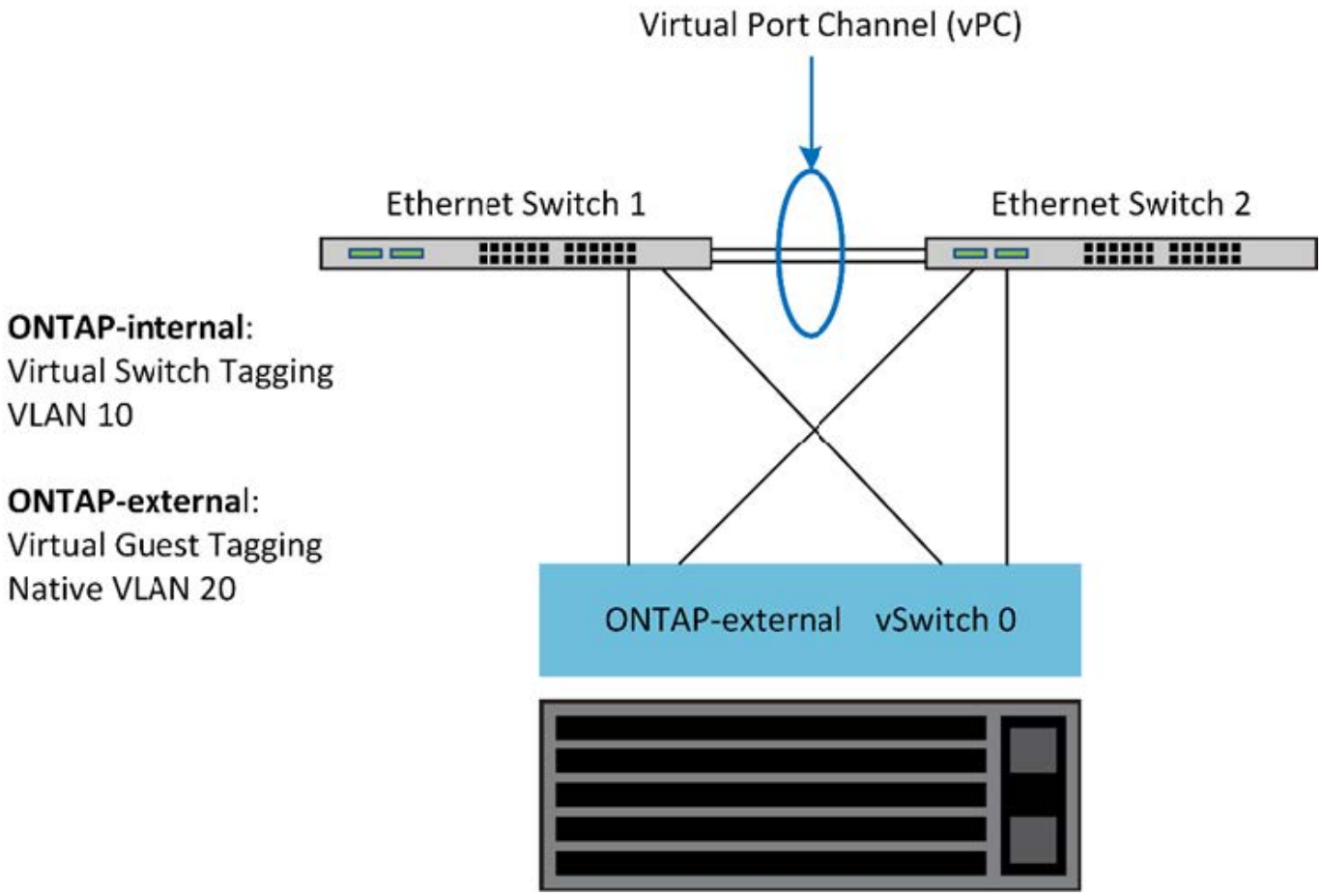
Dans cette configuration, le commutateur partagé devient un point unique de défaillance. Si possible, plusieurs commutateurs doivent être utilisés pour éviter une défaillance matérielle physique susceptible d'entraîner une panne du réseau du cluster.

Commutateurs physiques multiples

Lorsque la redondance est nécessaire, il convient d'utiliser plusieurs commutateurs de réseau physique. La figure suivante présente une configuration recommandée utilisée par un nœud d'un cluster ONTAP Select multinœud. Les cartes réseau des groupes de ports internes et externes sont câblées en différents

commutateurs physiques, ce qui protège l'utilisateur contre une défaillance unique du commutateur matériel. Un canal de port virtuel est configuré entre les commutateurs pour éviter les problèmes liés à l'arborescence à ressources multiples.

Configuration réseau utilisant plusieurs commutateurs physiques



Séparation du trafic de données et de gestion

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

Le trafic réseau externe de ONTAP Select se définit comme le trafic de données (CIFS, NFS et iSCSI), de gestion et de réplication (SnapMirror). Dans un cluster ONTAP, chaque type de trafic utilise une interface logique distincte qui doit être hébergée sur un port réseau virtuel. Sur la configuration à plusieurs nœuds du ONTAP Select, ces ports sont désignés comme e0a et e0b/e0g. Sur la configuration à un seul nœud, ces ports sont désignés sous la forme e0a et e0b/e0c, tandis que les ports restants sont réservés aux services de clusters internes.

NetApp recommande d'isoler le trafic de données et le trafic de gestion dans des réseaux de couche 2 distincts. Dans l'environnement ONTAP Select, cette opération s'effectue à l'aide de balises VLAN. Pour ce faire, il convient d'attribuer un groupe de ports marqué 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), ainsi qu'aux clusters e0b et e0g (clusters à plusieurs nœuds) pour le trafic de données.

Si la solution VST décrite précédemment dans ce document n'est pas suffisante, il peut être nécessaire de colocaliser à la fois les LIF de données et de gestion sur le même port virtuel. Pour ce faire, utilisez un

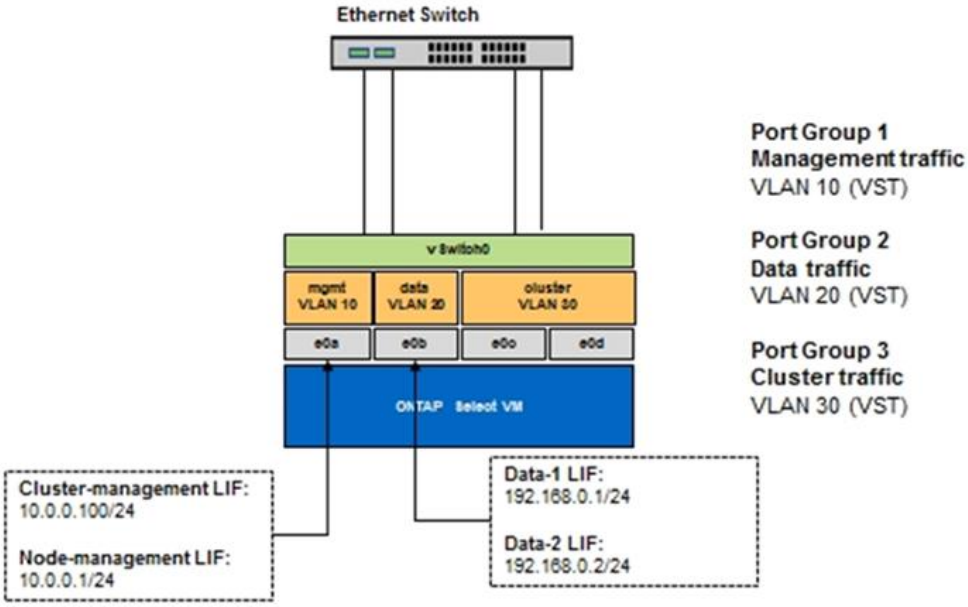
processus appelé VGT, dans lequel le marquage VLAN est effectué par la machine virtuelle.

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

Vous trouverez également un inconvénient lors de l'utilisation des clusters VGT et à deux nœuds. Dans les configurations en cluster à deux nœuds, l'adresse IP de gestion des nœuds permet d'établir la connectivité avec le médiateur avant que ONTAP soit entièrement disponible. Par conséquent, seul le marquage est et VST est pris en charge sur le groupe de ports mappé sur le LIF de gestion de nœud (port e0a). De plus, si la gestion et le trafic de données utilisent le même port group, seul l'outil est/VST est 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 présente le premier scénario VST, dans lequel le trafic est marqué au niveau de la couche vSwitch par le groupe de ports affecté. Dans cette configuration, les LIF de gestion du cluster et des nœuds sont attribuées au port ONTAP e0a et balisés avec l'ID VLAN 10 via le groupe de ports attribué. Les LIF de données sont attribuées au port e0b et au port e0c ou e0g, ainsi qu'à l'ID VLAN 20 donné, en utilisant un second groupe de ports. Les ports de cluster utilisent un troisième groupe de ports et se trouvent sur l'ID VLAN 30.

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



La figure suivante montre le deuxième scénario, VGT, dans lequel le trafic est marqué 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 sur les ports VM e0a et e0b. Cette configuration permet d'exécuter le balisage réseau directement dans ONTAP, plutôt que au niveau de la couche vSwitch. Les LIF de données et de gestion sont placées sur ces ports virtuels, ce qui permet d'ajouter une subdivision de niveau 2 dans un seul port d'ordinateur virtuel. Le VLAN du cluster (ID VLAN 30) est toujours marqué au niveau du port group.

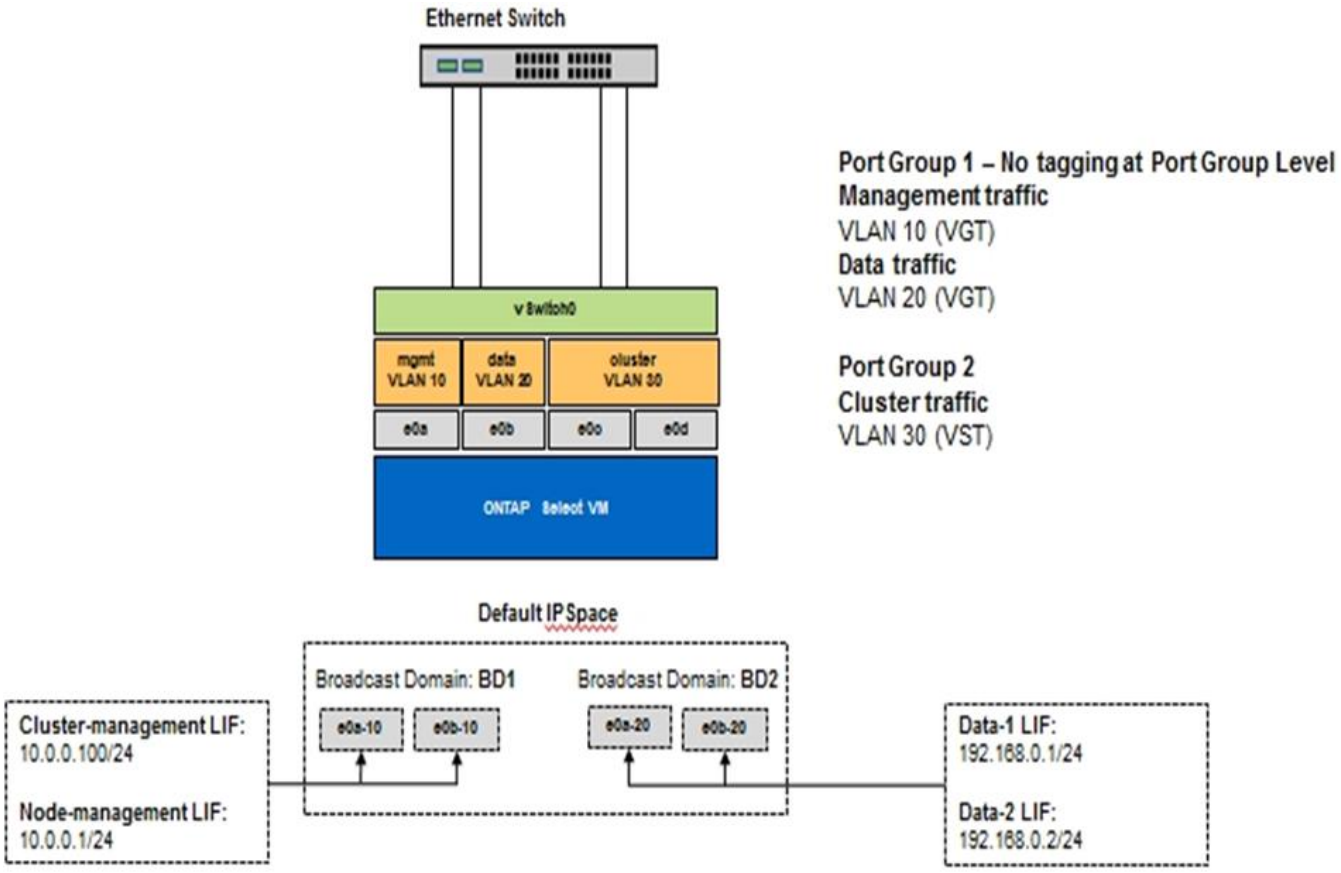
Notes:

- Ce style de configuration est particulièrement souhaitable lors de l'utilisation de plusieurs IPspaces. Regroupez les ports VLAN dans des IPspaces personnalisés distincts pour une isolation logique et une

colocation supplémentaires.

- Pour prendre en charge VGT, les cartes réseau hôte ESXi/ESX doivent être connectées aux ports de jonction sur le commutateur physique. L'ID VLAN des groupes de ports connectés au commutateur virtuel doit être défini sur 4095 pour activer la mise en circuit sur le groupe de ports.

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



Architecture haute disponibilité

Configurations haute disponibilité

Découvrez les options de haute disponibilité pour sélectionner le meilleur configuration de haute disponibilité pour votre environnement.

Même si les clients commencent à migrer leurs charges de travail applicatives entre des appliances de stockage haute performance et des solutions logicielles fonctionnant sur du matériel générique, les attentes et les besoins en matière de résilience et de tolérance aux pannes n'ont pas changé. Une solution haute disponibilité avec un objectif de point de récupération (RPO) zéro protège le client contre les pertes de données dues à une défaillance de composant de la pile d'infrastructure.

Une grande partie du marché du SDS est basée sur la notion de stockage sans partage, avec la réplication logicielle garantissant la résilience des données en stockant plusieurs copies des données utilisateur dans différents silos de stockage. ONTAP Select se développe sur ce site à l'aide des fonctions de réplication synchrone (RAID SyncMirror) fournies par ONTAP pour stocker une copie supplémentaire des données utilisateur au sein du cluster. Cela se produit dans le contexte d'une paire HA. Chaque paire haute disponibilité stocke deux copies des données utilisateur : une sur le stockage fourni par le nœud local et une sur le

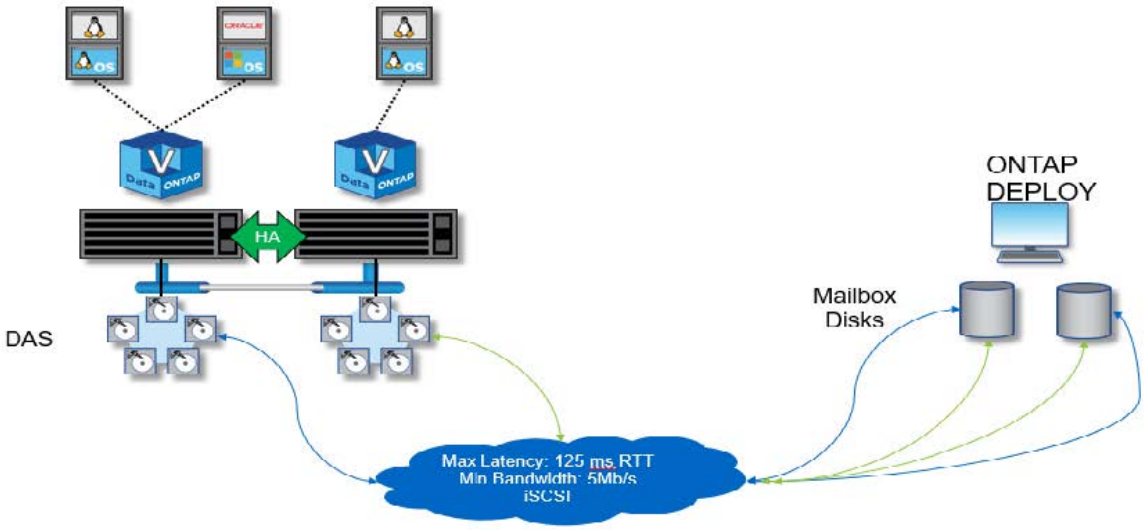
stockage fourni par le partenaire haute disponibilité. Dans un cluster ONTAP Select, la haute disponibilité et la réplication synchrone sont liées entre elles, et les fonctionnalités des deux ne peuvent pas être découplées ou utilisées indépendamment. Donc, la fonctionnalité de réplication synchrone n'est disponible que dans l'offre multinœud.

Dans un cluster ONTAP Select, la fonctionnalité de réplication synchrone fait partie de l'implémentation HA, et non pas à la place 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 haute disponibilité ONTAP Select : les clusters multinœuds (quatre, six ou huit nœuds) et les clusters à deux nœuds. La fonction utile d'un cluster ONTAP Select à deux nœuds est l'utilisation d'un service de médiateur externe afin de résoudre des scénarios split-brain. ONTAP Deploy VM sert de médiateur par défaut dans toutes les paires HA à deux nœuds qu'elle configure.

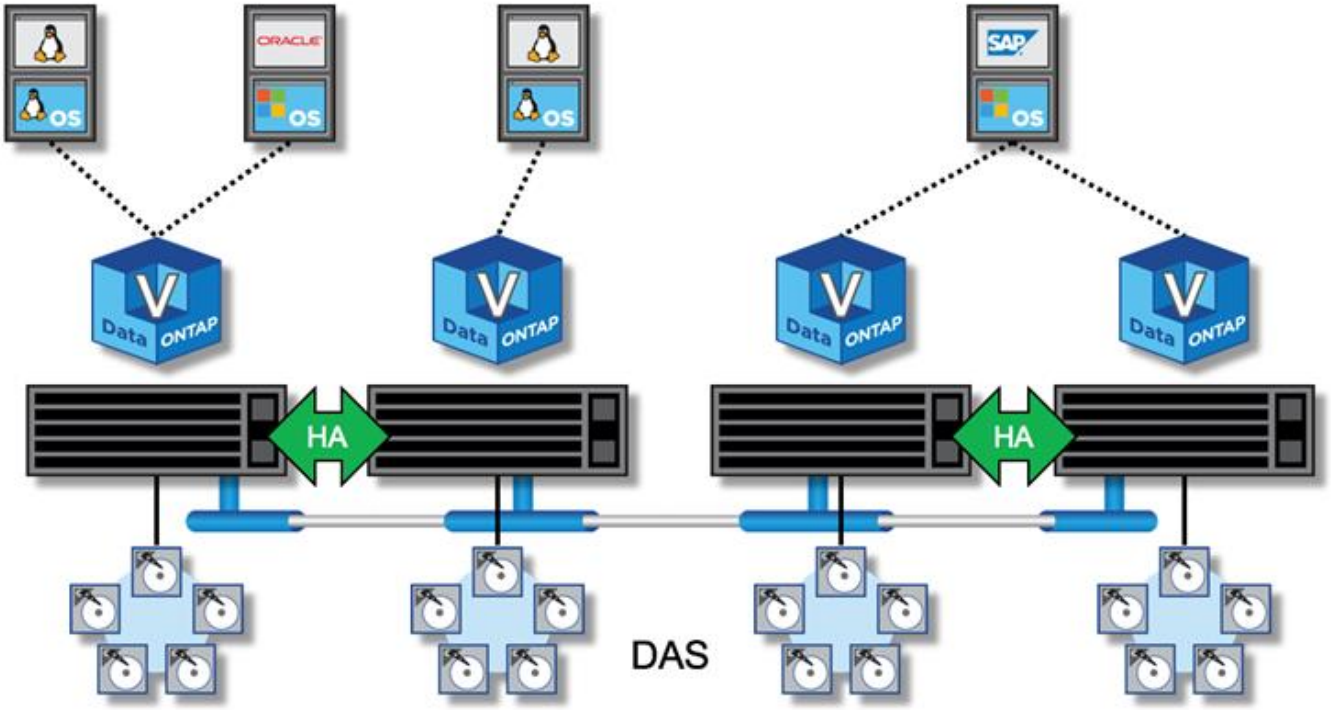
Les deux architectures sont représentées dans les figures suivantes.

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



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

Configuration HA à 2 nœuds par rapport à une configuration HA à plusieurs nœuds

Contrairement aux baies FAS, les nœuds ONTAP Select d'une paire haute disponibilité communiquent exclusivement sur le réseau IP. Cela signifie que le réseau IP est un point de défaillance unique, 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 des défaillances de nœud unique, car le quorum du cluster peut être établi par les trois nœuds survivants ou plus. Le cluster à deux nœuds s'appuie sur le service médiateur hébergé par ONTAP Deploy VM afin d'obtenir les mêmes résultats.

Le trafic réseau Heartbeat entre les nœuds ONTAP Select et le service médiateur ONTAP Deploy est minime et résilient. ONTAP Deploy VM peut être hébergé dans un data Center différent de celui du cluster à deux nœuds ONTAP Select.

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

Si ONTAP déployer une machine virtuelle agissant comme un médiateur est temporairement ou potentiellement indisponible de façon permanente, une machine virtuelle ONTAP Deploy secondaire peut être utilisée pour restaurer le quorum du cluster à deux nœuds. Ceci entraîne 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 de déploiement de ONTAP se fait à l'aide du protocole iSCSI sur IPv4. L'adresse IP de gestion de nœud ONTAP Select est l'initiateur et l'adresse IP de machine virtuelle de déploiement ONTAP 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 des boîtes aux lettres hébergées ONTAP Deploy sont automatiquement créés et masqués pour les adresses IP appropriées de gestion de nœuds ONTAP Select au moment de la création du cluster à deux nœuds. La configuration complète est automatiquement effectuée lors de l'installation et aucune autre action administrative n'est requise. L'instance ONTAP Deploy création du cluster est le médiateur par défaut de ce cluster.

Une action administrative est requise si l'emplacement initial du médiateur doit être modifié. Il est possible de restaurer le quorum d'un cluster même si la machine virtuelle ONTAP Deploy d'origine est perdue. Cependant, NetApp vous recommande de sauvegarder la base de données ONTAP Deploy après que chaque cluster à deux nœuds soit instancié.

Haute disponibilité à 2 nœuds par rapport à une haute disponibilité étendue à 2 nœuds (MetroCluster SDS)

Il est possible de stretch un cluster haute disponibilité actif-actif à deux nœuds sur de plus grandes distances et de placer chaque nœud dans un data Center différent. La seule distinction entre un cluster à deux nœuds et un cluster étendu à deux nœuds (ou MetroCluster SDS) est la distance de connectivité réseau entre les nœuds.

Le cluster à deux nœuds est défini comme un cluster pour lequel les deux nœuds sont situés dans le même data Center, à une distance de 300 M. En général, les deux nœuds ont des liaisons montantes vers le même commutateur réseau ou l'ensemble de commutateurs de réseau ISL (interswitch Link).

MetroCluster SDS à deux nœuds est défini comme un cluster dont les nœuds sont physiquement séparés (salles différentes, bâtiments différents et data centers différents) de plus de 300 M. En outre, les connexions ascendantes de chaque nœud sont connectées à des commutateurs réseau distincts. MetroCluster SDS ne nécessite pas de matériel dédié. Toutefois, l'environnement doit respecter les exigences de latence (maximum de 5 ms pour le temps de propagation aller-retour et 5 ms pour la fluctuation, pour un total de 10 ms) et la distance physique (maximum de 10 km).

Le SDS MetroCluster est une fonctionnalité Premium qui nécessite une licence Premium ou une licence Premium XL. La licence Premium prend en charge la création d'ordinateurs virtuels de petite et moyenne taille ainsi que de disques durs et SSD. La licence Premium XL prend également en charge la création de disques NVMe.

> MetroCluster SDS est pris en charge avec le stockage DAS et le stockage partagé (vNAS). Notez que les configurations vNAS offrent généralement une latence innée plus élevée en raison du réseau entre la machine virtuelle ONTAP Select et le stockage partagé. Les configurations MetroCluster SDS doivent offrir une latence maximale de 10 ms entre les nœuds, y compris la latence de stockage partagé. En d'autres termes, seule la mesure de la latence entre les VM Select est appropriée, car la latence du stockage partagé n'est pas négligeable pour ces configurations.

RSM HAUTE DISPONIBILITÉ et agrégats en miroir

Évitez toute perte de données à l'aide de RAID SyncMirror (RSM), des agrégats en miroir

et du chemin d'écriture.

Réplication synchrone

Le modèle haute disponibilité de ONTAP repose sur le concept de partenaires haute disponibilité. ONTAP Select étend cette architecture à l'univers des serveurs génériques non partagés en utilisant la fonctionnalité RAID SyncMirror (RSM) présente dans ONTAP pour répliquer les blocs de données entre les nœuds du 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 data centers. Pour plus d'informations, reportez-vous à la section "Les bonnes pratiques de la HA étendue à deux nœuds (MetroCluster SDS)".

Agrégats en miroir

Un cluster ONTAP Select se compose de deux à huit nœuds. Chaque paire haute disponibilité contient deux copies des données utilisateur, mises en miroir de façon synchrone sur les nœuds sur un réseau IP. Cette mise en miroir est transparente pour l'utilisateur, et il s'agit d'une propriété de l'agrégat de données, automatiquement configurée lors du processus de création de l'agrégat de données.

Tous les agrégats d'un cluster ONTAP Select doivent être mis en miroir afin d'assurer la disponibilité des données en cas de basculement d'un nœud et d'éviter ainsi un point de défaillance unique en cas de défaillance matérielle. Les agrégats d'un cluster ONTAP Select sont conçus à partir de disques virtuels fournis depuis 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 haute disponibilité du nœud actuel)

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

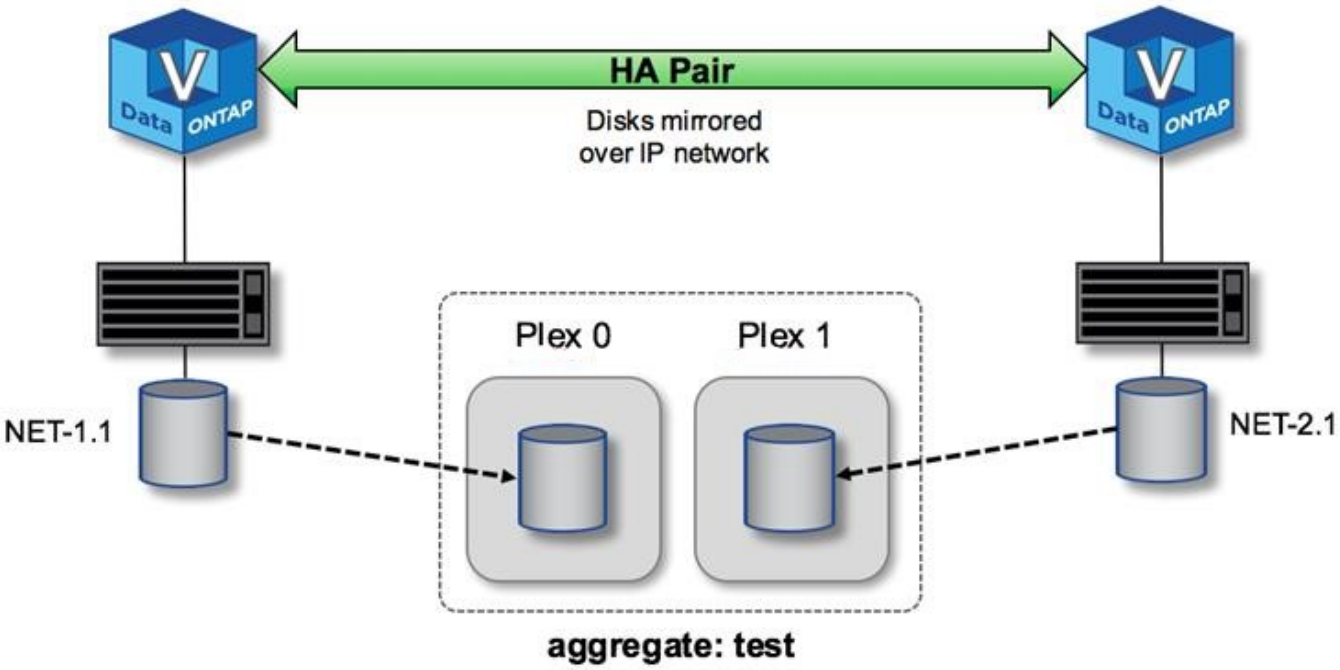
Cette approche est fondamentalement différente de celle que fonctionnent les clusters ONTAP standard. Cela s'applique à tous les disques root et de données du cluster ONTAP Select. L'agrégat contient des copies de données locales et en miroir. Par conséquent, un agrégat qui contient N disques virtuels offre la valeur de N/2 disques de stockage unique, parce que la deuxième copie de données réside sur ses propres disques uniques.

La figure suivante montre une paire HA au sein d'un cluster ONTAP Select à quatre nœuds. Au sein de ce cluster, il s'agit d'un seul agrégat (test) qui utilise le stockage des deux partenaires HA. Cet agrégat de données se compose de deux ensembles de disques virtuels : un ensemble local, fourni par le nœud de cluster propriétaire ONTAP Select (Plex 0) et un ensemble distant, fourni par le partenaire de basculement (Plex 1).

Plex 0 est le godet qui contient tous les disques locaux. Le plex 1 est un compartiment qui contient des disques en miroir ou des disques responsables du stockage d'une seconde copie répliquée des données utilisateur. Le nœud propriétaire de l'agrégat assure la contribution des disques à Plex 0, et le partenaire HA de ce nœud assure la contribution des disques à Plex 1.

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

Agrégat ONTAP Select mis en miroir



Lorsqu'un cluster ONTAP Select est déployé, tous les disques virtuels présents sur le système sont automatiquement affectés au plex approprié, ce qui ne nécessite aucune étape supplémentaire de la part de l'utilisateur concernant l'affectation de disque. Cela empêche l'assignation accidentelle des disques sur un plex incorrect et assure une configuration optimale des disques en miroir.

Chemin d'écriture

La mise en miroir synchrone des blocs de données entre les nœuds de cluster, afin d'éviter toute perte de données en cas de défaillance système, a un impact significatif sur le chemin d'une écriture entrante au fur et à mesure de sa propagation via un cluster ONTAP Select. Ce processus se compose de deux étapes :

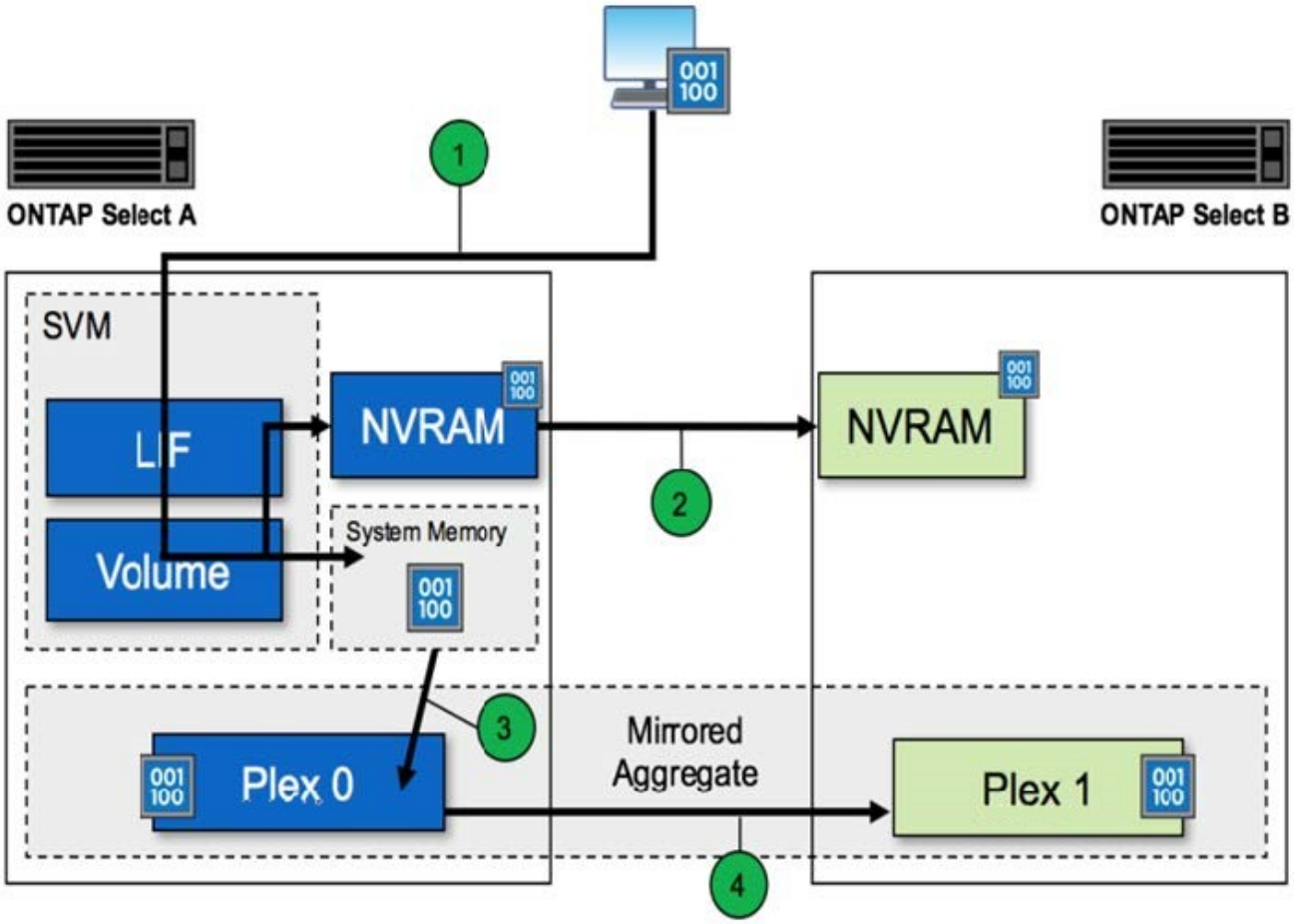
- Accusé de réception
- Transfert

Les écritures sur un volume cible sont effectuées au niveau d'une LIF de données et sont allouées à la partition NVRAM virtualisée présente sur un disque système du nœud ONTAP Select, avant d'être réreconnues au client. Dans une configuration haute disponibilité, une étape supplémentaire se produit, car ces écritures NVRAM sont immédiatement mises en miroir sur le partenaire de haute disponibilité du propriétaire du volume cible avant d'être reconnues. Ce processus garantit la cohérence du système de fichiers sur le nœud partenaire haute disponibilité en cas de panne matérielle sur le nœud d'origine.

Une fois l'écriture effectuée sur la mémoire NVRAM, ONTAP déplace régulièrement le contenu de cette partition vers le disque virtuel approprié, processus appelé déchargement. Ce processus n'a lieu qu'une seule fois sur le nœud du cluster possédant le volume cible, et ne se produit pas sur le partenaire de haute disponibilité.

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

Flot de travaux du chemin d'écriture ONTAP Select



L'acquittement d'écriture entrant comprend les étapes suivantes :

- Les écritures sont envoyées au système via une interface logique appartenant au nœud ONTAP Select A.
- Les écritures sont validées dans la mémoire NVRAM du nœud A et sont mises en miroir sur le partenaire haute disponibilité, le nœud B.
- Une fois la requête d'E/S présente sur les deux nœuds HA, la requête est ensuite réreconnue pour le client.

Le transfert par ONTAP Select de la mémoire NVRAM vers l'agrégat de données (ONTAP CP) inclut les étapes suivantes :

- Les écritures sont déchargées de la mémoire NVRAM virtuelle vers l'agrégat de données virtuelles.
- Le moteur de mise en miroir réplique de manière synchrone les blocs sur les deux plexes.

HA de détails supplémentaires

Battement des disques DE HAUTE DISPONIBILITÉ, boîte aux lettres haute disponibilité, battement des cœur de haute disponibilité, basculement haute disponibilité et rétablissement afin d'améliorer la protection des données.

Les pulsations du disque

Bien que l'architecture ONTAP Select HA tire parti de la plupart des chemins de code utilisés par les baies FAS

traditionnelles, il existe quelques exceptions. L'une de ces exceptions est la mise en œuvre d'une méthode de communication sans réseau utilisée par les nœuds de cluster pour empêcher l'isolement du réseau de provoquer un comportement split-brain. Un scénario split-brain est le résultat du partitionnement du cluster, généralement causé par des pannes réseau, où chaque côté estime que l'autre est en panne et tente de reprendre les ressources du cluster.

Les implémentations haute disponibilité de classe entreprise doivent gérer aisément ce type de scénario. Pour ce faire, ONTAP utilise une méthode d'cœur personnalisée, basée sur disque. Il s'agit du travail de la boîte aux lettres haute disponibilité, emplacement sur le stockage physique utilisé par les nœuds du cluster pour transmettre les messages de signal de détection. Cela permet au cluster de déterminer la connectivité et donc de définir le quorum en cas de basculement.

Sur les baies FAS, qui utilisent une architecture haute disponibilité du stockage partagé, ONTAP résout les problèmes « split-brain » de la manière suivante :

- Réservations persistantes SCSI
- Des métadonnées haute disponibilité persistantes
- État HAUTE DISPONIBILITÉ envoyé via l'interconnexion haute disponibilité

Toutefois, dans l'architecture sans partage d'un cluster ONTAP Select, un nœud ne peut voir que son propre stockage local et non celui du partenaire haute disponibilité. Par conséquent, lorsque le partitionnement réseau isole chaque côté d'une paire haute disponibilité, les méthodes ci-dessus permettant de déterminer le quorum du cluster et le comportement de basculement sont indisponibles.

Bien que la méthode existante de détection et d'évitement du split-brain ne puisse pas être utilisée, une méthode de médiation est toujours nécessaire, qui s'inscrit dans les contraintes d'un environnement sans partage. ONTAP Select étend encore davantage l'infrastructure de la boîte aux lettres existante, ce qui lui permet d'agir comme une méthode de médiation en cas de partitionnement du réseau. Étant donné que le stockage partagé n'est pas disponible, la médiation est effectuée par l'accès aux disques de boîte aux lettres via NAS. Ces disques sont répartis dans tout le cluster, notamment le médiateur dans un cluster à deux nœuds, à l'aide du protocole iSCSI. Il est donc possible de prendre des décisions intelligentes en matière de basculement par un nœud de cluster en fonction de l'accès à ces disques. Si un nœud peut accéder aux disques de boîte aux lettres d'autres nœuds en dehors de son partenaire de haute disponibilité, il est probable qu'il soit fonctionnel et prêt à être fonctionnement.

L'architecture de la boîte aux lettres et la méthode de heartbeef basée sur disque pour résoudre les problèmes de quorum du cluster et de split-brain sont les raisons pour lesquelles la variante à plusieurs nœuds de ONTAP Select nécessite quatre nœuds distincts ou un médiateur pour un cluster à deux nœuds.

Publication de boîte aux lettres HAUTE DISPONIBILITÉ

L'architecture de la boîte aux lettres haute disponibilité utilise un modèle d'envoi de messages. À intervalles réguliers, les nœuds du cluster publient des messages sur tous les autres disques de boîte aux lettres du cluster, notamment le médiateur, indiquant que le nœud est opérationnel. Dans un cluster sain à tout moment, un seul disque de boîte aux lettres sur un nœud de cluster contient des messages publiés de tous les autres nœuds de cluster.

Connecté à chaque nœud Select cluster est un disque virtuel utilisé spécifiquement pour l'accès aux boîtes aux lettres partagées. Ce disque est appelé disque de boîte aux lettres médiateur, car sa fonction principale est de servir de méthode de médiation de cluster en cas de défaillance de nœud ou de partitionnement de réseau. Ce disque de boîte aux lettres contient des partitions pour chaque nœud de cluster et est monté sur un réseau iSCSI par d'autres nœuds de cluster Select. Ces nœuds postont régulièrement des États d'intégrité sur

la partition appropriée du disque de la boîte aux lettres. L'utilisation de disques de boîtes aux lettres accessibles en réseau répartis dans tout le cluster vous permet de déduire l'état de santé des nœuds via une matrice de réaccessibilité. Par exemple, les nœuds de cluster A et B peuvent envoyer à la boîte aux lettres du nœud D du cluster, mais pas à la boîte aux lettres du nœud C. De plus, le nœud D du cluster ne peut pas envoyer la boîte aux lettres du nœud C, il est donc probable que le nœud C soit en panne ou qu'il soit isolé du réseau et qu'il doit être pris en charge.

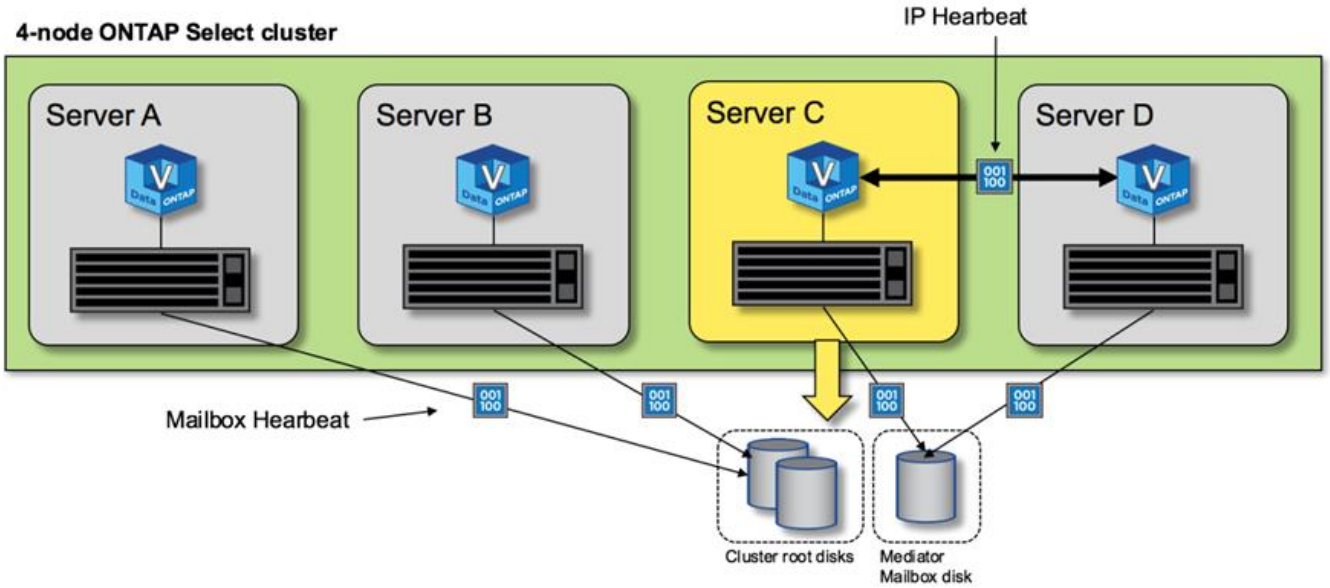
Heartbeef

Comme pour les plateformes NetApp FAS, ONTAP Select envoie régulièrement des messages de signal de détection haute disponibilité sur l'interconnexion haute disponibilité. Dans le cluster ONTAP Select, cette opération s'effectue sur une connexion réseau TCP/IP entre les partenaires haute disponibilité. Par ailleurs, les messages de signal de détection sur disque sont transmis à tous les disques de boîte aux lettres HA, notamment les disques de boîte aux lettres médiateur. Ces messages sont transmis toutes les quelques secondes et lus régulièrement. La fréquence d'envoi et de réception de ces dernières permet au cluster ONTAP Select de détecter les événements de panne haute disponibilité en 15 secondes environ, dans la même fenêtre que sur les plateformes FAS. Lorsque les messages de pulsation ne sont plus lus, un événement de basculement est déclenché.

La figure ci-dessous illustre le processus d'envoi et de réception de messages de signal de détection via l'interconnexion haute disponibilité et les disques médiateurs du point de vue d'un seul nœud de cluster ONTAP Select, nœud C.

Les pulsations du réseau sont transmises via l'interconnexion haute disponibilité au partenaire de haute disponibilité, le nœud D, tandis que les pulsations des disques utilisent des disques de toutes les boîtes aux lettres sur l'ensemble des nœuds du cluster, A, B, C et D.

Battement de cœur HA dans un cluster à quatre nœuds : état stable



Basculement et rétablissement HAUTE DISPONIBILITÉ

Lors d'une opération de basculement, le nœud survivant prend en charge le service des données de son nœud homologue à l'aide de la copie locale des données de son partenaire HA. Les E/S client peuvent continuer à être interrompues, mais les modifications apportées à ces données doivent être répliquées à nouveau avant que le rétablissement ne puisse avoir lieu. Notez que ONTAP Select ne prend pas en charge un rétablissement forcé, car cela entraîne la perte des modifications stockées sur le nœud survivant.

L'opération de resynchronisation arrière est automatiquement déclenchée lorsque le nœud redémarré rejoint le cluster. Le temps nécessaire à la synchronisation inverse dépend de plusieurs facteurs. Ces facteurs incluent le nombre de modifications à répliquer, la latence du réseau entre les nœuds et la vitesse des sous-systèmes de disques sur chaque nœud. Il est possible que le temps requis pour la synchronisation inverse dépasse la fenêtre de réaffichage automatique de 10 minutes. Dans ce cas, un rétablissement manuel est nécessaire après la synchronisation inverse. La progression de la resynchronisation arrière peut être surveillée à l'aide de la commande suivante :

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

Performance

Performance

Les performances varient en fonction de la configuration matérielle.

Les performances d'un cluster ONTAP Select peuvent considérablement varier en fonction des caractéristiques du matériel et de la configuration sous-jacents. La configuration matérielle spécifique est le principal facteur de performances d'une instance ONTAP Select particulière. Voici certains facteurs qui affectent les performances d'une instance ONTAP Select spécifique :

- **Fréquence principale**. En général, une fréquence plus élevée est préférable.
- **Socket simple contre multisocket**. ONTAP Select n'utilise pas de fonctions multisockets, mais la surcharge de l'hyperviseur pour la prise en charge des configurations multisockets représente un certain écart de performances totales.
- **Configuration de la carte RAID et pilote d'hyperviseur associé**. Le pilote par défaut fourni par l'hyperviseur peut être remplacé par le pilote du fournisseur de matériel.
- **Type de disque et nombre de disques dans le(s) groupe(s) RAID**.
- **Version de l'hyperviseur et niveau de patch**.

Performances : stockage SSD DAS haute disponibilité premium

Informations sur les performances de la plate-forme de référence.

Plate-forme de référence

Matériel ONTAP Select (Premium XL) (par nœud)

- FUJITSU PRIMERGY RX2540 M4 :
  - Processeur Intel® Xeon® Gold 6142b à 2.6 GHz
  - 32 cœurs physiques (16 x 2 sockets), 64 cœurs logiques
  - 256 GO DE RAM
  - Disques par hôte : 24 SSD de 960 Go
  - ESX 6.5U1

Matériel client

- 5 clients IBM 3550m4 NFSv3

Informations de configuration

- SW RAID 1 x 9 + 2 RAID-DP (11 lecteurs)
- 22+1 RAID-5 (RAID-0 dans ONTAP) / NVRAM de cache RAID
- Aucune fonctionnalité d'efficacité du stockage utilisée (compression, déduplication, copies Snapshot, SnapMirror, etc.)

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

Ces résultats sont basés sur ONTAP Select 9.6.

Résultats des performances pour un cluster ONTAP Select à un seul nœud (faisant partie d'une instance moyenne à quatre nœuds) sur un SSD DAS (Direct-Attached Storage), avec RAID logiciel et RAID matériel

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

Lecture séquentielle de 64 Ko

Détails :

- E/S directes activées
- 2 nœuds
- 2 cartes d'interface réseau de données par nœud
- 1 agrégat de données par nœud (RAID matériel de 2 To), (RAID logiciel de 8 To)
- 64 processus SIO, 1 thread par proc
- 32 volumes par nœud

- 1 x fichiers par proc ; les fichiers sont de 12000 Mo chacun

Écriture séquentielle de 64 Ko

Détails :

- E/S directes activées
- 2 nœuds
- 2 cartes d'interface réseau de données (NIC) par nœud
- 1 agrégat de données par nœud (RAID matériel de 2 To), (RAID logiciel de 4 To)
- 128 processus SIO, 1 thread par proc
- Volumes par nœud : 32 (RAID matériel), 16 (RAID logiciel)
- 1 x fichiers par proc ; les fichiers sont de 30720 Mo chacun

Lecture aléatoire 8K

Détails :

- E/S directes activées
- 2 nœuds
- 2 cartes réseau de données par nœud
- 1 agrégat de données par nœud (RAID matériel de 2 To), (RAID logiciel de 4 To)
- 64 processus SIO, 8 threads par processeur
- Volumes par nœud : 32
- 1 x fichiers par proc ; les fichiers sont de 122 28 Mo chacun

Écriture aléatoire 8K

Détails :

- E/S directes activées
- 2 nœuds
- 2 cartes réseau de données par nœud
- 1 agrégat de données par nœud (RAID matériel de 2 To), (RAID logiciel de 4 To)
- 64 processus SIO, 8 threads par processeur
- Volumes par nœud : 32
- 1 x fichiers par proc ; les fichiers sont de 8192Mo chacun

50 % aléatoire 8K lecture 50 %

Détails :

- E/S directes activées
- 2 nœuds
- 2 cartes réseau de données par nœud

- 1 agrégat de données par nœud (RAID matériel de 2 To), (RAID logiciel de 4 To)
- 64 threads SIO proc208 par proc
- Volumes par nœud : 32
- 1 x fichiers par proc ; les fichiers sont de 122 28 Mo chacun

Informations sur le copyright

Copyright © 2025 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.