Conception de réseau vSAN

VMware vSphere 8.0
VMware vSAN 8.0

Vous trouverez la documentation technique la plus récente sur le site Web de VMware, à l'adresse :

https://docs.vmware.com/fr/

VMware, Inc.
3401 Hillview Ave.
Palo Alto, CA 94304
www.vmware.com

VMware France SAS.
Tour Franklin
100-101 Terrasse Boieldieu
92042 Paris La Défense 8 Cedex
France
www.vmware.com/fr

Table des matières

À propos de la conception de réseau vSAN

1

Le guide de *conception de réseau vSAN* décrit les exigences réseau, la conception du réseau et les pratiques de configuration pour le déploiement d'un cluster vSAN hautement disponible et évolutif.

vSAN est une solution de stockage distribué. Comme pour toute solution distribuée, le réseau est un composant important de la conception. Pour des résultats optimaux, vous devez respecter les instructions fournies dans ce document, car un matériel et des conceptions de mise en réseau incorrects peuvent engendrer des résultats défavorables.

VMware prend l'intégration au sérieux. Afin de promouvoir ce principe pour notre client, nos partenaires et la communauté interne, nous créons du contenu à l'aide du langage inclusif.

Public cible

Ce guide est destiné à toute personne responsable de la conception, du déploiement et de la gestion d'un cluster vSAN. Les informations contenues dans ce guide sont destinées aux administrateurs réseau expérimentés qui maîtrisent la conception et la configuration du réseau, la gestion des machines virtuelles et les opérations de centre de données virtuel. Ce guide suppose également que vous soyez familiarisé avec VMware vSphere, notamment VMware ESXi, vCenter Server et vSphere Client.

Documents connexes

Parallèlement à ce guide, vous pouvez consulter les guides suivants pour en savoir plus sur la mise en réseau de vSAN :

- *Guide de planification et de déploiement de vSAN*, pour en savoir plus sur la création de clusters vSAN
- *Administration de VMware vSAN*, pour configurer un cluster vSAN et en savoir plus sur les fonctionnalités de vSAN
- *Guide de surveillance et de dépannage de vSAN*, pour surveiller et dépanner les clusters vSAN

Présentation du réseau vSAN 2

Vous pouvez utiliser vSAN pour provisionner le stockage partagé dans vSphere. vSAN regroupe des périphériques de stockage locaux ou à connexion directe d'un cluster hôte et crée un pool de stockage unique partagé sur tous les hôtes du cluster vSAN.

vSAN est une solution de stockage distribué et partagé qui dépend d'un réseau hautement disponible et correctement configuré pour le trafic de stockage vSAN. Un réseau performant et disponible est essentiel pour un déploiement vSAN réussi. Ce guide contient des recommandations sur la conception et la configuration d'un réseau vSAN.

vSAN dispose d'une architecture distribuée qui repose sur un réseau hautes performances, évolutif et résilient. Tous les nœuds hôtes d'un cluster vSAN communiquent sur le réseau IP. Tous les hôtes doivent conserver la connectivité monodiffusion IP pour pouvoir communiquer sur un réseau de couche 2 ou de couche 3. Pour de plus amples informations sur la communication monodiffusion, reportez-vous à la section Chapitre 4 Utilisation de la monodiffusion dans le réseau vSAN.

Les versions antérieures à vSAN 6.6 requièrent la multidiffusion IP. Si possible, utilisez toujours la dernière version de vSAN. Pour de plus amples informations sur la multidiffusion, reportez-vous à la section Chapitre 13 Utilisation de la multidiffusion dans un réseau vSAN.

Termes et définitions de la mise en réseau vSAN

vSAN introduit des termes et des définitions spécifiques importants à comprendre. Avant d'initier la conception de votre réseau vSAN, consultez les termes et les définitions vSAN clés de la mise en réseau.

Termes	Définitions
CLOM	Le gestionnaire d'objets de niveau cluster (CLOM) est chargé de veiller à ce que la configuration d'un objet corresponde à sa stratégie de stockage. Le CLOM vérifie si une quantité suffisante de groupes de disques est disponible pour satisfaire cette stratégie. Il décide où placer les composants et les témoins dans un cluster.
CMMDS	Le service de surveillance, d'appartenance et d'annuaire de cluster (CMMDS) est responsable de la récupération et de la maintenance d'un cluster de membres de nœuds en réseau. Il gère l'inventaire des éléments tels que les nœuds hôtes, les périphériques et les réseaux. Il stocke également les informations de métadonnées, telles que les stratégies et la configuration RAID pour les objets vSAN.
DOM	Le Distributed Object Manager (DOM) est responsable de la création des composants et de leur répartition dans le cluster. Après la création d'un objet DOM, l'un des nœuds (hôte) est désigné comme propriétaire DOM de cet objet. Cet hôte traite toutes les IOPS de cet objet DOM en localisant les composants enfants respectifs dans le cluster et en redirigeant les E/S vers les composants respectifs sur le réseau vSAN. Les objets DOM incluent les objets suivants : vdisk, snapshot, vmnamespace, vmswap, vmem, etc.
LSOM	Le gestionnaire d'objets à structure journalisée (LSOM, Log-Structured Object Manager) est responsable du stockage local des données sur le système de fichiers vSAN en tant que composant vSAN ou LSOM-Objet (composant de données ou composant témoin).
Association de cartes réseau	L'association de cartes d'interface réseau (NIC) peut être définie comme la configuration d'au moins deux adaptateurs réseau (NIC) en tant qu'« équipe » pour la haute disponibilité et l'équilibrage de charge.
NIOC	Network I/O Control (NIOC) détermine la bande passante accordée à différents types de trafics réseau sur un vSphere Distributed Switch. La distribution de bande passante est un paramètre pouvant être configuré par l'utilisateur. Lorsque NIOC est activé, le trafic du commutateur distribué est divisé en pools de ressources réseau prédéfinis : trafic à tolérance de panne, trafic iSCSI, trafic vMotion, trafic de gestion, trafic vSphere Replication, trafic NFS et trafic de machine virtuelle.

Termes	Définitions
Objets et composants	Chaque objet est constitué d'un ensemble de composants, déterminé par les capacités utilisées dans la stratégie de stockage de machine virtuelle. Une banque de données vSAN contient plusieurs types d'objets : ■ **Espace de noms de base de la VM** : l'espace de noms de base de la machine virtuelle est un répertoire de base de machine virtuelle dans lequel tous les fichiers de configuration de la machine virtuelle sont stockés. Cela inclut des fichiers tels que .vmx, les fichiers journaux, les disques de machine virtuelle et les fichiers de description de fichier de disque delta de snapshot. ■ **VMDK** : un VMDK est un disque de machine virtuelle ou un fichier .vmdk qui stocke le contenu du lecteur de disque dur d'une machine virtuelle. ■ **Objet de permutation de VM** : les objets d'échange de VM sont créés lorsqu'une machine virtuelle est sous tension. ■ **VMDK delta de snapshot** : les VMDK delta de snapshot sont créés lorsque des snapshots de machine virtuelle sont pris. ■ **Objet de mémoire** : les objets de mémoire sont créés lorsque l'option de mémoire de snapshot est sélectionnée lors de la création ou de l'interruption d'une machine virtuelle.
RDT	Le protocole RDT (Reliable Data Transport) est utilisé pour la communication entre les hôtes sur les ports VMkernel de vSAN. Il utilise le protocole TCP sur la couche de transport et est responsable de la création et de la destruction des connexions TCP (sockets) à la demande. Il est optimisé pour envoyer des fichiers volumineux.
SPBM	La gestion basée sur une stratégie de stockage (SPBM) fournit une structure de stratégie de stockage qui sert de panneau de contrôle unifié unique sur une grande variété de services de données et de solutions de stockage. Cette infrastructure vous permet d'aligner le stockage sur les impératifs des applications de vos machines virtuelles.
VASA	Les API de stockage vSphere pour la détection de stockage (VASA) est un ensemble d'interfaces de programmation d'application (API) qui permet à vCenter Server de reconnaître les capacités des baies de stockage. Les fournisseurs VASA communiquent avec le système vCenter Server pour déterminer la topologie de stockage, la capacité et les informations d'état qui prennent en charge la gestion basée sur des stratégies, la gestion des opérations et la fonctionnalité DRS.

Termes	Définitions
VLAN	Un VLAN permet à un segment LAN physique unique d'être davantage segmenté de sorte que des groupes de ports sont isolés les uns des autres comme s'ils se trouvaient sur des segments physiquement différents.
Composant témoin	Un témoin est un composant contenant uniquement des métadonnées et non des données d'application réelles. Il sert d'arbitre en cas de décision à prendre concernant la disponibilité des composants de banque de données restants, après une panne potentielle. Un témoin consomme environ 2 Mo d'espace pour les métadonnées sur la banque de données vSAN lors de l'utilisation du format sur disque 1.0 et 4 Mo pour le format sur disque version 2.0 et version ultérieure.

Présentation de la mise en réseau vSAN

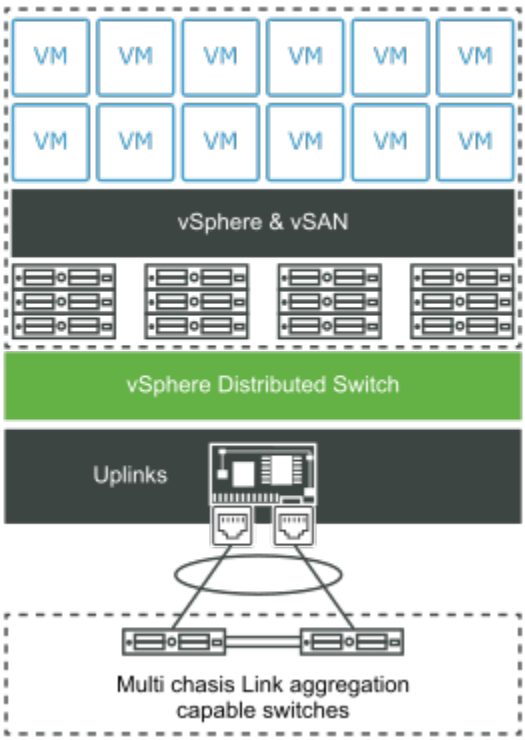
3

Un réseau vSAN facilite la communication entre les hôtes du cluster et doit garantir des performances rapides, ainsi qu'une disponibilité et une bande passante élevées.

vSAN utilise le réseau pour les communications entre les hôtes ESXi et pour les E/S de disque de machine virtuelle.

Les machines virtuelles (VM) sur les banques de données vSAN sont composées d'un ensemble d'objets, et chaque objet peut être constitué d'un ou de plusieurs composants. Ces composants sont distribués sur plusieurs hôtes pour une résilience accrue face aux défaillances du disque et de l'hôte. vSAN maintient et met à jour ces composants à l'aide du réseau vSAN.

Le diagramme suivant fournit un aperçu global du réseau vSAN :



Ce chapitre contient les rubriques suivantes :

- Caractéristiques du réseau vSAN
- Types de trafic ESXi

- Configuration de réseau requise pour vSAN

Caractéristiques du réseau vSAN

vSAN dépend du réseau. La présentation et la configuration des paramètres réseau vSAN appropriés est essentielle pour éviter les problèmes de performances et de stabilité.

Un réseau vSAN fiable et robuste présente les caractéristiques suivantes :

Monodiffusion

vSAN 6.6 et version ultérieures prennent en charge la communication monodiffusion. Le trafic monodiffusion est une transmission point à point des paquets IP d'un point du réseau vers un autre. La monodiffusion transmet le signal de pulsation envoyé de l'hôte principal à tous les autres hôtes à chaque seconde. Cela permet de garantir que les hôtes sont actifs et indique la participation des hôtes dans le cluster vSAN. Vous pouvez concevoir un réseau monodiffusion simple pour vSAN. Pour de plus amples informations sur la communication monodiffusion, reportez-vous à la section Chapitre 4 Utilisation de la monodiffusion dans le réseau vSAN.

Multidiffusion

Les versions antérieures à vSAN 6.6 utilisent la communication multidiffusion IP comme protocole de détection pour identifier les nœuds qui tentent de joindre un cluster vSAN.

Note Si possible, utilisez toujours la dernière version de vSAN.

La multidiffusion IP repose sur les protocoles de communication utilisés par les hôtes, les clients et les périphériques réseau pour participer aux communications basées sur la multidiffusion. Pour de plus amples informations sur la communication multidiffusion, reportez-vous à la section Chapitre 13 Utilisation de la multidiffusion dans un réseau vSAN.

Réseau de couche 2 et de couche 3

Tous les hôtes du cluster vSAN doivent être connectés à un réseau de couche 2 ou 3. Les versions de vSAN antérieures à vSAN 6.0 prennent uniquement en charge la mise en réseau de couche 2, alors que les versions suivantes incluent la prise en charge des protocoles de couche 2 et 3. Utilisez un réseau de couche 2 ou 3 pour garantir la communication entre les sites de données et le site témoin. Pour de plus amples informations sur les topologies de réseau de couche 2 et 3, reportez-vous à la section Déploiements standard.

Réseau VMkernel

Chaque hôte ESXi d'un cluster vSAN doit disposer d'un adaptateur réseau pour la communication avec vSAN. L'intégralité de la communication du nœud intra-cluster se produit via le port VMkernel vSAN. Les ports VMkernel fournissent des services de couche 2 et 3 à chaque hôte vSAN et aux machines virtuelles hébergées.

Trafic réseau vSAN

Plusieurs types de trafic différents sont disponibles dans le réseau vSAN, tel que le trafic de stockage et le trafic monodiffusion. Le calcul et le stockage d'une machine virtuelle peuvent se trouver sur le même hôte ou sur différents hôtes du cluster. Une machine virtuelle qui n'est pas configurée pour tolérer une panne peut s'exécuter sur un hôte et accéder à un objet de machine virtuelle ou à un composant résidant sur un hôte différent. Cela implique que toutes les E/S de la machine virtuelle passent par le réseau. Le trafic de stockage constitue la majeure partie du trafic d'un cluster vSAN.

La communication liée au cluster entre tous les hôtes ESXi crée du trafic dans le cluster vSAN. Ce trafic monodiffusion contribue également au trafic réseau vSAN.

Commutateur virtuel

vSAN prend en charge les types de commutateurs virtuels suivants :

- Le commutateur virtuel standard assure la connectivité entre les machines virtuelles et les ports VMkernel aux réseaux externes. Ce commutateur est local pour chaque hôte ESXi.
- Un vSphere Distributed Switch permet de centraliser le contrôle de l'administration du commutateur virtuel sur plusieurs hôtes ESXi. Un commutateur distribué fournit également des fonctionnalités de mise en réseau, telles que Network I/O Control (NIOC), qui peuvent vous aider à définir les niveaux de qualité de service (QoS) sur vSphere ou un réseau virtuel. vSAN inclut vSphere Distributed Switch quelle que soit la version de vCenter Server.

Bande passante

Le trafic vSAN peut partager des adaptateurs réseau physiques avec d'autres types de trafic système, tels que le trafic vSphere vMotion, le trafic vSphere HA et le trafic de machine virtuelle. Il offre également une bande passante supplémentaire pour les configurations réseau partagées où vSAN, la gestion vSphere, le trafic vSphere vMotion, etc., se trouvent sur le même réseau physique. Pour garantir la quantité de bande passante requise par vSAN, utilisez vSphere Network I/O Control dans le commutateur distribué.

Dans vSphere Network I/O Control, vous pouvez configurer la réservation et les partages pour le trafic vSAN sortant :

- Définissez une réservation de manière à ce que Network I/O Control puisse garantir la disponibilité d'une bande passante minimale sur l'adaptateur physique pour vSAN.
- Définissez la valeur de partage sur 100 afin que la bande passante soit en partie disponible pour vSAN en cas de saturation de l'adaptateur physique attribué à vSAN. Par exemple, l'adaptateur physique peut devenir saturé lorsqu'un autre adaptateur physique de l'équipe est défaillant et que la totalité du trafic dans le groupe de ports est transféré sur les autres adaptateurs de l'équipe.

Pour de plus amples informations sur l'utilisation de Network I/O Control pour configurer l'allocation de bande passante au trafic vSAN, consultez la documentation de *Mise en réseau de vSphere*.

Types de trafic ESXi

Les hôtes ESXi utilisent différents types de trafic réseau pour prendre en charge vSAN.

Les différents types de trafic que vous devez configurer pour vSAN sont présentés ci-dessous.

Tableau 3-1. Types de trafic réseau

Types de trafic	Description
Réseau de gestion	Le réseau de gestion est l'interface réseau principale qui utilise une pile TCP/IP VMkernel pour faciliter la connectivité et la gestion de l'hôte. Il peut également gérer le trafic système tel que vMotion, iSCSI, Système de fichiers du réseau (NFS), Fiber Channel over Ethernet (FCoE) et Tolérance de panne.
Réseau de machines virtuelles	Avec la mise en réseau virtuelle, vous pouvez mettre en réseau des machines virtuelles et construire des réseaux complexes au sein d'un hôte ESXi unique ou sur plusieurs hôtes ESXi.
Réseau vMotion	Type de trafic qui facilite la migration d'une machine virtuelle d'un hôte vers un autre. La migration avec vMotion exige des interfaces réseau correctement configurées sur des hôtes source et cible. Assurez-vous que le réseau vMotion est différent du réseau vSAN.
réseau vSAN	Un cluster vSAN requiert le réseau VMkernel pour l'échange de données. Chaque hôte ESXi du cluster vSAN doit disposer d'un adaptateur réseau VMkernel pour le trafic vSAN. Pour de plus amples informations, reportez-vous à Configurer un réseau VMkernel pour vSAN.

Configuration de réseau requise pour vSAN

vSAN est une solution de stockage distribué qui dépend du réseau pour la communication entre les hôtes. Avant le déploiement, assurez-vous que votre environnement vSAN répond à toutes les exigences de mise en réseau.

Configuration requise pour la carte réseau physique

Les cartes d'interface réseau (NIC) utilisées dans les hôtes vSAN doivent satisfaire à certaines exigences. vSAN fonctionne sur 10 Gbits/s, 25 Gbits/s, 40 Gbits/s, 50 Gbits/s et 100 Gbits/s.

Assurez-vous que vos hôtes respectent la configuration NIC minimale requise pour vSAN Original Storage Architecture (OSA) ou vSAN Express Storage Architecture (ESA).

Tableau 3-2. Configuration minimale requise et recommandations pour la carte réseau sur vSAN OSA

Topologie ou mode de déploiement	Architecture	Prise en charge de la carte réseau 1 GbE	Prise en charge de la carte réseau 10 GbE	Prise en charge des cartes réseau de capacité supérieure à 10 GbE	Latence inter-nœud	Bande passante ou latence de liaison intersite	Latence entre les nœuds et les hôtes témoins vSAN	Bande passante entre les nœuds et les hôtes témoins vSAN
Cluster standard	Cluster hybride	Oui (minimum)	Oui (recommandé)	Oui	RTT inférieur à 1 ms.	S/O	S/O	S/O
	Cluster intégralement Flash	Non	Oui	Oui (recommandé)				
Cluster étendu	Cluster hybride ou intégralement Flash	Non	Oui (minimum)	Oui	RTT inférieur à 1 ms sur chaque site.	La valeur recommandée est de 10 GbE (dépend de la charge de travail) et RTT de 5 ms maximum.	RTT inférieur à 200 ms. Jusqu'à 10 hôtes par site.	2 Mbits/s par 1 000 composants (100 Mbits/s maximum avec 45 000 composants).
							RTT inférieur à 100 ms. De 11 à 15 hôtes par site.	
Cluster à deux nœuds	Cluster hybride	Oui (jusqu'à 10 machines virtuelles)	Oui (recommandé)	Oui	RTT inférieur à 1 ms sur le même site.	La valeur recommandée est de 10 GbE et RTT de 5 ms maximum.	RTT inférieur à 500 ms.	2 Mbits/s par 1 000 composants (1,5 Mbits/s maximum).
	Cluster intégralement Flash	Non	Oui (minimum)					

Tableau 3-3. Configuration minimale requise et recommandations pour la carte réseau sur vSAN ESA

Type de déploiement	Prise en charge de la carte réseau 1 GbE	Prise en charge de la carte réseau 10 GbE	Prise en charge des cartes réseau de capacité supérieure à 10 GbE	Latence inter-nœud	Bande passante ou latence de liaison intersite	Latence entre les nœuds et les hôtes témoins vSAN	Bande passante entre les nœuds et les hôtes témoins vSAN
Cluster standard	Non	Non	Oui (25 GbE minimum)	RTT inférieur à 1 ms.	S/O	S/O	S/O
Cluster étendu	Non	Non	Oui (25 GbE minimum)	RTT inférieur à 1 ms sur chaque site.	La valeur recommandée est de 25 GbE (dépend de la charge de travail) et RTT de 5 ms maximum.	RTT inférieur à 200 ms. Jusqu'à 10 hôtes par site. RTT inférieur à 100 ms. De 11 à 15 hôtes par site.	2 Mbits/s par 1 000 composants (100 Mbits/s maximum avec 45 000 composants).
Cluster à deux nœuds	Non	Non	Oui (25 GbE minimum)	RTT inférieur à 1 ms sur le même site.	La valeur recommandée est de 25 GbE et RTT de 5 ms maximum.	RTT inférieur à 500 ms.	2 Mbits/s par 1 000 composants (1,5 Mbits/s maximum).

Note Ces exigences de carte réseau supposent que la perte de paquets ne dépasse pas 0,0001 % dans les environnements hyperconvergés. Il peut y avoir une incidence majeure sur les performances de vSAN en cas de dépassement de l'une de ces exigences.

Pour plus d'informations sur la configuration requise pour la carte réseau du cluster étendu, consultez le *Guide de cluster étendu vSAN*.

Configuration requise en matière de bande passante et de latence

Pour garantir des performances et une disponibilité élevées, les clusters vSAN doivent satisfaire la configuration requise en termes de bande passante et de latence réseau.

Les exigences de bande passante entre les sites principal et secondaire d'un cluster étendu vSAN dépendent de la charge de travail vSAN, de la quantité de données et de la manière dont vous souhaitez gérer les pannes. Pour de plus amples informations, consultez le *Guide de conception et de dimensionnement de VMware vSAN*.

Tableau 3-4. Configuration requise en matière de bande passante et de latence

Communication du site	Bande passante	Latence
Site vers le site	vSAN OSA : minimum de 10 Gbits/s vSAN ESA : minimum de 25 Gbits/s	Temps de latence inférieur à 5 ms RTT.
Site vers le témoin	2 Mbits/s par groupe de 1 000 composants vSAN	■ Temps de latence inférieur à 500 ms RTT pour 1 hôte par site. ■ Temps de latence inférieur à 200 ms RTT pour 10 hôtes maximum par site. ■ Temps de latence inférieur à 100 ms RTT pour 11 à 15 hôtes par site.

Prise en charge de la couche 2 et de la couche 3

VMware recommande la connectivité de couche 2 entre tous les hôtes vSAN partageant le sous-réseau.

vSAN prend également en charge les déploiements à l'aide de la connectivité de couche 3 routée entre les hôtes vSAN. Vous devez tenir compte du nombre de tronçons et de la latence supplémentaire encourus pendant l'acheminement du trafic.

Tableau 3-5. Prise en charge de la couche 2 et de la couche 3

Type de cluster	L2 pris en charge	L3 pris en charge	Considérations
Cluster hybride	Oui	Oui	L2 est recommandé et L3 est pris en charge.
Cluster 100 % Flash	Oui	Oui	L2 est recommandé et L3 est pris en charge.
Données de cluster étendu	Oui	Oui	L2 et L3 sont pris en charge entre les sites de données.
Témoin de cluster étendu	Non	Oui	L3 est pris en charge. L2 n'est pas pris en charge entre les données et les sites témoins.
Cluster à deux nœuds	Oui	Oui	L2 et L3 sont pris en charge entre les sites de données.

Configuration de routage et de commutation requise

Les trois sites d'un cluster étendu communiquent sur le réseau de gestion et sur le réseau vSAN. Les VM de tous les sites de données communiquent sur un réseau de machines virtuelles commun.

Les exigences de routage de cluster étendu vSANsont présentées ci-dessous.

Tableau 3-6. Configuration de routage requise

Communication du site	Modèle de déploiement	Couche	Routage
Site vers le site	Par défaut	Couche 2	Non requis
Site vers le site	Par défaut	Couche 3	Des routes statiques sont nécessaires.
Site vers le témoin	Par défaut	Couche 3	Des itinéraires statiques sont nécessaires.
Site vers le témoin	Séparation du trafic témoin	Couche 3	Des routes statiques sont nécessaires en cas d'utilisation d'une interface différente de l'interface de gestion (vmk0).
Site vers le témoin	Séparation du trafic témoin	Couche 2 pour cluster à deux hôtes	Des routes statiques ne sont pas nécessaires.

Configuration requise du commutateur virtuel

Vous pouvez créer un réseau vSAN avec un vSphere Standard Switch ou un vSphere Distributed Switch. Utilisez un commutateur distribué pour hiérarchiser la bande passante pour le trafic vSAN. vSAN utilise un commutateur distribué avec toutes les versions de vCenter Server.

Le tableau suivant compare les avantages et les bénéfices d'un commutateur distribué par rapport à un commutateur standard :

Tableau 3-7. Types de commutateurs virtuels

Condition de conception requise	Option 1 - vSphere Distributed Switch	Option 2 - vSphere Standard Switch	Description
Disponibilité	Aucun impact	Aucun impact	Vous pouvez utiliser l'une des options
Facilité de gestion	Impact positif	Impact négatif	Le commutateur distribué est géré de manière centralisée dans l'ensemble des hôtes, contrairement au commutateur standard qui est géré individuellement sur chaque hôte.
Performances	Impact positif	Impact négatif	Le commutateur distribué dispose de contrôles additionnels, tels que Network I/O Control, que vous pouvez utiliser pour garantir les performances pour le trafic vSAN.

Tableau 3-7. Types de commutateurs virtuels (suite)

Condition de conception requise	Option 1 - vSphere Distributed Switch	Option 2 - vSphere Standard Switch	Description
Facilité de récupération	Impact positif	Impact négatif	La configuration du commutateur distribué peut être sauvegardée et restaurée. Le commutateur standard ne dispose pas de cette fonctionnalité.
Sécurité	Impact positif	Impact négatif	Le commutateur distribué a ajouté des contrôles de sécurité intégrés pour faciliter la protection du trafic.

Exigences relatives au port réseau vSAN

Les déploiements de vSAN nécessitent des ports et des paramètres réseau spécifiques pour fournir un accès et des services.

vSAN envoie des messages sur certains ports de chaque hôte dans le cluster. Vérifiez que les pare-feu de l'hôte autorisent le trafic sur ces ports. Pour obtenir la liste de tous les ports et protocoles de vSAN pris en charge, reportez-vous au portail VMware Ports and Protocols à l'adresse https://ports.vmware.com/.

Considérations relatives au pare-feu

Lorsque vous activez vSAN sur un cluster, tous les ports requis sont ajoutés aux règles de pare-feu ESXi et configurés automatiquement. Un administrateur n'a pas besoin d'ouvrir des ports de pare-feu ou d'activer des services de pare-feu manuellement.

Vous pouvez afficher les ports ouverts pour les connexions entrantes et sortantes. Sélectionnez l'hôte ESXi, puis cliquez sur **Configurer > Profil de sécurité**.

Configuration de pare-feu de réseau requise

Lorsque vous configurez le pare-feu de réseau, déterminez la version de vSAN que vous déployez.

Lorsque vous activez vSAN sur un cluster, tous les ports nécessaires sont ajoutés aux règles de pare-feu ESXi et configurés automatiquement. Il n'est pas nécessaire d'ouvrir des ports de pare-feu ou d'activer des services de pare-feu manuellement. Vous pouvez afficher les ports ouverts pour les connexions entrantes et sortantes dans le profil de sécurité de l'hôte ESXi (**Configurer > Profil de sécurité**).

Règle de pare-feu vsanEncryption

Si votre cluster utilise le chiffrement vSAN, tenez compte de la communication entre les hôtes et le serveur KMS .

Le chiffrement vSAN requiert un serveur de gestion de clés (KMS) externe. vCenter Server obtient les ID de clés du KMS, et les distribue aux hôtes ESXi. Les serveurs KMS et les hôtes ESXi communiquent directement entre eux. Les serveurs KMS peuvent utiliser des numéros de port différents. La règle de pare-feu vsanEncryption vous permet donc de simplifier la communication entre chaque hôte vSAN et le serveur KMS. Cela permet à un hôte vSAN de communiquer directement avec n'importe quel port sur un serveur KMS (port TCP 0 à 65535).

Lorsqu'un hôte établit la communication avec un serveur KMS, les opérations suivantes se produisent.

- L'adresse IP du serveur KMS est ajoutée à la règle vsanEncryption et la règle de pare-feu est activée.
- La communication entre le nœud vSAN et le serveur KMS est établie lors de l'échange.
- Une fois la communication entre le nœud vSAN et le serveur KMS terminée, l'adresse IP est supprimée de la règle vsanEncryption, et la règle de pare-feu est désactivée à nouveau.

Les hôtes vSAN peuvent communiquer avec plusieurs hôtes KMS à l'aide de la même règle.

Utilisation de la monodiffusion dans le réseau vSAN

4

Le trafic monodiffusion fait référence à une transmission point à point d'un point du réseau à un autre. vSAN 6.6 et versions ultérieures utilisent la monodiffusion pour simplifier la conception et le déploiement du réseau.

Tous les hôtes ESXi utilisent le trafic monodiffusion et vCenter Server devient la source de l'appartenance au cluster. Les nœuds vSAN sont automatiquement mis à jour avec la dernière liste d'appartenance d'hôte fournie par vCenter. vSAN communique à l'aide de la monodiffusion pour les mises à jour de CMMDS.

Les versions de vSAN antérieures à 6.6 s'appuient sur la monodiffusion pour activer les signaux de pulsation et échanger des métadonnées entre les hôtes au sein du cluster. Si certains hôtes de votre cluster vSAN exécutent des versions antérieures du logiciel, un réseau multidiffusion est toujours nécessaire. Le basculement vers le réseau monodiffusion à partir de la multidiffusion garantit des performances et une prise en charge réseau améliorées. Pour de plus amples informations sur la multidiffusion, reportez-vous à la section Chapitre 13 Utilisation de la multidiffusion dans un réseau vSAN.

Ce chapitre contient les rubriques suivantes :

Comportement du groupe de disques antérieur à la version 5

La disponibilité d'un groupe de disques version 5 unique dans un groupe de disques vSAN version 6.6 déclenche le cluster pour qu'il communique de façon permanente en mode monodiffusion.

Les clusters vSAN version 6.6 rétablissent automatiquement la communication multidiffusion dans les cas suivants :

- Tous les hôtes du cluster exécutent vSAN 6.5 ou version antérieure.
- Tous les groupes de disques utilisent la version sur disque 3 ou antérieure.
- Un hôte non-vSAN 6.6, tel que vSAN 6.2 ou vSAN 6.5 est ajouté au cluster.

Par exemple, si un hôte exécutant vSAN 6.5 ou version antérieure est ajouté à un cluster vSAN 6.6 existant, le cluster revient en mode multidiffusion et inclut l'hôte 6.5 comme un nœud valide. Pour éviter ce comportement, utilisez la dernière version pour les hôtes ESXi et le format sur disque. Pour vous assurer que le cluster vSAN continue de communiquer en mode monodiffusion et ne revient pas à la multidiffusion, mettez à niveau les groupes de disques sur les hôtes vSAN 6.6 vers la version sur disque 5.0.

Note Évitez d'avoir un cluster en mode mixte dans lequel vSAN 6.5 ou version antérieure sont disponibles dans le même cluster avec vSAN version 6.6 ou ultérieure.

Comportement du groupe de disques version 5

La présence d'un groupe de disques version 5 unique dans un cluster vSAN version 6.6 déclenche le cluster pour communiquer de façon permanente en mode monodiffusion.

Dans un environnement où un cluster vSAN 6.6 utilise déjà une version sur disque 5 et un nœud vSAN 6.5 est ajouté au cluster, les événements suivants se produisent :

- Le nœud vSAN 6.5 forme sa propre partition réseau.
- Le nœud vSAN 6.5 continue de communiquer en mode multidiffusion, mais il ne peut pas communiquer avec les nœuds vSAN 6.6 lorsqu'ils utilisent le mode monodiffusion.

Un avertissement de résumé de cluster apparaît sur le format sur disque, indiquant qu'un nœud appartient à une version antérieure. Vous pouvez mettre le nœud à niveau vers la dernière version. Il n'est pas possible de mettre à niveau les versions de format de disque lorsqu'un cluster est en mode mixte.

Prise en charge de DHCP sur le réseau monodiffusion

Un déploiement vCenter Server sur un cluster vSAN 6.6 peut utiliser des adresses IP de protocole DHCP (DHCP) sans réservation.

Vous pouvez utiliser DHCP avec des réservations, car les adresses IP affectées sont liées aux adresses MAC des ports VMkernel.

Prise en charge d'IPv6 sur le réseau monodiffusion

vSAN 6.6 prend en charge IPv6 avec les communications monodiffusion.

Avec IPv6, l'adresse de liaison locale est automatiquement configurée sur toute interface utilisant le préfixe de liaison locale. Par défaut, vSAN n'ajoute pas l'adresse de liaison locale d'un nœud à d'autres nœuds de cluster voisins. vSAN 6.6 ne prend donc pas en charge les adresses locales de liaison IPv6 pour les communications monodiffusion.

Interroger la monodiffusion avec ESXCLI

Vous pouvez exécuter des commandes ESXCLI pour déterminer la configuration monodiffusion.

Afficher les modes de communication

À l'aide de la commande `esxcli vsan cluster get`, vous pouvez afficher le mode de CMMDS (monodiffusion ou multidiffusion) du nœud de cluster vSAN.

Procédure

- Exécutez la commande `esxcli vsan cluster get`.

Résultats

```
Cluster Information
   Enabled: true
   Current Local Time: 2020-04-09T18:19:52Z
   Local Node UUID: 5e8e3dc3-43ab-5452-795b-a03d6f88f022
   Local Node Type: NORMAL
   Local Node State: AGENT
   Local Node Health State: HEALTHY
   Sub-Cluster Master UUID: 5e8e3d3f-3015-9075-49b6-a03d6f88d426
   Sub-Cluster Backup UUID: 5e8e3daf-e5e0-ddb6-a523-a03d6f88dd4a
   Sub-Cluster UUID: 5282f9f3-d892-3748-de48-e2408dc34f72
   Sub-Cluster Membership Entry Revision: 11
   Sub_cluster Member Count: 5
   Sub-Cluster Member UUIDs: 5e8e3d3f-3015-9075-49b6-a03d6f88d426, 5e8e3daf-e5e0-ddb6-a523-
a03d6f88dd4a,
   5e8e3d73-6d1c-0b81-1305-a03d6f888d22, 5e8e3d33-5825-ee5c-013c-a03d6f88ea4c,
5e8e3dc3-43ab-5452-795b-a03d6f88f022
   Sub-Cluster Member HostNames: testbed-1.vmware.com, testbed2.vmware.com,
   testbed3.vmware.com, testbed4.vmware.com, testbed5.vmware.com
   Sub-Cluster Membership UUID: 0f438e5e-d400-1bb2-f4d1-a03d6f88d426
   Mode monodiffusion activé : true
   Maintenance Mode State: OFF
   Config Generation: ed845022-5c08-48d0-aa1d-6b62c0022222 7 2020-04-08T22:44:14.889
```

Vérifier les hôtes du cluster vSAN

Utilisez la commande `esxcli vsan cluster unicastagent list` pour vérifier si les hôtes du cluster vSAN fonctionnent en mode monodiffusion.

Procédure

- Exécutez la commande `esxcli vsan cluster unicastagent list`.

Résultats

```
NodeUuid                              IsWitness Supports Unicast IP Address  Port  Iface Name
Cert Thumbprint  SubClusterUuid
------------------------------------ --------- ---------------- ----------  ----- ----------
5e8e3d73-6d1c-0b81-1305-a03d6f888d22         0      true
10.198.95.10    12321
43:80:B7:A1:3F:D1:64:07:8C:58:01:2B:CE:A2:F5:DE:D6:B1:41:AB
5e8e3daf-e5e0-ddb6-a523-a03d6f88dd4a         0      true
10.198.94.240   12321
FE:39:D7:A5:EF:80:D6:41:CD:13:70:BD:88:2D:38:6C:A0:1D:36:69
5e8e3d3f-3015-9075-49b6-a03d6f88d426         0      true
10.198.94.244   12321
72:A3:80:36:F7:5D:8F:CE:B0:26:02:96:00:23:7D:8E:C5:8C:0B:E1
5e8e3d33-5825-ee5c-013c-a03d6f88ea4c         0      true
10.198.95.11    12321
5A:55:74:E8:5F:40:2F:2B:09:B5:42:29:FF:1C:95:41:AB:28:E0:57
```

La sortie inclut l'UUID du nœud vSAN, l'adresse IPv4, l'adresse IPv6, le port UDP avec lequel le nœud vSAN communique et indique si le nœud est un hôte de données (0) ou un hôte témoin (1). Vous pouvez utiliser ce résultat pour identifier les nœuds de cluster vSAN qui fonctionnent en mode monodiffusion et afficher les autres hôtes du cluster. vCenter Server conserve la liste de sortie.

Afficher les informations sur le réseau vSAN

Utilisez la commande `esxcli vsan network list` pour afficher les informations de réseau vSAN, telles que l'interface VMkernel utilisée par vSAN pour la communication, le port monodiffusion (12321) et le type de trafic (vSAN ou témoin) associées à l'interface vSAN.

Procédure

- Exécutez la commande `esxcli vsan network list`.

Résultats

```
Interface
   VmkNic Name: vmk1
   IP Protocol: IP
   Interface UUID: e290be58-15fe-61e5-1043-246e962c24d0
   Agent Group Multicast Address: 224.2.3.4
   Agent Group IPv6 Multicast Address: ff19::2:3:4
   Agent Group Multicast Port: 23451
   Master Group Multicast Address: 224.1.2.3
   Master Group IPv6 Multicast Address: ff19::1:2:3
   Master Group Multicast Port: 12345
   Host Unicast Channel Bound Port: 12321
   Multicast TTL: 5
   Traffic Type: vsan
```

Ce résultat affiche également les informations de multidiffusion.

Trafic intra-cluster

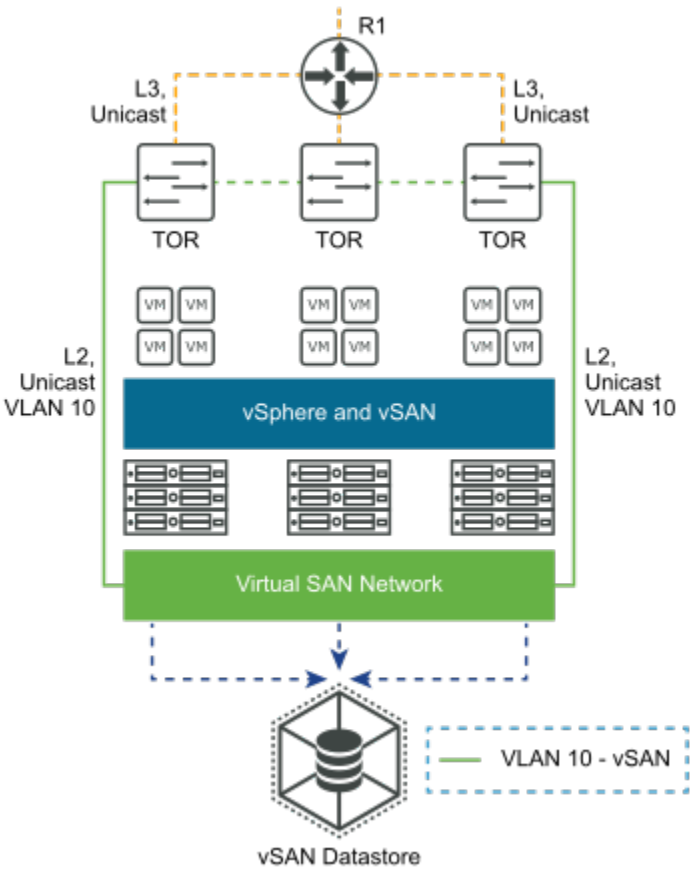
En mode monodiffusion, le nœud principal traite tous les nœuds du cluster, car il envoie le même message à tous les nœuds vSAN d'un cluster.

Par exemple, si N correspond au nombre de nœuds vSAN, le nœud principal envoie les messages N fois. Cela entraîne une légère augmentation du trafic de CMMDS vSAN. Vous ne remarquerez peut-être pas cette légère augmentation du trafic lors d'un fonctionnement normal en régime continu.

Trafic intra-cluster dans un seul rack

Si tous les nœuds d'un cluster vSAN sont connectés au même commutateur de haut de rack (TOR), l'augmentation totale du trafic se fait uniquement entre le nœud principal et le commutateur.

Si un cluster vSAN s'étend sur plusieurs commutateurs TOR, le trafic entre les commutateurs augmente. Si un cluster s'étend sur de nombreux racks, plusieurs TOR forment des domaines de pannes (FD) pour la détection des racks. Le nœud principal envoie N messages aux racks ou aux domaines de pannes, N correspondant au nombre d'hôtes dans chaque domaine de pannes.

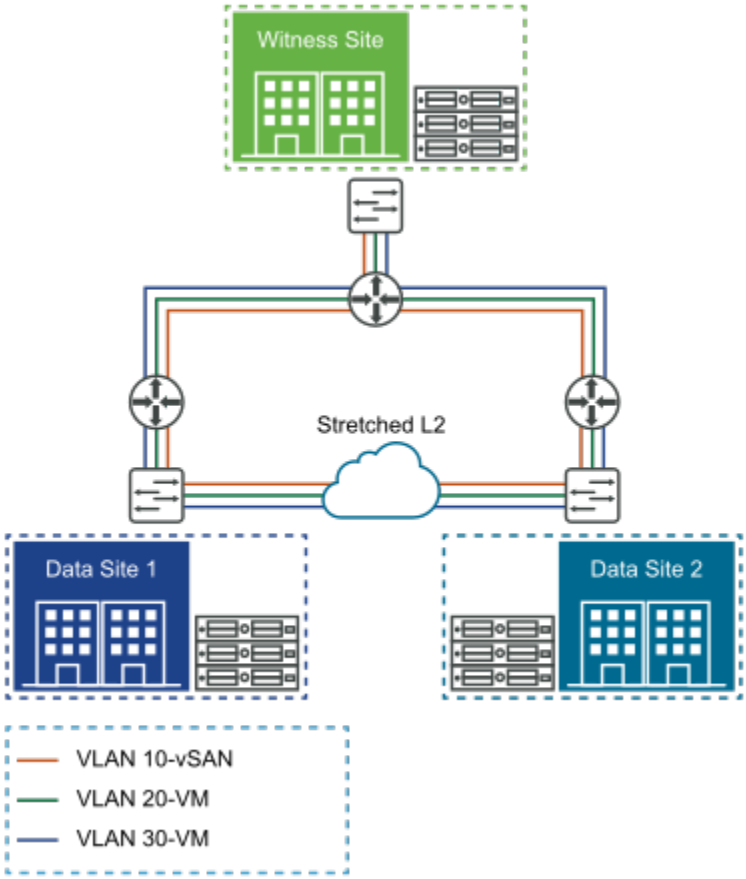


Trafic intra-cluster dans un cluster étendu

Dans un cluster étendu, le nœud principal se trouve sur le site préféré.

Dans un domaine de pannes, les données CMMDS doivent être communiquées depuis le site secondaire vers le site préféré. Pour calculer le trafic dans un cluster étendu, vous devez multiplier le nombre de nœuds d'un site secondaire par la taille du nœud CMMDS (en Mo) par le nombre de nœuds du site secondaire.

Trafic dans un cluster étendu = nombre de nœuds dans le site secondaire * taille du nœud CMMDS (en Mo) * nombre de nœuds dans le site secondaire.



Avec le trafic monodiffusion, il n'y a aucune modification des conditions de trafic du site témoin.

Configuration du transport réseau IP

5

Les protocoles de transport fournissent des services de communication sur le réseau. Ces services incluent le contrôle de la pile TCP/IP et de flux.

Ce chapitre contient les rubriques suivantes :

Piles TCP/IP vSphere

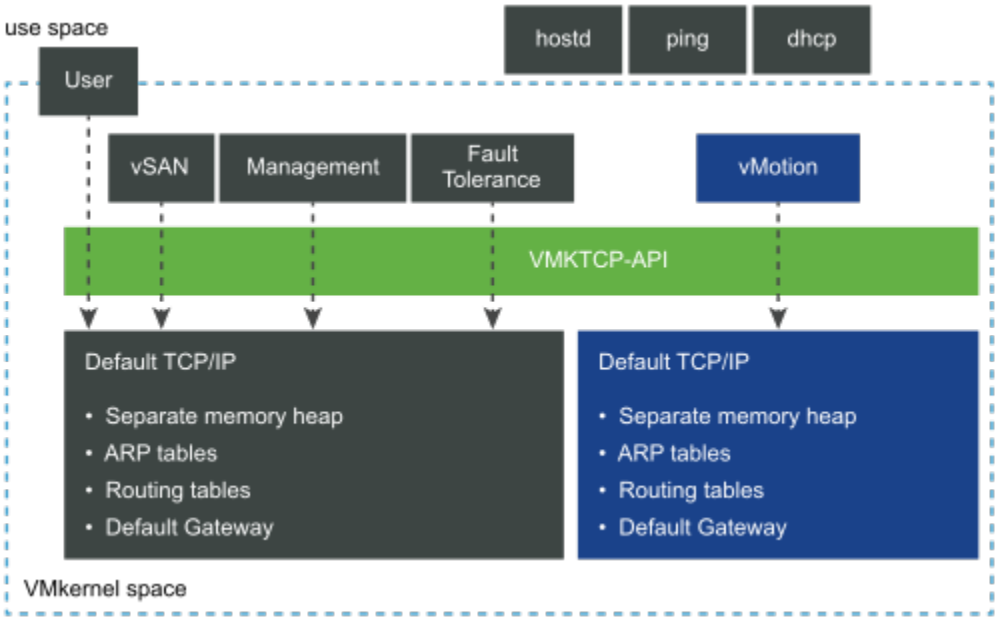
vSphere n'inclut pas de pile TCP/IP dédiée pour le service de trafic vSAN. Vous pouvez ajouter l'interface réseau VMkernel vSAN à la pile TCP/IP par défaut et définir des itinéraires statiques pour tous les hôtes du cluster vSAN.

vSphere ne prend pas en charge la création d'une pile TCP/IP vSAN personnalisée. Vous pouvez vérifier que le trafic vSAN dans les topologies de réseau de couche 3 sort par l'interface réseau VMkernel vSAN. Ajoutez l'interface réseau VMkernel vSAN à la pile TCP/IP par défaut et définissez des routes statiques pour tous les hôtes du cluster vSAN.

Note vSAN ne possède pas sa propre pile TCP/IP. Utilisez des routes statiques pour acheminer le trafic vSAN sur des réseaux L3.

vSphere 6.0 a introduit une nouvelle architecture de pile TCP/IP, qui peut utiliser plusieurs piles TPC/IP pour gérer différentes interfaces réseau VMkernel. Avec cette architecture, vous pouvez configurer des services de trafic tels que vMotion, la gestion et la tolérance de panne sur des piles TCP/IP isolées, qui peuvent utiliser plusieurs passerelles par défaut.

Pour les exigences relatives à l'isolation du trafic réseau et à la sécurité, déployez les différents services de trafic sur différents segments de réseau ou VLAN. Cela empêche les différents services de trafic de traverser la même passerelle par défaut.



Lorsque vous configurez les services de trafic sur des piles TCP/IP distinctes, déployez chaque type de service de trafic sur son propre segment de réseau. Les segments réseau sont accessibles via un adaptateur réseau physique avec la segmentation VLAN. Mappez chaque segment à des interfaces réseau VMkernel différentes avec les services de trafic respectifs activés.

Piles TCP/IP disponibles dans vSphere

vSphere fournit des piles TCP/IP qui prennent en charge les exigences relatives au trafic vSAN.

- **Pile TCP/IP par défaut**. Gérez les services de trafic liés à l'hôte. Cette pile partage une passerelle par défaut unique entre tous les services réseau configurés.
- **Pile TCP/IP vMotion**. Isole le trafic vMotion sur sa propre pile. L'utilisation de cette pile supprime ou désactive complètement le trafic vMotion de la pile TCP/IP par défaut.
- **Pile TCP/IP de provisionnement**. Isole certaines opérations liées à la machine virtuelle, telles que les migrations à froid, le clonage, le snapshot ou le trafic NFC.

Vous pouvez sélectionner une autre pile TCP/IP lors de la création d'une interface VMkernel.

Les environnements ayant des exigences réseau isolées pour les services de trafic vSphere ne peuvent pas utiliser la même passerelle par défaut pour diriger le trafic. L'utilisation de différentes piles TCP/IP simplifie la gestion de l'isolation du trafic, car vous pouvez utiliser différentes passerelles par défaut et éviter l'ajout d'itinéraires statiques. Utilisez cette technique lorsque vous devez acheminer le trafic vSAN vers un autre réseau qui n'est pas accessible sur la passerelle par défaut.

vSphere RDMA

vSAN 7.0 Update 2 et versions ultérieures prennent en charge l'accès direct à la mémoire à distance (RDMA, Remote Direct Memory Access).

RDMA offre un accès direct à la mémoire depuis celle d'un ordinateur vers la mémoire d'un autre ordinateur sans impliquer le système d'exploitation ni le CPU. Le transfert de la mémoire est déplacé vers les adaptateurs HCA (Host Channel Adapters) compatibles avec la technologie RDMA.

vSAN prend en charge le protocole RoCE v2. RoCE v2 nécessite un réseau configuré pour une opération sans perte.

Chaque hôte vSAN doit disposer d'une carte réseau certifiée vSAN prenant en charge RDMA, comme indiqué dans la section vSAN du Guide de compatibilité VMware. N'utilisez que les adaptateurs réseau du même modèle provenant du même fournisseur à chaque extrémité de la connexion.

Tous les hôtes du cluster doivent prendre en charge RDMA. Si la prise en charge de RDMA est interrompue pour n'importe quel hôte, le cluster vSAN entier bascule vers TCP.

vSAN avec RDMA prend en charge le basculement des cartes réseau, mais pas l'association de cartes réseau LACP ou basée sur le hachage/l'adresse IP.

Prise en charge d'IPv6

vSAN 6.2 et version ultérieure prend en charge IPv6.

vSAN prend en charge les versions IP suivantes.

- IPv4
- IPv6 (vSAN 6.2 et version ultérieure)
- IPv4/IPv6 mixte (vSAN 6.2 et version ultérieure)

Dans les versions antérieures à vSAN 6.2, seul IPv4 est pris en charge. Utilisez le mode mixte lors de la migration de votre cluster vSAN de IPv4 vers IPv6.

La multidiffusion IPv6 est également prise en charge. Toutefois, certaines restrictions s'appliquent avec l'écoute IPv6 et l'écoute IGMP sur Cisco ACI. Pour cette raison, n'implémentez pas IPv6 pour vSAN à l'aide de Cisco ACI.

Pour de plus amples informations sur l'utilisation d'IPv6, consultez votre fournisseur de réseau.

Routes statiques

Vous pouvez utiliser des routes statiques afin de permettre aux interfaces réseau vSAN des hôtes d'un sous-réseau de joindre les hôtes sur un autre réseau.

La plupart des organisations séparent le réseau vSAN du réseau de gestion, de sorte que le réseau vSAN ne comporte pas de passerelle par défaut. Dans un déploiement L3, les hôtes résidant sur des sous-réseaux différents ou des segments L2 différents ne peuvent pas se joindre mutuellement via la passerelle par défaut, qui est généralement associée au réseau de gestion.

Utilisez des *routes statiques* pour permettre aux interfaces réseau vSAN des hôtes d'un sous-réseau de joindre les réseaux vSAN sur les hôtes de l'autre réseau. Les itinéraires statiques indiquent à un hôte comment atteindre un réseau particulier sur une interface au lieu d'utiliser la passerelle par défaut.

L'exemple suivant illustre comment ajouter une route statique IPv4 à un hôte ESXi. Spécifiez la passerelle (-g) et le réseau (-n) que vous souhaitez atteindre via cette passerelle :

```
esxcli network ip route ipv4 add –g 172.16.10.253 –n 192.168.10.0/24
```

Lorsque les routes statiques ont été ajoutées, la connectivité du trafic vSAN est disponible sur tous les réseaux, en supposant que l'infrastructure physique le permet. Exécutez la commande `vmkping` pour tester et confirmer la communication entre les différents réseaux en exécutant une commande Ping sur l'adresse IP ou la passerelle par défaut du réseau distant. Vous pouvez également vérifier les différents paquets de taille (-s) et éviter la fragmentation (-d) du paquet.

```
vmkping –I vmk3 192.168.10.253
```

Trames jumbo

vSAN prend entièrement en charge les trames Jumbo sur le réseau vSAN.

Les trames Jumbo sont des trames Ethernet avec plus de 1 500 octets de charge utile. Les trames Jumbo comportent généralement jusqu'à 9 000 octets de charge utile, mais il existe des variations.

L'utilisation des trames Jumbo peut réduire l'utilisation du CPU et améliorer le débit.

Vous devez décider si ces gains sont plus importants que la surcharge occasionnée par l'implémentation des trames Jumbo dans le réseau. Dans les centres de données où les trames Jumbo sont déjà activées dans l'infrastructure réseau, vous pouvez les utiliser pour vSAN.
Le coût opérationnel de la configuration des trames Jumbo dans l'ensemble du réseau peut contrebalancer les avantages de CPU et de performances limités.

Utilisation de VMware NSX avec vSAN

6

vSAN et VMware NSX peuvent être déployés et coexister dans la même infrastructure vSphere.

NSX ne prend pas en charge la configuration du réseau de données vSAN sur une superposition VXLAN géré par NSX ou Geneve.

vSAN et NSX sont compatibles. vSAN et NSX ne dépendent pas l'un de l'autre pour fournir leurs fonctionnalités, ressources et services.

Cependant, vous ne pouvez pas placer le trafic réseau de vSAN sur une superposition VxLAN/Geneve gérée par NSX. NSX ne prend pas en charge la configuration du trafic réseau de données vSAN sur une superposition VxLAN/Geneve gérée par NSX.

Éviter toute dépendance circulaire entre les réseaux VMkernel et la superposition VxLAN qu'ils prennent en charge est l'une des raisons pour lesquelles le trafic VMkernel n'est pas pris en charge sur la superposition VxLAN gérée par NSX. Les réseaux logiques fournis avec la superposition VxLAN gérée par NSX sont utilisés par les machines virtuelles, qui exigent une mobilité et une flexibilité du réseau.

Lorsque vous implémentez le LACP/LAG dans NSX, le problème le plus important avec le LAG se produit lorsqu'il est utilisé dans un environnement Cisco Nexus qui définit les LAG en tant que canaux de ports virtuels (vPC). Disposer d'un VPC implique que vous ne pouvez pas exécuter un protocole de routage dynamique de périphériques Edge vers des commutateurs Cisco physiques, car Cisco ne prend pas en charge cette configuration.

Utilisation du contrôle de congestion et du contrôle de flux 7

Utilisez le contrôle de flux pour gérer le taux de transfert de données entre les expéditeurs et les récepteurs sur le réseau vSAN. Le contrôle de congestion gère l'encombrement du réseau.

Contrôle de flux

Vous pouvez utiliser le contrôle de flux pour gérer le taux de transfert de données entre deux périphériques.

Le contrôle de flux est configuré lorsque deux périphériques connectés physiquement effectuent une négociation automatique.

Un nœud réseau surchargé peut envoyer une trame de pause pour interrompre la transmission de l'expéditeur pendant une période spécifiée. Une trame avec une adresse de destination multidiffusion envoyée à un commutateur est transférée via tous les autres ports du commutateur. Les trames de pause possèdent une adresse de destination multidiffusion spéciale qui les distingue d'autres trafics multidiffusion. Un commutateur conforme ne transmet pas une trame de pause. Les trames envoyées à cette plage sont censées être traitées uniquement dans le commutateur. Les trames de pause ont une durée limitée et expirent après un certain laps de temps. Deux ordinateurs connectés via un commutateur ne s'envoient jamais de trames de pause entre eux, mais peuvent envoyer des trames de pause à un commutateur.

L'une des raisons d'utilisation des trames de pause consiste à prendre en charge les contrôleurs d'interface réseau (NIC) qui ne disposent pas d'une mémoire tampon suffisante pour gérer la réception à vitesse maximale. Ce problème n'est pas courant avec les progrès en matière de vitesses de bus et de taille de mémoire.

Contrôle de congestion

Le contrôle de congestion vous permet de contrôler le trafic sur le réseau.

Le contrôle de congestion s'applique principalement aux réseaux de commutation de paquets. La congestion du réseau au sein d'un commutateur peut être due à des liens inter-commutateur surchargés. Si les liaisons inter-commutateur surchargent la capacité sur la couche physique, le commutateur introduit des trames de pause pour se protéger.

Contrôle de flux prioritaire

Le contrôle de flux basé sur la priorité (PFC) vous permet d'éliminer la perte de trame en raison d'une congestion.

Le contrôle de flux basé sur la priorité (IEEE 802.1Qbb) est assuré par un mécanisme similaire à des trames suspendues, mais fonctionne sur des priorités individuelles. Le PFC est également appelé contrôle de flux basé sur la classe (CBFC) ou par pause de priorité (PPP).

Contrôle de flux et contrôle de surcharge

Le contrôle de flux est un mécanisme de bout en bout qui contrôle le trafic entre un expéditeur et un récepteur. Le contrôle de flux se produit dans la couche de liaison de données et la couche de transport.

Le contrôle de congestion est utilisé par un réseau pour contrôler l'encombrement du réseau. Ce problème n'est pas aussi courant dans les réseaux modernes ayant des progrès en matière de vitesses de bus et de taille de mémoire. Un scénario plus probable est la surcharge du réseau au sein d'un commutateur. Le contrôle de congestion est géré par la couche réseau et la couche de transport.

Considérations relatives à la conception du contrôle de flux

Par défaut, le contrôle de flux est activé sur toutes les interfaces réseau des hôtes ESXi.

La configuration du contrôle de flux sur une carte réseau est effectuée par le pilote. Lorsqu'une carte réseau est surchargée par le trafic réseau, la carte réseau envoie des trames de pause.

Les mécanismes de contrôle de flux, tels que les trames de pause, peuvent déclencher une latence globale dans les E/S invitées de machine virtuelle en raison de la latence accrue au niveau de la couche réseau vSAN. Certains pilotes réseau fournissent des options de module qui configurent la fonctionnalité de contrôle de flux dans le pilote. Certains pilotes réseau vous permettent de modifier les options de configuration à l'aide de l'utilitaire de ligne de commande `ethtool` sur la console de l'hôte ESXi. Utilisez les options de module ou les `ethtool`, selon les détails d'implémentation d'un pilote spécifique.

Pour plus d'informations sur la configuration du contrôle de flux sur les hôtes ESXi, consultez l'article 1013413 de la base de connaissances VMware.

Dans les déploiements de 1 Gbits/s, laissez le contrôle de flux activé sur les interfaces réseau ESXi (par défaut). Si des trames de pause posent problème, planifiez soigneusement la désactivation du contrôle de flux en conjonction avec la prise en charge du fournisseur de matériel ou les services VMware Global Support.

Pour découvrir comment vous pouvez reconnaître la présence de trames de pause envoyées d'un récepteur à un hôte ESXi, reportez-vous à la section Chapitre 12 Dépannage du réseau vSAN. Un grand nombre de trames de pause dans un environnement indique généralement un problème de réseau ou de transport sous-jacent à examiner.

Association de cartes réseau de base, basculement et équilibrage de charge

8

De nombreux environnements vSAN requièrent un certain niveau de redondance réseau.

Vous pouvez utiliser l'association de cartes réseau pour garantir la redondance du réseau. Vous pouvez configurer deux adaptateurs réseau (cartes réseau) ou plus en tant qu'équipe pour la haute disponibilité et l'équilibrage de charge. L'association de cartes réseau de base est disponible avec la mise en réseau vSphere et ces techniques peuvent affecter la conception et l'architecture de vSAN.

Plusieurs options d'association de cartes réseau sont disponibles. Évitez les stratégies d'association de cartes réseau qui nécessitent des configurations de commutateur physique ou qui requièrent une bonne compréhension des concepts de mise en réseau, tels que l'agrégation de liens. Les meilleurs résultats sont obtenus avec une configuration de base, simple et fiable.

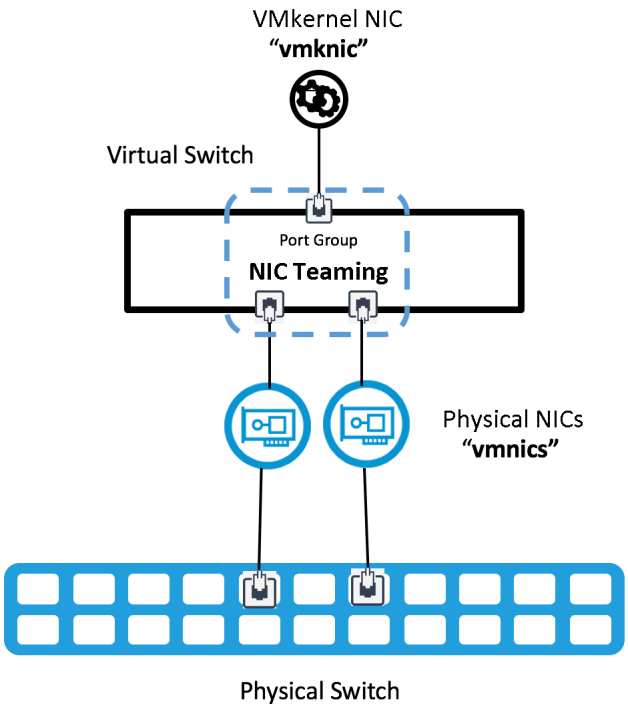
Si vous n'êtes pas certain des options d'association de cartes réseau, utilisez une configuration Actif/En veille avec basculement explicite.

Ce chapitre contient les rubriques suivantes :

Association de cartes réseau de base

L'association de cartes réseau de base utilise plusieurs liaisons montantes physiques, un vmknic et un commutateur unique.

L'association de cartes réseau vSphere utilise plusieurs adaptateurs de liaison montante, appelés vmnic, qui sont associés à un seul commutateur virtuel pour former une équipe. Il s'agit de l'option la plus simple. Vous pouvez la configurer à l'aide d'un commutateur vSphere Standard ou d'un vSphere Distributed Switch.



Basculement et redondance

vSAN peut utiliser l'association de cartes réseau et la stratégie de basculement de base fournies par vSphere.

L'association de cartes réseau sur un vSwitch peut avoir plusieurs liaisons montantes actives ou une configuration de liaison montante Actif/En veille. L'association de cartes réseau de base ne requiert aucune configuration spéciale au niveau de la couche de commutateur physique.

Note vSAN n'utilise pas l'association de cartes réseau pour l'équilibrage de charge.

Une configuration d'association de cartes réseau classique présente les paramètres suivants. Lorsque vous travaillez sur des commutateurs distribués, modifiez les paramètres du groupe de ports distribués utilisé pour le trafic vSAN.

- Équilibrage de charge : route basée sur le port virtuel d'origine
- Détection de panne réseau : état du lien uniquement
- Avertir les commutateurs : oui
- Retour arrière : oui

Équilibrage de charge du trafic vSAN.

- Équilibrage de charge : route basée sur le port virtuel d'origine
- Détection de panne réseau : état du lien uniquement
- Avertir les commutateurs : oui
- Retour arrière : oui

Configurer l'équilibrage de charge pour les associations de cartes réseau

Plusieurs techniques d'équilibrage de charge sont disponibles pour l'association de cartes réseau, et chaque technique a ses avantages et ses inconvénients.

Itinéraire basé sur le port virtuel d'origine

Dans les configurations de type Actif/Actif ou Actif/Passif, utilisez **Itinéraire basé sur le port virtuel d'origine** pour l'association des cartes réseau de base. Lorsque cette stratégie est effective, une seule carte réseau physique est utilisée par port VMkernel.

Avantages

- Il s'agit de la méthode d'association de cartes réseau la plus simple, qui requiert une configuration de commutateur physique minimale.
- Cette méthode n'exige qu'un seul port pour le trafic vSAN, ce qui simplifie le dépannage.

Inconvénients

- Une interface VMkernel unique est limitée à la bande passante d'une seule carte réseau physique. Étant donné que les environnements vSAN types utilisent un seul adaptateur VMkernel, une seule carte réseau physique est utilisée dans l'association.

Route basée sur la charge de carte réseau physique

La méthode **Route basée sur la charge de carte réseau physique** repose sur la méthode **Route basée sur le port virtuel d'origine**. Dans cette dernière, le commutateur virtuel vérifie la charge réelle des liaisons montantes et prend des mesures pour la réduire la charge sur les liaisons montantes surchargées. Cette méthode d'équilibrage de charge n'est disponible qu'avec un vSphere Distributed Switch, pas sur les commutateurs vSphere Standard.

Le commutateur distribué calcule les liaisons montantes pour chaque port VMkernel en utilisant l'ID de port et le nombre de liaisons montantes dans l'équipe de cartes réseau. Le Distributed Switch vérifie les liaisons montantes toutes les 30 secondes. Si la charge est supérieure à 75 %, l'ID de port du port VMkernel dont les E/S sont les plus élevées est déplacé vers une autre liaison montante.

Avantages

- Aucune configuration de commutateur physique n'est requise.
- Bien que vSAN possède un port VMkernel, les mêmes liaisons montantes peuvent être partagées par d'autres ports VMkernel ou services réseau. vSAN peut bénéficier de différentes liaisons montantes provenant d'autres services en conflit, tels que vMotion ou la gestion.

Inconvénients

- Étant donné que vSAN n'a généralement qu'un seul port VMkernel configuré, son efficacité est limitée.

- Le VMkernel ESXi réévalue la charge de trafic après chaque intervalle de temps, ce qui peut provoquer une surcharge de traitement.

Paramètres : détection de pannes réseau

Utilisez le paramètre par défaut : **État de lien seulement**. N'utilisez pas le sondage de balise pour la détection de défaillances de liaison. Le sondage de balise requiert au moins trois cartes réseau physiques pour éviter les scénarios split-brain. Pour de plus amples informations, consultez l'article 1005577 de la base de connaissances VMware.

Paramètres : avertir les commutateurs

Utilisez le paramètre par défaut : **Oui**. Les commutateurs physiques disposent de tables de transfert d'adresses MAC pour associer chaque adresse MAC à un port de commutateur physique. Lorsqu'une trame arrive, le commutateur détermine l'adresse MAC de destination dans la table et détermine le port physique approprié.

Si un basculement de carte réseau se produit, l'hôte ESXi doit informer les commutateurs réseau que quelque chose a changé ou le commutateur physique peut continuer à utiliser les anciennes informations et envoyer les trames au port incorrect.

Lorsque vous définissez Avertir les commutateurs sur **Oui**, si une carte réseau physique est défectueuse et que le trafic est réacheminé vers une autre carte réseau physique dans l'équipe, le commutateur virtuel envoie des notifications sur le réseau pour mettre à jour les tables de recherche sur les commutateurs physiques.

Ce paramètre n'intercepte pas les problèmes de configuration VLAN ou les pertes de liaison montante qui se produisent plus en amont dans le réseau. Le contrôle de santé des partitions de réseau vSAN peut détecter ces problèmes.

Paramètres : retour arrière

Cette option détermine le mode de retour en activité d'un adaptateur physique lors de la récupération après échec. Un événement de basculement déclenche le trafic réseau pour passer d'une carte réseau à une autre. Lorsqu'un état de **liaison montante** est détecté sur la carte réseau d'origine, le trafic revient automatiquement à l'adaptateur réseau d'origine lorsque le retour arrière est défini sur **Oui**. Lorsque le retour arrière est défini sur **Non**, un retour arrière manuel est nécessaire.

La définition du retour arrière sur **Non** peut être utile dans certains cas. Par exemple, après la récupération d'un port de commutateur physique suite à une panne, le port peut être actif, mais peut mettre plusieurs secondes à démarrer le transfert du trafic. Le retour arrière automatique a été connu pour provoquer des problèmes dans certains environnements qui utilisent le protocole STP (Spanning Tree Protocol). Pour de plus amples informations sur le protocole STP (Spanning Tree Protocol), consultez l'article 1003804 de la base de connaissances VMware.

Définition de l'ordre de basculement

L'ordre de basculement détermine les liens actifs pendant les opérations normales et les liens actifs en cas de basculement. Différentes configurations prises en charge sont possibles pour le réseau vSAN.

Liaisons montantes Actif/En veille : si une panne se produit lors d'une configuration Actif/En veille, le pilote de la carte réseau avertit vSphere d'un événement de liaison inactive sur la liaison montante 1. La liaison montante en veille 2 devient active et le trafic reprend sur la liaison montante 2.

Liaisons montantes Actif/Actif : si vous définissez l'ordre de basculement sur Actif/Actif, le port virtuel utilisé par le trafic vSAN ne peut pas utiliser simultanément les deux ports physiques.

Si la configuration d'association de cartes réseau pour la liaison montante 1 et la liaison montante 2 est active, il n'est pas nécessaire que la liaison montante en veille devienne active.

Note En cas d'utilisation d'une configuration de type Actif/Actif, assurez-vous que le retour arrière est défini sur **Non**. Pour de plus amples informations, consultez l'article 2072928 de la base de connaissances VMware.

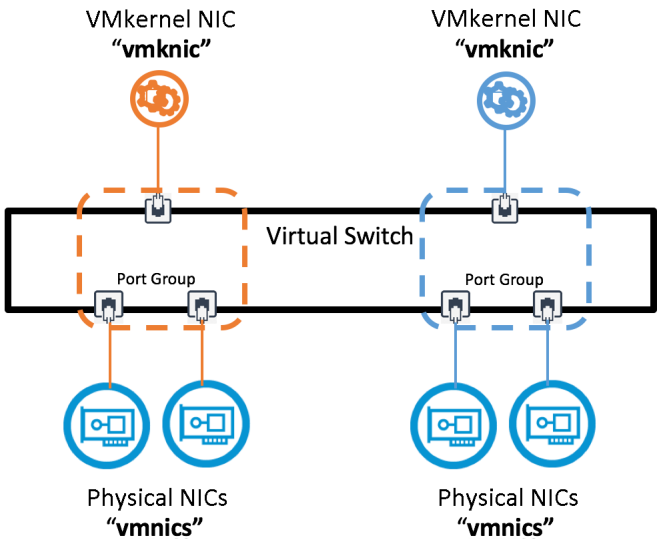
Association avancée de cartes réseau

9

Vous pouvez utiliser des méthodes d'association avancée de cartes réseau avec plusieurs adaptateurs VMkernel pour configurer le réseau vSAN. Si vous utilisez le protocole LAG/LACP (Link Aggregation Protocol), le réseau vSAN peut être configuré avec un seul adaptateur VMkernel.

Vous pouvez utiliser l'association avancée de cartes réseau pour implémenter un réseau isolé, de sorte qu'une panne survenant sur un chemin réseau n'a aucun impact sur le chemin d'accès au réseau. Si une partie d'un chemin réseau est défaillante, l'autre chemin réseau peut acheminer le trafic. Configurez plusieurs cartes réseau VMkernel pour vSAN sur des sous-réseaux différents, tels qu'un autre VLAN ou une infrastructure de réseau physique distincte.

vSphere et vSAN ne prennent pas en charge plusieurs adaptateurs VMkernel (vmknic) sur le même sous-réseau. Pour plus d'informations, consultez la base de connaissances VMware 2010877.



Ce chapitre contient les rubriques suivantes :

- Présentation du groupe d'agrégation de liens
- Présentation des séparations de réseau

- Avantages et inconvénients des configurations réseau isolées avec vSAN
- Exemples de configuration d'association de cartes réseau

Présentation du groupe d'agrégation de liens

En utilisant le protocole LACP, un périphérique réseau peut négocier un regroupement automatique de liens en envoyant des paquets LACP à un homologue.

Un groupe d'agrégation de liens (LAG) est défini par la norme IEEE 802.1 AX-2008, qui indique que l'agrégation de liens permet d'agréger un ou plusieurs liens pour former un groupe d'agrégation de liens.

Le LAG peut être configuré comme statique (manuel) ou dynamique en utilisant LACP pour négocier la formation du LAG. LACP peut être configuré comme suit :

Actif

Les périphériques envoient immédiatement des messages LACP lorsque le port apparaît. Les périphériques terminaux sur lesquels LACP est activé (par exemple, les hôtes ESXi et les commutateurs physiques) envoient et reçoivent mutuellement des trames appelées messages LACP afin de négocier la création d'un LAG.

Passif

Les périphériques placent un port dans un état de négociation passif, dans lequel le port répond uniquement aux messages LACP reçus, mais n'initie pas de négociation.

Note Si l'hôte et le commutateur sont tous deux en mode passif, le LAG ne s'initialise pas, car une partie active est nécessaire pour déclencher la liaison. Au moins l'un d'entre eux doit être actif.

Dans vSphere 5.5 et versions ultérieures, cette fonctionnalité est appelée **LACP amélioré**. Cette fonctionnalité est uniquement prise en charge sur vSphere Distributed Switch version 5.5 ou ultérieure.

Pour de plus amples informations sur la prise en charge du protocole LACP sur un vSphere Distributed Switch, consultez la documentation de mise en réseau de vSphere 6.

Note Le nombre de LAG pouvant être utilisés dépend de la capacité de l'environnement physique sous-jacent et de la topologie du réseau virtuel.

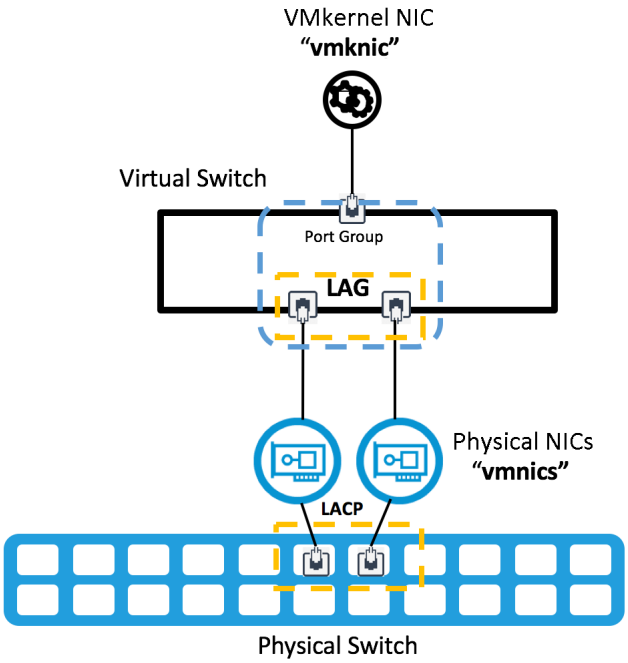
Pour plus d'informations sur les différentes options d'équilibrage de charge, consultez l'article 2051826 de la base de connaissances.

Agrégation de liens statique et dynamique

Vous pouvez utiliser LACP pour combiner et regrouper plusieurs connexions réseau.

Lorsque le protocole LACP est en mode **Actif** ou **Dynamique**, un commutateur physique envoie des messages LACP aux périphériques réseau, tels que les hôtes ESXi, pour négocier la création d'un groupe d'agrégation de liens (LAG).

Pour configurer l'agrégation de liens sur des hôtes à l'aide de vSphere Standard Switches (et de vSphere Distributed Switches antérieurs à v 5.5), configurez un groupe de canaux statiques sur le commutateur physique. Pour plus de détails, consultez la documentation de votre fournisseur.



Avantages et inconvénients de l'agrégation de liens dynamiques

Tenez compte des compromis liés à l'utilisation de l'agrégation de liens dynamique.

Avantages

Améliore les performances et la bande passante. Un hôte ou un port VMkernel vSAN peut communiquer avec de nombreux autres hôtes vSAN à l'aide de nombreuses options d'équilibrage de charge.

Garantit la redondance de l'adaptateur réseau. Si une carte réseau tombe en panne et que l'état du lien échoue, les autres cartes réseau de l'équipe continuent de transmettre le trafic.

Améliore l'équilibrage de charge du trafic. L'équilibrage de charge du trafic suite à des pannes est automatique et rapide.

Inconvénients

Flexibilité réduite. La configuration du commutateur physique exige que les ports de commutateur physique soient paramétrés dans une configuration de canal de port.

Complexité accrue. L'utilisation de plusieurs commutateurs pour produire une configuration de redondance physique complète est complexe. Les implémentations spécifiques du fournisseur compliquent la configuration.

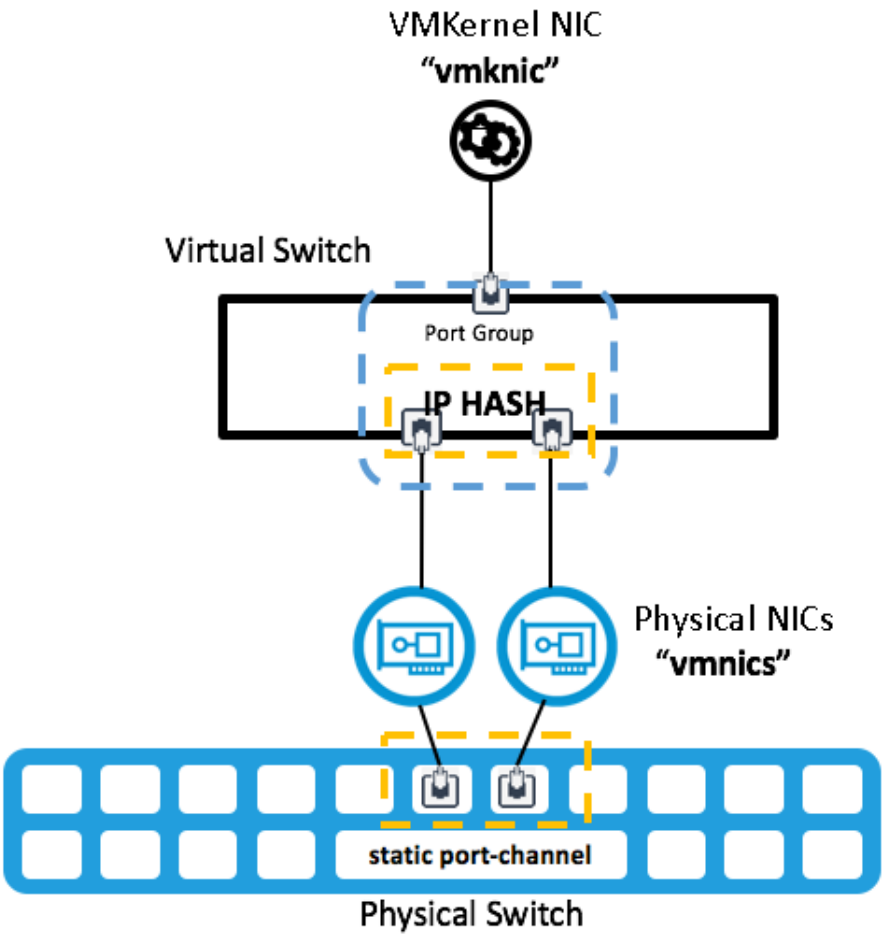
LACP statique avec route basée sur le hachage IP

Vous pouvez créer un cluster vSAN 6.6 à l'aide du protocole LACP statique avec une stratégie de hachage IP. Cette section porte sur les dispositifs vSphere Standard Switch. Toutefois, vous pouvez également utiliser des dispositifs vSphere Distributed Switches.

Vous pouvez utiliser l'itinéraire basé sur la stratégie d'équilibrage de charge à hachage IP. Pour de plus amples informations sur le hachage IP, consultez la documentation de vSphere.

Sélectionnez la stratégie d'équilibrage de charge par **Route basée sur le hachage IP** au niveau d'un vSwitch ou d'un groupe de ports. Définissez toutes les liaisons montantes affectées au groupe de canaux statiques sur la position de liaison montante active sur les stratégies d'association et de basculement au niveau du commutateur virtuel ou du groupe de ports.

Lorsque le hachage IP est configuré sur un groupe de ports vSphere, le groupe de ports utilise la stratégie **Route basée sur le hachage IP**. Le nombre de ports dans le canal de port doit être identique au nombre de liaisons montantes figurant dans l'équipe.



montantes figurant dans l'équipe. Avantages et inconvénients d'un LACP statique avec hachage IP

Tenez compte des compromis lors de l'utilisation du protocole LACP statique avec hachage IP.

Avantages

- **Améliore les performances et la bande passante**. Un hôte ou un port VMkernel vSAN peut communiquer avec de nombreux autres hôtes vSAN à l'aide de l'algorithme de hachage IP.
- **Garantit la redondance de l'adaptateur réseau**. Si une carte réseau tombe en panne et que l'état du lien échoue, les autres cartes réseau de l'équipe continuent de transmettre le trafic.
- **Accroît la flexibilité**. Vous pouvez utiliser le hachage IP avec vSphere Standard Switches et vSphere Distributed Switches.

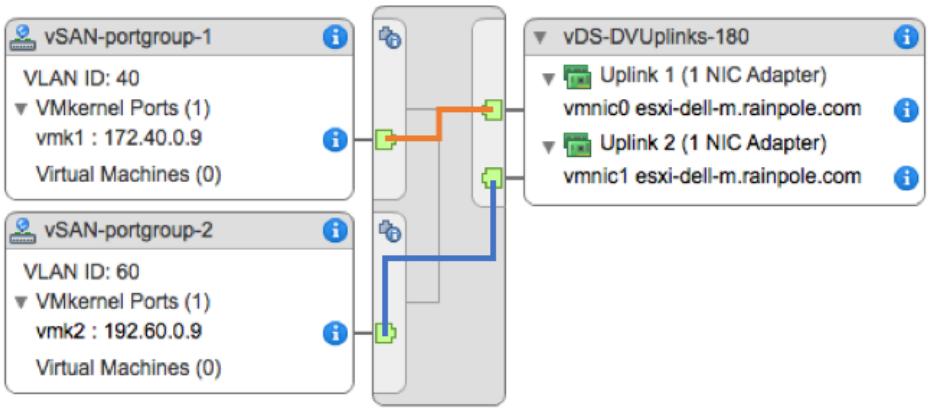
Inconvénients

- **La configuration du commutateur physique est moins flexible**. Les ports de commutateur physique doivent être paramétrés dans une configuration de canal de port statique.
- **Augmentation des risques de configuration incorrecte**. Formulaire de canaux de ports statiques sans vérification à l'une des extrémités (contrairement au canal de port dynamique LACP).
- **Complexité accrue**. L'introduction d'une configuration de redondance physique complète ajoute de la complexité lorsque plusieurs commutateurs sont utilisés. Les implémentations peuvent être très spécifiques au fournisseur.
- **Équilibrage de charge limité**. Si votre environnement ne contient que quelques adresses IP, le commutateur virtuel peut transmettre systématiquement le trafic via une liaison montante dans l'équipe. Cela peut être particulièrement le cas pour les clusters vSAN de petite taille.

Présentation des séparations de réseau

Vous pouvez utiliser des méthodes d'association de cartes réseau avancées pour créer une infrastructure de stockage isolée. Deux réseaux de stockage sont utilisés pour créer une topologie réseau de stockage redondante, chaque réseau de stockage étant physiquement et logiquement isolé de l'autre par une séparation.

Vous pouvez configurer un réseau isolé pour vSAN dans un environnement vSphere. Configurez plusieurs ports VMkernel par hôte vSAN. Associez chaque port VMkernel à des liaisons montantes physiques dédiées en utilisant un seul vSwitch ou plusieurs commutateurs virtuels, tels que vSphere Standard Switch ou vSphere Distributed Switch.



En règle générale, chaque liaison montante doit être connectée à une infrastructure physique entièrement redondante.

Cette topologie n'est pas idéale. La défaillance de composants, tels que des cartes réseau sur différents hôtes résidant sur le même réseau, peut provoquer une interruption des E/S de stockage. Pour éviter ce problème, mettez en œuvre la redondance de la carte réseau physique sur tous les hôtes et segments du réseau. L'exemple de configuration 2 traite de cette topologie en détail.

Ces configurations s'appliquent aux topologies L2 et L3, avec des configurations monodiffusion et multidiffusion.

Avantages et inconvénients des configurations réseau isolées avec vSAN

Les isolements de réseau peuvent être utiles pour séparer et isoler le trafic vSAN. Soyez prudent lorsque vous configurez cette topologie.

Avantages

- Séparation physique et logique du trafic vSAN.

Inconvénients

- vSAN ne prend pas en charge plusieurs adaptateurs VMkernel (cartes vmknic) sur le même sous-réseau. Pour de plus amples informations, consultez l'article 2010877 de la base de connaissances VMware.
- La configuration est complexe et sujette aux erreurs. Pour cette raison, le dépannage est plus complexe.
- La disponibilité du réseau n'est pas garantie avec plusieurs vmknic dans certaines pannes asymétriques, telles qu'une défaillance de carte réseau sur un hôte et une autre défaillance de carte réseau sur un autre hôte.
- Le trafic vSAN à équilibrage de charge entre les cartes réseau physiques n'est pas garanti.

- Augmentation des coûts des hôtes vSAN, car vous aurez peut-être besoin de plusieurs adaptateurs VMkernel (vmknic) pour protéger diverses cartes réseau physiques (vmnic). Par exemple, 2 x 2 vmnic peuvent être nécessaires pour garantir la redondance de deux vmknic vSAN.
- Les ressources logiques requises sont doublées, notamment les ports VMkernel, les adresses IP et les réseaux VLAN.
- vSAN n'implémente pas la liaison de port. Autrement dit, les techniques telles que les chemins d'accès multiples ne sont pas disponibles.
- Les topologies de couche 3 ne conviennent pas au trafic vSAN comptant plusieurs vmknic. Ces topologies peuvent ne pas fonctionner comme prévu.
- La configuration de l'hôte de ligne de commande peut être nécessaire pour modifier les adresses multidiffusion vSAN.

Le protocole LACP dynamique combine, ou regroupe, plusieurs connexions réseau en parallèle pour augmenter le débit et garantir la redondance. Lorsque l'association de cartes réseau est configurée avec LACP, l'équilibrage de charge du réseau vSAN sur plusieurs liaisons montantes se produit. Cet équilibrage de charge se produit au niveau de la couche réseau et n'est pas effectué par vSAN.

Note Les autres termes utilisés pour décrire l'agrégation de liens incluent la jonction de ports, le regroupement de liaisons, le couplage Ethernet/réseau/carte réseau, EtherChannel.

Cette section porte sur le protocole LACP (Link Aggregation Control Protocol). La norme IEEE est 802.3ad, mais certains fournisseurs disposent de fonctionnalités LACP exclusives, telles que PAgP (Port Aggregation Protocol). Suivez les meilleures pratiques recommandées par votre fournisseur.

Note La prise en charge du protocole LACP introduite dans vSphere Distributed Switch 5.1 prend uniquement en charge l'équilibrage de charge à hachage IP. vSphere Distributed Switch 5.5 et versions ultérieures prennent entièrement en charge le protocole LACP.

LACP est une norme de l'industrie utilisant des canaux de ports. De nombreux algorithmes de hachage sont disponibles. La stratégie de groupe de ports vSwitch et la configuration du canal de port doivent être compatibles et correspondre.

Exemples de configuration d'association de cartes réseau

Les configurations d'association de cartes réseau suivantes illustrent les scénarios types de mise en réseau vSAN.

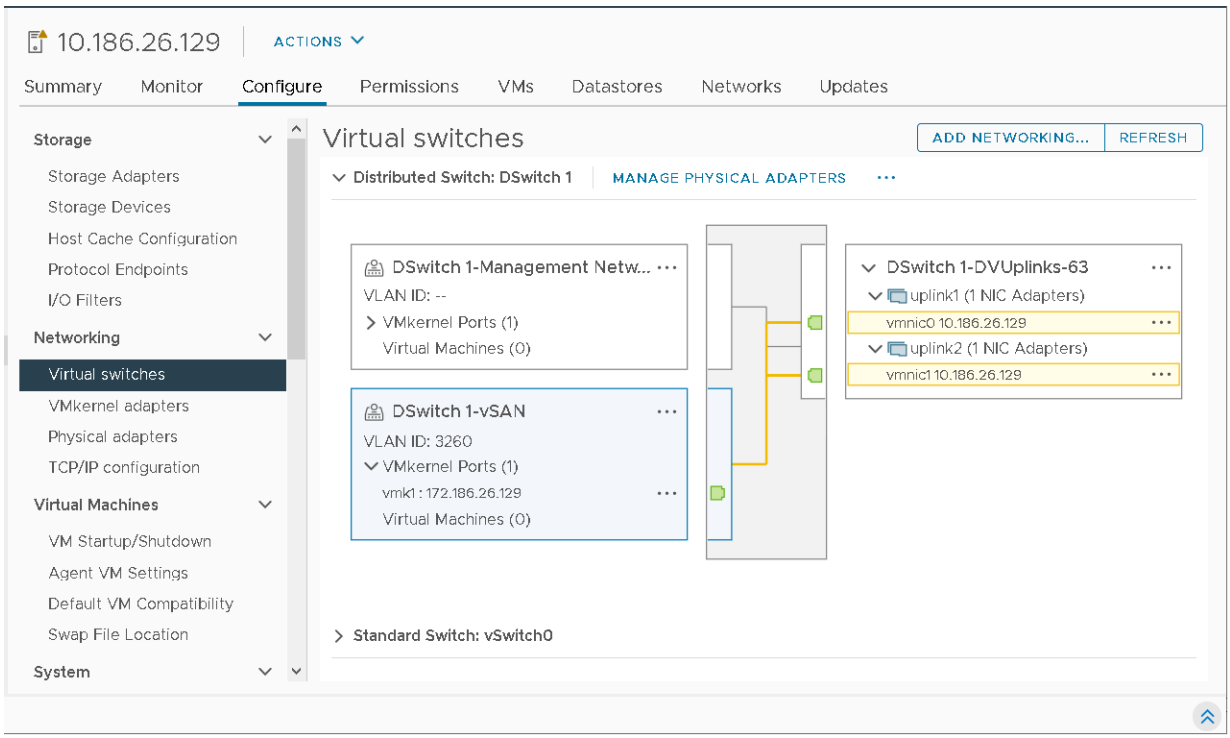
Configuration 1 : vmknic unique, route basée sur la charge de carte réseau physique

Vous pouvez configurer l'association de cartes réseau Actif/Actif de base avec la stratégie **Itinéraire basé sur la charge de carte réseau physique** pour les hôtes vSAN. Utilisez un commutateur vSphere Distributed Switch (vDS).

Dans cet exemple, le vDS doit disposer de deux liaisons montantes configurées pour chaque hôte. Un groupe de ports distribués est désigné pour le trafic vSAN et isolé sur un réseau VLAN spécifique. Les trames Jumbo sont déjà activées sur le vDS avec une valeur de MTU de 9 000.

Configurez l'association et le basculement pour le trafic du groupe de ports distribués pour le trafic vSAN comme suit :

- Stratégie d'équilibrage de charge définie sur **Itinéraire basé sur la charge de carte réseau physique**.
- La Détection de panne réseau est définie sur **État de lien seulement**.
- Avertir les commutateurs défini sur **Oui**.
- Retour arrière défini sur **Non**. Vous pouvez définir le retour arrière sur **Oui**, mais pas dans cet exemple.
- Assurez-vous que les deux liaisons montantes se trouvent dans la position **Liaisons montantes actives**.



Perte de redondance de la liaison montante du réseau

Lorsque l'état inactif du lien est détecté, la charge de travail passe d'une liaison montante à une autre. Il n'existe aucun impact notable sur le cluster vSAN et sur la charge de travail de la VM.

Récupération et retour arrière

Lorsque vous définissez **Retour arrière** sur **Non**, le trafic n'est pas rétabli sur le vmnic d'origine. Si **Retour arrière** est défini sur **Oui**, le trafic est rétabli sur le vmnic d'origine lors de la récupération.

Équilibrage de charge

Étant donné qu'il s'agit d'une carte réseau VMkernel unique, il n'y a aucun avantage à utiliser **Route basée sur la charge physique** en termes de performances.

Une seule carte réseau physique est utilisée à la fois. L'autre carte réseau physique est inactive.

Configuration 2 : plusieurs vmknic, Route basée sur l'ID de port d'origine

Vous pouvez utiliser deux réseaux VLAN non routables logiquement et physiquement séparés afin de produire une topologie isolée.

Cet exemple fournit les étapes de configuration pour un vSphere Distributed Switch. Cependant, vous pouvez également utiliser des commutateurs vSphere Standard. Cela nécessite deux cartes réseau physiques de 10 Go et les sépare logiquement sur la couche de mise en réseau vSphere.

Créez deux groupes de ports distribués pour chaque vmknic VMkernel vSAN. Chaque groupe de ports dispose d'une balise VLAN distincte. Pour la configuration du VMkernel vSAN, deux adresses IP sur les deux réseaux VLAN sont nécessaires pour le trafic vSAN.

Note Les implémentations pratiques utilisent généralement quatre liaisons montantes physiques pour une redondance complète.

Pour chaque groupe de ports, la stratégie d'association et de basculement utilise les paramètres par défaut.

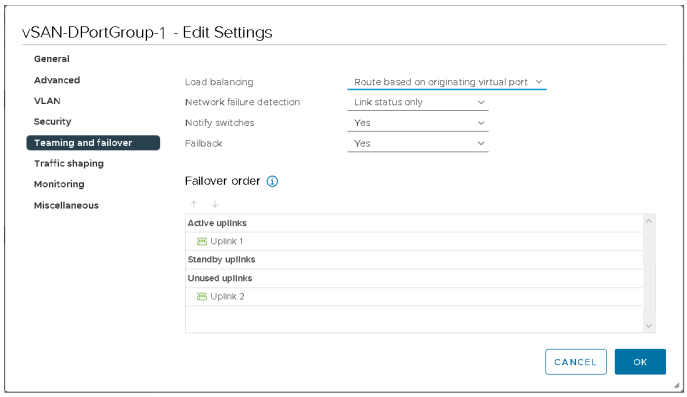
- Équilibrage de charge défini sur **Route basée sur l'ID de port d'origine**
- Détection de panne réseau définie sur **État de lien seulement**
- Avertir les commutateurs défini sur la valeur par défaut **Oui**
- Le retour arrière est défini sur la valeur par défaut **Oui**
- La configuration de liaison montante présente une liaison montante dans la position **Actif** et une liaison montante dans la position **Inutilisé**.

Un réseau est complètement isolé de l'autre réseau.

Groupe de ports 1 vSAN

Cet exemple utilise un groupe de ports distribués appelé **vSAN-DPortGroup-1**. **VLAN 3266** est marqué pour ce groupe de ports avec la stratégie d'association et de basculement suivante :

- Trafic sur le groupe de ports marqué avec VLAN 3266
- Équilibrage de charge défini sur **Route basée sur l'ID de port d'origine**
- Détection de panne réseau définie sur **État de lien seulement**
- Avertir les commutateurs défini sur la valeur par défaut **Oui**
- Retour arrière défini sur la valeur par défaut **Oui**
- La configuration de liaison montante indique le paramètre **liaison montante 1** défini dans la position **Actif** et le paramètre **liaison montante 2** défini dans la position **Inutilisé**.



Groupe de ports 2 vSAN

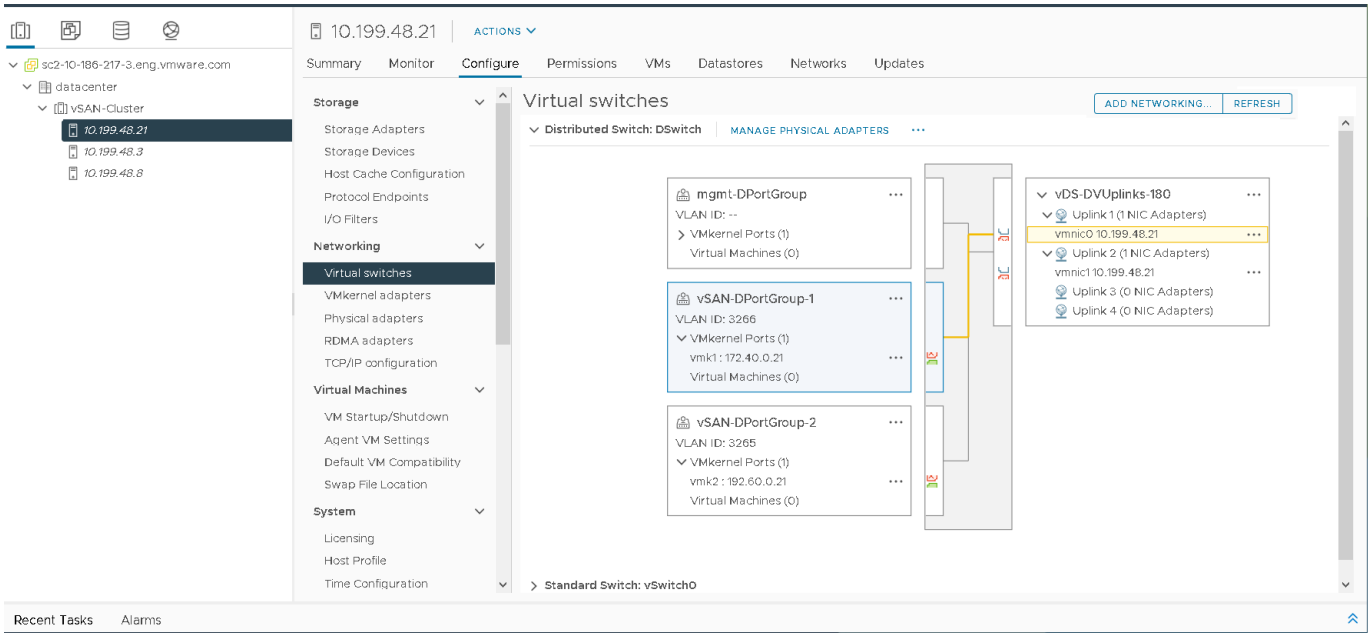
Pour compléter le groupe de ports 1 vSAN, configurez un deuxième groupe de ports distribués appelé **vSAN-portgroup-2** avec les différences suivantes :

- Trafic sur le groupe de ports marqué avec VLAN 3265
- La configuration de liaison montante indique le paramètre **Liaison montante 2** défini dans la position **Actif** et le paramètre **Liaison montante 1** défini dans la position **Inutilisé** .

Configuration du port VMkernel vSAN

Créez deux interfaces VMkernel vSAN et sur les deux groupes de ports. Dans cet exemple, les groupes de ports sont nommés **vmk1** et **vmk2**.

- **vmk1** est associé à VLAN 3266 (172.40.0.xx) et par conséquent, au groupe de ports **vSAN-DPortGroup-1**.
- **vmk2** est associé à VLAN 3265 (192.60.0.xx) et, par conséquent, au groupe de ports **vSAN-DPortGroup-2**.



Équilibrage de charge

vSAN ne dispose pas d'un mécanisme d'équilibrage de charge pour différencier plusieurs vmknic. Le chemin des E/S vSAN choisi n'est donc pas déterministe sur l'ensemble des cartes réseau physiques. Les graphiques de performances de vSphere montrent qu'une carte réseau physique est souvent plus utilisée que l'autre. Un test d'E/S simple effectué dans nos laboratoires, utilisant 120 machines virtuelles avec un taux de lecture/écriture de 70:30 avec une taille de bloc de 64 000 sur un cluster vSAN100 % Flash à quatre hôtes, a révélé un déséquilibrage de la charge sur les cartes réseau.

Les graphiques de performances de vSphere affichent une charge non équilibrée entre les cartes réseau.

Perte de redondance de la liaison montante du réseau

Supposez une panne de réseau introduite dans cette configuration. vmnic1 n'est pas activé sur un hôte vSAN spécifique. Par conséquent, le port **vmk2** est affecté. Une carte réseau défaillante déclenche les alarmes de connectivité réseau et les alarmes de redondance.

Pour vSAN, ce processus de basculement se déclenche approximativement **10 secondes** après que CMMDS (services de surveillance, d'appartenance et d'annuaire de cluster) a détecté une panne. Lors du basculement et de la récupération, vSAN arrête toutes les connexions actives sur le réseau défaillant et tente de rétablir les connexions sur le réseau fonctionnel restant.

Étant donné que deux ports VMkernel vSAN distincts communiquent sur des réseaux VLAN isolés, des pannes de contrôle de santé vSAN peuvent être déclenchées. Cela est normal, car **vmk2** ne peut plus communiquer avec ses homologues sur VLAN 3265.

Les graphiques de performances montrent que la charge de travail affectée a été redémarrée sur vmnic0, car vmnic1 présente une panne. Ce test illustre une distinction importante entre l'association de cartes réseau vSphere et cette topologie. vSAN tente de rétablir ou de redémarrer les connexions sur le réseau restant.

Cependant, dans certains scénarios de panne, la récupération des connexions affectées peut nécessiter jusqu'à **90 secondes** pour être exécutée en raison de l'expiration du délai de connexion à TCP ESXi. Les tentatives de connexion ultérieures peuvent échouer, mais les tentatives de connexion expirent à 5 secondes et les tentatives passent en revue toutes les adresses IP possibles. Ce comportement peut affecter les E/S invitées de la machine virtuelle. Par conséquent, les E/S d'application et de machine virtuelle devront éventuellement être réessayées.

Par exemple, sur les machines virtuelles Windows Server 2012, les ID d'événement 153 (réinitialisation de périphérique) et 129 (événements de nouvelle tentative) peuvent être consignés pendant le processus de basculement et de récupération. Dans l'exemple, l'ID d'événement 129 a été consigné pendant environ 90 secondes jusqu'à ce que les E/S aient été récupérées.

Vous devrez peut-être modifier les paramètres de délai d'expiration de disque de certains systèmes d'inactivité invités pour vous assurer qu'ils ne sont pas sérieusement affectés. Les valeurs de délai d'expiration de disque peuvent varier en fonction de la présence de VMware Tools, ainsi que du type et de la version spécifiques du système d'exploitation invité. Pour de plus amples informations sur la modification des valeurs de délai d'expiration du disque du système d'exploitation invité, accédez à l'article 1009465 de la base de connaissances VMware.

Récupération et retour arrière

Lorsque le réseau est réparé, les charges de travail ne sont pas automatiquement rééquilibrées, sauf si une autre erreur d'imposition de la charge de travail se produit, en raison d'une autre panne. Dès que le réseau affecté est récupéré, il devient disponible pour les nouvelles connexions TCP.

Configuration 3 : LACP dynamique

Vous pouvez configurer un canal de port LACP à deux ports sur un commutateur et un groupe d'agrégation de liens à deux liaisons montantes sur un vSphere Distributed Switch.

Dans cet exemple, utilisez la mise en réseau 10 Gb avec deux liaisons montantes physiques par serveur.

Note vSAN sur RDMA ne prend pas en charge cette configuration.

Configurer le commutateur réseau

Configurez le vSphere distributed switch avec les paramètres suivants.

- Identifiez les ports en question sur lesquels l'hôte vSAN se connectera.
- Créez un canal de port.
- Si vous utilisez plusieurs réseaux VLAN, joignez le réseau VLAN approprié au canal de port.

- Configurez les options de distribution ou d'équilibrage de charge souhaitées (hachage).
- Configuration du mode LACP sur Actif/Dynamique.
- Vérifiez la configuration du MTU.

Configurer vSphere

Configurez le réseau vSphere avec les paramètres suivants.

- Configurez vDS avec le MTU approprié.
- Ajoutez des hôtes au vDS.
- Créez un LAG avec le nombre approprié de liaisons montantes et des attributs correspondant au canal de port.
- Attribuez des liaisons montantes physiques au LAG.
- Créez un groupe de ports distribués pour le trafic vSAN et attribuez un réseau VLAN approprié.
- Configurez des ports VMkernel pour vSAN avec un MTU correct.

Configurer le commutateur physique

Configurez le commutateur physique avec les paramètres suivants. Pour obtenir des instructions sur cette configuration sur les serveurs Dell, rendez-vous sur : http://www.dell.com/support/article/fr/fr/19/HOW10364.

Configurez un LAG à deux liaisons montantes :

- Utilisez les ports de commutateur 36 et 18.
- Cette configuration utilise la jonction VLAN, afin que le canal de port soit en mode de jonction VLAN avec les VLAN joints appropriés.
- Utilisez la méthode suivante pour l'équilibrage de charge ou la distribution de charge : **Adresses IP source et de destination, port TCP/UDP et VLAN**
- Vérifiez que le mode LACP est **Actif** (Dynamique).

Utilisez les commandes suivantes pour configurer un canal de port individuel sur un commutateur Dell :

- Créez un canal de port.

```
#interface port-channel 1
```

- Définissez le canal de port sur le mode de jonction VLAN.

```
#switchport mode trunk
```

- Autorisez l'accès au VLAN.

```
#switchport trunk allowed vlan 3262
```

- Configurez l'option d'équilibrage de charge.

```
#hashing-mode 6
```

- Attribuez les ports appropriés au canal de port et définissez le mode sur Actif.
- Vérifiez que le canal de port est correctement configuré.

```
#show interfaces port-channel 1


Channel Ports Ch-Type Hash Type Min-links Local Prf

------- ---------------------------- -------- --------- --------- ----------

Po1 Active: Te1/0/36, Te1/0/18 Dynamic 6 1 Disabled

Hash Algorithm Type

1 - Source MAC, VLAN, EtherType, source module and port Id

2 - Destination MAC, VLAN, EtherType, source module and port Id

3 - Source IP and source TCP/UDP port

4 - Destination IP and destination TCP/UDP port

5 - Source/Destination MAC, VLAN, EtherType, source MODID/port

6 - Source/Destination IP and source/destination TCP/UDP port

7 - Enhanced hashing mode
#interface range Te1/0/36, Te1/0/18

#channel-group 1 mode active
```

Configuration complète :

```
#interface port-channel 1

#switchport mode trunk

#switchport trunk allowed vlan 3262

#hashing-mode 6

#exit

#interface range Te1/0/36,Te1/018

#channel-group 1 mode active

#show interfaces port-channel 1
```

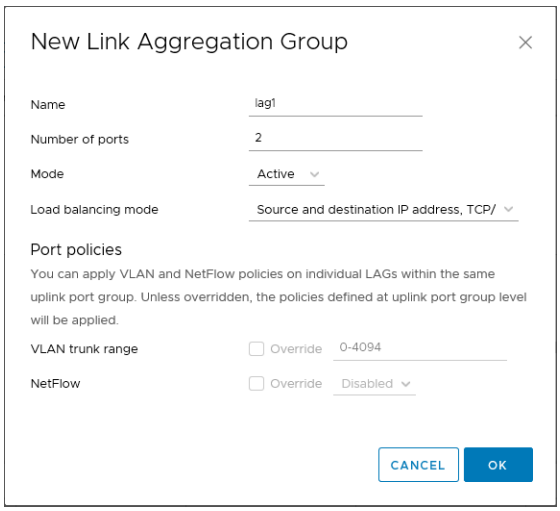
Note Répétez cette procédure sur tous les ports de commutateur participants qui sont connectés aux hôtes vSAN.

Configurer vSphere Distributed Switch

Avant de commencer, vérifiez que le vDS est mis à niveau vers une version prenant en charge le protocole LACP. Pour cela, cliquez avec le bouton droit sur le vDS et vérifiez si l'option de mise à niveau est disponible. Vous devrez peut-être mettre à niveau le vDS vers une version prenant en charge LACP.

Créer un LAG sur le vDS

Pour créer un LAG sur un Distributed Switch, sélectionnez le vDS, cliquez sur l'onglet **Configurer**, puis sélectionnez **LACP**. Ajoutez un nouveau LAG.



Configurez le LAG avec les propriétés suivantes :

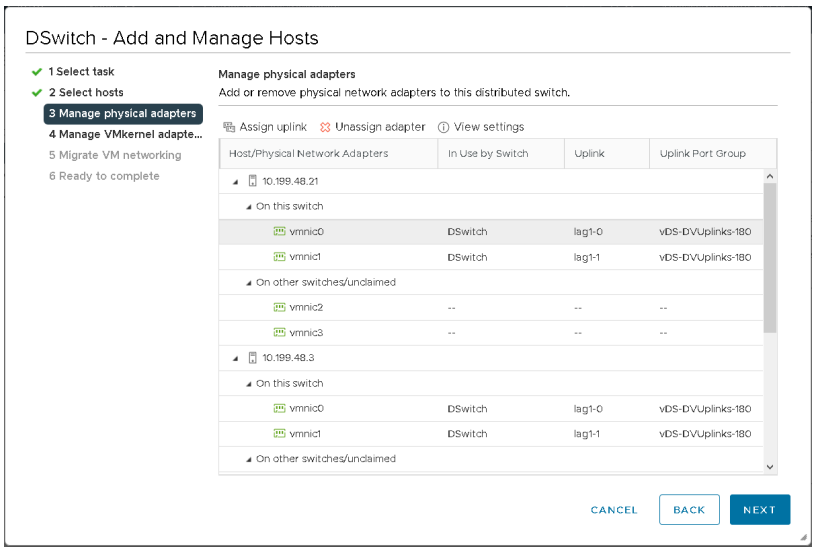
- Nom du LAG : **lag1**
- Nombre de ports : **2** (pour associer le canal de port sur le commutateur)
- Mode : **Actif** pour associer le commutateur physique.
- Mode d'équilibrage de charge : **Adresses IP source et de destination, port TCP/UDP et VLAN**

Ajouter des liaisons montantes physiques au LAG

vSAN hôtes ont été ajoutés au vDS. Attribuez la vmnic individuelle aux ports LAG appropriés.

- Cliquez avec le bouton droit sur le vDS et sélectionnez **Ajouter et gérer des hôtes...**
- Sélectionnez **Gérer la mise en réseau de l'hôte** et ajoutez vos hôtes attachés.
- Dans **Gérer les adaptateurs physiques**, sélectionnez les adaptateurs appropriés et attribuez-les au port LAG.
- Migrez vmnic0 de la position de la liaison montante 1 au port 0 sur LAG1.

Répétez la procédure pour vmnic1 sur la deuxième position de port de LAG, lag1-1.



Configurer la stratégie d'association et de basculement d'un groupe de ports distribués

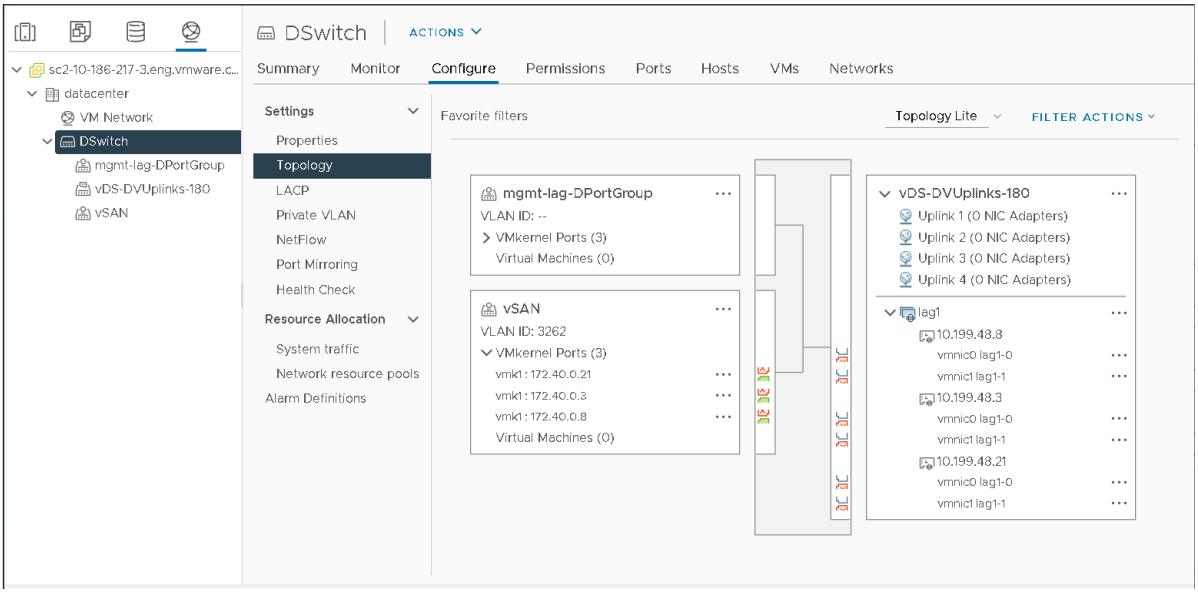
Affectez le groupe LAG en tant que **liaison montante active** sur la stratégie d'association et de basculement d'un groupe de ports distribués. Sélectionnez ou créez le groupe de ports distribués désigné pour le trafic vSAN. Cette configuration utilise un groupe de ports vSAN appelé **vSAN** avec l'ID de VLAN 3262 marqué. Modifiez le groupe de ports et configurez la stratégie d'association et de basculement pour refléter la nouvelle configuration du LAG.

Assurez-vous que le groupe LAG **lag1** se trouve dans la position Liaisons montantes actives et vérifiez que les autres liaisons montantes sont dans la position **Inutilisé**.

Note Lorsqu'un groupe d'agrégation de liens (LAG) est sélectionné comme unique liaison montante active, le mode d'équilibrage de charge du LAG remplace le mode d'équilibrage de charge du groupe de ports. Par conséquent, la stratégie suivante ne joue aucun rôle : **Route basée sur le port virtuel d'origine**.

Créer les interfaces VMkernel

L'étape finale consiste à créer les interfaces VMkernel pour utiliser le nouveau groupe de ports distribués, en veillant à ce qu'ils soient marqués pour le trafic vSAN. Notez que chaque vmknic vSAN peut communiquer sur vmnic0 et vmnic1 sur un groupe LAG pour garantir l'équilibrage de charge et le basculement.



Configurer l'équilibrage de charge

Du point de vue de l'équilibrage de charge, il n'existe pas d'équilibrage systématique du trafic dans l'ensemble des hôtes sur tous les vmnic dans cette configuration LAG, mais la cohérence est supérieure par rapport à la méthode **Itinéraire basé sur la charge de carte réseau physique** utilisée dans la configuration 1 et la méthode de vmknic isolés/multiples utilisée dans la configuration 2.

Le graphique des performances vSphere des hôtes individuels révèle un meilleur équilibrage de charge.

Perte de redondance de la liaison montante du réseau

Lorsque vmnic1 n'est pas activé sur un hôte vSAN spécifique, une alarme de redondance du réseau est déclenchée.

Aucune alarme de santé de vSAN n'est déclenchée et l'impact sur les E/S du client est minime par rapport à la configuration multi-vmknic isolée. Il n'est pas nécessaire que cette configuration arrête les sessions TCP avec LACP configuré.

Récupération et retour arrière

Dans un scénario de retour arrière, le comportement est différent entre l'association basée sur la charge, plusieurs vmknic et LACP dans un environnement vSAN. Après la récupération de vmnic1, le trafic est automatiquement équilibré sur les deux liaisons montantes actives. Ce comportement peut être avantageux pour le trafic vSAN.

Le retour arrière est-il défini sur Oui ou sur Non ?

Une stratégie d'équilibrage de charge du LAG remplace la stratégie d'association et de basculement pour les groupes de ports distribués vSphere. Tenez également compte des recommandations relatives à la valeur de retour arrière. Les tests de laboratoire n'affichent aucune différence de comportement visible entre le retour arrière défini sur **Oui** ou sur **Non** avec LACP. Les paramètres du LAG sont prioritaires sur les paramètres du groupe de ports.

Note Les valeurs de détection de panne réseau demeurent **état de lien seulement**, car la détection de balise n'est pas prise en charge avec LACP. Consultez l'article Présentation de l'équilibrage de charge par hachage IP (2006129) de la base de connaissances VMware

Configuration 4 : LACP statique – Route basée sur le hachage IP

Vous pouvez utiliser un canal de port statique LACP à deux ports sur un commutateur et deux liaisons montantes actives sur un vSphere Standard Switch.

Dans cette configuration, utilisez la mise en réseau 10 Gb avec deux liaisons montantes physiques par serveur. Une interface VMkernel unique (vmknic) pour vSAN existe sur chaque hôte.

Pour de plus amples informations sur les exigences de l'hôte et les exemples de configuration, consultez les articles suivants de la base de connaissances VMware :

- Exigences de l'hôte pour l'agrégation de liens pour ESXi et ESX (1001938)
- Exemple de configuration d'EtherChannel/Link Aggregation Control Protocol (LACP) avec les commutateurs ESXi/ESX et Cisco/HP (article 1004048 de la base de connaissances)

Note vSAN sur RDMA ne prend pas en charge cette configuration.

Configurer le commutateur physique

Configurez un canal de port statique à deux liaisons montantes comme suit :

- Ports de commutateur 43 et 44
- Jonction VLAN, le canal de port est par conséquent en mode de jonction VLAN, avec les réseaux VLAN joints appropriés.
- Ne spécifiez pas la stratégie d'équilibrage de charge sur le groupe de canaux de ports.

Ces étapes peuvent être utilisées pour configurer un canal de port individuel sur le commutateur :

Étape 1 : créez un canal de port.

```
#interface port-channel 13
```

Étape 2 : définissez le canal de port sur le mode de jonction VLAN.

```
#switchport mode trunk
```

Étape 3 : autorisez les réseaux VLAN appropriés.

```
#switchport trunk allowed vlan 3266
```

Étape 4 : affectez les ports appropriés au canal de port et définissez le mode sur Actif.

```
#interface range Te1/0/43, Te1/0/44

#channel-group 1 mode on
```

Étape 5 : vérifiez que le canal de port est configuré en tant que canal de port statique.

```
#show interfaces port-channel 13


Channel Ports Ch-Type Hash Type Min-links Local Prf

------- ---------------------------- -------- --------- --------- --

Po13 Active: Te1/0/43, Te1/0/44 Static 7 1 Disabled

Hash Algorithm Type

1 - Source MAC, VLAN, EtherType, source module and port Id

2 - Destination MAC, VLAN, EtherType, source module and port Id

3 - Source IP and source TCP/UDP port

4 - Destination IP and destination TCP/UDP port

5 - Source/Destination MAC, VLAN, EtherType, source MODID/port

6 - Source/Destination IP and source/destination TCP/UDP port

7 - Enhanced hashing mode
```

Configurer le vSphere Standard Switch

Cet exemple part du principe que vous comprenez la configuration et la création de vSphere Standard Switches.

Cet exemple utilise la configuration suivante :

- Hôtes vSAN identiques
- Liaisons montantes nommées vmnic0 et vmnic1
- VLAN 3266 lié aux ports de commutateur et au canal de port
- Trames jumbo

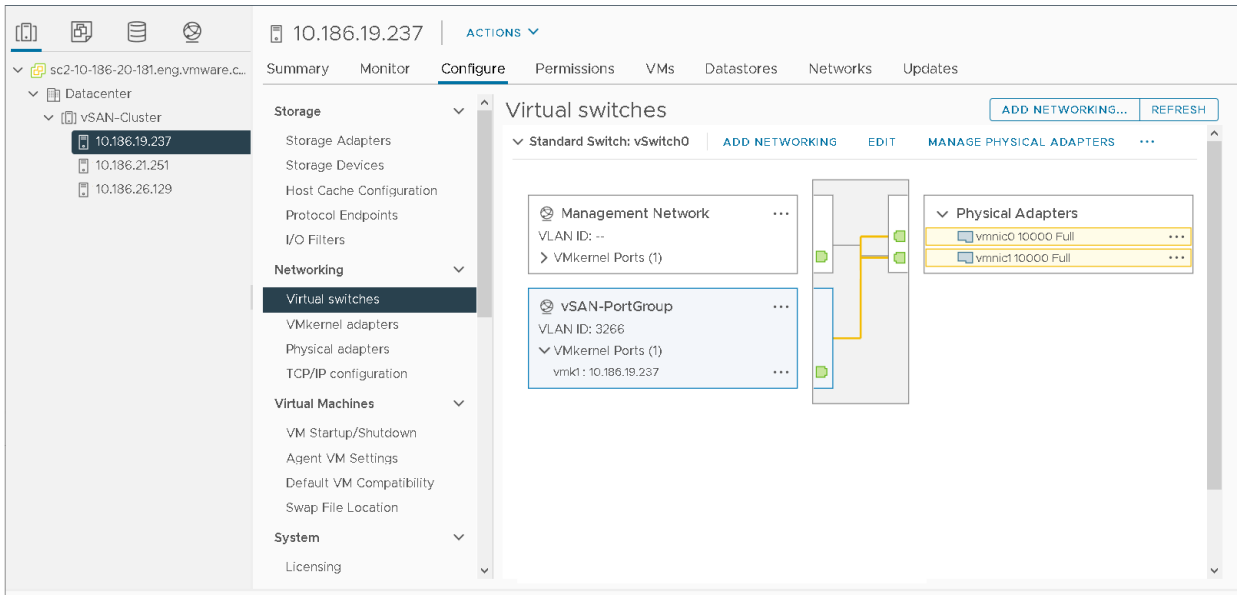
Sur chaque hôte, créez un **vSwitch1** avec le MTU défini sur 9 000 et vmnic0 et vmnic1 ajoutés au vSwitch. Dans la stratégie d'association et de basculement, définissez les deux adaptateurs dans la position **Actif**. Définissez la stratégie d'équilibrage de charge sur **Route basée sur le hachage IP**.

Configurez l'association et le basculement pour le trafic du groupe de ports distribués pour le trafic vSAN comme suit :

- Stratégie d'équilibrage de charge définie sur **Route basée sur le hachage IP**.
- La Détection de panne réseau est définie sur **État de lien seulement**.

- Avertir les commutateurs défini sur **Oui**.
- Le Retour arrière est défini sur **Oui**.
- Assurez-vous que les deux liaisons montantes se trouvent dans la position **Liaisons montantes actives**.

Utilisez les valeurs par défaut pour la détection du réseau, les commutateurs et le retour arrière. Tous les groupes de ports héritent de la stratégie d'association et de basculement qui a été définie au niveau du vSwitch. Vous pouvez remplacer les stratégies d'association et de basculement d'un groupe de ports individuels pour les différencier du vSwitch parent, mais veillez à utiliser le même ensemble de liaisons montantes pour l'équilibrage de charge à hachage IP pour tous les groupes de ports.



Configurer l'équilibrage de charge

Même si les deux liaisons montantes physiques sont utilisées, il n'existe pas d'équilibrage homogène du trafic sur tous les vmnic physiques. La figure illustre que seul le trafic actif est le trafic vSAN, qui représentait essentiellement quatre vmknic ou adresses IP. Le comportement peut être dû au faible nombre d'adresses IP et de hachages possibles. Cependant, dans certaines situations, le commutateur virtuel peut transmettre systématiquement le trafic via une liaison montante dans l'équipe. Pour de plus amples informations sur l'algorithme de hachage IP, consultez la documentation de vSphere officielle à propos de la *Route basée sur le hachage IP.*

Redondance du réseau

Dans cet exemple, vmnic1 est connecté à un port qui a été désactivé du commutateur, pour que nous puissions nous concentrer sur le comportement de panne et de redondance. Notez qu'une alarme de redondance de liaison montante de réseau a été déclenchée.

Aucune alarme de santé de vSAN n'a été déclenchée. Les composants de cluster et de machine virtuelle ne sont pas affectés et les E/S de stockage invitées ne sont pas interrompues par cette panne.

Récupération et retour arrière

Une fois vmnic1 récupéré, le trafic est automatiquement équilibré sur les deux liaisons montantes actives.

Network I/O Control

10

Utilisez vSphere Network I/O Control pour définir les niveaux de qualité de service (QoS) sur le trafic réseau.

vSphere Network I/O Control est une fonctionnalité disponible avec les dispositifs vSphere Distributed Switches. Utilisez-la pour mettre en œuvre la qualité de service (QoS) sur le trafic réseau. Cela peut être utile pour vSAN lorsque le trafic vSAN doit partager la carte réseau physique avec d'autres types de trafic, tels que vMotion, la gestion, les machines virtuelles.

Réservations, partages et limites

Vous pouvez définir une **réservation** afin que Network I/O Control s'assure qu'une bande passante minimale est disponible sur l'adaptateur physique pour vSAN.

Les réservations peuvent être utiles lorsque le trafic *par salve*, comme vMotion ou l'évacuation d'hôte complète, peut affecter le trafic vSAN. Les réservations sont appelées uniquement en cas de contention de la bande passante réseau. L'un des inconvénients des réservations dans Network I/O Control est que la bande passante de réservation inutilisée ne peut pas être allouée au trafic de la machine virtuelle. La quantité totale de bande passante réservée sur tous les types de trafic système ne peut pas dépasser 75 % de la bande passante fournie par l'adaptateur réseau physique de plus faible capacité.

Meilleures pratiques de vSAN pour les réservations. Étant donné que le trafic réservé à vSAN ne peut pas être alloué au trafic de la machine virtuelle, évitez d'utiliser des réservations NIOC dans les environnements vSAN.

La définition de **parts** rend une certaine quantité de bande passante disponible pour vSAN lorsque l'adaptateur physique affecté à la vSAN est saturé. Cela empêche vSAN d'utiliser la capacité intégrale de l'adaptateur physique lors des opérations de reconstruction et de synchronisation. Par exemple, l'adaptateur physique peut devenir saturé lorsqu'un autre adaptateur physique de l'équipe est défaillant et que la totalité du trafic dans le groupe de ports est transféré sur les autres adaptateurs de l'équipe. L'option **parts** veille à ce qu'aucun autre trafic n'ait d'impact sur le réseau vSAN.

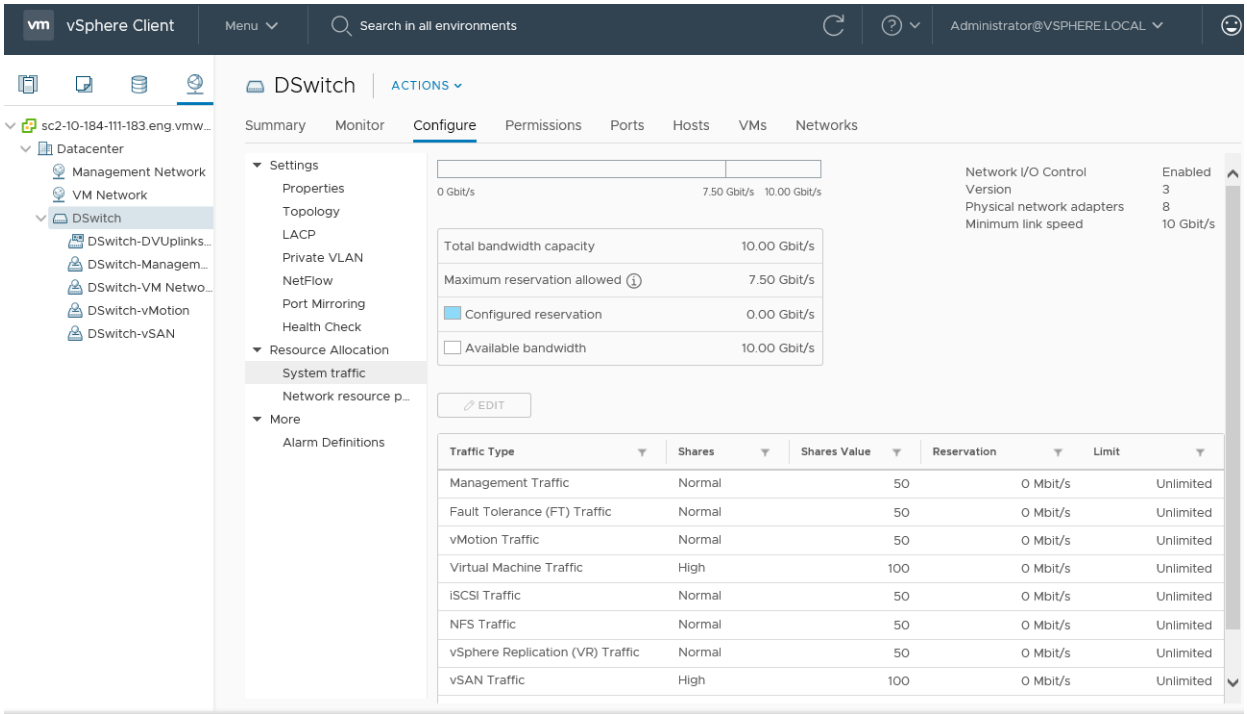
Recommandation de vSAN relative aux parts. Il s'agit de la technique la plus juste d'allocation de bande passante dans NIOC. Cette technique est préférable pour une utilisation dans les environnements vSAN.

La définition de **limites** définit la bande passante maximale pouvant être utilisée par un certain type de trafic sur un adaptateur. Si aucune autre personne n'utilise la bande passante supplémentaire, le type de trafic avec la limite ne peut pas non plus l'utiliser.

Recommandation de vSAN relative aux limites. Étant donné que les types de trafic avec des limites ne peuvent pas utiliser la bande passante supplémentaire, évitez d'utiliser des limites NIOC dans les environnements vSAN.

Pools de ressources réseau

Vous pouvez afficher tous les types de trafic système pouvant être contrôlés par Network I/O Control. Si vous disposez de plusieurs réseaux de machines virtuelles, vous pouvez attribuer une certaine bande passante au trafic de la machine virtuelle. Utilisez des pools de ressources réseau pour consommer des parties de cette bande passante en fonction du groupe de ports de la machine virtuelle.



Activation de Network I/O Control

Vous pouvez activer Network I/O Control dans les propriétés de configuration du vDS. Cliquez avec le bouton droit sur le vDS dans vSphere Client, puis choisissez **Paramètres > Modifier les paramètres** dans le menu.

Note Network I/O Control n'est disponible que sur les dispositifs vSphere Distributed Switches, pas sur les dispositifs vSwitch standard.

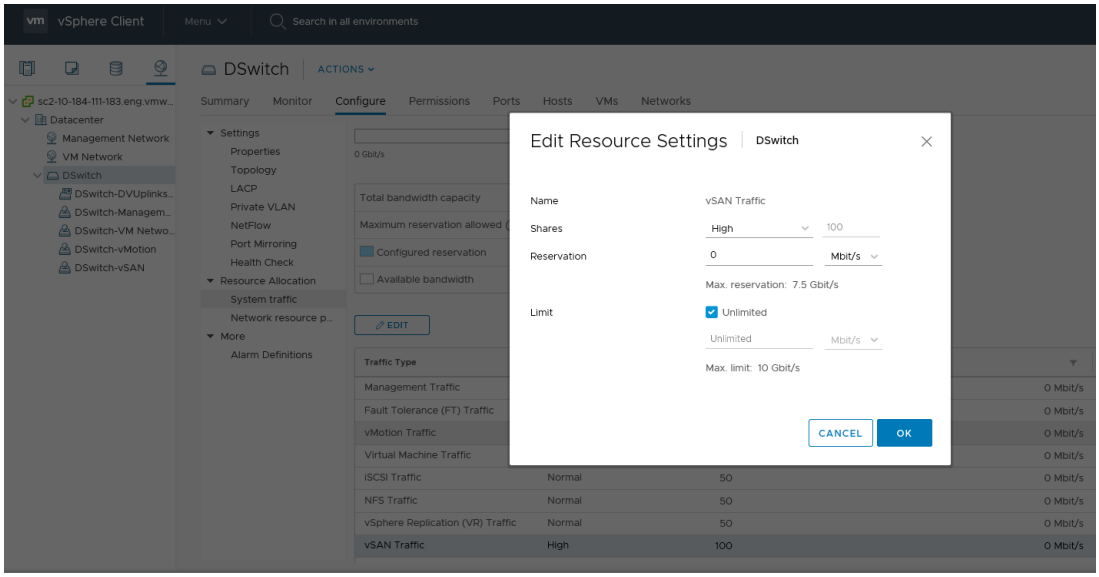
Vous pouvez utiliser Network I/O Control pour réserver la bande passante au trafic réseau en fonction de la capacité des adaptateurs physiques sur un hôte. Par exemple, si le trafic vSAN utilise des adaptateurs réseau physiques 10 GbE et que ces adaptateurs sont partagés avec d'autres types de trafic système, vous pouvez utiliser vSphere Network I/O Control pour garantir une certaine quantité de bande passante pour vSAN. Cela peut être utile lorsque le trafic, comme vSphere vMotion, vSphere HA et le trafic de machine virtuelle, partagent la même carte réseau physique que le réseau vSAN.

Ce chapitre contient les rubriques suivantes :

Exemple de configuration Network I/O Control

Vous pouvez configurer Network I/O Control pour un cluster vSAN.

Envisagez un cluster vSAN avec un seul adaptateur physique 10 GbE. Cette carte réseau gère le trafic pour les vSAN, les vSphere vMotion et les machines virtuelles. Pour modifier la valeur des parts d'un type de trafic, sélectionnez ce type de trafic dans la vue Trafic système (**VDS > Configurer > Allocation de ressources > Trafic système**), puis cliquez sur **Modifier**. La valeur des parts pour le trafic vSAN a été modifiée de la valeur par défaut Normal/50 à Haut/100.



Modifiez les autres types de trafic afin qu'ils correspondent aux valeurs de partage indiquées dans le tableau.

Tableau 10-1. Exemples de paramètres NIOC

Type de trafic	Parts	Valeur
vSAN	Haut	100
vSphere vMotion	Bas	25

Tableau 10-1. Exemples de paramètres NIOC (suite)

Machine virtuelle	Normal	50
iSCSI/NFS	Bas	25

Si l'adaptateur 10 GbE est saturé, Network I/O Control alloue 5 Gbits/s à vSAN sur l'adaptateur physique, 3,5 Gbits/s à vMotion et 1,5 Gbit/s au trafic de la machine virtuelle. Utilisez ces valeurs comme point de départ pour paramétrer la configuration NIOC sur votre réseau vSAN. Assurez-vous que vSAN présente la priorité la plus élevée de n'importe quel protocole.

Pour de plus amples informations sur les différents paramètres d'allocation de bande passante, consultez la documentation de *Mise en réseau vSphere*.

Avec chacune des éditions de vSphere pour vSAN, VMware fournit un dispositif vSphere Distributed Switch dans le cadre de l'édition. Network I/O Control peut être configuré avec n'importe quelle édition de vSAN.

Présentation des topologies réseau vSAN

11

L'architecture vSAN prend en charge différentes topologies réseau. Ces topologies ont un impact sur le déploiement et la gestion globale de vSAN.

L'introduction de la prise en charge de la monodiffusion dans vSAN 6.6 simplifie la conception du réseau.

Ce chapitre contient les rubriques suivantes :

Déploiements standard

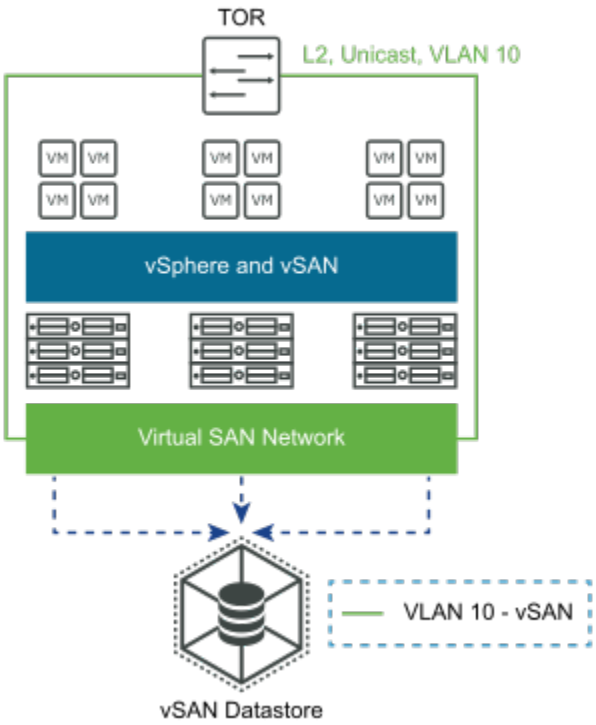
vSAN prend en charge plusieurs types de déploiements mono-site.

Couche 2, site unique, rack unique

Cette topologie réseau est responsable du transfert de paquets via des périphériques de couche 2 intermédiaires tels que des hôtes, des ponts ou des commutateurs.

La topologie réseau de couche 2 offre les conditions d'implémentation et de gestion de vSAN les plus simples. VMware recommande l'utilisation et la configuration de l'écoute IGMP pour éviter d'envoyer un trafic multidiffusion inutile sur le réseau. Dans ce premier exemple, nous examinons un site unique et peut-être même un rack unique de serveurs utilisant vSAN 6.5 ou version antérieure. Cette version utilise la multidiffusion. Par conséquent, activez l'écoute IGMP. Étant donné que tous les éléments se trouvent sur le même réseau L2, il n'est pas nécessaire de configurer le routage pour le trafic multidiffusion.

Les implémentations de couche 2 sont davantage simplifiées dans vSAN 6.6 et version ultérieure, qui introduit la prise en charge de la monodiffusion. L'écoute IGMP n'est pas nécessaire.



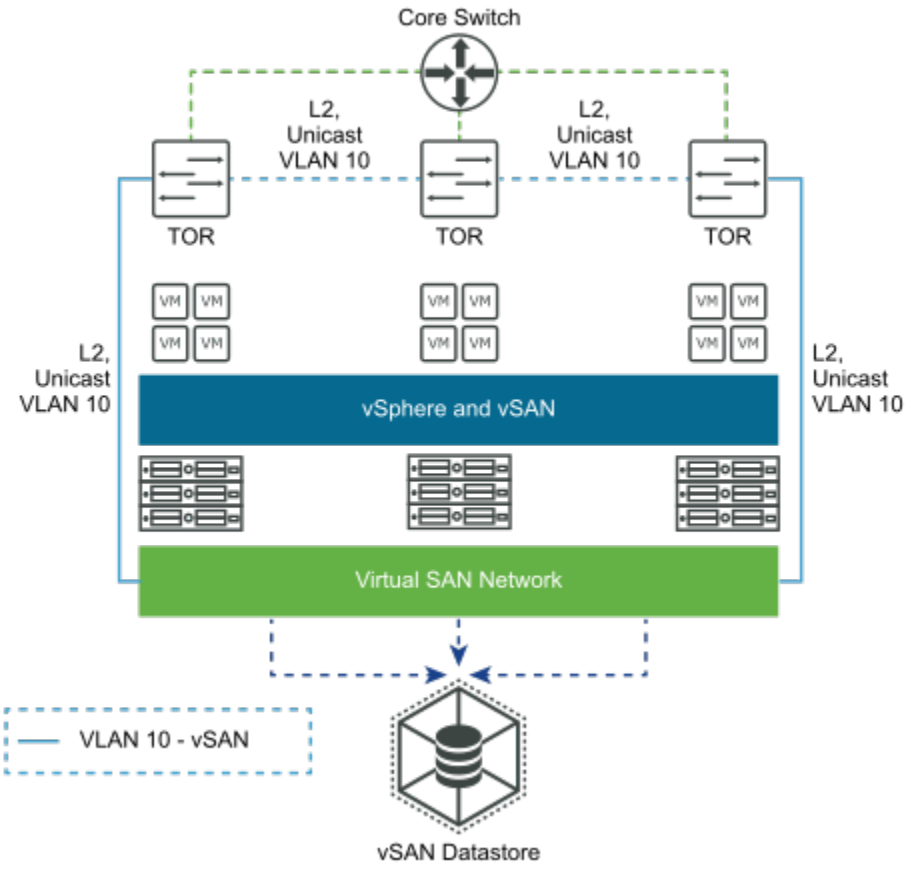
Couche 2, site unique, racks multiples

Cette topologie réseau fonctionne avec l'implémentation de couche 2 comptant plusieurs racks, et plusieurs commutateurs top-of-rack, ou TOR, connectés à un commutateur principal.

Dans les figures suivantes, la ligne bleue en pointillés entre les TOR indique que le réseau vSAN est disponible et accessible à tous les hôtes du cluster vSAN. Cependant, les hôtes des différents racks communiquent entre eux par le biais de la couche 3, ce qui implique d'utiliser PIM pour acheminer le trafic multidiffusion entre les hôtes. Les TOR ne sont pas connectés physiquement les uns aux autres.

VMware recommande de configurer tous les TOR pour l'écoute IGMP afin d'éviter tout trafic multidiffusion inutile sur le réseau. Dans la mesure où le trafic n'est pas acheminé, il n'est pas nécessaire de configurer PIM pour acheminer le trafic multidiffusion.

Cette implémentation est plus simple dans vSAN 6.6 et version ultérieure car le trafic vSAN est monodiffusion. Avec le trafic monodiffusion, il n'est pas nécessaire de configurer l'écoute IGMP sur les commutateurs.



Couche 3, site unique, racks multiples

Cette topologie réseau fonctionne pour les déploiements de vSAN où la couche 3 est utilisée pour acheminer le trafic vSAN.

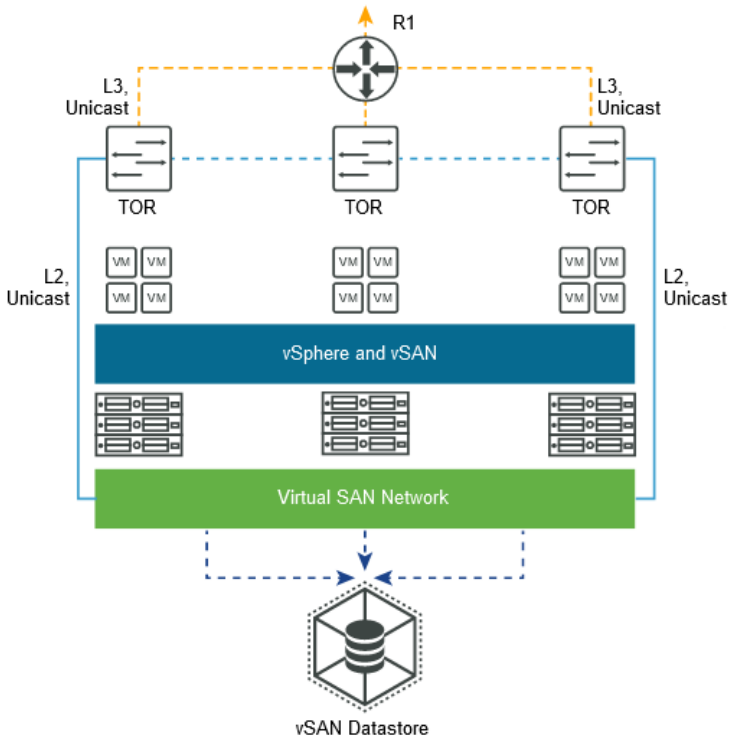
Cette topologie réseau de couche 3 simple utilise plusieurs racks dans le même centre de données, chacun avec son propre commutateur TOR. Acheminez le réseau vSAN entre les différents racks sur L3 afin de permettre à tous les hôtes du cluster vSAN de communiquer. Placez les ports VMkernel pour vSAN sur différents sous-réseaux ou VLAN, et utilisez un sous-réseau ou un VLAN distinct pour chaque rack.

Cette topologie réseau achemine les paquets via des périphériques compatibles couche 3 intermédiaires, tels que des routeurs et des commutateurs compatibles couche 3. Chaque fois que des hôtes sont déployés sur différents segments de réseau de couche 3, le résultat est une topologie réseau acheminée.

Pour vSAN 6.5 et version antérieure, VMware recommande l'utilisation et la configuration de l'écoute IGMP, car ces déploiements exigent la multidiffusion. Configurez PIM sur les commutateurs physiques pour faciliter le routage du trafic multidiffusion.

vSAN 6.6 et version ultérieure simplifie cette topologie. Dans la mesure où il n'existe aucun trafic multidiffusion, il n'est pas nécessaire de configurer l'écoute IGMP. Vous n'avez pas besoin de configurer PIM pour acheminer le trafic multidiffusion.

Vous trouverez ci-dessous un exemple de déploiement de vSAN 6.6 sur L3. Il n'existe aucune condition requise pour l'écoute IGMP ou PIM car il n'existe aucun trafic multidiffusion.



Déploiements de cluster étendu

vSAN prend en charge les déploiements de cluster étendus qui couvrent deux emplacements.

Dans vSAN 6.5 et version antérieure, le trafic vSAN entre les sites de données est **multidiffusion** pour les métadonnées et **monodiffusion** pour les E/S.

Dans vSAN 6.6 et version ultérieure, l'intégralité du trafic est en **monodiffusion**. Dans toutes les versions de vSAN, le trafic témoin entre un site de données et l'hôte témoin est en monodiffusion.

Couche 2 partout

Vous pouvez configurer un cluster étendu vSAN dans un réseau de couche 2, mais cette configuration n'est pas recommandée.

Envisagez une conception dans laquelle le cluster étendu vSAN est configuré dans une conception de couche 2 de grande taille. Les machines virtuelles sont déployées sur le site de données 1 et le site 2. Le site 3 contient l'hôte témoin.

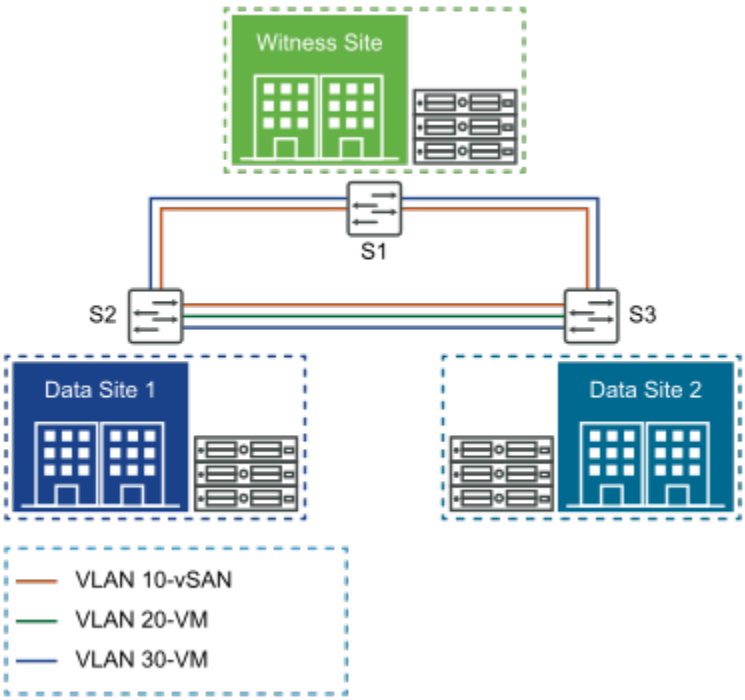
Note Pour de meilleurs résultats, évitez d'utiliser un réseau de couche 2 étendu sur tous les sites.

Pour illustrer la couche 2 partout aussi simplement que possible, nous utilisons des commutateurs (et non des routeurs) dans les topologies.

Les réseaux de couche 2 ne peuvent pas avoir de boucles (chemins multiples). Par conséquent, les fonctionnalités telles que le protocole STP (Spanning Tree Protocol) sont nécessaires pour bloquer l'une des connexions entre le site 1 et le site 2. Considérons à présent une situation dans laquelle le lien entre le site 2 et le site 3 est rompu (le lien entre le site 1 et le site 2). Le trafic réseau est désormais basculé du site 1 vers le site 2 via l'hôte témoin sur le site 3. Étant donné que VMware prend en charge une bande passante beaucoup plus faible et une latence plus élevée pour l'hôte témoin, vous observez une diminution importante des performances si le trafic du réseau de données passe par un site témoin doté de spécifications inférieures.

Dans les cas où le basculement du trafic entre des sites de données via le site témoin n'a pas d'incidence sur la latence des applications et que la bande passante est acceptable, une configuration L2 étendue entre les sites est possible. Dans la plupart des cas, une telle configuration n'est pas faisable et accroît la complexité des exigences de mise en réseau.

Avec vSAN 6.5 ou version antérieure, qui utilise le trafic multidiffusion, vous devez configurer l'écoute IGMP sur les commutateurs. Cette opération n'est pas nécessaire dans vSAN 6.6 et version ultérieure. PIM n'est pas nécessaire, car il n'y a aucun routage du trafic multidiffusion.



Configurations de cluster étendu prises en charge

vSAN prend en charge les configurations de cluster étendu.

La configuration suivante empêche l'acheminement du trafic du site 1 vers le site 2 via l'hôte témoin, en cas de défaillance sur l'un des réseaux des sites de données. Cette configuration permet d'éviter une dégradation des performances. Pour vous assurer que le trafic de données n'est pas commuté via l'hôte témoin, utilisez la topologie réseau suivante.

Entre le site 1 et le site 2, implémentez une configuration étendue de couche 2 commutée ou une configuration de couche 3 routée. Les deux configurations sont prises en charge.

Entre le site 1 et l'hôte témoin, implémentez une configuration de couche 3 routée.

Entre le site 2 et l'hôte témoin, implémentez une configuration de couche 3 routée.

Ces configurations (L2+L3 et L3 partout) sont décrites avec une attention portée au trafic multidiffusion dans vSAN 6.5 et version antérieure, et le trafic monodiffusion uniquement, qui est disponible dans vSAN 6.6. Le trafic multidiffusion introduit des étapes de configuration supplémentaires pour l'écoute IGMP, et PIM pour le routage du trafic multidiffusion.

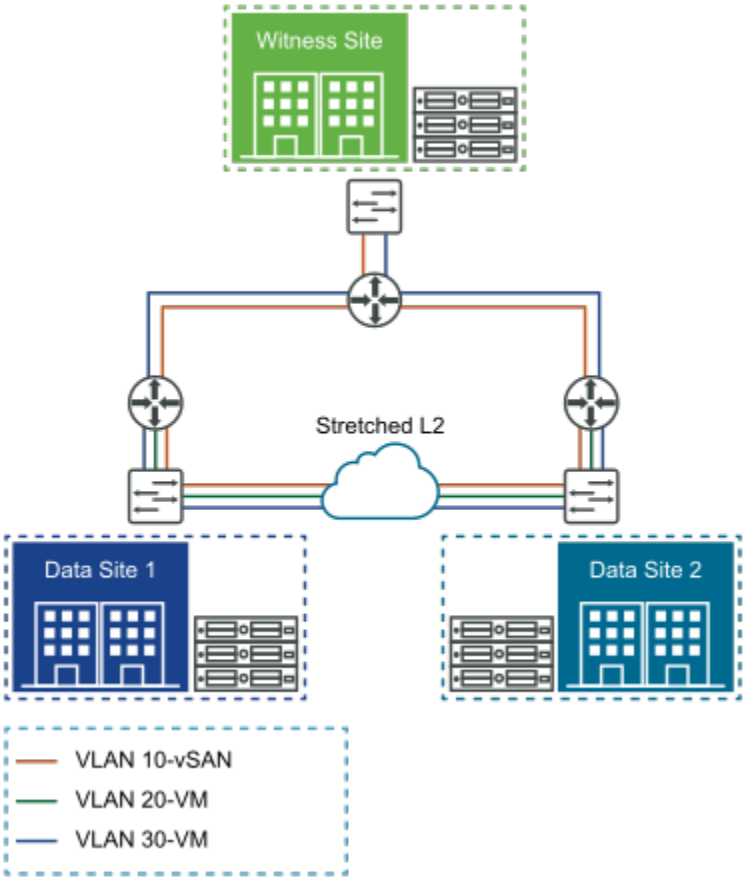
Nous examinerons un réseau étendu de couche 2 entre les sites de données et un réseau routé de couche 3 vers le site témoin. Pour démontrer une combinaison de couche 2 et de couche 3 aussi simplement que possible, utilisez une combinaison de commutateurs et de routeurs dans les topologies.

Couche 2 étendue entre les sites de données, couche 3 vers l'hôte témoin

vSAN prend en charge les configurations étendues de couche 2 entre les sites de données.

Dans ce cas, le seul trafic acheminé est le trafic témoin. Dans vSAN 6.5 et version antérieure, qui utilise la multidiffusion, utilisez l'écoute IGMP pour le trafic multidiffusion sur le vSAN L2 étendu entre les sites de données. Cependant, dans la mesure où le trafic témoin est en monodiffusion, il n'est pas nécessaire d'implémenter PIM sur les segments de couche 3.

Dans vSAN 6.6, qui utilise la monodiffusion, il n'est pas nécessaire d'envisager l'écoute IGMP ou PIM.



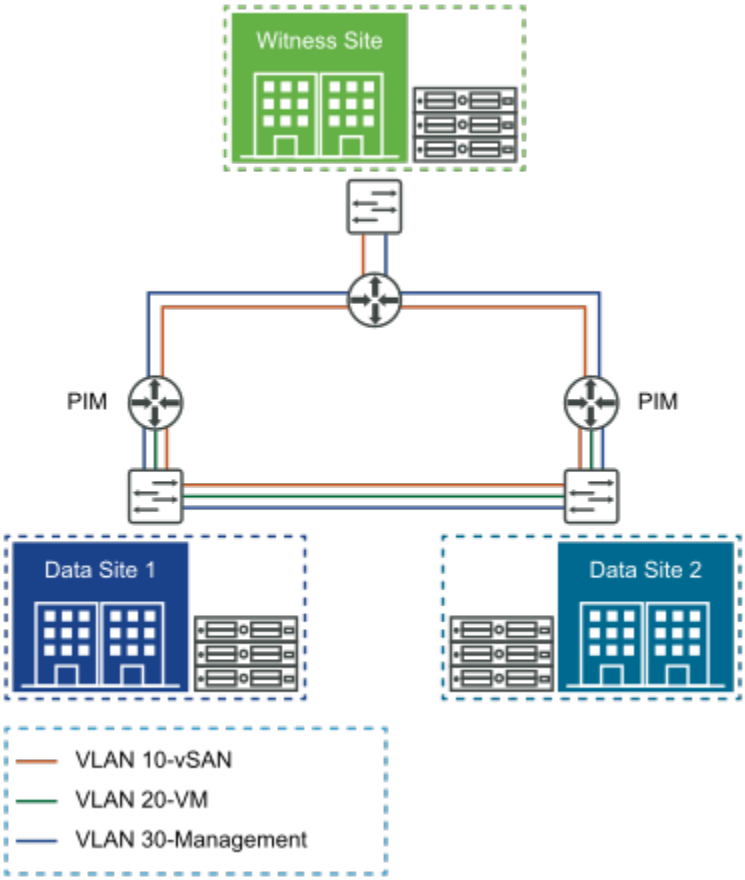
Couche 3 partout

Dans cette configuration de cluster étendu vSAN, le trafic de données est acheminé entre les sites de données et l'hôte témoin.

Pour implémenter la couche 3 partout aussi simplement que possible, utilisez des routeurs ou des commutateurs de routage dans les topologies.

Par exemple, envisagez un environnement avec vSAN 6.5 ou version antérieure, qui utilise le trafic multidiffusion. Dans ce cas, configurez l'écoute IGMP sur les commutateurs du site de données pour gérer le volume de trafic multidiffusion sur le réseau. Cela n'est pas nécessaire sur l'hôte témoin, car le trafic témoin est en monodiffusion. Le trafic multidiffusion est acheminé entre les sites de données, c'est pourquoi vous devez configurez PIM pour autoriser le routage du trafic multidiffusion.

Dans vSAN 6.6 et versions ultérieures, ni l'écoute IGMP ni PIM ne sont nécessaires, car la totalité du trafic acheminé est monodiffusion.



Séparation du trafic témoin sur les clusters étendus vSAN

vSAN prend en charge la séparation du trafic témoin sur les clusters étendus.

Dans vSAN 6.5 et versions ultérieures, vous pouvez séparer le trafic témoin du trafic vSAN dans les configurations à deux nœuds. Cela signifie que les deux hôtes vSAN peuvent être connectés directement sans commutateur 10 Go.

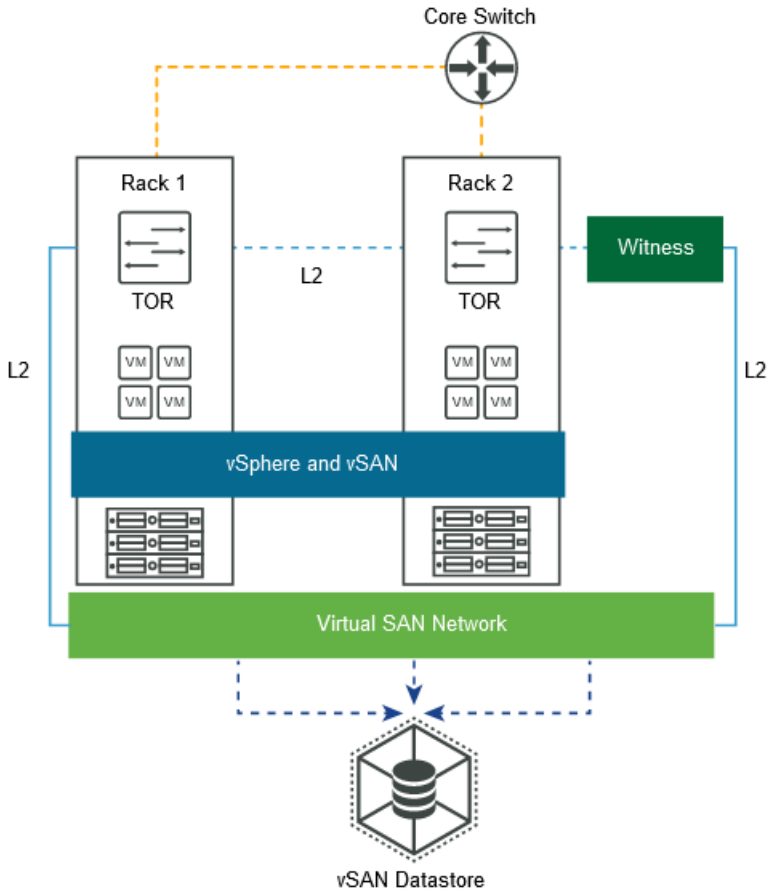
Cette séparation du trafic témoin est uniquement prise en charge sur les déploiements à deux nœuds dans vSAN 6.6. La séparation du trafic témoin sur des clusters étendus vSAN est prise en charge dans vSAN 6.7 et versions ultérieures.

Utilisation d'un cluster étendu pour permettre la détection des racks

Avec les clusters étendus, vSAN garantit la détection des racks sur un site unique.

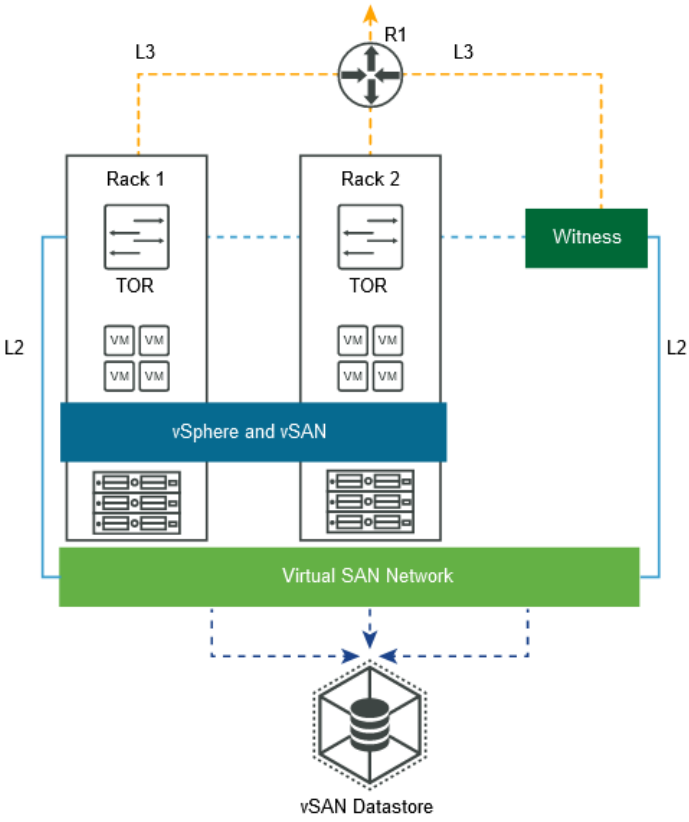
Si vous disposez de deux racks d'hôtes vSAN, vous pouvez continuer à exécuter votre cluster vSAN après une défaillance de rack complète. Dans ce cas, la disponibilité des charges de travail de la VM est fournie par le rack restant et un hôte témoin distant.

Note Pour que cette configuration soit prise en charge, ne placez pas l'hôte témoin dans les deux racks des hôtes vSAN.



Dans cet exemple, si le rack 1 fait défaut, le rack 2 et l'hôte témoin garantissent la disponibilité de la VM. Cette configuration est un environnement antérieur àvSAN 6.6 et requiert que la multidiffusion soit configurée sur le réseau. L'hôte témoin doit se trouver sur le réseau vSAN. Le trafic témoin est monodiffusion. Dans vSAN 6.6 et version ultérieure, l'intégralité du trafic est en monodiffusion.

Cette topologie est également prise en charge sur L3. Placez les ports VMkernel pour vSAN sur différents sous-réseaux ou VLAN, et utilisez un sous-réseau ou un VLAN distinct pour chaque rack.



Cette topologie prend en charge les déploiements avec deux racks pour assurer la détection des racks (domaines de pannes) avec un cluster étendu vSAN. Cette solution utilise un hôte témoin externe au cluster.

Déploiements de vSAN à deux nœuds

vSAN prend en charge les déploiements à deux nœuds. Les déploiements de vSAN à deux nœuds sont utilisés pour les bureaux distants/succursales (ROBO) disposant d'un faible nombre de charges de travail, mais exigeant une haute disponibilité.

Les déploiements de vSAN à deux nœuds utilisent un troisième hôte témoin, qui peut être situé à distance de la succursale. Le témoin est souvent conservé dans la succursale, conjointement avec les composants de gestion, tels que vCenter Server.

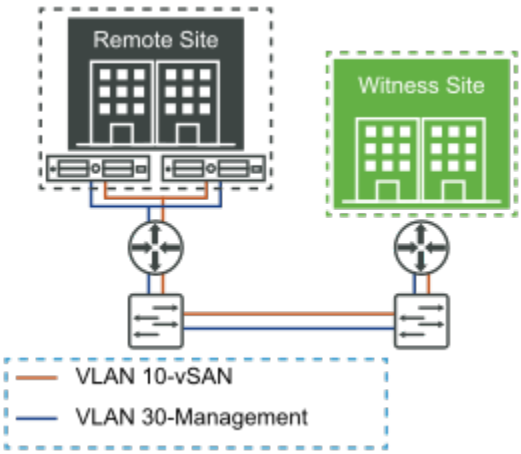
Versions de déploiements de vSAN à deux nœuds antérieures à vSAN 6.5

Les versions de vSAN antérieures à la version 6.5 qui prennent en charge les déploiements à deux nœuds nécessitent un commutateur physique sur le site distant.

Les systèmes vSAN à deux nœuds précoces ont besoin d'inclure un commutateur physique de 10 Go sur le site distant. Si les seuls serveurs sur ce site distant étaient les hôtes vSAN, cette solution peut s'avérer inefficace.

Avec ce déploiement, en l'absence d'autres périphériques utilisant le commutateur de 10 Go, aucune attention particulière ne doit être portée à l'écoute IGMP. Si d'autres périphériques sur le site distant partagent le commutateur de 10 Go, utilisez l'écoute IGMP pour éviter un trafic multidiffusion excessif et inutile.

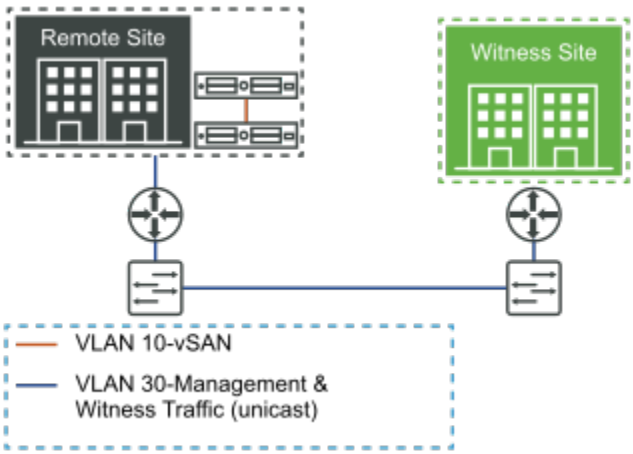
PIM n'est pas nécessaire, car le seul trafic acheminé est le trafic témoin, qui est monodiffusion.



Déploiements à deux nœuds pour vSAN 6.5 et versions ultérieures

vSAN 6.5 et versions ultérieures prend en charge les déploiements à deux nœuds.

Avec vSAN 6.5 et versions ultérieures, cette implémentation de vSAN à deux nœuds est beaucoup plus simple. vSAN 6.5 et version ultérieure permet aux deux hôtes du site de données d'être connectés directement.

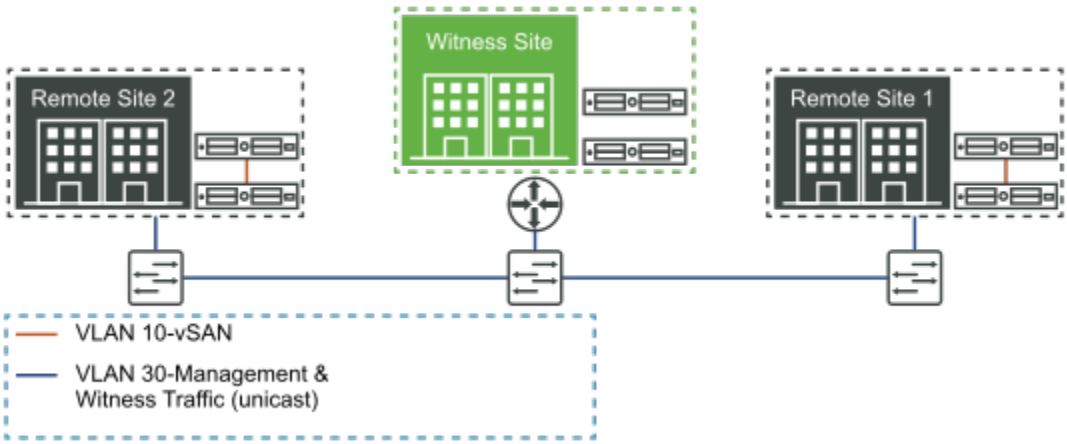


Pour activer cette fonctionnalité, le trafic témoin est entièrement distinct du trafic de données vSAN. Le flux du trafic de données vSAN peut avoir lieu entre les deux nœuds lors de la connexion directe, alors que le trafic témoin peut être acheminé vers le site témoin sur le réseau de gestion.

Le dispositif témoin peut être situé à distance de la succursale. Par exemple, le témoin peut s'exécuter à nouveau dans le centre de données principal, parallèlement à l'infrastructure de gestion (vCenter Server, vROps, Log Insight, etc.). Le témoin peut également résider à distance à partir de la succursale dans vCloud Air.

Dans cette configuration, il n'existe aucun commutateur sur le site distant. Par conséquent, il n'est pas nécessaire de configurer la prise en charge du trafic multidiffusion sur le réseau dos-à-dos vSAN. Il n'est pas nécessaire de tenir compte du trafic multidiffusion sur le réseau de gestion, car le trafic témoin est monodiffusion.

Les versions vSAN 6.6 et ultérieures utilisent toutes le trafic monodiffusion, il n'y a donc aucune remarque sur la multidiffusion. Plusieurs déploiements à deux nœuds de bureaux distants/succursales sont également pris en charge, dès lors que chacun dispose de son propre témoin unique.



Considérations communes pour les déploiements de vSAN à deux nœuds.

Les déploiements de vSAN à deux nœuds prennent en charge d'autres topologies. Cette section décrit les configurations courantes.

Pour plus d'informations sur les configurations à deux nœuds et les considérations détaillées de déploiement hors du réseau, reportez-vous à la documentation de base de vSAN.

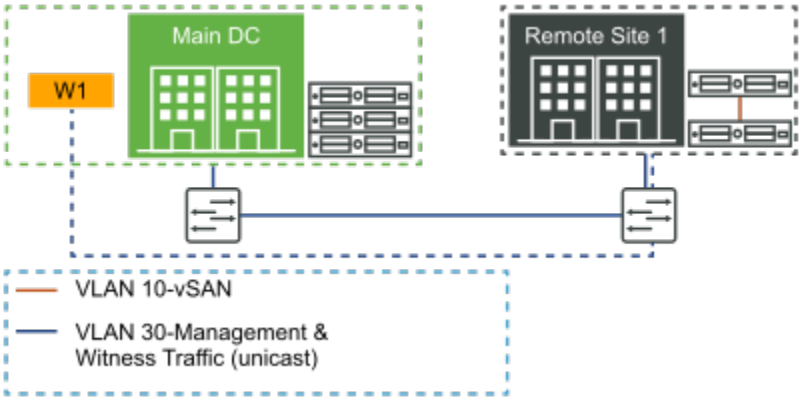
Exécution du témoin sur un autre cluster à deux nœuds

vSAN ne prend pas en charge l'exécution du témoin sur un autre cluster à deux nœuds.

Témoin exécuté sur un autre déploiement de vSAN standard

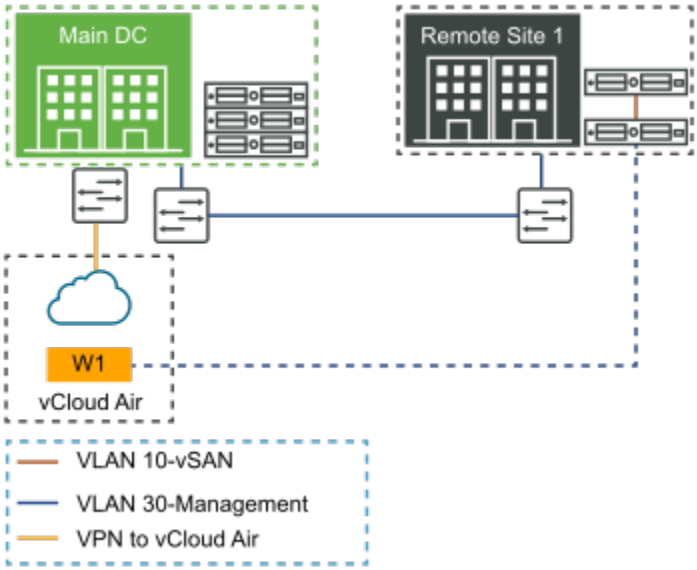
vSAN prend en charge l'exécution du témoin sur un autre déploiement de vSAN standard.

Cette configuration est prise en charge. Toute panne du système vSAN à deux hôtes sur le site distant n'a aucun effet sur la disponibilité de l'environnement vSAN standard au niveau du centre de données principal.



Témoin en cours d'exécution dans vCloud Air

vSAN vous permet d'exécuter un témoin dans vCloud Air.



Configuration du réseau des sites de données vers l'hôte témoin

Les interfaces d'hôte dans les sites de données communiquent avec l'hôte témoin sur le réseau vSAN. Différentes options de configuration sont disponibles.

Cette rubrique explique comment implémenter ces configurations. Elle indique comment les interfaces sur les hôtes des sites de données, qui communiquent entre eux sur le réseau vSAN, interagissent avec l'hôte témoin.

Option 1 : témoin ESXi physique connecté sur L3 avec des routes statiques

Les sites de données peuvent être connectés sur un réseau L2 étendu. Utilisez également cette option pour le réseau de gestion des sites de données, le réseau vSAN, le réseau vMotion et le réseau de machines virtuelles.

Le routeur de réseau physique dans cette infrastructure réseau ne transfère pas automatiquement le trafic des hôtes des sites de données (site 1 et site 2) vers l'hôte du site témoin (site 3). Afin de configurer le cluster étendu vSAN, tous les hôtes du cluster doivent communiquer. Il est possible de déployer un cluster étendu dans cet environnement.

La solution consiste à utiliser des *routes statiques* configurées sur les hôtes ESXi, afin que le trafic vSAN du site 1 et du site 2 puisse accéder à l'hôte témoin sur le site 3. Dans le cas des hôtes ESXi sur les sites de données, ajoutez une route statique à l'interface vSAN, qui redirige le trafic vers l'hôte témoin sur le site 3 sur une passerelle spécifiée pour ce réseau. Dans le cas de l'hôte témoin, une route statique, qui redirige le trafic vSAN destiné aux hôtes des sites de données, doit être ajoutée à l'interface vSAN. Utilisez la commande suivante pour ajouter une route statique sur chaque hôte ESXi dans le cluster étendu : **`esxcli network ip route ipv4 add -g`** *<gateway>* **`-n`** *<network>*

Note vCenter Server doit être en mesure de gérer les hôtes ESXi sur les sites de données et sur le site témoin. Tant qu'il existe une connectivité directe entre l'hôte témoin et vCenter Server, il n'y a aucun problème supplémentaire concernant le réseau de gestion.

Il n'est pas nécessaire de configurer un réseau vMotion ou un réseau de VM, ou d'ajouter des routes statiques pour ces réseaux dans le contexte d'un cluster étendu vSAN. Les machines virtuelles ne sont jamais migrées ni déployées vers l'hôte témoin vSAN. Son objectif est de conserver uniquement les objets témoins et ne nécessite aucun de ceux deux réseaux pour cette tâche.

Option 2 : dispositif témoin ESXi virtuel connecté sur L3 avec des routes statiques

Étant donné que l'hôte témoin est une machine virtuelle déployée sur un hôte physique ESXi, qui ne fait pas partie du cluster vSAN, cet hôte ESXi physique doit disposer d'au moins un réseau de VM préconfiguré. Ce réseau de VM doit accéder au réseau de gestion et au réseau vSAN partagé par les hôtes ESXi sur les sites de données.

Note L'hôte témoin ne doit pas être nécessairement un hôte dédié. Il peut être utilisé pour de nombreuses autres charges de travail de VM, tout en hébergeant simultanément le témoin. De nombreuses organisations choisissent leur infrastructure de gestion pour héberger le dispositif témoin, qui doit se trouver sur les hôtes ou le cluster dans lesquels s'exécutent vCenter Server, vRealize Operations et Log Insight.

Une autre option consiste à disposer de deux réseaux de machines virtuelles préconfigurés sur l'hôte ESXi physique sous-jacent, l'un pour le réseau de gestion et l'autre pour le réseau vSAN. Lorsque le témoin ESXi virtuel est déployé sur cet hôte ESXi physique, le réseau doit être attaché et configuré en conséquence.

Une fois que vous avez déployé l'hôte témoin ESXi virtuel, configurez la route statique. Supposons que les sites de données sont connectés sur un réseau L2 étendu. Utilisez également cette option pour le réseau de gestion des sites de données, le réseau vSAN, le réseau vMotion et le réseau de machines virtuelles. Le trafic vSAN n'est pas acheminé depuis les hôtes des sites de données (site 1 et site 2) vers l'hôte du site témoin (site 3) sur la passerelle par défaut. Pour configurer le cluster étendu vSAN, tous les hôtes du cluster nécessitent des routes statiques, afin que le trafic vSAN du site 1 et du site 2 puisse accéder à l'hôte témoin du site 3. Utilisez la commande **`esxcli network ip route`** pour ajouter une route statique sur chaque hôte ESXi.

Déploiements complexes

Il est possible de déployer vSAN dans des configurations inhabituelles ou complexes.

Ces topologies inhabituelles exigent des précautions spéciales.

Trois emplacements, Aucun cluster étendu, Hôtes témoins distribués

Vous pouvez déployer des vSAN dans plusieurs salles, bâtiments ou sites, au lieu de déployer une configuration de cluster étendu.

Cette configuration est prise en charge. L'une des conditions requises est que la latence entre les sites doit être au même niveau que la latence attendue pour un déploiement de vSAN normal dans le même centre de données. La latence doit être **< 1 ms** entre tous les hôtes. Si la latence est supérieure à cette valeur, envisagez un cluster étendu qui tolère une latence de 5 ms. Dans vSAN 6.5 ou version antérieure, des aspects de multidiffusion supplémentaires relatifs à la multidiffusion doivent être pris en compte.

Pour de meilleurs résultats, conservez une configuration uniforme sur tous les sites dans une telle topologie. Pour garantir la disponibilité des machines virtuelles, configurez des domaines de pannes dans lesquels les hôtes de chaque espace, bâtiment ou site sont placés dans le même domaine de pannes. Évitez le partitionnement asymétrique du cluster, où l'hôte A ne peut pas communiquer avec l'hôte B, mais où l'hôte B peut communiquer avec l'hôte A.

Deux nœuds déployés en tant que cluster étendu de 1+1+W

Vous pouvez déployer une configuration à deux nœuds en tant que configuration de cluster étendu en plaçant chaque hôte dans des salles, des bâtiments ou des sites différents.

La tentative d'augmentation du nombre d'hôtes sur chaque site échoue avec une erreur liée à la licence. Pour tout cluster dont la taille est supérieure à deux hôtes et utilisant la fonctionnalité de dispositif/hôte témoin dédié (N+N+Témoin, où N>1), la configuration est considérée comme un cluster étendu vSAN.

Dépannage du réseau vSAN

12

vSAN vous permet d'examiner et de résoudre les différents types de problèmes qui découlent d'un réseau vSAN incorrectement configuré.

Les opérations de vSAN dépendent de la configuration, de la fiabilité et des performances du réseau. De nombreuses demandes de support proviennent d'une configuration de réseau incorrecte ou du fait que le réseau ne fonctionne pas comme prévu.

Utilisez le service de santé vSAN pour résoudre les problèmes liés au réseau. Les contrôles de santé du réseau peuvent vous orienter vers un article approprié de la base de connaissances, en fonction des résultats du contrôle de santé. L'article de la base de connaissances contient des instructions pour résoudre le problème de réseau.

Contrôles de santé du réseau

Le service de santé inclut une catégorie pour les contrôles de santé de mise en réseau.

À chaque contrôle de santé est associé un lien **Demander à VMware**. En cas d'échec d'un contrôle de santé, cliquez sur **Demander à VMware** et lisez l'article de la base de connaissances VMware associé pour obtenir de plus amples détails et des conseils sur la procédure de résolution du problème.

Les contrôles de santé de mise en réseau suivants fournissent des informations utiles sur votre environnement vSAN.

- vSAN **: vérification de la connectivité (monodiffusion) de base**. Ce contrôle permet de vérifier que la connectivité IP existe entre tous les hôtes ESXi du cluster vSAN, en exécutant une commande Ping sur chaque hôte ESXi du réseau vSAN à partir de chaque autre hôte ESXi.
- **vMotion : vérification de la connectivité (monodiffusion) de base**. Ce contrôle permet de vérifier qu'une connectivité IP existe entre tous les hôtes ESXi du cluster vSAN sur lesquels vMotion est configuré. Chaque hôte ESXi sur le réseau vMotion envoie une commande Ping à tous les autres hôtes ESXi.
- **Une vmknic vSAN est configurée pour tous les hôtes**. Ce contrôle permet de vous assurer que chaque hôte ESXi du cluster vSAN dispose d'une carte réseau VMkernel configurée pour le trafic vSAN.
- Tous les hôtes ont des paramètres multidiffusion correspondants. Ce contrôle permet de vous assurer qu'une adresse multidiffusion est correctement configurée pour chaque hôte.

- **Tous les hôtes ont des sous-réseaux correspondants**. Ce contrôle permet de vérifier que tous les hôtes ESXi d'un cluster vSAN ont été configurés de manière à ce que toutes les cartes réseau VMkernel vSAN se trouvent sur le même sous-réseau IP.
- **Hôtes déconnectés de VC**. Ce contrôle permet de vérifier que le vCenter Server dispose d'une connexion active sur tous les hôtes ESXi du cluster vSAN.
- **Hôtes présentant des problèmes de connectivité**. Ce contrôle fait référence aux cas où vCenter Server indique que l'hôte est connecté, mais où les appels API de vCenter sur l'hôte échouent. Il peut mettre en évidence les problèmes de connectivité entre un hôte et le vCenter Server.
- **Latence du réseau**. Ce contrôle procède à une vérification de la latence du réseau des hôtes vSAN. Si le seuil est supérieur à 100 ms, un avertissement s'affiche. Si le seuil de latence est supérieur à 200 ms, une erreur est générée.
- **vMotion : contrôles de MTU (test Ping avec des paquets de grande taille)**. Ce contrôle complète la vérification de la connectivité Ping de base de vMotion. La taille maximale de l'unité de transmission est augmentée pour améliorer les performances du réseau. Les MTU configurés de manière incorrecte peuvent ne pas apparaître comme un problème de configuration du réseau, mais peuvent provoquer des problèmes de performances.
- **Partition du cluster vSAN**. Ce contrôle de santé examine le cluster pour identifier le nombre de partitions existantes. Il affiche une erreur si le cluster vSAN contient plusieurs partitions.
- **Évaluation de la multidiffusion basée sur d'autres contrôles**. Ce contrôle de santé regroupe les données de tous les contrôles de santé du réseau. Si ce contrôle échoue, cela indique que la multidiffusion est probablement la cause principale d'une partition réseau.

Commandes de vérification du réseau

Lorsque le réseau vSAN a été configuré, utilisez ces commandes pour vérifier son état. Vous pouvez vérifier l'adaptateur VMkernel (vmknic) utilisé pour vSAN et les attributs qu'il contient.

Utilisez les commandes ESXCLI et RVC pour vérifier que le réseau est entièrement fonctionnel et pour résoudre tous les problèmes de réseau liés à vSAN.

Vous pouvez vérifier que le vmknic utilisé pour le réseau vSAN est correctement configuré sur tous les hôtes, que la multidiffusion est fonctionnelle et que les hôtes participant au cluster vSAN peuvent communiquer correctement entre eux.

esxcli vsan network list

Cette commande vous permet d'identifier l'interface VMkernel utilisée par le réseau vSAN.

Le résultat ci-dessous indique que le réseau vSAN utilise vmk2. Cette commande continue de fonctionner, même si vSAN a été désactivé et que les hôtes ne participent plus à vSAN.

Il convient également de contrôler la multidiffusion du groupe d'agents et la multidiffusion du groupe de maîtres.

```
[root@esxi-dell-m:~] esxcli vsan network list
Interface
   VmkNic Name: vmk1
   IP Protocol: IP
   Interface UUID: 32efc758-9ca0-57b9-c7e3-246e962c24d0
   Agent Group Multicast Address: 224.2.3.4
   Agent Group IPv6 Multicast Address: ff19::2:3:4
   Agent Group Multicast Port: 23451
   Master Group Multicast Address: 224.1.2.3
   Master Group IPv6 Multicast Address: ff19::1:2:3
   Master Group Multicast Port: 12345
   Host Unicast Channel Bound Port: 12321
   Multicast TTL: 5
   Traffic Type: vsan
```

Cela fournit des informations utiles, telles que l'interface VMkernel utilisée pour le trafic vSAN. Dans ce cas, il s'agit de **vmk1**. Toutefois, les adresses multidiffusion sont également affichées. Cette information peut s'afficher même lorsque le cluster est en cours d'exécution en mode monodiffusion. L'adresse et le port multidiffusion du groupe sont présents. Le port 23451 est utilisé pour le signal de pulsation, envoyé toutes les secondes par l'hôte principal, et est visible sur tous les autres hôtes du cluster. Le port 12345 est utilisé pour les mises à jour de CMMDS entre les hôtes principal et de sauvegarde.

esxcli network ip interface list

Cette commande vous permet de vérifier des éléments tels que vSwitch ou le commutateur distribué.

Utilisez cette commande pour vérifier le vSwitch ou le commutateur distribué auquel il est associé, ainsi que la taille du MTU, qui peut être utile si les trames Jumbo ont été configurées dans l'environnement. Dans ce cas, le MTU a la valeur par défaut 1500.

```
[root@esxi-dell-m:~] esxcli network ip interface list
vmk0
   Name: vmk0
   <<truncated>>
vmk1
   Name: vmk1
   MAC Address: 00:50:56:69:96:f0
   Enabled: true
   Portset: DvsPortset-0
   Portgroup: N/A
   Netstack Instance: defaultTcpipStack
   VDS Name: vDS
   VDS UUID: 50 1e 5b ad e3 b4 af 25-18 f3 1c 4c fa 98 3d bb
   VDS Port: 16
   VDS Connection: 1123658315
   Opaque Network ID: N/A
```

```
   Opaque Network Type: N/A
   External ID: N/A
   MTU: 9000
   TSO MSS: 65535
   Port ID: 50331814
```

La taille maximale de l'unité de transmission indiquée est de 9000. Ce port VMkernel est donc configuré pour les trames Jumbo, qui nécessitent un MTU d'environ 9000. VMware ne fournit aucune recommandation concernant l'utilisation des trames Jumbo. Cependant, les trames Jumbo peuvent être utilisées avec vSAN.

esxcli network ip interface ipv4 get –i vmk2

Cette commande affiche des informations telles que l'adresse IP et le masque de réseau de l'interface VMkernel vSAN.

Avec ces informations, un administrateur peut à présent commencer à utiliser d'autres commandes disponibles sur la ligne de commande pour vérifier que le réseau vSAN fonctionne correctement.

```
[root@esxi-dell-m:~] esxcli network ip interface ipv4 get -i vmk1
Name  IPv4 Address  IPv4 Netmask   IPv4 Broadcast  Address Type  Gateway  DHCP DNS
----  ------------  -------------  --------------  ------------  -------  --------
vmk1  172.40.0.9   255.255.255.0   172.40.0.255   STATIC         0.0.0.0 false
```

vmkping

La commande `vmkping` vérifie si tous les autres hôtes ESXi du réseau répondent à vos demandes Ping.

```
~ # vmkping -I vmk2 172.32.0.3 -s 1472 -d
 PING 172.32.0.3 (172.32.0.3): 56 data bytes
 64 bytes from 172.32.0.3: icmp_seq=0 ttl=64 time=0.186 ms
 64 bytes from 172.32.0.3: icmp_seq=1 ttl=64 time=2.690 ms
 64 bytes from 172.32.0.3: icmp_seq=2 ttl=64 time=0.139 ms

 --- 172.32.0.3 ping statistics ---
 3 packets transmitted, 3 packets received, 0% packet loss
 round-trip min/avg/max = 0.139/1.005/2.690 ms
```

Même s'il ne vérifie pas la fonctionnalité de multidiffusion, il peut faciliter l'identification d'un hôte ESXi non autorisé qui présente des problèmes de réseau. Vous pouvez également examiner les délais de réponse pour vérifier l'existence d'une latence anormale sur le réseau vSAN.

Si les trames Jumbo sont configurées, cette commande ne signale aucun problème si la taille du MTU de trame Jumbo est incorrecte. Par défaut, cette commande utilise une taille de MTU de 1500. S'il est nécessaire de vérifier que les trames Jumbo fonctionnent correctement de bout en bout, utilisez vmkping avec une option de taille de paquet supérieure (-s) comme suit :

```
~ # vmkping -I vmk2 172.32.0.3 -s 8972 -d
PING 172.32.0.3 (172.32.0.3): 8972 data bytes
9008 bytes from 172.32.0.3: icmp_seq=0 ttl=64 time=0.554 ms
9008 bytes from 172.32.0.3: icmp_seq=1 ttl=64 time=0.638 ms
9008 bytes from 172.32.0.3: icmp_seq=2 ttl=64 time=0.533 ms

--- 172.32.0.3 ping statistics ---
3 packets transmitted, 3 packets received, 0% packet loss
round-trip min/avg/max = 0.533/0.575/0.638 ms
~ #
```

Pensez à ajouter-d à la commande vmkping pour déterminer si des paquets peuvent être envoyés sans fragmentation.

esxcli network ip neighbor list

Cette commande permet de vérifier si tous les hôtes vSAN se trouvent sur le même segment de réseau.

Dans cette configuration, nous disposons d'un cluster à quatre hôtes, et cette commande renvoie les entrées ARP (Address Resolution Protocol) des trois autres hôtes, notamment leurs adresses IP et leur vmknic (vSAN est configuré pour utiliser vmk1 sur tous les hôtes de ce cluster).

```
[root@esxi-dell-m:~] esxcli network ip neighbor list -i vmk1
Neighbor     Mac Address        Vmknic   Expiry  State  Type
-----------  -----------------  ------  -------  -----  -------
172.40.0.12  00:50:56:61:ce:22  vmk1    164 sec         Unknown
172.40.0.10  00:50:56:67:1d:b2  vmk1    338 sec         Unknown
172.40.0.11  00:50:56:6c:fe:c5  vmk1    162 sec         Unknown
[root@esxi-dell-m:~]
```

esxcli network diag ping

Cette commande vérifie les doublons sur le réseau et les temps aller-retour.

Pour obtenir de plus amples détails sur la connectivité réseau vSAN entre les différents hôtes, ESXCLI propose une commande de diagnostic de réseau puissante. Vous trouverez ci-dessous un exemple de ce type de résultat, où l'interface VMkernel se trouve sur vmk1 et où l'adresse IP du réseau vSAN distant d'un autre hôte sur le réseau est 172.40.0.10

```
[root@esxi-dell-m:~] esxcli network diag ping -I vmk1 -H 172.40.0.10
   Trace:
         Received Bytes: 64
         Host: 172.40.0.10
         ICMP Seq: 0
```

```
          TTL: 64
          Round-trip Time: 1864 us
          Dup: false
          Detail:

          Received Bytes: 64
          Host: 172.40.0.10
          ICMP Seq: 1
          TTL: 64
          Round-trip Time: 1834 us
          Dup: false
          Detail:

          Received Bytes: 64
          Host: 172.40.0.10
          ICMP Seq: 2
          TTL: 64
          Round-trip Time: 1824 us
          Dup: false
          Detail:
    Summary:
          Host Addr: 172.40.0.10
          Transmitted: 3
          Recieved: 3
          Duplicated: 0
          Packet Lost: 0
          Round-trip Min: 1824 us
          Round-trip Avg: 1840 us
          Round-trip Max: 1864 us
[root@esxi-dell-m:~]
```

vsan.lldpnetmap

Cette commande RVC affiche des informations sur le port de liaison montante.

S'il existe des commutateurs non Cisco avec LLDP (Link Layer Discovery Protocol) activé dans l'environnement, il existe une commande RVC pour afficher des informations sur la liaison montante <-> le commutateur <-> le port de commutateur. Pour plus d'informations sur RVC, reportez-vous au Guide de commandes RVC.

Cela vous permet de déterminer les hôtes et les commutateurs auxquels ils sont reliés lorsque le cluster vSAN s'étend sur plusieurs commutateurs. Il peut faciliter l'isolement d'un problème sur un commutateur spécifique lorsque seul un sous-ensemble des hôtes du cluster est concerné.

```
> vsan.lldpnetmap 02013-08-15 19:34:18 -0700: This operation will take
30-60 seconds ...+---------------+---------------------------+| Host          | LLDP
info                 |+---------------+---------------------------+| 10.143.188.54 | w2r13-
vsan-x650-2: vmnic7 ||               | w2r13-vsan-x650-1: vmnic5 |+---------------
+---------------------------+
```

Cette fonction est uniquement disponible avec les commutateurs prenant en charge LLDP. Pour le configurer, connectez-vous au commutateur et exécutez ce qui suit :

```
switch# config t
Switch(Config)# feature lldp
```

Pour vérifier que LLDP est activé :

```
switch(config)#do show running-config lldp
```

Note LLDP fonctionne en mode d'envoi et de réception, par défaut. Vérifiez les paramètres des propriétés du vDS si les informations sur le commutateur physique ne sont pas détectées. Par défaut, le vDS est créé avec le protocole de détection défini sur CDP, le protocole de détection Cisco. Pour résoudre ce problème, définissez le protocole de détection sur LLDP et définissez l'opération pour **les deux** sur le vDS.

Vérification des communications multidiffusion

Les configurations multidiffusion peuvent entraîner des problèmes de déploiement initial de vSAN.

L'une des méthodes les plus simples pour vérifier si la multidiffusion fonctionne correctement dans votre environnement vSAN consiste à utiliser la commande `tcpdump-uw`. Cette commande est disponible à partir de la ligne de commande des hôtes ESXi.

Cette commande `tcpdump-uw` indique si l'hôte principal envoie correctement les paquets multidiffusion (port et informations IP) et si tous les autres hôtes du cluster les reçoivent.

Sur l'hôte principal, cette commande affiche les paquets envoyés à l'adresse multidiffusion. Sur tous les autres hôtes, les mêmes paquets sont visibles (de l'hôte principal à l'adresse multidiffusion). S'ils ne sont pas visibles, la multidiffusion ne fonctionne pas correctement. Exécutez la commande `tcpdump-uw` indiquée ici sur n'importe quel hôte du cluster, et les signaux de pulsation de l'hôte principal sont visibles. Dans ce cas, l'hôte principal se trouve à l'adresse IP 172.32.0.2. Le **-v** désignant le niveau de détails est facultatif.

```
[root@esxi-hp-02:~] tcpdump-uw -i vmk2 multicast -v
tcpdump-uw: listening on vmk2, link-type EN10MB (Ethernet), capture size 96 bytes
11:04:21.800575 IP truncated-ip - 146 bytes missing! (tos 0x0, ttl 5, id 34917, offset 0,
flags [none], proto UDP (17), length 228)
    172.32.0.4.44824 > 224.1.2.3.12345: UDP, length 200
11:04:22.252369 IP truncated-ip - 234 bytes missing! (tos 0x0, ttl 5, id 15011, offset 0,
flags [none], proto UDP (17), length 316)
    172.32.0.2.38170 > 224.2.3.4.23451: UDP, length 288
11:04:22.262099 IP truncated-ip - 146 bytes missing! (tos 0x0, ttl 5, id 3359, offset 0,
flags [none], proto UDP (17), length 228)
    172.32.0.3.41220 > 224.2.3.4.23451: UDP, length 200
11:04:22.324496 IP truncated-ip - 146 bytes missing! (tos 0x0, ttl 5, id 20914, offset 0,
flags [none], proto UDP (17), length 228)
    172.32.0.5.60460 > 224.1.2.3.12345: UDP, length 200
11:04:22.800782 IP truncated-ip - 146 bytes missing! (tos 0x0, ttl 5, id 35010, offset 0,
flags [none], proto UDP (17), length 228)
```

```
    172.32.0.4.44824 > 224.1.2.3.12345: UDP, length 200
11:04:23.252390 IP truncated-ip - 234 bytes missing! (tos 0x0, ttl 5, id 15083, offset 0,
flags [none], proto UDP (17), length 316)
    172.32.0.2.38170 > 224.2.3.4.23451: UDP, length 288
11:04:23.262141 IP truncated-ip - 146 bytes missing! (tos 0x0, ttl 5, id 3442, offset 0,
flags [none], proto UDP (17), length 228)
    172.32.0.3.41220 > 224.2.3.4.23451: UDP, length 200
```

Même si ce résultat peut sembler un peu confus, il va sans dire que le résultat affiché ici indique que les quatre hôtes du cluster obtiennent un signal de pulsation de l'hôte principal. Cette commande **`tcpdump-UW`** doit être exécutée sur chaque hôte pour vérifier qu'ils reçoivent tous le signal de pulsation. Cela permet de vérifier que l'hôte principal envoie les signaux de pulsation et que tous les autres hôtes du cluster les reçoivent, indiquant ainsi que la multidiffusion fonctionne.

Si certains des hôtes vSAN ne parviennent pas à détecter les signaux de pulsation d'une seconde de l'hôte principal, l'administrateur réseau doit vérifier la configuration multidiffusion de leurs commutateurs.

Pour éviter l'affichage du message ennuyeux **`IP tronqué – 146 octets manquants !`**, utilisez l'option **`-s0`** sur la même commande pour interrompre la troncature de paquets :

```
[root@esxi-hp-02:~] tcpdump-uw -i vmk2 multicast -v -s0
tcpdump-uw: listening on vmk2, link-type EN10MB (Ethernet), capture size 65535 bytes
11:18:29.823622 IP (tos 0x0, ttl 5, id 56621, offset 0, flags [none], proto UDP (17), length
228)
    172.32.0.4.44824 > 224.1.2.3.12345: UDP, length 200
11:18:30.251078 IP (tos 0x0, ttl 5, id 52095, offset 0, flags [none], proto UDP (17), length
228)
    172.32.0.3.41220 > 224.2.3.4.23451: UDP, length 200
11:18:30.267177 IP (tos 0x0, ttl 5, id 8228, offset 0, flags [none], proto UDP (17), length
316)
    172.32.0.2.38170 > 224.2.3.4.23451: UDP, length 288
11:18:30.336480 IP (tos 0x0, ttl 5, id 28606, offset 0, flags [none], proto UDP (17), length
228)
    172.32.0.5.60460 > 224.1.2.3.12345: UDP, length 200
11:18:30.823669 IP (tos 0x0, ttl 5, id 56679, offset 0, flags [none], proto UDP (17), length
228)
    172.32.0.4.44824 > 224.1.2.3.12345: UDP, length 200
```

La commande **`tcpdump`** est associée à l'appartenance IGMP (Internet Group Management Protocol). Les hôtes (et les périphériques réseau) utilisent IGMP pour établir l'appartenance au groupe multidiffusion.

Chaque hôte ESXi du cluster vSAN envoie des rapports d'adhésion IGMP réguliers (jonction).

La commande **`tcpdump`** affiche les rapports de membres IGMP d'un hôte :

```
[root@esxi-dell-m:~] tcpdump-uw -i vmk1 igmp
tcpdump-uw: verbose output suppressed, use -v or -vv for full protocol decode
listening on vmk1, link-type EN10MB (Ethernet), capture size 262144 bytes
15:49:23.134458 IP 172.40.0.9 > igmp.mcast.net: igmp v3 report, 1 group record(s)
15:50:22.994461 IP 172.40.0.9 > igmp.mcast.net: igmp v3 report, 1 group record(s)
```

La sortie montre que les rapports IGMP v3 sont établis, ce qui indique que l'hôte ESXi met régulièrement à jour son appartenance. Si un administrateur réseau a des doutes sur l'exécution correcte d'IGMP par les hôtes vSAN ESXi, l'exécution de cette commande sur chaque hôte ESXi du cluster et l'affichage de cette trace peuvent être utilisés pour la vérification.

Si vous disposez de communications multidiffusion, utilisez IGMP v3.

En effet, la commande suivante peut être utilisée pour examiner simultanément le trafic multidiffusion et IGMP :

```
[root@esxi-hp-02:~] tcpdump-uw -i vmk2 multicast or igmp -v -s0
```

Il arrive fréquemment que le cluster vSAN soit configuré sur plusieurs commutateurs physiques, mais alors que la multidiffusion a été activée sur un commutateur en particulier, elle n'a pas été activée sur tous les commutateurs. Dans ce cas, les formulaires de cluster avec deux hôtes ESXi dans une partition et un autre hôte ESXi (connecté à l'autre commutateur) ne peuvent pas joindre ce cluster. En lieu et place, il forme son propre cluster vSAN dans une autre partition. La commande `vsan.lldpnetmap` présentée précédemment peut vous aider à déterminer la configuration du réseau, ainsi que les hôtes et les commutateurs auxquels ils sont attachés.

Pendant qu'un cluster vSAN se forme, des indicateurs montrent que la multidiffusion peut poser problème.

Supposons que la liste de contrôle du sous-réseau, du VLAN et de la MTU a été suivie et que chaque hôte du cluster peut exécuter la commande `vmkping` sur chaque hôte du cluster.

En cas de problème de multidiffusion lors de la création du cluster, un symptôme commun est que chaque hôte ESXi forme son propre cluster vSAN, lui-même constituant l'hôte principal. Si chaque hôte possède un ID de partition réseau unique, ce symptôme suggère qu'il n'existe aucune multidiffusion entre les hôtes.

Cependant, si dans un cas particulier, un sous-ensemble des hôtes ESXi forme un cluster, qu'un autre sous-ensemble forme un autre cluster et que chacun d'entre eux possède des partitions uniques avec leur propre l'hôte principal, une sauvegarde, voire des hôtes d'agent, la multidiffusion est alors activée sur le commutateur, mais pas dans l'ensemble des commutateurs. vSAN affiche les hôtes sur le premier commutateur physique formant leur propre partition de cluster et les hôtes sur le second commutateur physique formant leur propre partition de cluster, chacun avec son propre l'hôte principal. Si vous pouvez vérifier les commutateurs auxquels les hôtes du cluster se connectent et que les hôtes d'un cluster sont connectés au même commutateur, le problème se situe probablement à ce niveau.

Vérification des performances du réseau vSAN

Assurez-vous que la bande passante est suffisante entre vos hôtes ESXi. Cet outil peut vous aider à vérifier si votre réseau vSAN fonctionne de manière optimale.

Pour vérifier les performances du réseau vSAN, vous pouvez utiliser l'outil `iperf` pour mesurer la bande passante et la latence TCP maximales. Ce dernier se trouve dans `/usr/lib/vmware/vsan/bin/iperf.copy`. L'exécuter avec `--help` pour afficher les différentes options. Utilisez cet outil pour vérifier la bande passante et la latence réseau entre les hôtes ESXi participant à un cluster vSAN.

L'article 2001003 de la base de connaissances VMware peut vous assister dans les tâches de configuration et de test.

Cela est très utile lorsqu'un cluster vSAN est mis en service. L'exécution des tests de **iperf** sur le réseau vSAN lorsque le cluster est déjà en production peut nuire aux performances des machines virtuelles exécutées sur le cluster.

Vérification des limites du réseau vSAN

La commande **`vsan.check.limits`** vérifie qu'aucun des seuils vSAN n'est dépassé.

```
> ls
0 /
1 vcsa-04.rainpole.com/
> cd 1
/vcsa-04.rainpole.com> ls
0 Datacenter (datacenter)
/vcsa-04.rainpole.com> cd 0
/vcsa-04.rainpole.com/Datacenter> ls
0 storage/
1 computers [host]/
2 networks [network]/
3 datastores [datastore]/
4 vms [vm]/
/vcsa-04.rainpole.com/Datacenter> cd 1
/vcsa-04.rainpole.com/Datacenter/computers> ls
0 Cluster (cluster): cpu 155 GHz, memory 400 GB
1 esxi-dell-e.rainpole.com (standalone): cpu 38 GHz, memory 123 GB
2 esxi-dell-f.rainpole.com (standalone): cpu 38 GHz, memory 123 GB
3 esxi-dell-g.rainpole.com (standalone): cpu 38 GHz, memory 123 GB
4 esxi-dell-h.rainpole.com (standalone): cpu 38 GHz, memory 123 GB
/vcsa-04.rainpole.com/Datacenter/computers> vsan.check_limits 0
2017-03-14 16:09:32 +0000: Querying limit stats from all hosts ...
2017-03-14 16:09:34 +0000: Fetching vSAN disk info from esxi-dell-m.rainpole.com (may take a
moment) ...
2017-03-14 16:09:34 +0000: Fetching vSAN disk info from esxi-dell-n.rainpole.com (may take a
moment) ...
2017-03-14 16:09:34 +0000: Fetching vSAN disk info from esxi-dell-o.rainpole.com (may take a
moment) ...
2017-03-14 16:09:34 +0000: Fetching vSAN disk info from esxi-dell-p.rainpole.com (may take a
moment) ...
2017-03-14 16:09:39 +0000: Done fetching vSAN disk infos
+--------------------------+--------------------
+-------------------------------------------------------------------+
| Host                     | RDT
| Disks                                                             |
```

```
+--------------------------+--------------------
+-------------------------------------------------------------------+
| esxi-dell-m.rainpole.com |
Assocs: 1309/45000 | Components: 485/9000                                                |
|                          | Sockets:
89/10000  | naa.500a075113019b33: 0% Components: 0/0                          |
|                          | Clients:
136       | naa.500a075113019b37: 40% Components: 81/47661                    |
|                          | Owners:
138        | t10.ATA_____Micron_P420m2DMTFDGAR1T4MAX_____ 0% Components: 0/0 |
|                          |                     |
naa.500a075113019b41: 37% Components: 80/47661                   |
|                          |                     |
naa.500a07511301a1eb: 38% Components: 81/47661                   |
|                          |                     |
naa.500a075113019b39: 39% Components: 79/47661                   |
|                          |                     |
naa.500a07511301a1ec: 41% Components: 79/47661                   |
<<truncated>>
```

Du point de vue du réseau, le nombre d'associations RDT (assoc.) et de sockets constituent les aspects les plus importants. Il existe 45 000 associations par hôte dans vSAN 6.0 et version ultérieure. Une association RDT est utilisée pour suivre l'état du réseau poste à poste dans vSAN. vSAN est dimensionné de sorte qu'il n'est jamais à court d'associations RDT. vSAN limite également le nombre de sockets TCP qu'il est autorisé à utiliser et vSAN est dimensionné de sorte que son allocation de sockets TCP n'arrive jamais à échéance. Il existe une limite de 10 000 sockets par hôte.

Un **client** vSAN représente l'accès de l'objet dans le cluster vSAN. Le client représente généralement une machine virtuelle s'exécutant sur un hôte. Le client et l'objet ne se trouvent pas sur le même hôte. Il n'existe pas de limite définie, mais cette mesure s'affiche pour vous aider à comprendre comment les clients s'équilibrent entre les hôtes.

Il n'y a qu'un seul **propriétaire** vSAN pour un objet vSAN spécifique, généralement colocalisé avec le client vSAN accédant à cet objet. Les propriétaires de vSAN coordonnent tous les accès à l'objet vSAN et mettent en œuvre des fonctionnalités comme la mise en miroir et la répartition. Il n'existe pas de limite définie, mais cette mesure est à nouveau affichée afin de mieux comprendre comment les propriétaires s'équilibrent entre les hôtes.

Utilisation de la multidiffusion dans un réseau vSAN

13

La multidiffusion est une technique de communication réseau qui envoie des paquets d'information à un groupe de destinations sur un réseau IP.

Les versions de vSAN antérieures à 6.6 prennent en charge la multidiffusion IP et la communication multidiffusion IP utilisée comme protocole de détection pour identifier les nœuds qui tentent de joindre un cluster vSAN. Les versions de vSAN antérieures à la version 6.6 dépendent de la communication multidiffusion IP lors de l'entrée dans les groupes de clusters et la sortie de ces derniers, et pendant d'autres opérations de communication intra-cluster. Assurez-vous d'activer et de configurer la multidiffusion IP dans les segments de réseau IP pour distribuer le service de trafic vSAN.

Une adresse multidiffusion IP est appelée groupe multidiffusion (MG). La multidiffusion IP envoie des paquets source à plusieurs récepteurs sous forme de transmission groupée. La multidiffusion IP repose sur les protocoles de communication que les hôtes, les clients et les périphériques réseau utilisent pour participer aux communications basées sur la multidiffusion. Les protocoles de communication comme IGMP (Internet Group Management Protocol) et PIM (Protocol Independent Multicast) sont les principaux composants et dépendances pour l'utilisation des communications multidiffusion IP.

Lors de la création d'un cluster vSAN, une adresse multidiffusion par défaut est affectée à chaque cluster vSAN. Le service de trafic vSAN affecte automatiquement les paramètres d'adresse multidiffusion par défaut à chaque hôte. Cette adresse multidiffusion envoie des trames à un groupe multidiffusion par défaut et à un agent du groupe multidiffusion.

Lorsque plusieurs clusters vSAN résident sur le même réseau de couche 2, VMware recommande de modifier l'adresse multidiffusion par défaut dans les clusters vSAN supplémentaires. Ceci permet d'éviter que plusieurs clusters reçoivent tous les flux multidiffusion. Pour de plus amples informations sur la modification de l'adresse multidiffusion vSAN par défaut, consultez l'article 2075451 de la base de connaissances VMware.

Ce chapitre contient les rubriques suivantes :

Protocole de gestion de groupes Internet

Vous pouvez utiliser le protocole IGMP (Internet Group Management Protocol) pour ajouter des récepteurs à l'appartenance du groupe multidiffusion IP dans les domaines de couche 2.

IGMP permet aux récepteurs d'envoyer des demandes aux groupes multidiffusion qu'ils souhaitent joindre. Devenir membre d'un groupe multidiffusion permet aux routeurs de transférer le trafic des groupes multidiffusion sur le segment de couche 3 où le récepteur est connecté au port de commutateur.

Vous pouvez utiliser l'écoute IGMP pour limiter les ports de commutateur physique participant au groupe multidiffusion aux liaisons montantes de port VMkernel vSAN uniquement. L'écoute IGMP est configurée avec une requête d'écoute IGMP. La nécessité de configurer une requête d'écoute IGMP pour prendre en charge l'écoute IGMP varie selon le fournisseur du commutateur. Consultez votre fournisseur de commutateur spécifique pour la configuration de l'écoute IGMP.

vSAN prend en charge IGMP version 2 et IGMP version 3. Lorsque vous effectuez le déploiement de vSAN sur des segments de réseau de couche 3, vous pouvez configurer un périphérique compatible de couche 3 tel qu'un routeur ou un commutateur avec une connexion et un accès aux mêmes segments de réseau de couche 3.

Tous les ports VMkernel du réseau vSAN s'abonnent à un groupe multidiffusion via IGMP pour éviter une saturation du trafic multidiffusion de tous les ports réseau.

Note Vous pouvez désactiver l'écoute IGMP si vSAN se trouve sur un réseau VLAN non acheminé ou lié que vous pouvez étendre aux ports vSAN de tous les hôtes du cluster.

Multidiffusion indépendante du protocole

Le protocole PIM (Protocol Independent Multicast) est constitué de protocoles de routage multidiffusion de couche 3.

Il fournit différentes techniques de communication pour le trafic multidiffusion IP afin d'atteindre les récepteurs qui se trouvent dans des segments de couche 3 différents des sources de groupes multidiffusion. Pour un cluster vSAN antérieur à la version 6.6, vous devez utiliser PIM pour permettre l'acheminement du trafic multidiffusion sur différents sous-réseaux. Consultez votre fournisseur de réseau pour l'implémentation de PIM.

Considérations de mise en réseau pour iSCSI sur vSAN

14

Le service cible iSCSI vSAN permet aux hôtes et aux charges de travail physiques qui résident en dehors du cluster vSAN d'accéder à la banque de données vSAN. Cette fonctionnalité active un initiateur iSCSI sur un hôte distant afin de transporter les données de niveau bloc vers une cible iSCSI sur un périphérique de stockage au sein d'un cluster vSAN.

Les cibles iSCSI sur vSAN sont gérées à l'aide de SPBM (Storage Policy Based Management), comme pour d'autres objets vSAN. Cela permet d'économiser de l'espace et de garantir la sécurité des LUN iSCSI via la déduplication et la compression, ainsi que le chiffrement. Pour améliorer la sécurité, le service cible iSCSI vSAN utilise le protocole CHAP (Challenge Handshake Authentication Protocol) et l'authentification CHAP mutuelle.

vSAN identifie chaque cible iSCSI en fonction de son nom qualifié iSCSI (IQN) unique. La cible iSCSI est présentée à un initiateur iSCSI distant à l'aide de l'IQN, afin que l'initiateur puisse accéder au LUN de la cible. Le service cible iSCSI vSAN permet de créer des groupes d'initiateurs iSCSI. Le groupe d'initiateurs iSCSI limite l'accès aux initiateurs membres du groupe uniquement.

Ce chapitre contient les rubriques suivantes :

Caractéristiques d'un réseau vSAN iSCSI

Les caractéristiques d'un réseau vSAN iSCSI sont présentées ci-dessous :

- Routage iSCSI : les initiateurs iSCSI peuvent établir des connexions routées vers des cibles iSCSI vSAN sur un réseau L3.
- IPv4 et IPv6 : le réseau vSAN iSCSI prend en charge les protocoles IPv4 et IPv6.
- Sécurité IP : le protocole IPSec sur le réseau iSCSI vSAN offre une sécurité accrue.

  Note Les hôtes ESXi prennent en charge IPsec via IPv6 uniquement.

- Trames Jumbo : les trames Jumbo sont prises en charge sur le réseau vSAN iSCSI.
- Association de cartes réseau : toutes les configurations d'association de cartes réseau sont prises en charge sur le réseau vSAN iSCSI.
- Plusieurs connexions par session (MCS) : l'implémentation d'iSCSI vSAN ne prend pas en charge MCS.

Migration du commutateur standard vers le commutateur distribué

15

Vous pouvez migrer un vSphere Standard Switch vers un vSphere Distributed Switch et utiliser Network I/O Control. Cela vous permet de hiérarchiser la QoS (Quality of Service) sur le trafic vSAN.

Avertissement Il est préférable d'avoir accès aux hôtes ESXi, même si vous n'en avez pas besoin. En cas de problème, vous pouvez accéder à la console des hôtes ESXi.

Notez la configuration du vSwitch d'origine. En particulier, notez les paramètres d'équilibrage de charge et d'association de cartes réseau sur la source. Assurez-vous que la configuration de la destination correspond à la source.

Créer un commutateur distribué

Créez le vSwitch distribué et attribuez-lui un nom.

1 Dans la vue Hôte et Clusters de vSphere Client, cliquez avec le bouton droit sur un centre de données et sélectionnez **Nouveau commutateur distribué**.

2 Entrez un nom.

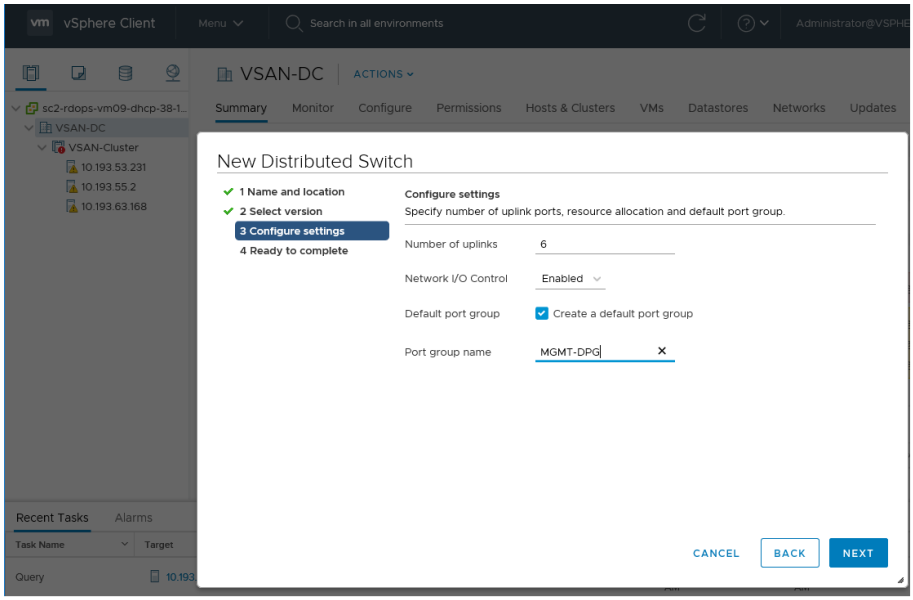
3 Sélectionnez la version du vSphere Distributed Switch. Dans cet exemple, la version 6.6.0 est utilisée pour la migration.

4 Ajoutez les paramètres. Déterminez le nombre de liaisons montantes que vous utilisez actuellement pour la mise en réseau. Cet exemple en compte six : gestion, vMotion, machines virtuelles et trois pour vSAN (une configuration LAG). Entrez 6 pour le nombre de liaisons montantes. Votre environnement peut être différent, mais vous pouvez le modifier ultérieurement.

   À ce stade, vous pouvez créer un groupe de ports par défaut, mais des groupes de ports supplémentaires sont nécessaires.

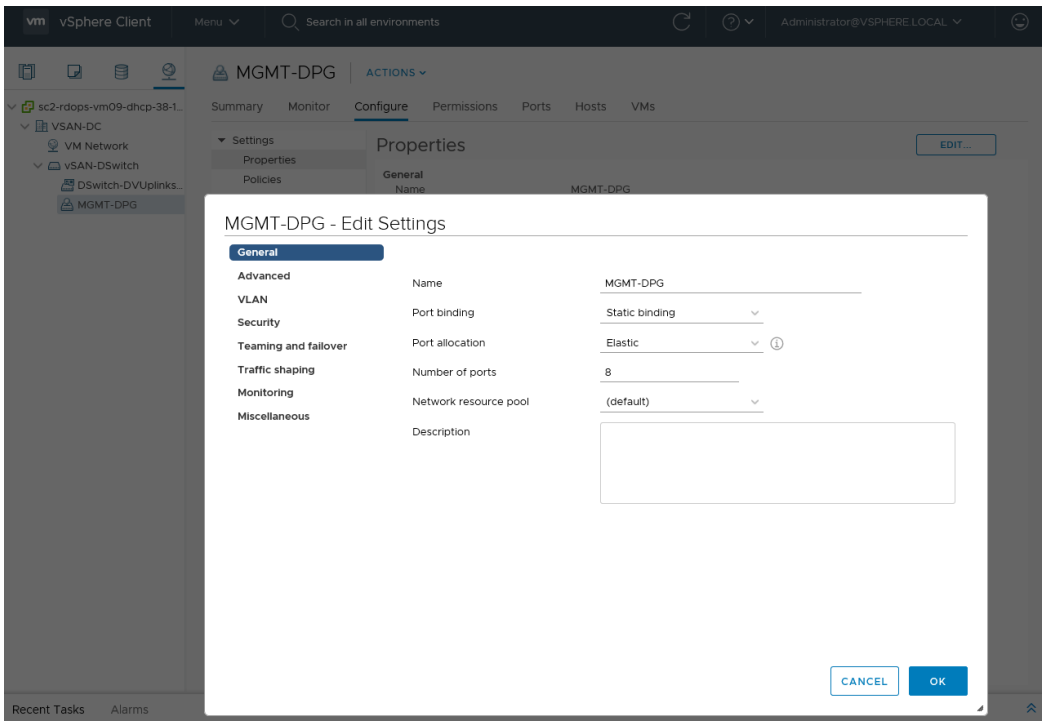
5 Terminez la configuration du vSwitch distribué.

La prochaine étape consiste à configurer et à créer les groupes de ports supplémentaires.



Créer des groupes de ports

Un groupe de ports par défaut unique a été créé pour le réseau de gestion. Modifiez ce groupe de ports pour vous assurer qu'il dispose de toutes les caractéristiques du groupe de ports de gestion sur le vSwitch Standard, tel que l'association de VLAN et de cartes réseau, et les paramètres de basculement.



Configurez le groupe de ports de gestion.

1 Dans la vue Mise en réseau de vSphere Client, sélectionnez le groupe de ports distribués, puis cliquez sur **Modifier**.

2 Pour certains groupes de ports, vous devez modifier le réseau VLAN. Comme VLAN 51 est le VLAN de gestion, marquez le groupe de ports distribués en conséquence.

3 Cliquez sur **OK**.

Créez des groupes de ports distribués pour vMotion, la mise en réseau de machines virtuelles et la mise en réseau de vSAN.

1 Cliquez avec le bouton droit sur le vSphere Distributed Switch et sélectionnez **Groupe de ports distribués > Nouveau groupe de ports distribués** dans le menu.

2 Pour cet exemple, créez un groupe de ports pour le réseau vMotion.

Créez tous les groupes de ports distribués sur le vSwitch distribué. Migrez ensuite les liaisons montantes, la mise en réseau VMkernel et la mise en réseau des machines virtuelles vers le vSwitch distribué et les groupes de ports distribués associés.

Avertissement Migrez les liaisons montantes et les réseaux en mode pas à pas pour continuer sans incident et avec précaution.

Migrer le réseau de gestion

Migrez le réseau de gestion (vmk0) et la liaison montante associée (vmnic0) à partir du vSwitch standard vers le vSwitch distribué (vDS).

1 Ajoutez des hôtes au vDS.

   a Cliquez avec le bouton droit sur le vDS et sélectionnez **Ajouter et gérer des hôtes** dans le menu.

   b Ajoutez des hôtes au vDS. Cliquez sur l'icône verte Ajouter (+) et ajoutez tous les hôtes du cluster.

2 Configurez les adaptateurs physiques et les adaptateurs VMkernel.

   a Cliquez sur **Gérer les adaptateurs physiques** pour migrer les adaptateurs physiques et les adaptateurs VMkernel, vmnic0 et vmk0 vers le vDS.

   b Sélectionnez une liaison montante appropriée sur le vDS pour l'adaptateur physique vmnic0. Dans cet exemple, utilisez Uplink1. L'adaptateur physique est sélectionné et une liaison montante est choisie.

3 Migrez le réseau de gestion sur vmk0 depuis le vSwitch standard vers le vSwitch distribué. Effectuez ces étapes sur chaque hôte.

   a Sélectionnez vmk0, et cliquez sur **Affecter un groupe de ports**.

   b Affectez le groupe de ports distribués créé au préalable pour le réseau de gestion.

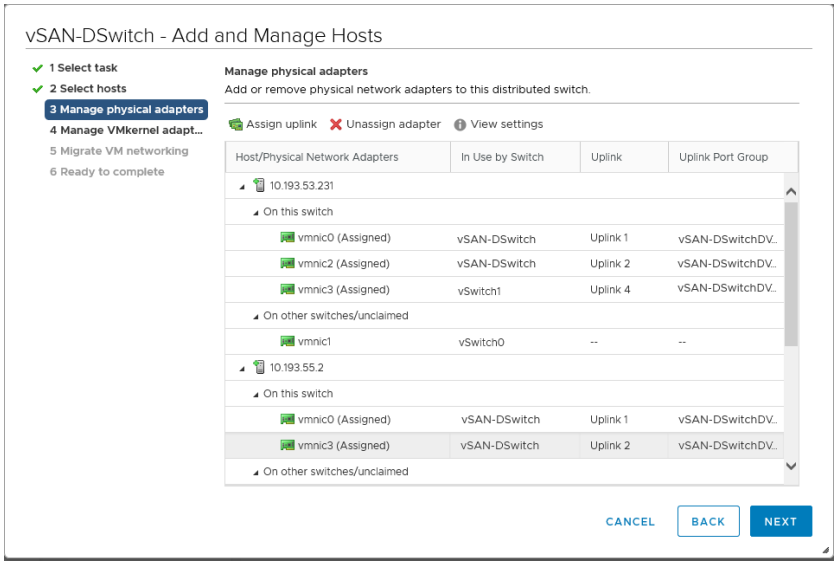
4 Terminez la configuration.

   a Vérifiez les modifications pour vous assurer que vous ajoutez quatre hôtes, quatre liaisons montantes (vmnic0 depuis chaque hôte) et quatre adaptateurs VMkernel (vmk0 depuis chaque hôte).

   b Cliquez sur **Terminer**.

Lorsque vous examinez la configuration de mise en réseau de chaque hôte, vérifiez les paramètres du commutateur, ceux-ci devant intégrer une liaison montante (vmnic0) et le port de gestion vmk0 sur chaque hôte.

Répétez ce processus pour les autres réseaux.



Migrer vMotion

Pour migrer le réseau vMotion, suivez les mêmes étapes que celles utilisées pour le réseau de gestion.

Avant de commencer, assurez-vous que le groupe de ports distribués pour le réseau vMotion possède les mêmes attributs que le groupe de ports sur le vSwitch standard. Migrez ensuite la liaison montante utilisée pour vMotion (vmnic1), avec l'adaptateur VMkernel (vmk1).

Migrer le réseau vSAN

Si vous disposez d'une seule liaison montante pour le réseau vSAN, utilisez le même processus qu'auparavant. Cependant, si vous utilisez plusieurs liaisons montantes, des étapes supplémentaires sont nécessaires.

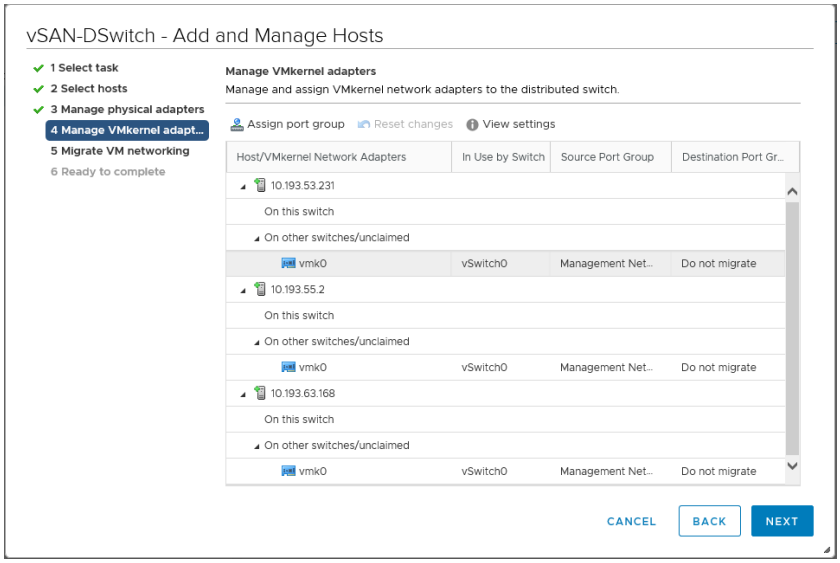
Si le réseau vSAN utilise l'agrégation de liens (LACP) ou qu'il se trouve sur un réseau VLAN différent des autres réseaux VMkernel, placez certaines liaisons montantes dans un état inutilisé pour certains adaptateurs VMkernel.

Par exemple, l'adaptateur VMkernel vmk2 est utilisé pour vSAN. Cependant, les liaisons montantes vmnic3, 4 et 5 sont utilisées pour vSAN et elles se trouvent dans une configuration LACP. Par conséquent, pour vmk2, tous les autres vmnic (0, 1 et 2) doivent être placés dans un état non utilisé. De même, pour l'adaptateur de gestion (vmk0) et l'adaptateur vMotion (vmk0), placez les liaisons montantes/vmnic vSAN dans un état non utilisé.

Modifiez les paramètres du groupe de ports distribués et modifiez la stratégie de chemin d'accès et les paramètres de basculement. Sur la page **Gérer l'adaptateur réseau physique**, procédez comme suit pour plusieurs adaptateurs.

Affectez l'adaptateur VMkernel (vmk2) vSAN au groupe de ports distribués pour vSAN.

Note Si vous ne migrez que les liaisons montantes pour le réseau vSAN, vous ne pourrez peut-être modifier les paramètres du groupe de ports distribués qu'après la migration. Pendant ce temps, vSAN peut présenter des problèmes de communication. Une fois la migration effectuée, passez aux paramètres du groupe de ports distribués, et effectuez les modifications de stratégie et marquez toutes les liaisons montantes qui ne doivent pas être utilisées. La mise en réseau de vSAN revient ensuite à la normale lorsque cette tâche est terminée. Utilisez le service de santé vSAN pour vérifier que tout fonctionne correctement.



Migrer le réseau de VM

La dernière tâche nécessaire à la migration du réseau à partir d'un vSwitch standard vers un vSwitch distribué consiste à migrer le réseau de la machine virtuelle.

Gérer la mise en réseau des hôtes

1 Cliquez avec le bouton droit sur le vDS et sélectionnez **Ajouter et gérer des hôtes** dans le menu.

2 Sélectionnez tous les hôtes du cluster pour migrer la mise en réseau de machines virtuelles pour tous les hôtes vers le vSwitch distribué.

Ne déplacez pas de liaisons montantes. Cependant, si la mise en réseau de machines virtuelles sur vos hôtes utilisait une autre liaison montante, migrez la liaison montante à partir du vSwitch standard.

3 Sélectionnez les machines virtuelles à migrer à partir d'un réseau de machines virtuelles sur le vSwitch standard vers le groupe de ports distribués de la machine virtuelle sur le vSwitch distribué. Cliquez sur **Affecter un groupe de ports**, puis sélectionnez le groupe de ports distribués.

4 Vérifiez les modifications et cliquez sur **Terminer**. Dans cet exemple, vous passez à des machines virtuelles. Tous les modèles utilisant le réseau de machines virtuelles vSwitch standard d'origine doivent être convertis en machines virtuelles et modifiés. Le nouveau groupe de ports distribués pour les machines virtuelles doit être sélectionné en tant que réseau. Cette étape ne peut pas être obtenue via l'Assistant de migration.

Étant donné que le vSwitch standard n'a plus de liaisons montantes ou de groupes de ports, il peut être supprimé en toute sécurité.

La migration d'un vSphere Standard Switch vers un vSphere Distributed Switch est à présent terminée.

Résumé de la liste de contrôle pour le réseau vSAN

16

Utilisez le résumé de la liste de contrôle pour vérifier la configuration requise pour votre réseau vSAN.

- Vérifiez si vous utilisez une carte réseau partagée de 10 Go ou une carte réseau dédiée de 1 Go. Les clusters intégralement Flash requièrent des cartes réseau de 10 Go.
- Vérifiez que les connexions d'association de cartes réseau redondantes sont configurées.
- Vérifiez si le contrôle de flux est activé sur les cartes réseau de l'hôte ESXi.
- Vérifiez que le port VMkernel pour le trafic réseau vSAN est configuré sur chaque hôte.
- Vérifiez si vous disposez d'un réseau VLAN, d'un MTU et d'un sous-réseau identiques sur toutes les interfaces.
- Vérifiez que vous pouvez exécuter **`vmkping`** correctement entre tous les hôtes. Utilisez le service de santé pour la vérification.
- Si vous utilisez des trames Jumbo, vérifiez que vous pouvez exécuter correctement **`vmkping`** avec un paquet de taille 9 000 entre tous les hôtes. Utilisez le service de santé pour la vérification.
- Si votre version de vSAN est antérieure à la version 6.6, vérifiez si le trafic multidiffusion est activé sur le réseau.
- Si votre version de vSAN est antérieure à la version 6.6 et que plusieurs clusters vSAN résident sur le même réseau, configurez le trafic multidiffusion de sorte à utiliser des adresses multidiffusion uniques.
- Si votre version de vSAN est antérieure à la version 6.6 et s'étend sur plusieurs commutateurs, vérifiez si le trafic multidiffusion est configuré sur l'ensemble des commutateurs.
- Si votre version de vSAN est antérieure à la version 6.6 et est acheminée, vérifiez si PIM est configuré pour autoriser le routage du trafic multidiffusion.
- Assurez-vous que le commutateur physique peut satisfaire les conditions requises vSAN (multidiffusion, contrôle de flux, interopérabilité des fonctionnalités).
- Vérifiez que le réseau ne rencontre pas de problèmes de performances, comme un nombre excessif de paquets ignorés ou de trames de pause.
- Vérifiez que les limites du réseau se situent dans des marges acceptables.

- Testez les performances du réseau vSAN avec **`iperf`** et vérifiez qu'il satisfait aux attentes.