Cluster de stockage vSphere Metro avec ONTAP

Enterprise applications

NetApp
May 09, 2024

Sommaire

Cluster de stockage vSphere Metro avec ONTAP

Cluster de stockage vSphere Metro avec ONTAP

L'hyperviseur vSphere de pointe de VMware peut être déployé en tant que cluster étendu appelé vMSC (vSphere Metro Storage Cluster).

Les solutions VMSC sont prises en charge avec NetApp® MetroCluster™ et la synchronisation active SnapMirror (anciennement appelée SnapMirror Business Continuity ou SMBC) et assurent une continuité de l'activité avancée si un ou plusieurs domaines à défaillance subissent une panne totale. La résilience aux différents modes de défaillance dépend des options de configuration que vous choisissez.

Disponibilité continue pour les environnements vSphere

L'architecture ONTAP est une plateforme de stockage flexible et évolutive qui fournit des services SAN (FCP, iSCSI et NVMe-of) et NAS (NFS v3 et v4.1) pour les datastores. Les systèmes de stockage NetApp AFF, ASA et FAS utilisent le système d'exploitation ONTAP pour offrir des protocoles supplémentaires pour l'accès au stockage invité comme S3 et SMB/CIFS.

NetApp MetroCluster utilise la fonction HA (basculement du contrôleur ou CFO) de NetApp pour se protéger contre les défaillances du contrôleur. Elle inclut également la technologie SyncMirror locale, le basculement de cluster en cas d'incident (basculement du contrôleur à la demande ou CFOD), la redondance matérielle et la séparation géographique pour atteindre des niveaux élevés de disponibilité. SyncMirror met en miroir les données de manière synchrone sur les deux moitiés de la configuration MetroCluster en écrivant les données sur deux plexes : le plex local (sur le tiroir local) assure activement le service des données et le plex distant (sur le tiroir distant) n'assure généralement pas le service des données. La redondance matérielle est mise en place pour tous les composants MetroCluster, tels que les contrôleurs, le stockage, les câbles, les commutateurs (utilisés avec Fabric MetroCluster) et les adaptateurs.

La synchronisation active NetApp SnapMirror offre une protection granulaire des datastores avec les protocoles SAN FCP et iSCSI, ce qui vous permet de protéger de manière sélective uniquement les workloads prioritaires. Il offre un accès actif/actif aux sites locaux et distants, contrairement à NetApp MetroCluster, qui est une solution de secours actif. Actuellement, la synchronisation active est une solution asymétrique où l'un des côtés est préféré à l'autre, offrant de meilleures performances. Pour ce faire, la fonctionnalité ALUA (Asymmetric Logical Unit Access) informe automatiquement l'hôte ESXi des contrôleurs qui lui préfèrent. Cependant, NetApp a annoncé qu'une synchronisation active permettra bientôt un accès totalement symétrique.

Pour créer un cluster VMware HA/DRS sur deux sites, les hôtes ESXi sont utilisés et gérés par une appliance vCenter Server (VCSA). Les réseaux de gestion vSphere, vMotion® et machine virtuelle sont connectés via un réseau redondant entre les deux sites. Le serveur vCenter gérant le cluster HA/DRS peut se connecter aux hôtes ESXi sur les deux sites et doit être configuré à l'aide de vCenter HA.

Reportez-vous à la section "Comment créer et configurer des clusters dans le client vSphere" Pour configurer vCenter HA.

Reportez-vous également à la section "Bonnes pratiques pour VMware vSphere Metro Storage Cluster".

Qu'est-ce que le cluster de stockage vSphere Metro ?

VSphere Metro Storage Cluster (vMSC) est une configuration certifiée qui protège les machines virtuelles et les conteneurs contre les défaillances. Pour y parvenir, les concepts de stockage étendus ainsi que les

clusters d'hôtes ESXi sont répartis sur différents domaines à défaillance, tels que les racks, les bâtiments, les campus ou même les villes. Les technologies de stockage avec synchronisation active NetApp MetroCluster et SnapMirror assurent respectivement une protection RPO=0 ou RPO=0 aux clusters hôtes. La configuration vMSC est conçue pour assurer la disponibilité continue des données, même en cas de défaillance d'un « site » physique ou logique complet. Un périphérique de stockage faisant partie de la configuration vMSC doit être certifié après avoir suivi un processus de certification vMSC réussi. Tous les périphériques de stockage pris en charge sont disponibles dans le "Guide de compatibilité du stockage VMware".

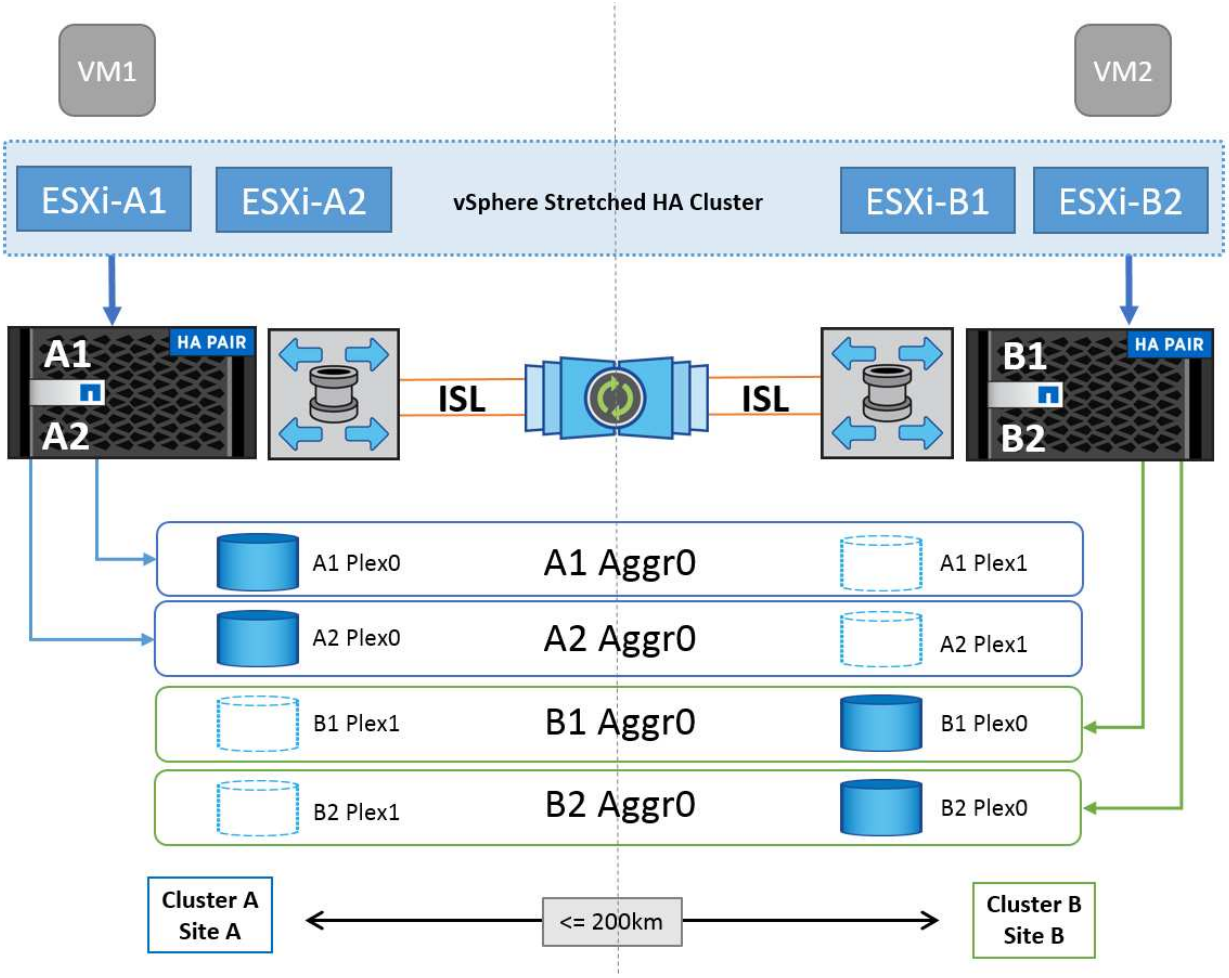
Pour plus d'informations sur les conseils de conception pour vSphere Metro Storage Cluster, reportez-vous à la documentation suivante :

- "Prise en charge de VMware vSphere avec NetApp MetroCluster"
- "Prise en charge de VMware vSphere avec la continuité de l'activité NetApp SnapMirror" (Maintenant appelé synchronisation active SnapMirror)

Selon les considérations relatives à la latence, NetApp MetroCluster peut être déployé dans deux configurations différentes pour une utilisation avec vSphere :

- MetroCluster extensible
- MetroCluster de structure

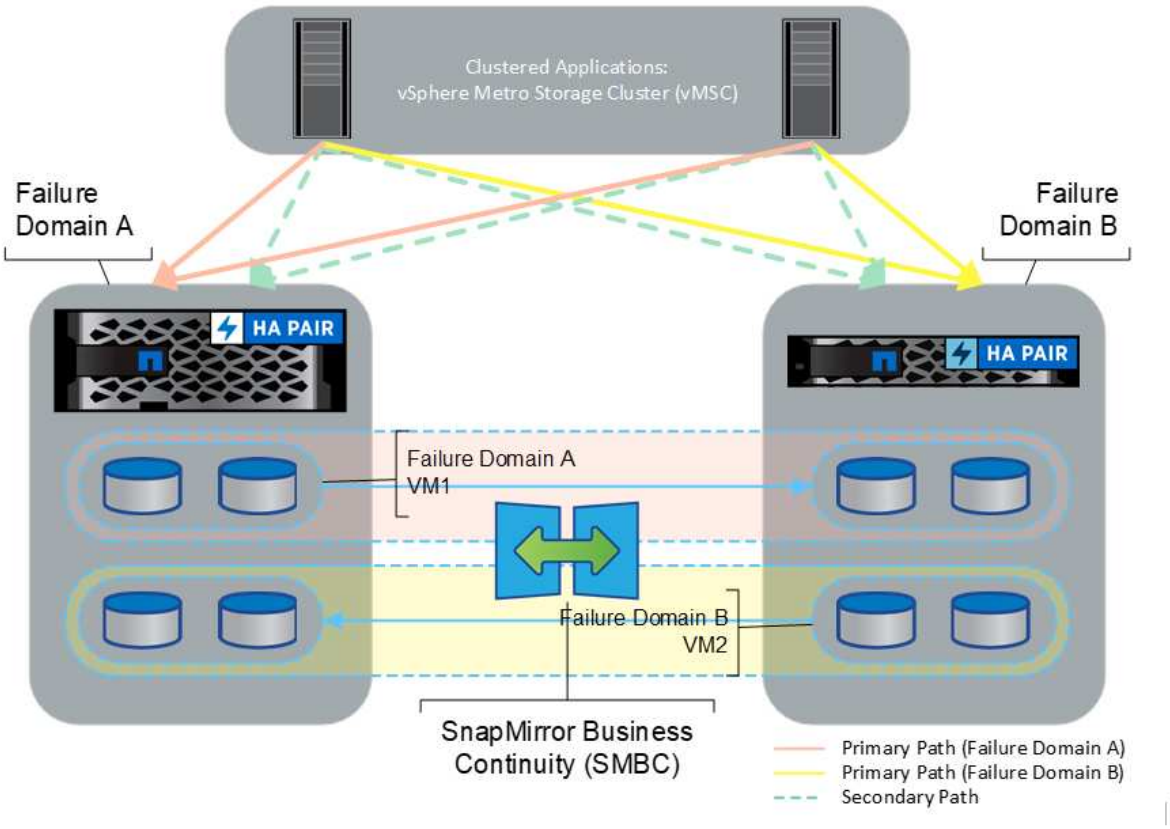
Voici une illustration de la topologie générale d'Stretch MetroCluster.



Reportez-vous à la section "Documentation MetroCluster" Pour obtenir des informations spécifiques sur la conception et le déploiement de MetroCluster.

La synchronisation active SnapMirror peut également être déployée de deux manières différentes.

- Asymétrique
- Symétrique (préversion privée dans ONTAP 9.14.1)



Reportez-vous à la section "Documents NetApp" Pour des informations spécifiques sur le design et le déploiement de SnapMirror active Sync.

Présentation de la solution VMware vSphere

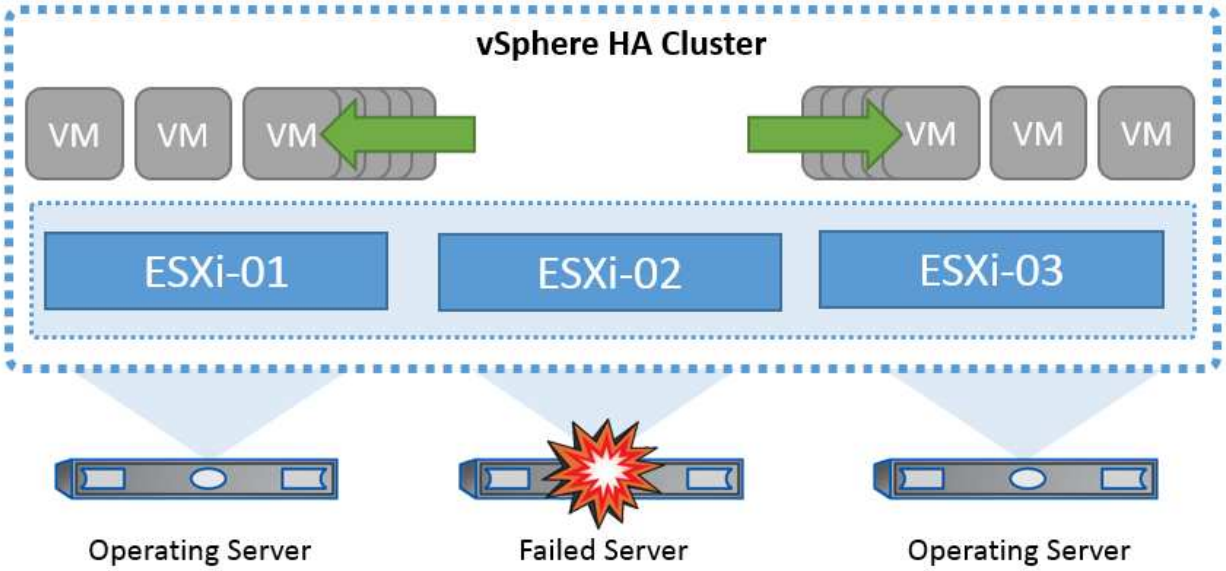
VCenter Server Appliance (VCSA) est le puissant système de gestion centralisée et une interface unique pour vSphere qui permet aux administrateurs d'exploiter efficacement les clusters ESXi. Cet outil facilite les fonctions clés telles que le provisionnement des machines virtuelles, les opérations vMotion, la haute disponibilité (HA), Distributed Resource Scheduler (DRS), Tanzu Kubernetes Grid et bien plus encore. Elle constitue un composant essentiel des environnements clouds VMware et doit être conçue en tenant compte de la disponibilité du service.

Haute disponibilité vSphere

La technologie de cluster de VMware regroupe les serveurs ESXi en pools de ressources partagées pour les machines virtuelles et fournit la haute disponibilité (HA) vSphere. VSphere HA offre une haute disponibilité et une simplicité d'utilisation pour les applications qui s'exécutent sur des machines virtuelles. Lorsque la fonction de haute disponibilité est activée sur le cluster, chaque serveur ESXi maintient la communication avec les autres hôtes de sorte que si un hôte ESXi ne répond plus ou est isolé, le cluster de haute disponibilité peut

négocier la restauration des machines virtuelles qui s'exécutaient sur cet hôte ESXi parmi les hôtes survivants du cluster. En cas de défaillance d'un système d'exploitation invité, vSphere HA redémarre la machine virtuelle concernée sur le même serveur physique. La haute disponibilité vSphere permet de réduire les temps d'indisponibilité planifiés, d'éviter les temps d'indisponibilité non planifiés et de restaurer rapidement les données en cas de panne.

Cluster vSphere HA qui récupère les machines virtuelles à partir d'un serveur défaillant.



Il est important de comprendre que VMware vSphere ne connaît pas la synchronisation active NetApp MetroCluster ou SnapMirror et que tous les hôtes ESXi du cluster vSphere sont identifiés comme des hôtes éligibles pour les opérations de cluster haute disponibilité selon les configurations d'affinité de l'hôte et du groupe de machines virtuelles.

Détection de défaillance de l'hôte

Dès la création du cluster HA, tous les hôtes du cluster participent à des élections et l'un des hôtes devient maître. Chaque esclave exécute une pulsation réseau vers le maître, et le maître effectue à son tour une pulsation réseau sur tous les hôtes esclaves. L'hôte maître d'un cluster vSphere HA est responsable de la détection de la défaillance des hôtes esclaves.

En fonction du type de défaillance détecté, les machines virtuelles exécutées sur les hôtes peuvent avoir besoin d'être basculées.

Dans un cluster vSphere HA, trois types de défaillance d'hôte sont détectés :

- Défaillance - Un hôte cesse de fonctionner.
- Isolation - Un hôte devient isolé du réseau.
- Partition : Un hôte perd la connectivité réseau avec l'hôte maître.

L'hôte maître surveille les hôtes esclaves du cluster. Cette communication s'effectue par échange de battements de cœur réseau toutes les secondes. Lorsque l'hôte maître cesse de recevoir ces battements de cœur d'un hôte esclave, il vérifie la livenité de l'hôte avant de déclarer l'échec de l'hôte. La vérification de la livenité effectuée par l'hôte maître consiste à déterminer si l'hôte esclave échange des pulsations avec l'un des datastores. En outre, l'hôte maître vérifie si l'hôte répond aux requêtes ping ICMP envoyées à ses adresses IP de gestion pour détecter s'il est simplement isolé de son nœud maître ou complètement isolé du réseau. Pour

ce faire, il exécute une commande ping sur la passerelle par défaut. Une ou plusieurs adresses d'isolement peuvent être spécifiées manuellement pour améliorer la fiabilité de la validation de l'isolement.

Meilleure pratique

NetApp recommande de spécifier au moins deux adresses d'isolement supplémentaires, et que chacune de ces adresses soit site-local. Cela améliorera la fiabilité de la validation de l'isolement.

Réponse d'isolation de l'hôte

Isolation Response est un paramètre de vSphere HA qui détermine l'action déclenchée sur les machines virtuelles lorsqu'un hôte d'un cluster vSphere HA perd ses connexions réseau de gestion mais continue à s'exécuter. Il existe trois options pour ce paramètre, « Désactivé », « Arrêter et redémarrer les machines virtuelles » et « Arrêter et redémarrer les machines virtuelles ».

Il est préférable d'arrêter le système plutôt que de le mettre hors tension, qui ne vide pas les dernières modifications apportées au disque ou ne commet pas les transactions. Si les machines virtuelles ne s'arrêtent pas dans les 300 secondes, elles sont éteintes. Pour modifier le temps d'attente, utilisez l'option avancée das.isolashutdowntimeout.

Avant que la haute disponibilité ne lance la réponse d'isolation, elle vérifie d'abord si l'agent principal vSphere HA possède le datastore qui contient les fichiers de configuration de la machine virtuelle. Si ce n'est pas le cas, l'hôte ne déclenchera pas la réponse d'isolation, car il n'y a pas de maître pour redémarrer les machines virtuelles. L'hôte vérifie régulièrement l'état du datastore pour déterminer s'il est demandé par un agent vSphere HA qui détient le rôle principal.

Meilleure pratique

NetApp recommande de définir la « réponse d'isolation de l'hôte » sur Désactivé.

Une condition de split-brain peut se produire si un hôte est isolé ou partitionné à partir de l'hôte maître vSphere HA et que le maître ne peut pas communiquer via des datastores heartbeat ou par ping. Le maître déclare l'hôte isolé comme étant mort et redémarre les machines virtuelles sur les autres hôtes du cluster. Une condition de split-brain existe maintenant parce qu'il y a deux instances de la machine virtuelle en cours d'exécution, dont une seule peut lire ou écrire les disques virtuels. Il est désormais possible d'éviter les conditions de split-brain en configurant VM Component protection (VMCP).

Protection des composants VM (VMCP)

L'une des améliorations de vSphere 6, concernant la haute disponibilité, est VMCP. VMCP offre une protection améliorée contre les conditions de tous les chemins d'accès (APD) et de perte permanente de périphérique (PDL) pour le stockage bloc (FC, iSCSI, FCoE) et de fichiers (NFS).

Perte permanente de périphérique (PDL)

PDL est une condition qui se produit lorsqu'un périphérique de stockage tombe en panne de manière permanente ou est supprimé administrativement et ne devrait pas revenir. La baie de stockage NetApp émet un code de détection SCSI pour ESXi déclarant que le périphérique est définitivement perdu. Dans la section Conditions de défaillance et réponse de la machine virtuelle de vSphere HA, vous pouvez configurer la réponse après la détection d'une condition PDL.

Meilleure pratique

NetApp recommande de définir la "réponse du datastore avec PDL" sur "**éteindre et redémarrer les**

machines virtuelles". Lorsque cette condition est détectée, une machine virtuelle est redémarrée instantanément sur un hôte sain dans le cluster vSphere HA.

Tous les chemins en panne (APD)

L'APD est une condition qui se produit lorsqu'un périphérique de stockage devient inaccessible à l'hôte et qu'aucun chemin vers la matrice n'est disponible. ESXi considère cela comme un problème temporaire avec le périphérique et s'attend à ce qu'il redevienne disponible.

Lorsqu'une condition APD est détectée, une minuterie démarre. Au bout de 140 secondes, la condition APD est officiellement déclarée et le périphérique est marqué comme étant hors délai APD. Lorsque les 140 secondes sont écoulées, la haute disponibilité commence à compter le nombre de minutes spécifié dans le délai d'attente pour le basculement de machine virtuelle. Une fois le délai spécifié écoulé, la haute disponibilité redémarre les machines virtuelles impactées. Vous pouvez configurer VMCP pour qu'il réponde différemment si vous le souhaitez (désactivé, événements de problème ou mise hors tension et redémarrage des machines virtuelles).

Meilleure pratique

NetApp recommande de configurer la « réponse pour le datastore avec APD » sur « * mettre hors tension et redémarrer les machines virtuelles (conservatrices)* ».

Conservateur fait référence à la probabilité que la haute disponibilité soit capable de redémarrer les machines virtuelles. Si elle est définie sur conservateur, la haute disponibilité ne redémarrera la machine virtuelle concernée par l'APD que si elle sait qu'un autre hôte peut la redémarrer. Dans le cas d'un environnement agressif, la haute disponibilité essaiera de redémarrer la machine virtuelle même si elle ne connaît pas l'état des autres hôtes. Cela peut entraîner le redémarrage des machines virtuelles si aucun hôte n'a accès au datastore sur lequel elles se trouvent.

Si le statut APD est résolu et que l'accès au stockage est restauré avant le délai d'expiration, la haute disponibilité ne redémarrera pas inutilement la machine virtuelle, sauf si vous la configurez explicitement pour le faire. Si une réponse est souhaitée, même lorsque l'environnement a récupéré de la condition APD, la réponse pour la restauration APD après le délai APD doit être configurée pour réinitialiser les machines virtuelles.

Meilleure pratique

NetApp recommande de configurer la réponse pour la récupération APD après le délai APD sur Désactivé.

Implémentation de VMware DRS pour NetApp MetroCluster

VMware DRS est une fonctionnalité qui regroupe les ressources hôtes dans un cluster et est principalement utilisée pour équilibrer la charge au sein d'un cluster dans une infrastructure virtuelle. VMware DRS calcule principalement les ressources CPU et mémoire pour effectuer l'équilibrage de charge dans un cluster. Étant donné que vSphere ne connaît pas la mise en cluster étendue, il prend en compte tous les hôtes des deux sites lors de l'équilibrage de charge. Pour éviter le trafic intersite, NetApp recommande de configurer des règles d'affinité DRS pour gérer une séparation logique des machines virtuelles. Cela permet de garantir que, sauf en cas de défaillance complète du site, les systèmes HA et DRS n'utilisent que les hôtes locaux.

Si vous créez une règle d'affinité DRS pour votre cluster, vous pouvez spécifier comment vSphere applique cette règle lors du basculement d'une machine virtuelle.

Vous pouvez spécifier deux types de règles pour le basculement de vSphere HA :

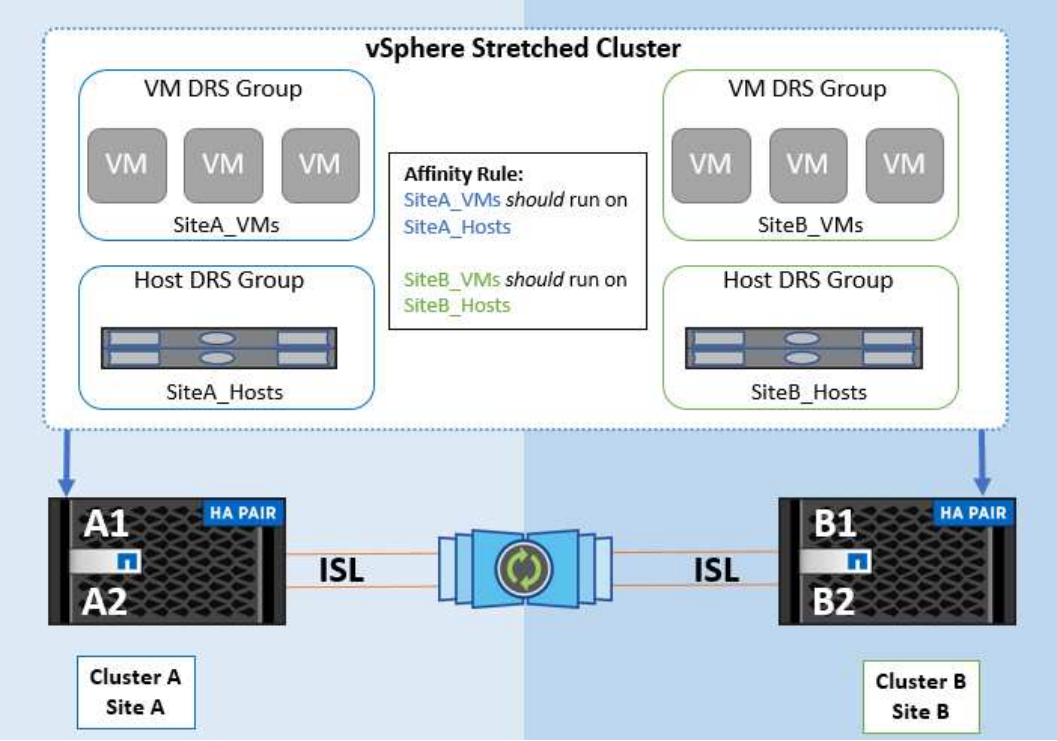
- Les règles d'anti-affinité pour les machines virtuelles forcent les machines virtuelles spécifiées à rester

séparées pendant les opérations de basculement.

- Les règles d'affinité des hôtes VM placent les machines virtuelles spécifiées sur un hôte particulier ou un membre d'un groupe défini d'hôtes lors des actions de basculement.

En utilisant les règles d'affinité pour les hôtes de machine virtuelle dans VMware DRS, il est possible d'avoir une séparation logique entre le site A et le site B, de sorte que la machine virtuelle s'exécute sur l'hôte au même site que la baie configurée comme contrôleur de lecture/écriture principal pour un datastore donné. De plus, les règles d'affinité des hôtes de VM permettent aux machines virtuelles de rester locales au stockage, ce qui à son tour ascert la connexion de la machine virtuelle en cas de défaillances réseau entre les sites.

Voici un exemple de groupes d'hôtes de machine virtuelle et de règles d'affinité.



Meilleure pratique

NetApp recommande de mettre en place des règles « a » plutôt que des règles « a », car elles sont violées par vSphere HA en cas de défaillance. L'utilisation de règles « must » peut entraîner des interruptions de service.

La disponibilité des services doit toujours prévaloir sur les performances. Lorsqu'un data Center complet tombe en panne, les règles « must » doivent choisir les hôtes du groupe d'affinité des hôtes de la machine virtuelle et, lorsque le data Center n'est pas disponible, les machines virtuelles ne redémarrent pas.

Implémentation de VMware Storage DRS avec NetApp MetroCluster

La fonction VMware Storage DRS permet l'agrégation de datastores en une seule unité et équilibre les disques de la machine virtuelle lorsque les seuils de contrôle d'E/S du stockage sont dépassés.

Le contrôle des E/S du stockage est activé par défaut sur les clusters DRS compatibles avec Storage DRS. Le contrôle des E/S du stockage permet à un administrateur de contrôler la quantité d'E/S de stockage allouée aux serveurs virtuels pendant les périodes d'encombrement des E/S. Ainsi, les serveurs virtuels plus importants sont préfénables aux serveurs virtuels moins importants pour l'allocation des ressources d'E/S.

Storage DRS utilise Storage vMotion pour migrer les machines virtuelles vers différents datastores au sein

d'un cluster de datastores. Dans un environnement NetApp MetroCluster, la migration des machines virtuelles doit être contrôlée dans les datastores de ce site. Par exemple, la machine virtuelle A, qui s'exécute sur un hôte du site A, doit idéalement migrer au sein des datastores du SVM sur le site A. Si ce n'est pas le cas, la machine virtuelle continue à fonctionner mais avec des performances dégradées, puisque la lecture/l'écriture du disque virtuel se fera à partir du site B via des liens inter-sites.

Meilleure pratique

NetApp recommande de créer des clusters de datastores en fonction de l'affinité avec les sites de stockage. En d'autres termes, les datastores avec affinité pour le site A ne doivent pas être associés à des clusters de datastores avec affinité pour le site B.

Lorsqu'une machine virtuelle est nouvellement provisionnée ou migrée à l'aide de Storage vMotion, NetApp recommande de mettre à jour manuellement toutes les règles VMware DRS spécifiques à ces machines virtuelles en conséquence. Cela permet de vérifier l'affinité de la machine virtuelle au niveau du site pour l'hôte et le datastore et de réduire ainsi la surcharge réseau et stockage.

Directives de conception et de mise en œuvre VMSC

Ce document présente les lignes directrices en matière de conception et d'implémentation pour vMSC avec systèmes de stockage ONTAP.

Configuration du stockage NetApp

Les instructions d'installation de NetApp MetroCluster (appelées « configuration MCC ») sont disponibles à l'adresse "Documentation MetroCluster". Des instructions pour la synchronisation active SnapMirror sont également disponibles à l'adresse "Présentation de la continuité de l'activité SnapMirror".

Une fois que vous avez configuré MetroCluster, son administration revient à gérer un environnement ONTAP traditionnel. Vous pouvez configurer des machines virtuelles de stockage (SVM) à l'aide de divers outils tels que l'interface de ligne de commande (CLI), System Manager ou Ansible. Une fois les SVM configurés, créez des interfaces logiques (LIF), des volumes et des LUN sur le cluster qui seront utilisés pour les opérations normales. Ces objets seront automatiquement répliqués sur l'autre cluster à l'aide du réseau de peering de cluster.

Si vous n'utilisez pas MetroCluster, vous pouvez utiliser la synchronisation active SnapMirror qui offre une protection granulaire du datastore et un accès actif-actif sur plusieurs clusters ONTAP dans différents domaines de défaillance. La synchronisation active SnapMirror utilise des groupes de cohérence pour assurer la cohérence de l'ordre d'écriture dans un ou plusieurs datastores. Vous pouvez également créer plusieurs groupes de cohérence selon les besoins de vos applications et de vos datastores. Les groupes de cohérence sont particulièrement utiles pour les applications qui nécessitent une synchronisation des données entre plusieurs datastores. La synchronisation active SnapMirror prend également en charge les mappages de périphériques Raw Device (RDM) et le stockage connecté par l'invité avec les initiateurs iSCSI invités. Pour en savoir plus sur les groupes de cohérence, consultez la page "Présentation des groupes de cohérence".

La gestion d'une configuration vMSC avec SnapMirror Active Sync est différente de celle d'un MetroCluster. Tout d'abord, il s'agit d'une configuration SAN uniquement. Les datastores NFS ne peuvent pas être protégés avec la synchronisation active SnapMirror. Ensuite, vous devez mapper les deux copies des LUN sur vos hôtes ESXi afin qu'elles puissent accéder aux datastores répliqués dans les deux domaines de défaillance.

Haute disponibilité VMware vSphere

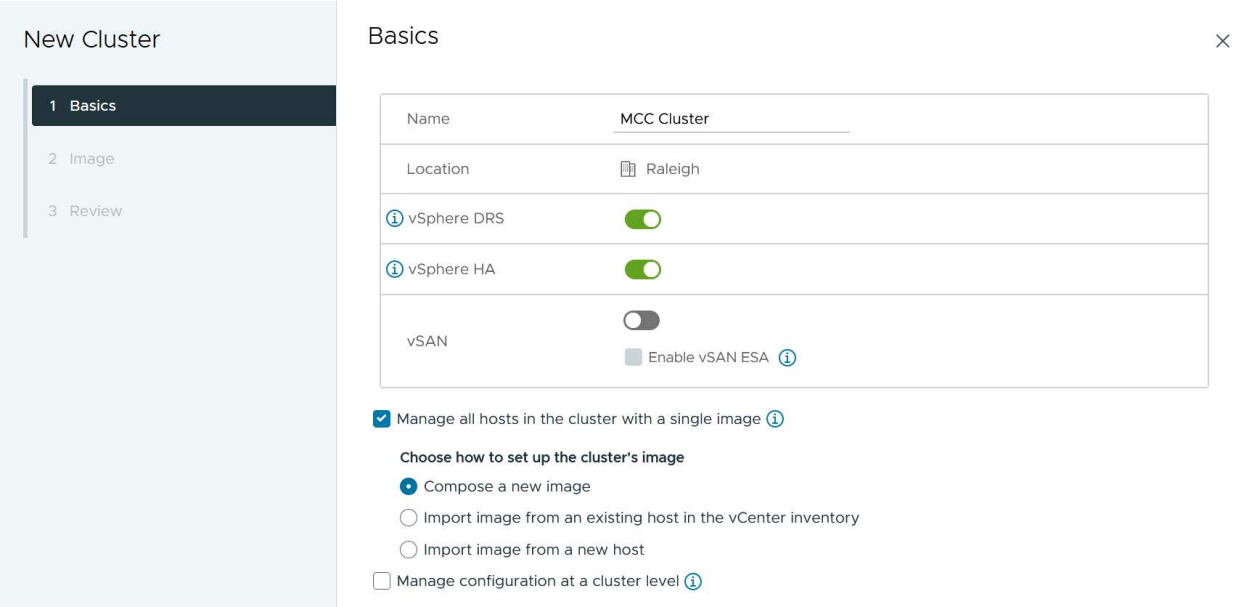
Créer un cluster haute disponibilité vSphere

La création d'un cluster vSphere HA est un processus en plusieurs étapes entièrement documenté à l'adresse "Comment créer et configurer des clusters dans vSphere client sur docs.vmware.com". En bref, vous devez d'abord créer un cluster vide, puis, à l'aide de vCenter, vous devez ajouter des hôtes et spécifier les paramètres vSphere HA et autres du cluster.

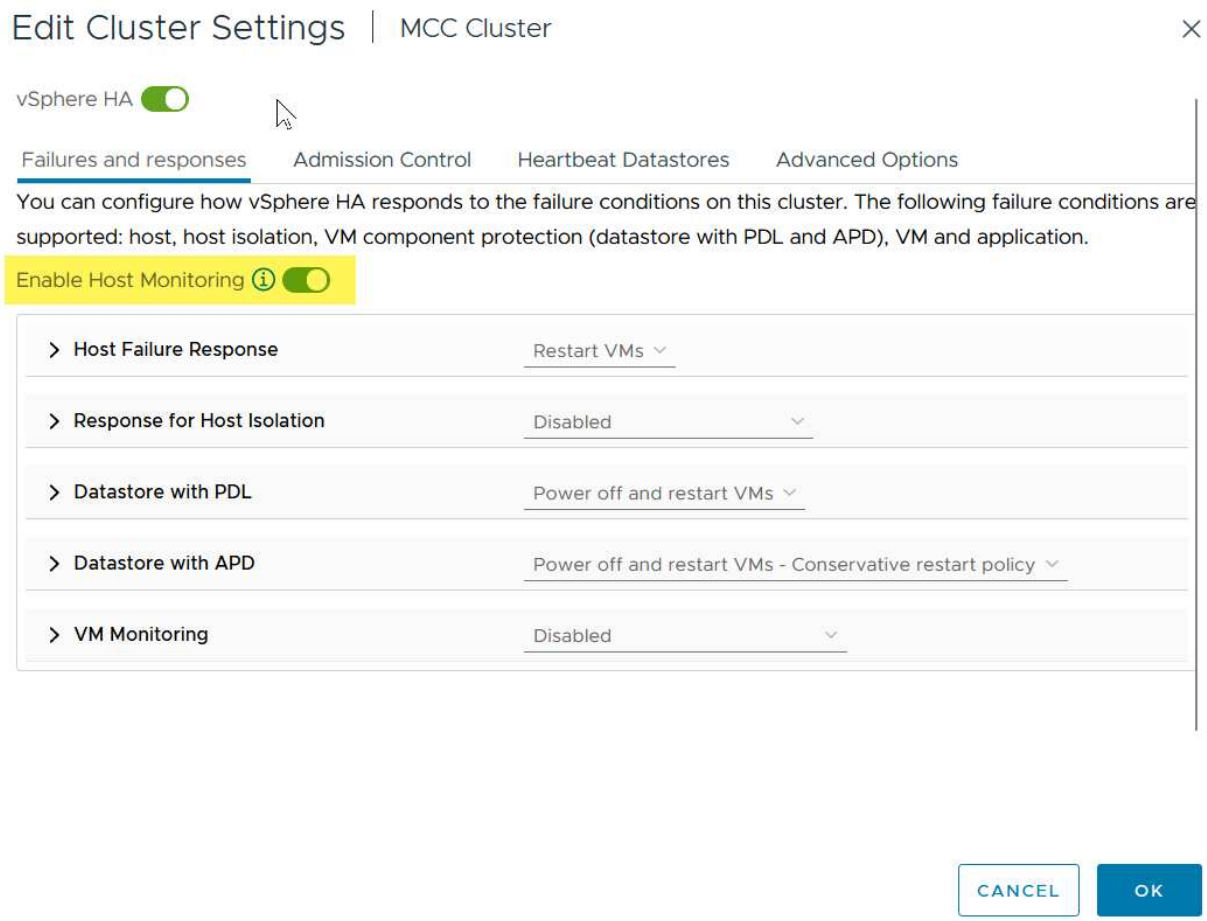
Note: rien dans ce document ne remplace "Bonnes pratiques pour VMware vSphere Metro Storage Cluster"

Pour configurer un cluster HA, effectuez les étapes suivantes :

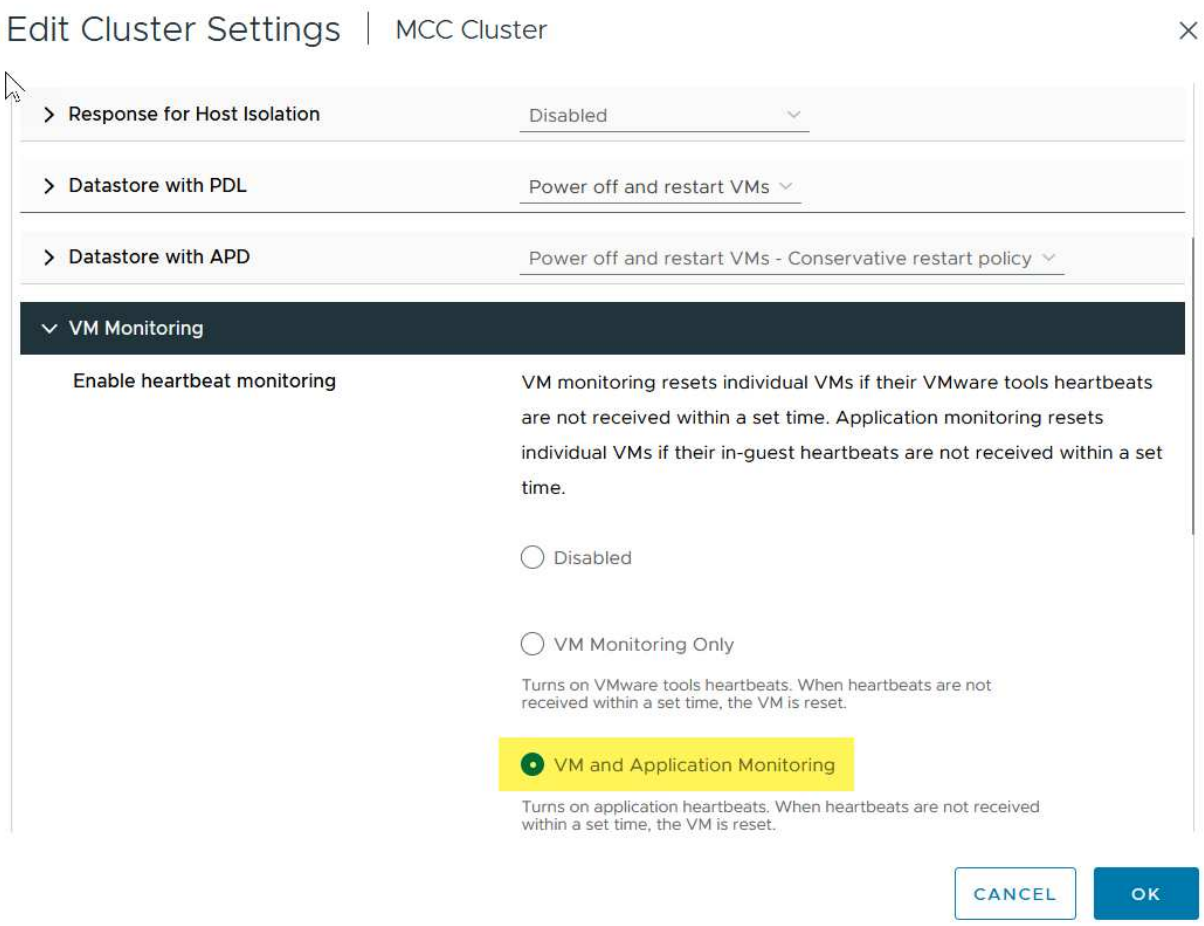
1. Connectez-vous à l'interface utilisateur vCenter.
2. Dans hôtes et clusters, accédez au data Center où vous souhaitez créer votre cluster haute disponibilité.
3. Cliquez avec le bouton droit de la souris sur l'objet de data Center et sélectionnez Nouveau cluster. Dans les notions de base, assurez-vous d'avoir activé vSphere DRS et vSphere HA. Suivez l'assistant.



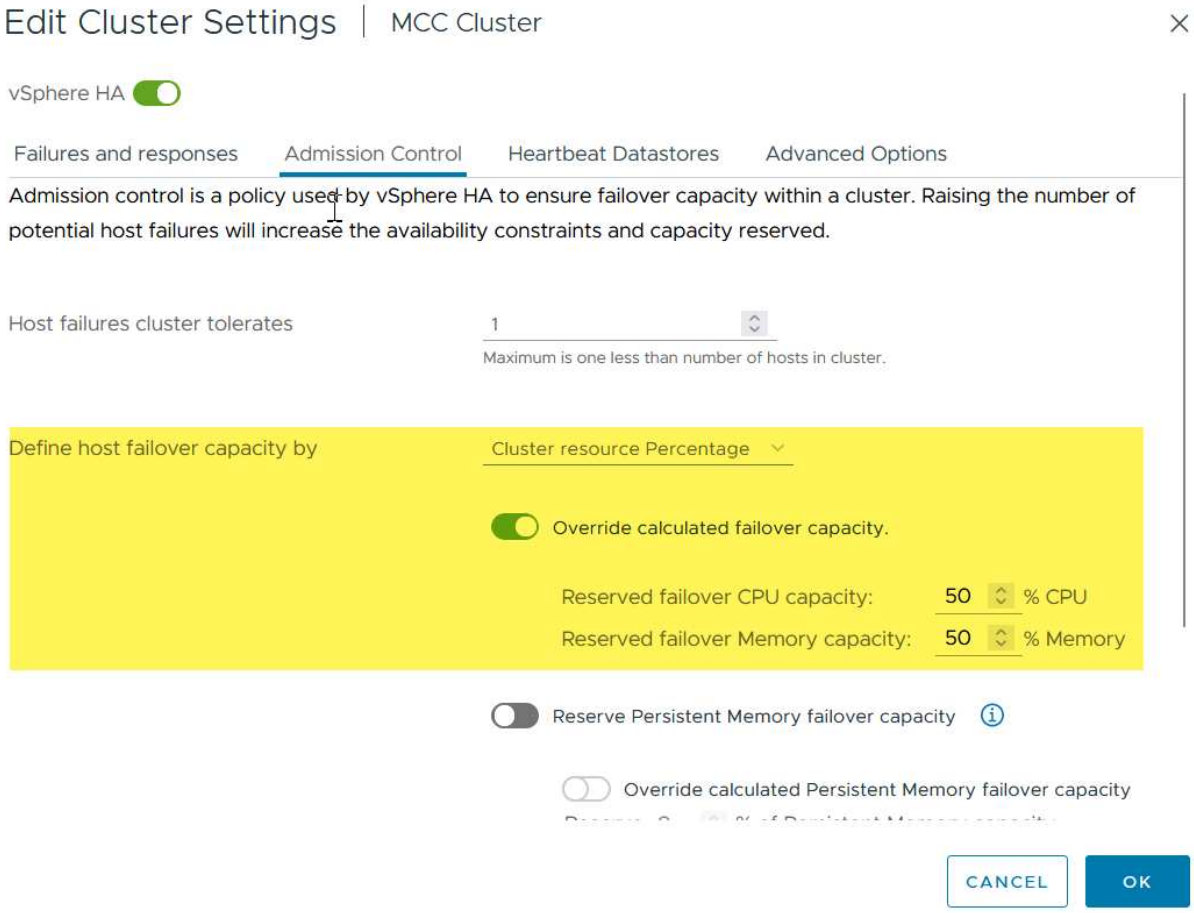
1. Sélectionnez le cluster et accédez à l'onglet configure. Sélectionnez vSphere HA et cliquez sur Edit.
2. Sous surveillance de l'hôte, sélectionnez l'option Activer la surveillance de l'hôte.



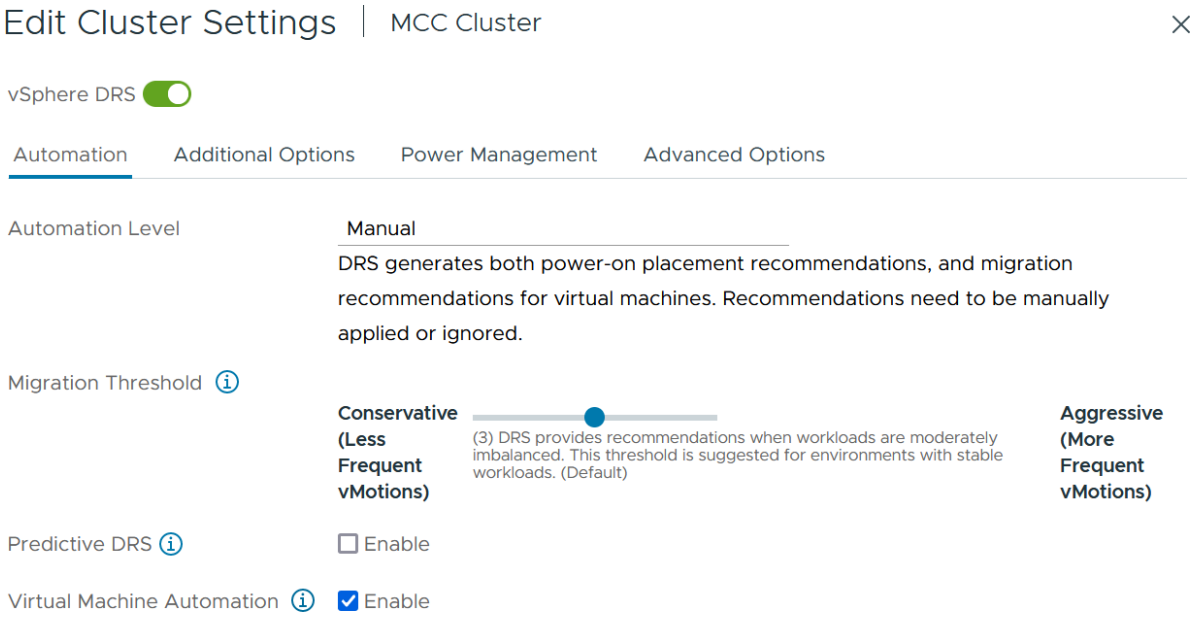
1. Toujours sous l'onglet défaillances et réponses, sous surveillance VM, sélectionnez l'option VM Monitoring Only ou VM and application Monitoring.



1. Sous contrôle d'admission, définissez l'option de contrôle d'admission HA sur réserve de ressources de cluster ; utilisez 50 % CPU/MEM.



1. Cliquez sur OK.
2. Sélectionnez DRS et cliquez sur EDIT.
3. Définissez le niveau d'automatisation sur manuel, sauf si vos applications en ont besoin.



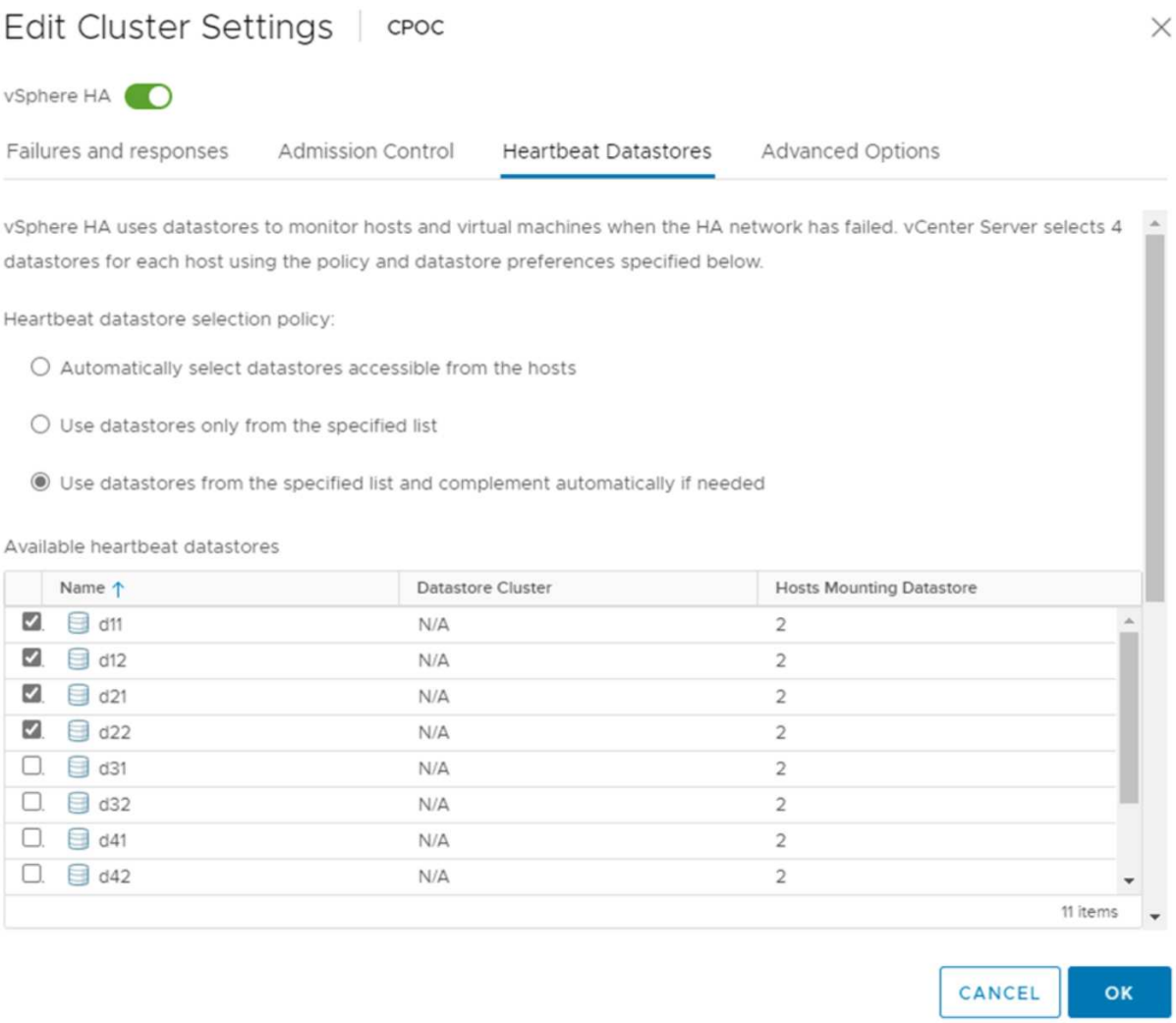
1. Activer la protection des composants VM, voir "docs.vmware.com".
2. Les paramètres vSphere HA supplémentaires suivants sont recommandés pour vMSC avec MCC :

Panne	Réponse
Défaillance d'hôte	Redémarrage des machines virtuelles
Isolation de l'hôte	Désactivé
Datastore avec perte de périphérique permanente (PDL)	Mettez les machines virtuelles hors tension et redémarrez-les
Datastore avec tous les chemins en panne (APD)	Mettez les machines virtuelles hors tension et redémarrez-les
Client qui ne bat pas	Réinitialiser les VM
Règle de redémarrage de machine virtuelle	Déterminé par l'importance de la machine virtuelle
Réponse pour l'isolation de l'hôte	Arrêtez et redémarrez les machines virtuelles
Réponse pour datastore avec PDL	Mettez les machines virtuelles hors tension et redémarrez-les
Réponse pour le datastore avec APD	Mise hors tension et redémarrage des machines virtuelles (prudent)
Délai de basculement de machine virtuelle pour APD	3 minutes
Réponse pour la restauration APD avec délai d'expiration APD	Désactivé
Sensibilité de surveillance des machines virtuelles	Présélection haute

Configurez les datastores pour Heartbeat

VSphere HA utilise les datastores pour surveiller les hôtes et les machines virtuelles en cas de panne du réseau de gestion. Vous pouvez configurer la façon dont vCenter sélectionne les datastores Heartbeat. Pour configurer des datastores pour les pulsations, procédez comme suit :

1. Dans la section pulsation du datastore, sélectionnez utiliser les datastores dans la liste spécifiée et complétez automatiquement si nécessaire.
2. Sélectionnez les datastores que vCenter doit utiliser sur les deux sites et appuyez sur OK.

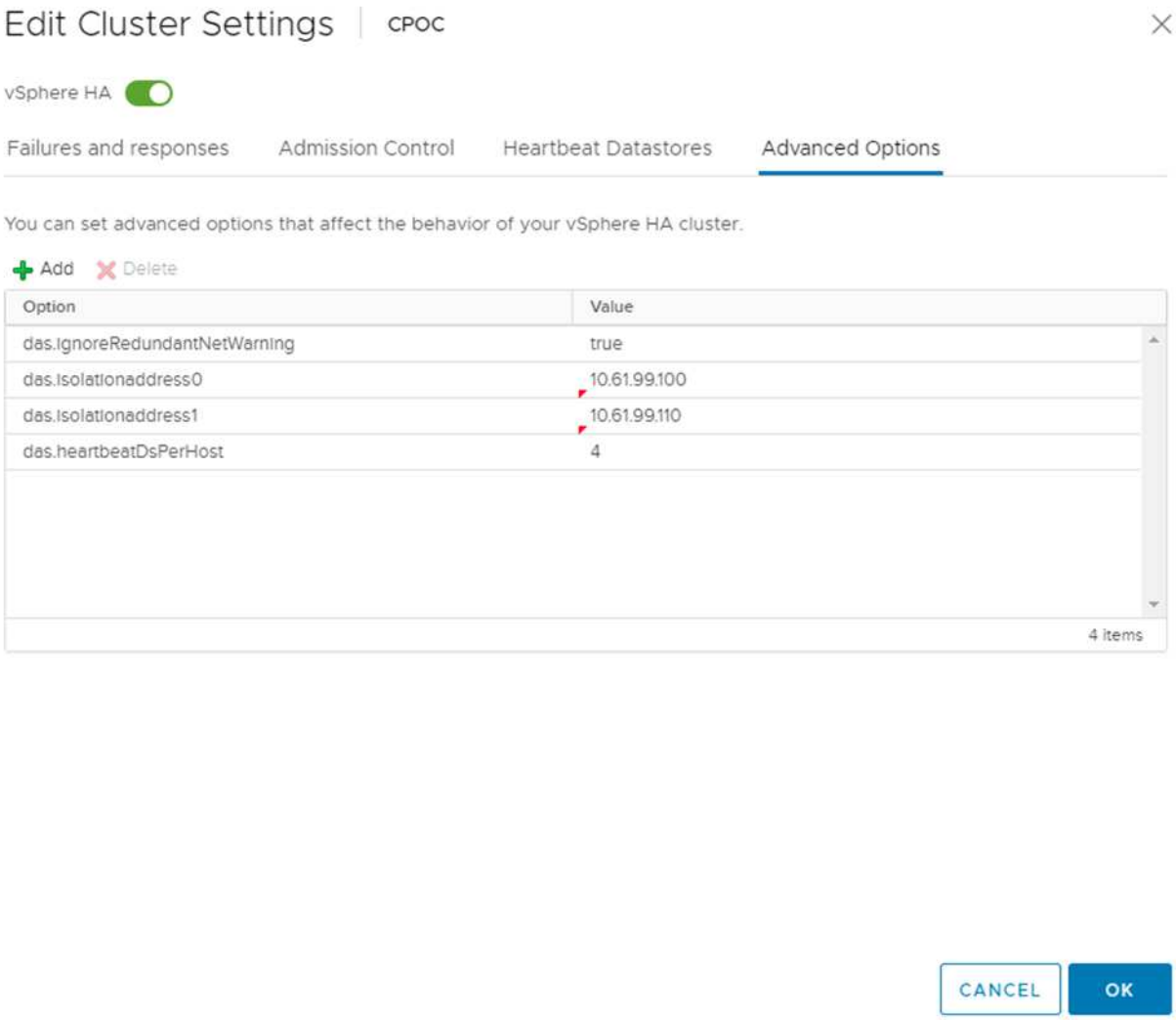


Configurer les options avancées

Détection de défaillance de l'hôte

Les événements d'isolation se produisent lorsque les hôtes d'un cluster haute disponibilité perdent la connectivité au réseau ou à d'autres hôtes du cluster. Par défaut, vSphere HA utilise la passerelle par défaut de son réseau de gestion comme adresse d'isolation par défaut. Toutefois, vous pouvez spécifier des adresses d'isolement supplémentaires pour que l'hôte puisse envoyer une requête ping afin de déterminer si une réponse d'isolement doit être déclenchée. Ajoutez deux adresses IP d'isolation pouvant être ping, une par site. N'utilisez pas l'adresse IP de la passerelle. Le paramètre avancé de vSphere HA utilisé est das.isolaaddress. Vous pouvez utiliser des adresses IP ONTAP ou Mediator à cette fin.

Reportez-vous à la section "core.vmware.com" pour plus d'informations_.__



L'ajout d'un paramètre avancé appelé das.heartbeatDsPerHost peut augmenter le nombre de datastores de pulsation. Utilisez quatre datastores de pulsation (DSS HB)—deux par site. Utilisez l'option « Sélectionner dans la liste mais compléter ». Ceci est nécessaire car si un site tombe en panne, vous avez toujours besoin de deux DSS HB. Toutefois, ceux-ci n'ont pas à être protégés avec la synchronisation active MCC ou SnapMirror.

Reportez-vous à la section "core.vmware.com" pour plus d'informations_.__

Affinité avec VMware DRS pour NetApp MetroCluster

Dans cette section, nous créons des groupes DRS pour les machines virtuelles et les hôtes pour chaque site/cluster dans l'environnement MetroCluster. Ensuite, nous configurons les règles VM/Host pour aligner l'affinité des hôtes VM avec les ressources de stockage locales. Par exemple, les machines virtuelles du site A appartiennent au groupe de machines virtuelles sitea_VM et les hôtes du site A appartiennent au groupe d'hôtes sitea_hosts. Ensuite, dans VM\Host Rules, nous faisons état que sitea_vm doit s'exécuter sur les hôtes de sitea_hosts.

Meilleure pratique

- NetApp recommande vivement la spécification **devrait s'exécuter sur les hôtes du groupe** plutôt que la spécification **doit s'exécuter sur les hôtes du groupe**. En cas de défaillance d'un hôte sur un site, les machines virtuelles Du site A doivent être redémarrées sur les hôtes du site B via vSphere HA, mais cette

dernière spécification ne permet pas à HA de redémarrer les machines virtuelles sur le site B, car il s'agit d'une règle stricte. Il s'agit d'une règle souple qui ne sera pas respectée en cas de haute disponibilité, garantissant ainsi la disponibilité plutôt que la performance.

Remarque : vous pouvez créer une alarme basée sur des événements qui est déclenchée lorsqu'une machine virtuelle viole une règle d'affinité VM-Host. Dans le client vSphere, ajoutez une nouvelle alarme pour la machine virtuelle et sélectionnez « VM viole VM-Host Affinity Rule » comme déclencheur d'événement. Pour plus d'informations sur la création et la modification d'alarmes, reportez-vous à la section "Surveillance et performances vSphere" documentation :

Créer des groupes d'hôtes DRS

Pour créer des groupes d'hôtes DRS spécifiques au site A et au site B, procédez comme suit :

1. Dans le client Web vSphere, cliquez avec le bouton droit de la souris sur le cluster dans l'inventaire et sélectionnez Paramètres.
2. Cliquez sur VM\Host Groups.
3. Cliquez sur Ajouter.
4. Saisissez le nom du groupe (par exemple, sitea_hosts).
5. Dans le menu Type, sélectionnez Groupe d'hôtes.
6. Cliquez sur Ajouter et sélectionnez les hôtes souhaités sur le site A, puis cliquez sur OK.
7. Répétez ces étapes pour ajouter un autre groupe d'hôtes pour le site B.
8. Cliquez sur OK.

Créer des groupes VM DRS

Pour créer des groupes VM DRS spécifiques au site A et au site B, procédez comme suit :

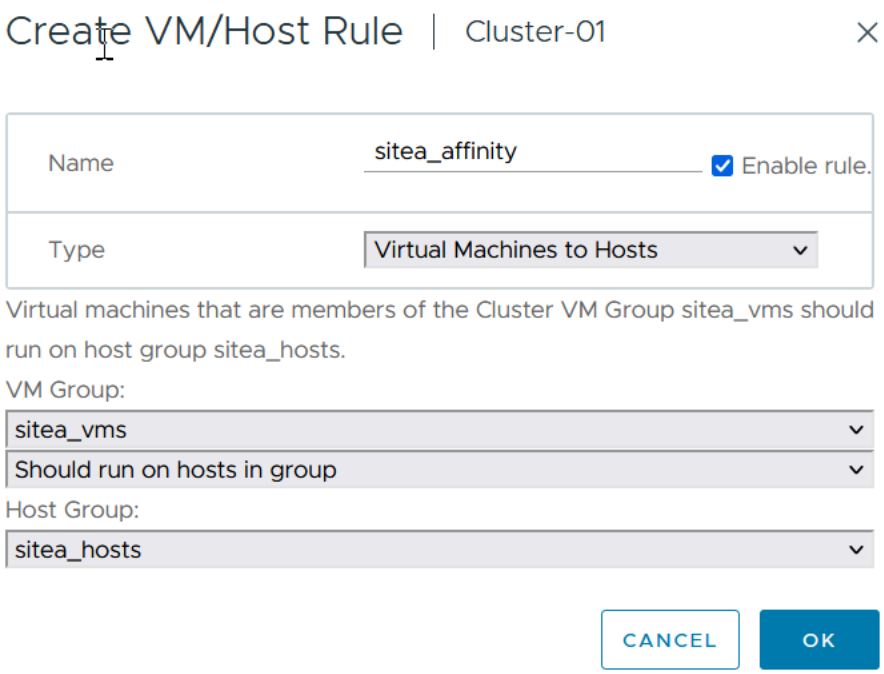
1. Dans le client Web vSphere, cliquez avec le bouton droit de la souris sur le cluster dans l'inventaire et sélectionnez Paramètres.
2. Cliquez sur VM\Host Groups.
3. Cliquez sur Ajouter.
4. Saisissez le nom du groupe (par exemple, sitea_vm).
5. Dans le menu Type, sélectionnez VM Group.
6. Cliquez sur Ajouter, sélectionnez les machines virtuelles souhaitées sur le site A, puis cliquez sur OK.
7. Répétez ces étapes pour ajouter un autre groupe d'hôtes pour le site B.
8. Cliquez sur OK.

Créer des règles d'hôte VM

Pour créer des règles d'affinité DRS spécifiques au site A et au site B, procédez comme suit :

1. Dans le client Web vSphere, cliquez avec le bouton droit de la souris sur le cluster dans l'inventaire et sélectionnez Paramètres.
2. Cliquez sur VM\Host Rules.
3. Cliquez sur Ajouter.
4. Tapez le nom de la règle (par exemple, sitea_affinité).

5. Vérifiez que l'option Activer la règle est cochée.
6. Dans le menu Type, sélectionnez ordinateurs virtuels vers hôtes.
7. Sélectionnez le groupe VM (par exemple, sitea_vm).
8. Sélectionnez le groupe Host (par exemple, sitea_hosts).
9. Répétez ces étapes pour ajouter une autre règle VM\Host pour le site B.
10. Cliquez sur OK.



VMware vSphere Storage DRS pour NetApp MetroCluster

Créer des clusters de datastores

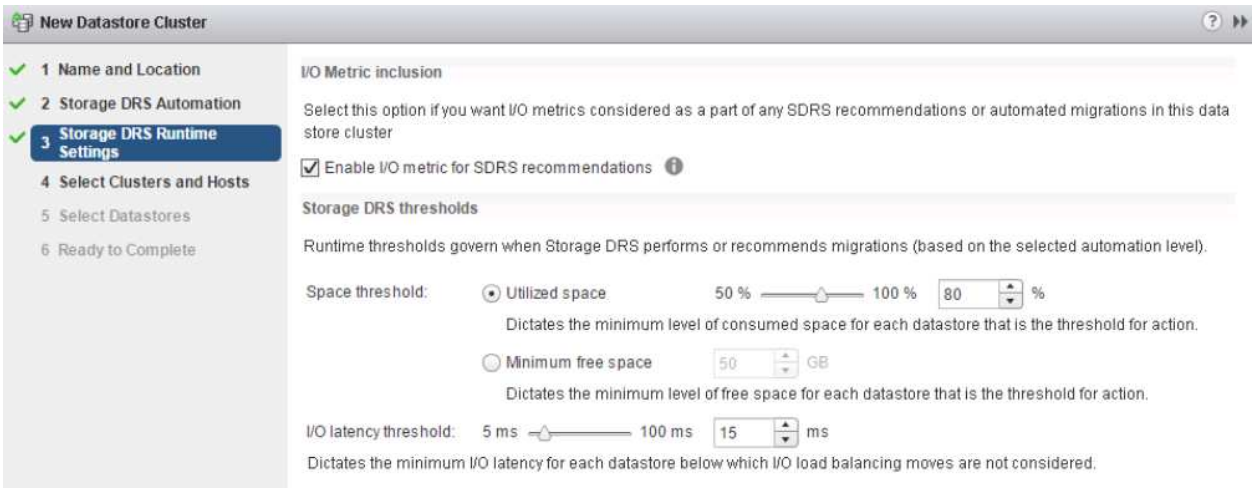
Pour configurer un cluster de datastore pour chaque site, procédez comme suit :

1. À l'aide du client web vSphere, accédez au data Center où réside le cluster HA sous Storage.
2. Cliquez avec le bouton droit de la souris sur l'objet datacenter et sélectionnez Storage > New datastore Cluster.
3. Sélectionnez l'option ACTIVER Storage DRS et cliquez sur Suivant.
4. Définissez toutes les options sur pas d'automatisation (mode manuel) et cliquez sur Suivant.

Meilleure pratique

- NetApp recommande de configurer Storage DRS en mode manuel, afin que l'administrateur puisse décider et contrôler les opérations de migration.

1. Vérifiez que la case Activer les mesures d'E/S pour les recommandations SDRS est cochée ; les paramètres de mesure peuvent être laissés avec les valeurs par défaut.



1. Sélectionnez le cluster HA et cliquez sur Next.

1. Sélectionnez les datastores appartenant au site A et cliquez sur Suivant.

1. Vérifiez les options et cliquez sur Terminer.
2. Répétez ces étapes pour créer le cluster de datastore du site B et vérifier que seuls les datastores du site B sont sélectionnés.

Disponibilité du serveur vCenter

Vos appliances vCenter Server (VCSA) doivent être protégées avec vCenter HA. VCenter HA vous permet de

déployer deux VCSA dans une paire haute disponibilité actif-passif. Un dans chaque domaine de défaillance. Pour en savoir plus sur vCenter HA, rendez-vous sur "docs.vmware.com".

Résilience pour les événements planifiés et non planifiés

NetApp MetroCluster et la synchronisation active SnapMirror sont des outils puissants qui améliorent la haute disponibilité et la continuité de l'activité du matériel NetApp et du logiciel ONTAP®.

Ces outils assurent une protection à l'échelle du site pour l'ensemble de l'environnement de stockage, garantissant ainsi la disponibilité permanente de vos données. Que vous utilisiez des serveurs autonomes, des clusters à haute disponibilité, des conteneurs Docker ou des serveurs virtualisés, la technologie NetApp assure la disponibilité du stockage de manière transparente en cas de panne totale due à une coupure d'alimentation, à des problèmes de climatisation, de connectivité réseau, à l'arrêt des baies de stockage ou à une erreur de fonctionnement.

La synchronisation active MetroCluster et SnapMirror propose trois méthodes de base pour la continuité des données en cas d'événements planifiés ou non :

- Des composants redondants pour une protection contre les défaillances d'un seul composant
- Basculement de haute disponibilité locale en cas d'événements affectant un contrôleur unique
- Protection complète du site – reprise rapide du service en déplaçant le stockage et l'accès client du cluster source vers le cluster de destination

Cela signifie que les opérations se poursuivent en toute transparence en cas de défaillance d'un seul composant et reviennent automatiquement au fonctionnement redondant lorsque le composant défectueux est remplacé.

Tous les clusters ONTAP, à l'exception des clusters à un seul nœud (en général, les versions Software-defined, telles que ONTAP Select, par exemple), disposent de fonctionnalités haute disponibilité intégrées appelées Takeover et giveback. Chaque contrôleur du cluster est couplé à un autre contrôleur, formant une paire haute disponibilité. Ces paires garantissent que chaque nœud est connecté localement au stockage.

Le basculement est un processus automatisé qui consiste à prendre le contrôle du stockage d'un nœud pour assurer les services de données. Le rétablissement est le processus inverse qui restaure le fonctionnement normal. Le basculement peut être planifié, par exemple lors de la maintenance matérielle ou des mises à niveau ONTAP, ou non planifié, suite à une panne matérielle ou de panique sur un nœud.

Lors d'un basculement, les interfaces logiques NAS dans les configurations MetroCluster basculent automatiquement. Toutefois, les LIF SAN (Storage Area Network) ne basculent pas ; elles continuent d'utiliser le chemin direct vers les LUN (Logical Unit Numbers).

Pour plus d'informations sur le basculement et le rétablissement HA, reportez-vous au "Présentation de la gestion des paires HAUTE DISPONIBILITÉ". Notez que cette fonctionnalité n'est pas spécifique à la synchronisation active MetroCluster ou SnapMirror.

Le basculement de site avec MetroCluster a lieu lorsqu'un site est hors ligne ou lors d'une activité planifiée pour la maintenance à l'échelle du site. Le site restant assume la propriété des ressources de stockage (disques et agrégats) du cluster hors ligne, et les SVM sur le site en panne sont mis en ligne et redémarrés sur le site en cas de sinistre, tout en préservant leur identité complète pour l'accès des clients et des hôtes.

Avec la synchronisation active SnapMirror, dans la mesure où les deux copies sont activement utilisées simultanément, vos hôtes existants continueront de fonctionner. Le médiateur NetApp est nécessaire pour

garantir que le basculement de site se produit correctement.

Scénarios de panne pour vMSC avec MCC

Les sections suivantes décrivent les résultats attendus de différents scénarios de défaillance avec les systèmes vMSC et NetApp MetroCluster.

Défaillance d'un seul chemin de stockage

Dans ce scénario, si des composants tels que le port HBA, le port réseau, le port du commutateur de données frontal ou un câble FC ou Ethernet échouent, ce chemin particulier vers le périphérique de stockage est marqué comme mort par l'hôte ESXi. Si plusieurs chemins sont configurés pour le périphérique de stockage en fournissant la résilience au niveau du port HBA/réseau/commutateur, ESXi effectue idéalement un basculement de chemin. Pendant cette période, les ordinateurs virtuels restent en fonctionnement sans être affectés, car la disponibilité du stockage est assurée par plusieurs chemins vers le périphérique de stockage.

Note: il n'y a pas de changement dans le comportement de MetroCluster dans ce scénario, et tous les datastores continuent d'être intacts de leurs sites respectifs.

Meilleure pratique

Dans les environnements dans lesquels les volumes NFS/iSCSI sont utilisés, NetApp recommande de configurer au moins deux liaisons montantes réseau pour le port vmkernel NFS dans le vSwitch standard et la même pour le groupe de ports où l'interface vmkernel NFS est mappée pour le vSwitch distribué. Le regroupement de cartes réseau peut être configuré en mode actif-actif ou actif-veille.

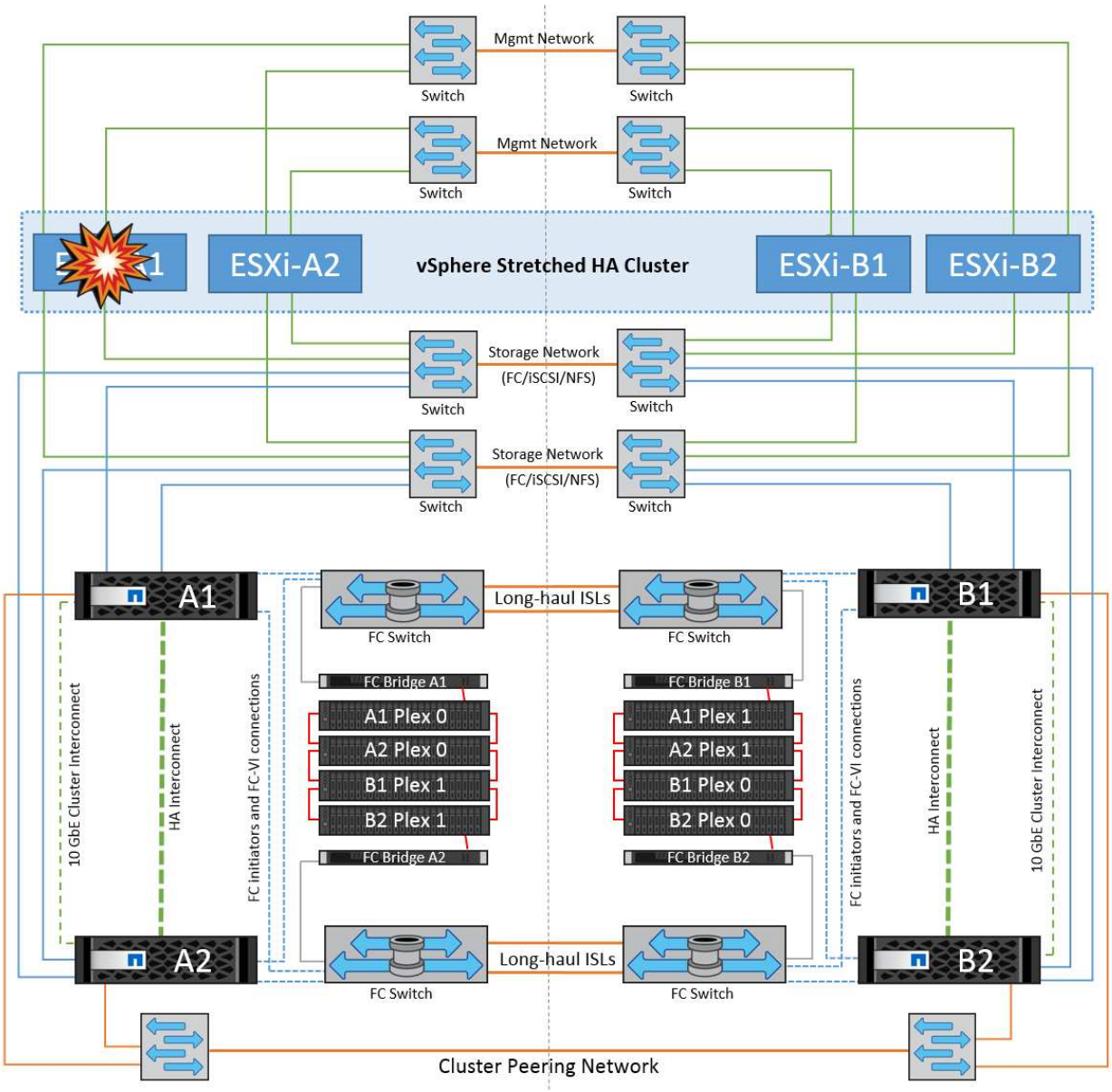
En outre, pour les LUN iSCSI, les chemins d'accès multiples doivent être configurés en liant les interfaces vmkernel aux adaptateurs réseau iSCSI. Pour plus d'informations, reportez-vous à la documentation sur le stockage vSphere.

Meilleure pratique

Dans les environnements dans lesquels des LUN Fibre Channel sont utilisées, NetApp recommande d'avoir au moins deux HBA, ce qui garantit la résilience au niveau des HBA/ports. NetApp recommande également la segmentation entre un initiateur unique et une seule cible comme meilleure pratique pour la configuration de la segmentation.

Virtual Storage Console (VSC) doit être utilisé pour définir des règles de chemins d'accès multiples, car il définit des règles pour tous les périphériques de stockage NetApp, nouveaux ou existants.

Défaillance d'un hôte ESXi unique



Dans ce scénario, en cas de défaillance de l'hôte ESXi, le nœud maître du cluster VMware HA détecte la panne de l'hôte, car il ne reçoit plus de pulsations réseau. Pour déterminer si l'hôte est réellement en panne ou uniquement une partition réseau, le nœud maître surveille les pulsations du datastore et, s'il est absent, il effectue une vérification finale en envoyant une requête ping aux adresses IP de gestion de l'hôte en panne. Si toutes ces vérifications sont négatives, le nœud maître déclare cet hôte comme étant en panne et toutes les machines virtuelles qui s'exécutaient sur cet hôte en panne sont redémarrées sur l'hôte survivant du cluster.

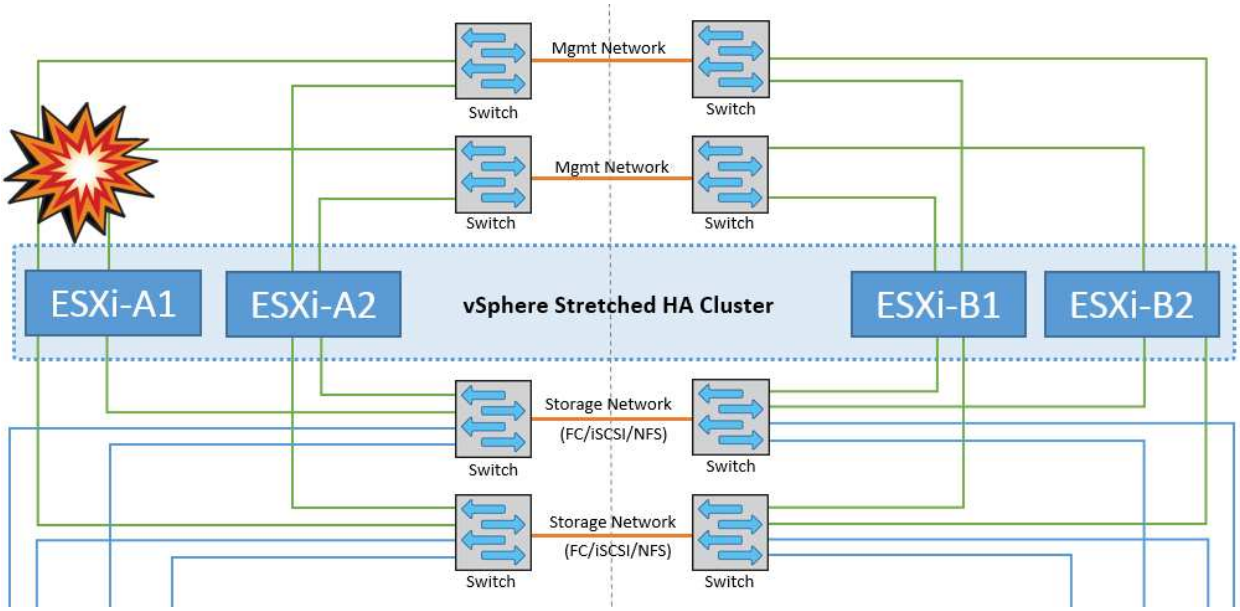
Si les règles d'affinité des machines virtuelles DRS et des hôtes ont été configurées (les machines virtuelles du groupe de machines virtuelles sitea_vm doivent exécuter des hôtes dans le groupe d'hôtes sitea_hosts), le maître haute disponibilité vérifie d'abord les ressources disponibles sur le site A. Si aucun hôte n'est disponible sur le site A, le maître tente de redémarrer les machines virtuelles sur les hôtes du site B.

Il est possible que les machines virtuelles soient démarrées sur les hôtes ESXi de l'autre site s'il existe une contrainte de ressource sur le site local. Cependant, les règles d'affinité VM et hôte DRS définies seront correctes si des règles sont enfreintes en migrant les machines virtuelles vers des hôtes ESXi survivants sur le site local. Dans les cas où DRS est défini sur manuel, NetApp recommande d'invoquer DRS et d'appliquer les recommandations pour corriger le positionnement de la machine virtuelle.

Dans ce scénario, le comportement de MetroCluster n'a pas changé et tous les datastores sont toujours

intacts sur leurs sites respectifs.

Isolation de l’hôte ESXi

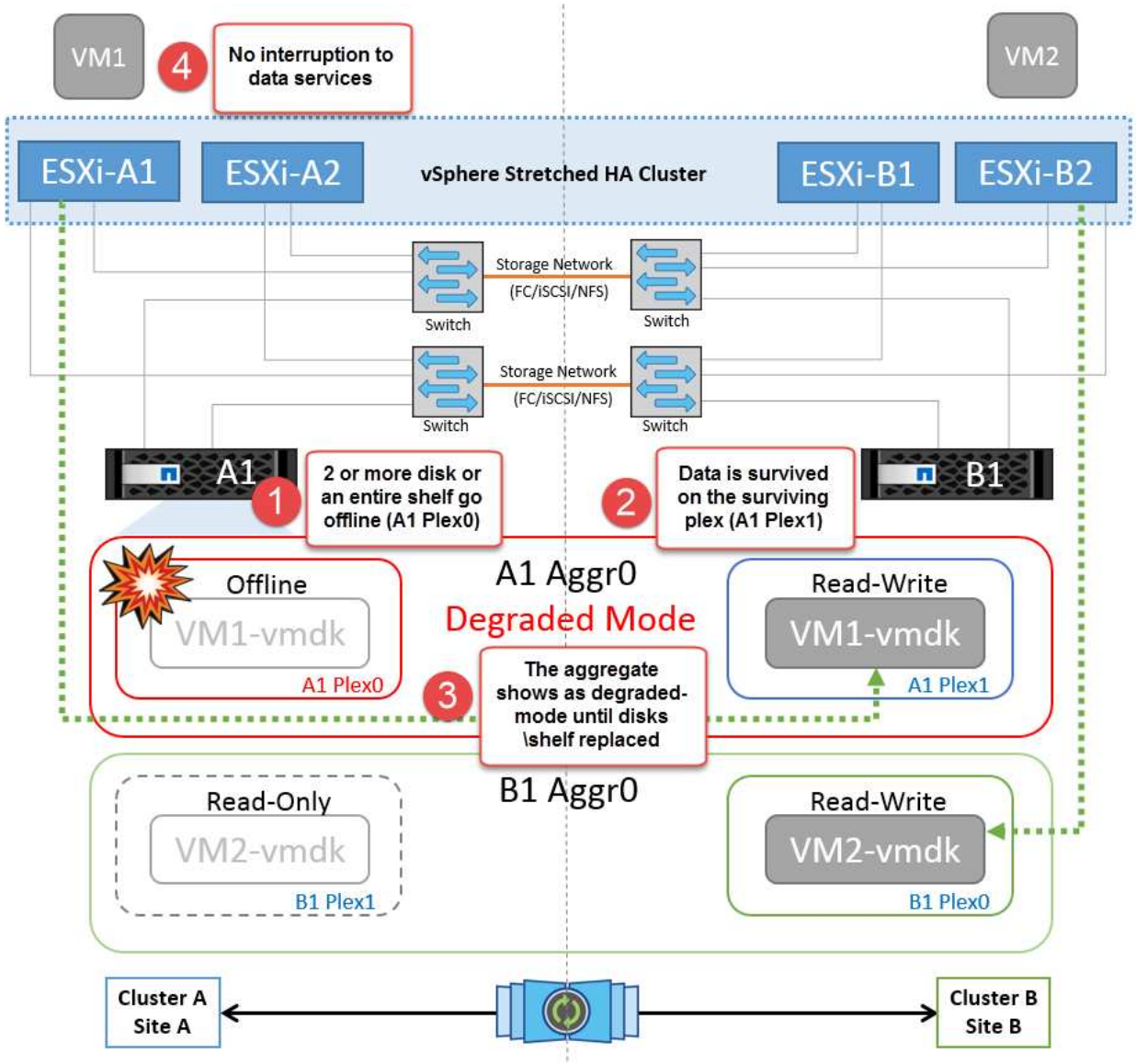


Dans ce scénario, si le réseau de gestion de l’hôte ESXi est en panne, le nœud principal du cluster HA ne recevra aucun battement de cœur. Cet hôte est donc isolé dans le réseau. Pour déterminer s’il a échoué ou s’il est isolé uniquement, le nœud maître commence à surveiller le battement de cœur du datastore. S’il est présent, l’hôte est déclaré isolé par le nœud maître. Selon la réponse d’isolement configurée, l’hôte peut choisir de mettre hors tension, d’arrêter les machines virtuelles ou même de laisser les machines virtuelles sous tension. L’intervalle par défaut pour la réponse d’isolement est de 30 secondes.

Dans ce scénario, le comportement de MetroCluster n’a pas changé et tous les datastores sont toujours intacts sur leurs sites respectifs.

Panne de tiroir disque

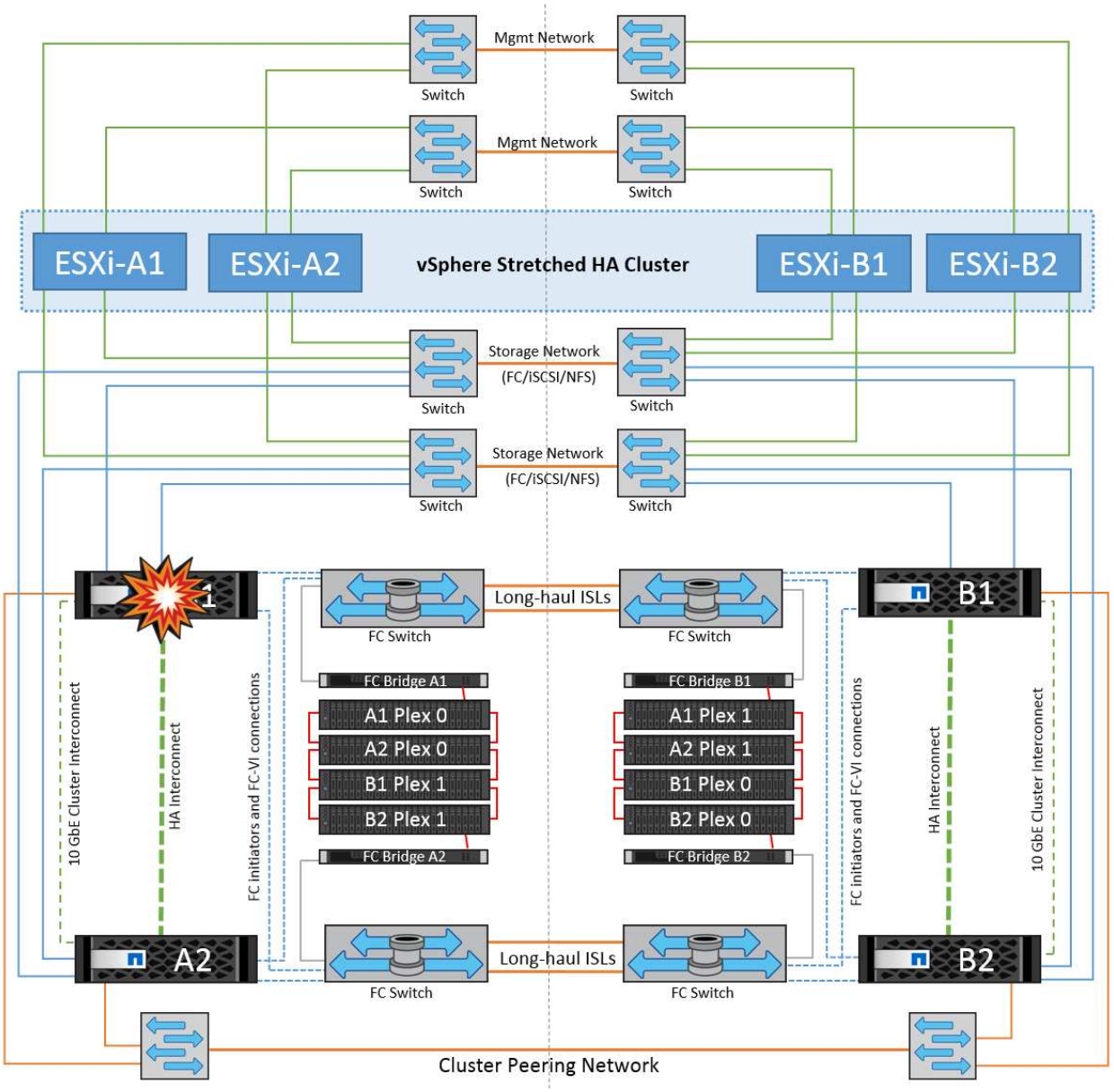
Dans ce scénario, il y a une panne de plus de deux disques ou d’un tiroir entier. Les données sont servies depuis le plex opérationnel sans interruption des services de données. La défaillance de disque peut affecter un plex local ou distant. Les agrégats s’affichent en mode dégradé, car un seul plex est actif. Une fois les disques défaillants remplacés, les agrégats affectés resynchroniseront automatiquement pour reconstruire les données. Après la resynchronisation, les agrégats reviennent automatiquement en mode miroir normal. Si plus de deux disques au sein d’un même groupe RAID sont défaillants, le plex doit être reconstruit à partir de zéro.



Remarque : au cours de cette période, il n'y a pas d'impact sur les opérations d'E/S de la machine virtuelle, mais les performances sont dégradées car les données sont accessibles depuis le tiroir disque distant via les liaisons ISL.

Panne d'un seul contrôleur de stockage

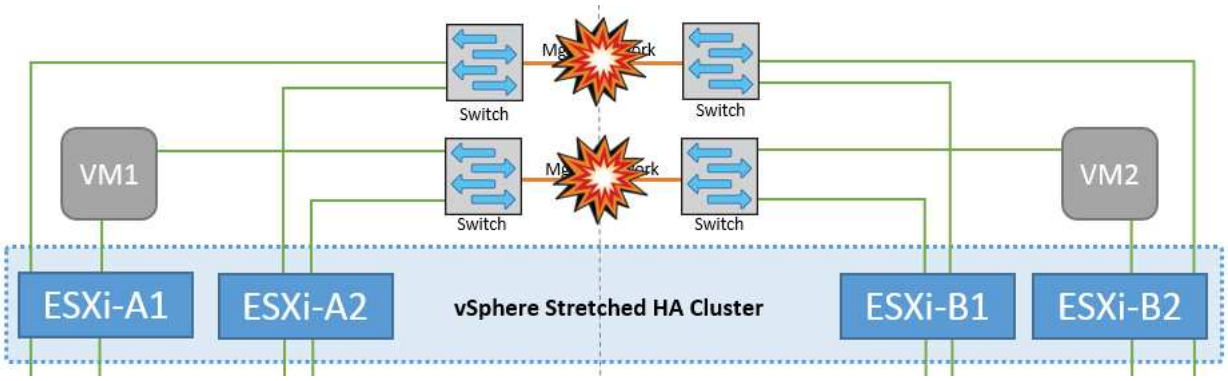
Dans ce scénario, l'un des deux contrôleurs de stockage tombe en panne sur un site. Comme il existe une paire haute disponibilité sur chaque site, la panne d'un nœud entraîne le basculement vers l'autre nœud de manière transparente et automatique. Par exemple, si le nœud A1 tombe en panne, son stockage et ses charges de travail sont automatiquement transférés vers le nœud A2. Les machines virtuelles ne seront pas affectées, car tous les plexes restent disponibles. Les nœuds du second site (B1 et B2) ne sont pas affectés. En outre, vSphere HA ne prendra aucune action, car le nœud maître du cluster recevra toujours les battements de cœur du réseau.



Si le basculement fait partie d'un incident en cours (le nœud A1 bascule vers A2) et qu'il y a une panne ultérieure de A2, ou la panne complète du site A, le basculement après un incident peut se produire sur le site B.

Défaillances de liaison entre commutateurs

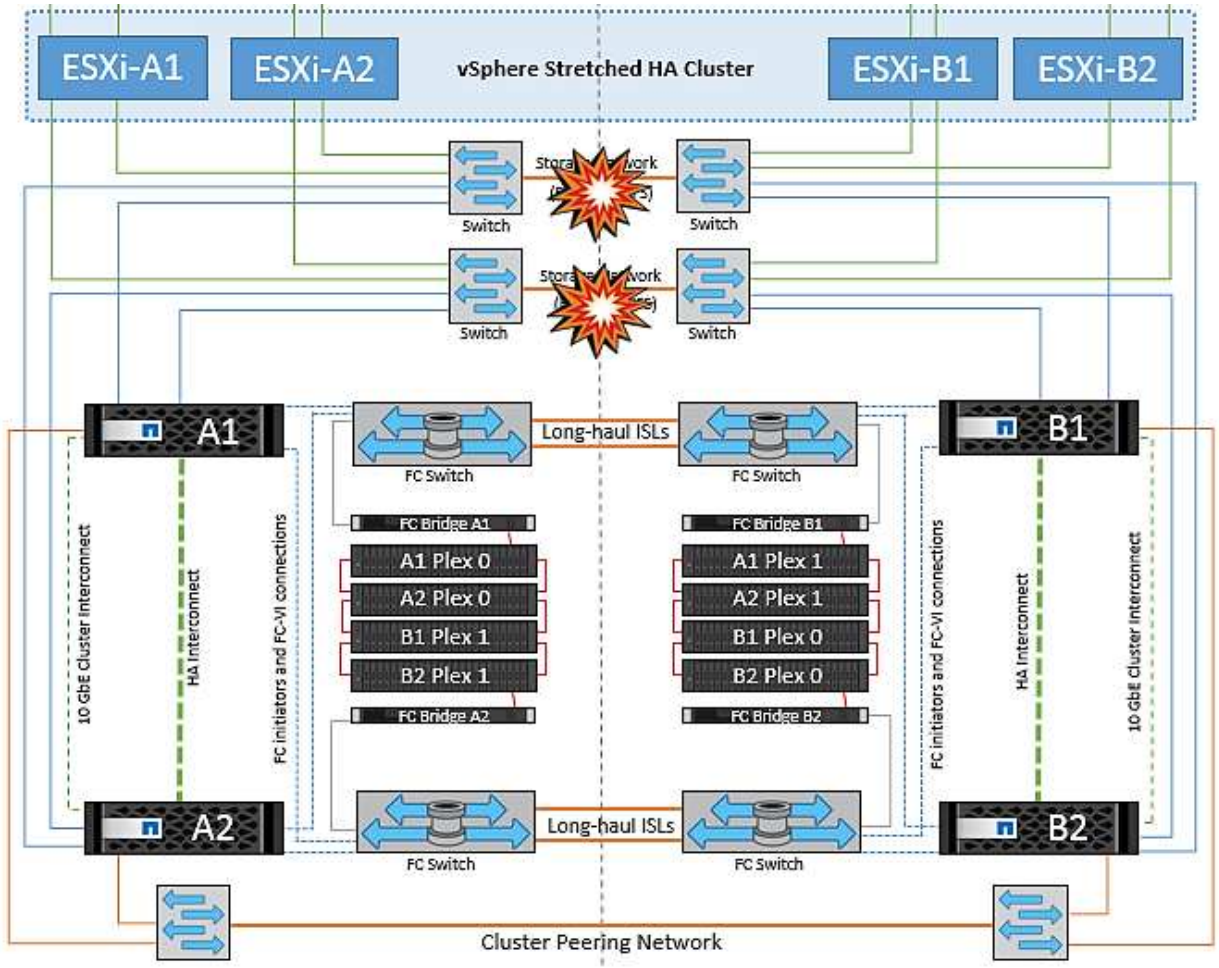
Défaillance de la liaison inter-commutateur sur le réseau de gestion



Dans ce scénario, si les liaisons ISL du réseau de gestion de l'hôte frontal tombent en panne, les hôtes ESXi du site A ne pourront pas communiquer avec les hôtes ESXi du site B. Cela entraîne une partition réseau, car les hôtes ESXi d'un site particulier ne peuvent pas envoyer les battements de cœur du réseau au nœud maître du cluster HA. Ainsi, il y aura deux segments de réseau en raison de la partition et il y aura un nœud maître dans chaque segment qui protégera les machines virtuelles des défaillances de l'hôte au sein du site particulier.

Remarque : pendant cette période, les machines virtuelles restent en cours d'exécution et il n'y a pas de changement dans le comportement de MetroCluster dans ce scénario. Tous les datastores sont toujours intacts sur leurs sites respectifs.

Défaillance de la liaison intercommutateur sur le réseau de stockage

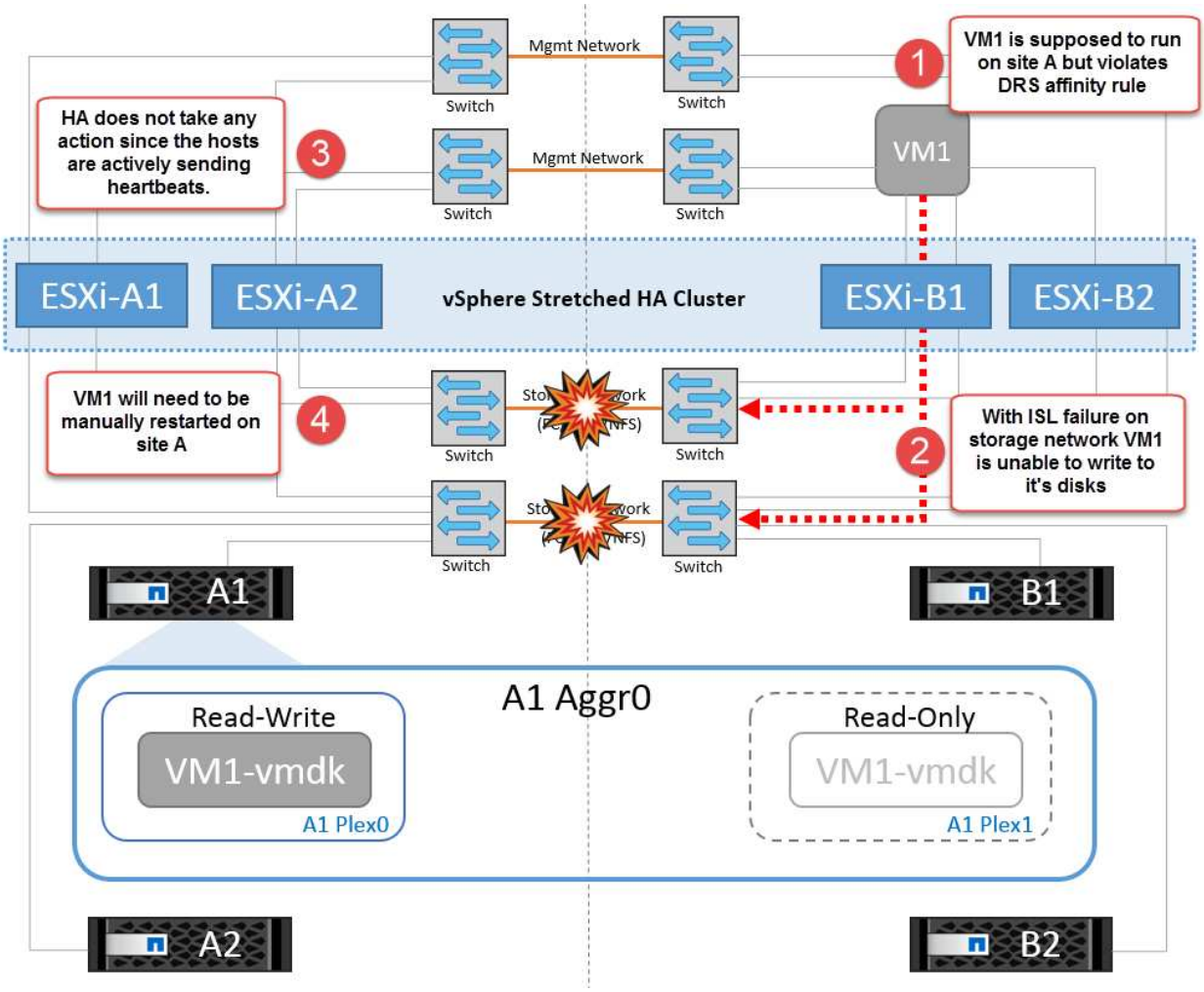


Dans ce scénario, si les liaisons ISL du réseau de stockage back-end tombent en panne, les hôtes du site A

perdront l'accès aux volumes de stockage ou aux LUN du cluster B sur le site B et vice versa. Les règles VMware DRS sont définies de manière à ce que l'affinité entre l'hôte et le site de stockage facilite l'exécution des machines virtuelles sans impact sur le site.

Pendant cette période, les machines virtuelles restent en cours d'exécution sur leurs sites respectifs et le comportement de MetroCluster n'a pas changé dans ce scénario. Tous les datastores sont toujours intacts sur leurs sites respectifs.

Si, pour une raison quelconque, la règle d'affinité a été enfreinte (par exemple, VM1, qui était censé s'exécuter à partir du site A où ses disques résident sur les nœuds du cluster A local, s'exécute sur un hôte du site B), le disque de la machine virtuelle est accessible à distance via des liens ISL. En raison d'une défaillance de la liaison ISL, VM1 exécuté sur le site B ne pouvait pas écrire sur ses disques, car les chemins vers le volume de stockage sont en panne et cette machine virtuelle est en panne. Dans ce cas, VMware HA ne prend aucune action, car les hôtes envoient activement des battements du cœur. Ces machines virtuelles doivent être manuellement désactivées et activées sur leurs sites respectifs. La figure suivante illustre une machine virtuelle violant une règle d'affinité DRS.

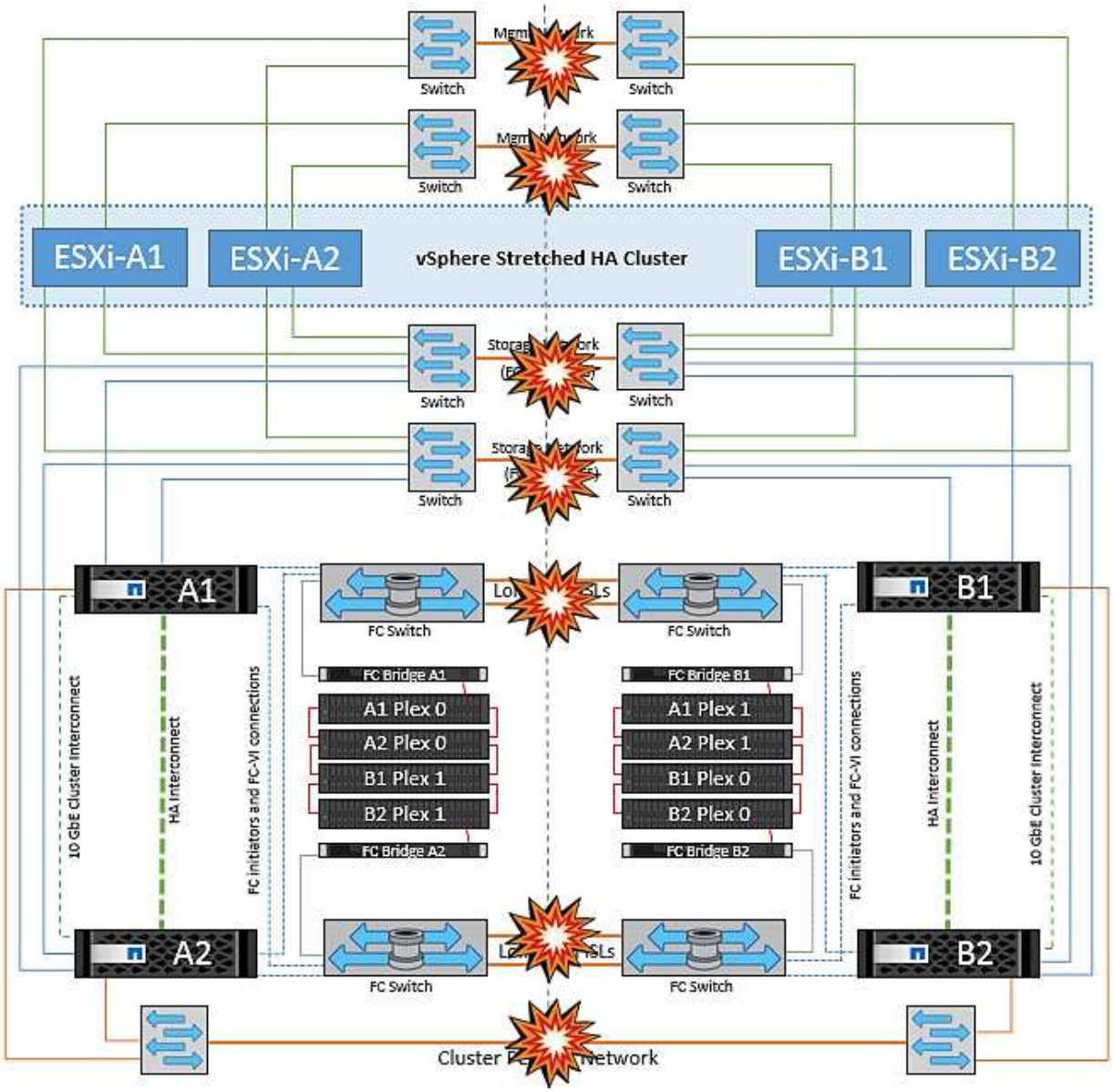


Défaillance de tous les commutateurs ou partition complète du centre de données

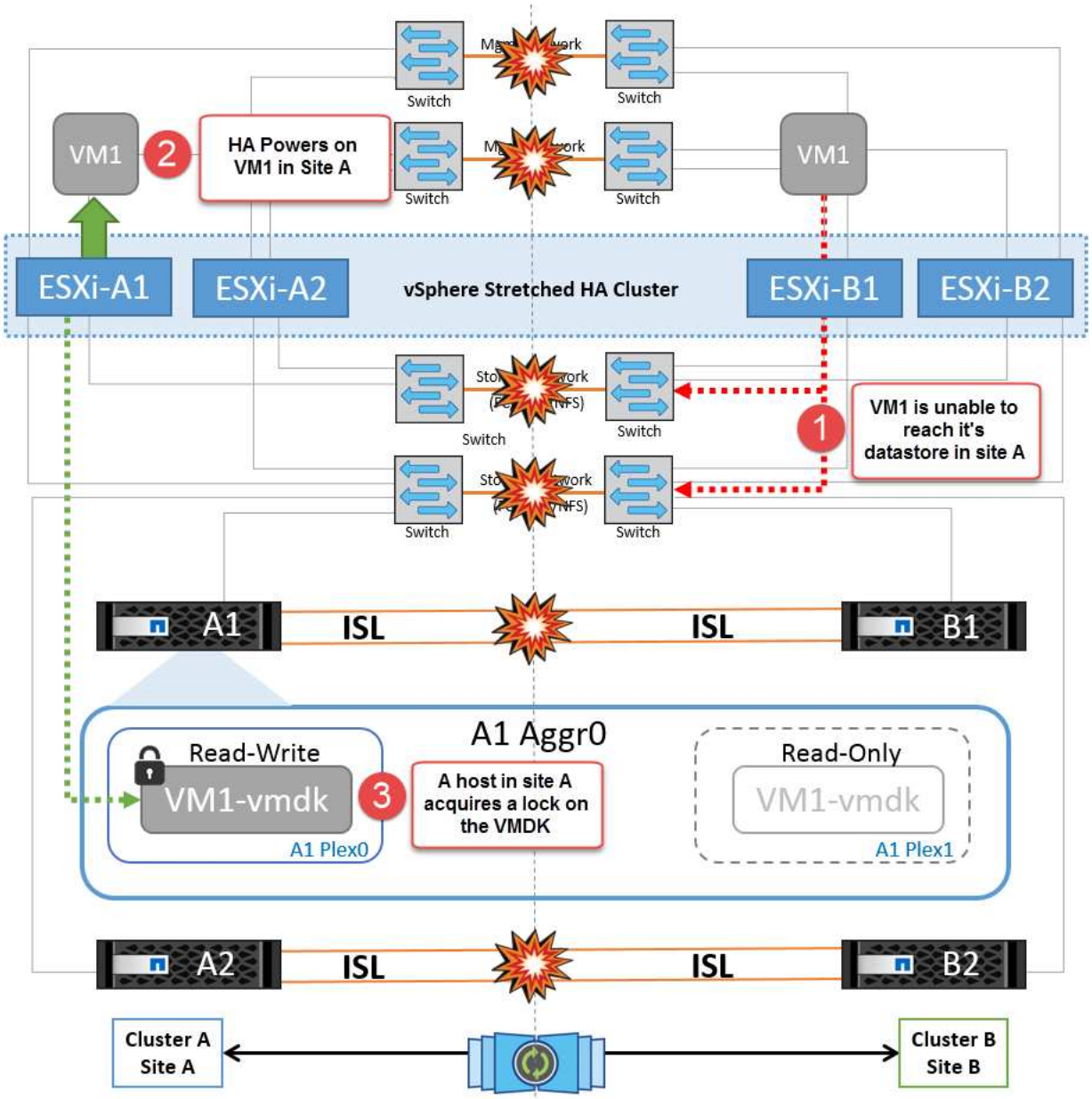
Dans ce scénario, toutes les liaisons ISL entre les sites sont en panne et les deux sites sont isolés les uns des autres. Comme nous l'avons vu dans les scénarios précédents, tels que la défaillance des liens ISL au niveau du réseau de gestion et du réseau de stockage, les machines virtuelles ne sont pas affectées par la défaillance complète des liens ISL.

Une fois les hôtes ESXi partitionnés entre les sites, l'agent vSphere HA vérifie la présence de battements de

cœur du datastore et, sur chaque site, les hôtes ESXi locaux pourront mettre à jour les battements de cœur du datastore vers leur volume/LUN de lecture/écriture respectif. Les hôtes du site A partent du principe que les autres hôtes ESXi du site B ont échoué car il n'y a pas de pulsations réseau/datastore. VSphere HA sur le site A tentera de redémarrer les machines virtuelles du site B, ce qui finira par échouer car les datastores du site B ne seront pas accessibles en raison d'une panne de lien ISL du stockage. Une situation similaire est répétée sur le site B.



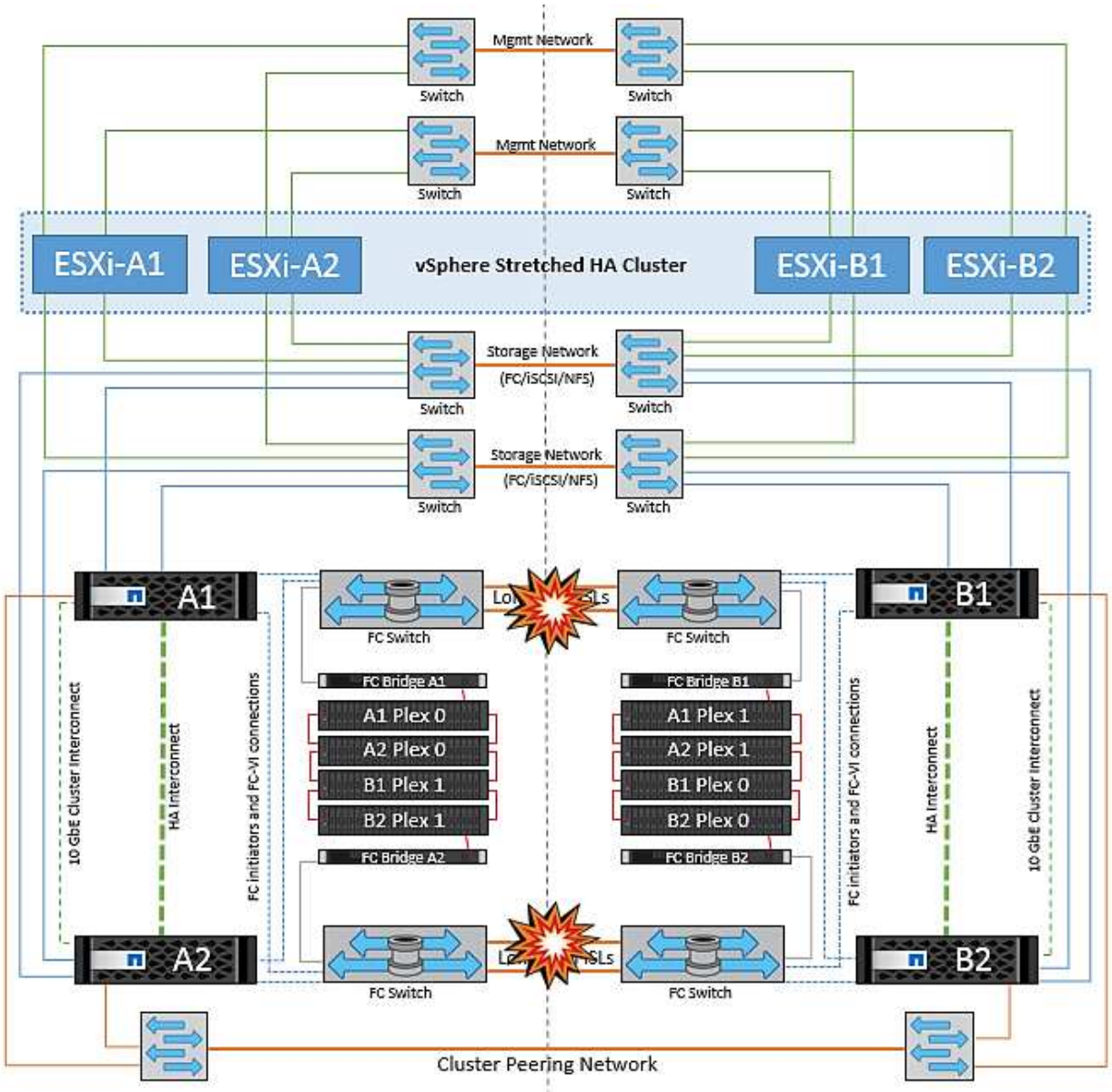
NetApp recommande de déterminer si une machine virtuelle a enfreint les règles DRS. Toutes les machines virtuelles exécutées à partir d'un site distant sont en panne, car elles ne pourront pas accéder au datastore. VSphere HA redémarrera cette machine virtuelle sur le site local. Une fois les liens ISL de nouveau en ligne, la machine virtuelle qui s'exécutait sur le site distant est arrêtée, car il ne peut pas y avoir deux instances de machines virtuelles fonctionnant avec les mêmes adresses MAC.



Défaillance de la liaison inter-commutateur sur les deux fabriques dans NetApp MetroCluster

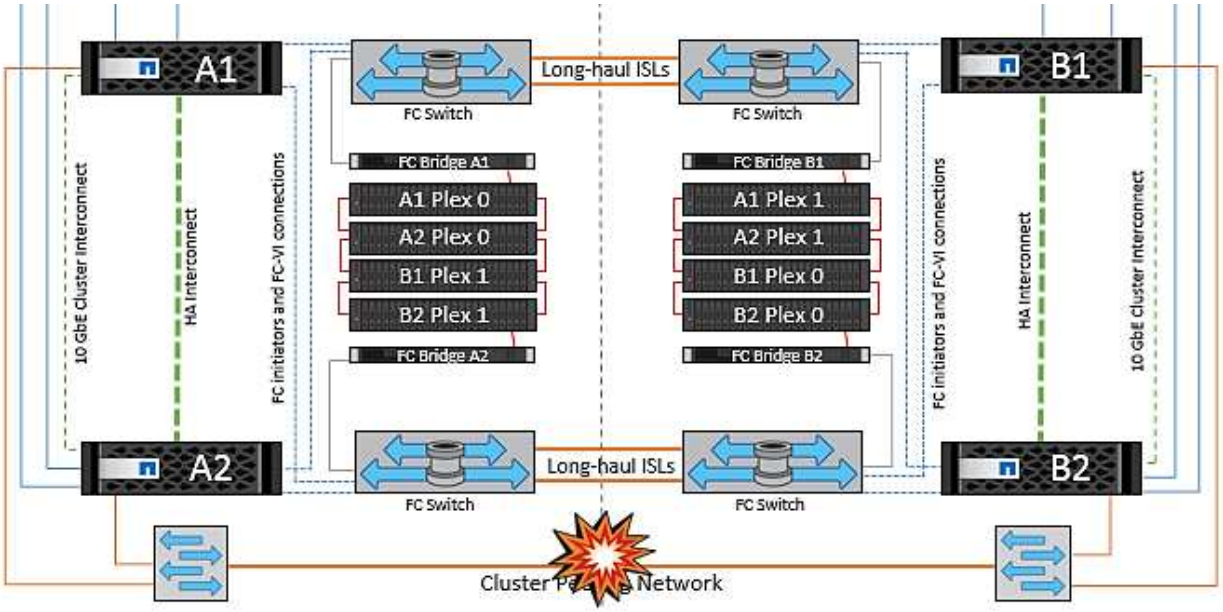
Dans le cas d'une défaillance d'un ou de plusieurs liens ISL, le trafic continue à travers les liens restants. Si toutes les liaisons ISL des deux structures échouent, de sorte qu'il n'y ait pas de liaison entre les sites pour le stockage et la réplication NVRAM, chaque contrôleur continue de transmettre ses données locales. Lors de la restauration d'un ISL au moins, la resynchronisation de tous les plexes se fera automatiquement.

Toute écriture effectuée après l'arrêt de toutes les ISL ne sera pas mise en miroir sur l'autre site. Un basculement sur incident, dans cet état, entraînerait la perte des données non synchronisées. Dans ce cas, une intervention manuelle est requise pour la restauration après le basculement. S'il est probable qu'aucune ISL ne soit disponible pendant une période prolongée, l'administrateur peut choisir de fermer tous les services de données afin d'éviter tout risque de perte de données en cas de basculement en cas d'incident. L'exécution de cette action doit être comparée à la probabilité d'un incident nécessitant un basculement avant qu'au moins un lien ISL ne soit disponible. Sinon, si les liens ISL échouent dans un scénario en cascade, un administrateur peut déclencher un basculement planifié vers l'un des sites avant que tous les liens n'aient échoué.



Défaillance du lien de peering de cluster

Dans le cas d'une défaillance de liaison de cluster peering, les liens ISL de la structure sont toujours actifs, les services de données (lectures et écritures) continuent sur les deux sites vers les deux plexes. Toute modification de la configuration du cluster (par exemple, ajout d'un SVM, provisionnement d'un volume ou d'une LUN dans un SVM existant) ne peut pas être propagée à l'autre site. Ils sont conservés dans les volumes de métadonnées CRS locaux et automatiquement propagés à l'autre cluster lors de la restauration du lien du cluster peering. Si un basculement forcé est nécessaire avant la restauration de la liaison de cluster peering, les modifications de la configuration du cluster en attente seront automatiquement lues à partir de la copie répliquée à distance des volumes de métadonnées sur le site survivant dans le cadre du processus de basculement.



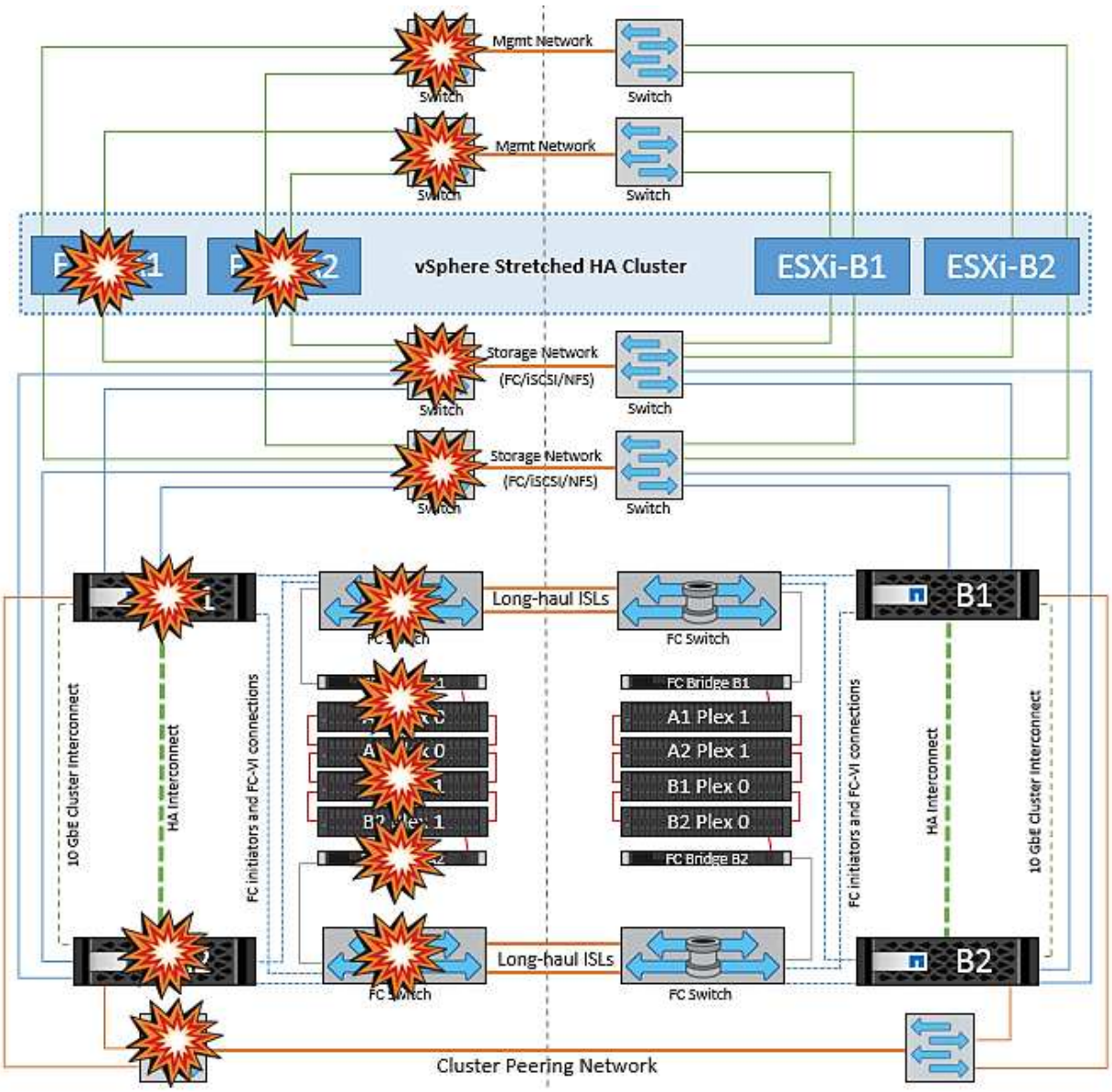
Défaillance complète du site

Dans un scénario de défaillance de site complet A, les hôtes ESXi du site B n'obtiennent pas la pulsation réseau des hôtes ESXi du site A car ils sont en panne. Le maître haute disponibilité sur le site B vérifie que les pulsations du datastore ne sont pas présentes, déclare que les hôtes du site A sont en panne et tente de redémarrer le site A des machines virtuelles sur le site B. Pendant cette période, l'administrateur du stockage effectue un basculement pour reprendre les services des nœuds défaillants sur le site survivant, ce qui restaure tous les services de stockage du site A sur le site B. Une fois que les volumes ou les LUN du site A sont disponibles sur le site B, l'agent principal de haute disponibilité tente de redémarrer le site A des machines virtuelles sur le site B.

Si la tentative de redémarrage d'une machine virtuelle par l'agent principal vSphere HA (qui implique son enregistrement et sa mise sous tension) échoue, le redémarrage est relancé après un délai. Le délai entre les redémarrages peut être configuré jusqu'à un maximum de 30 minutes. VSphere HA tente ces redémarrages au maximum pour un nombre maximal de tentatives (six tentatives par défaut).

Remarque : le maître HA ne lance pas les tentatives de redémarrage tant que le gestionnaire de placement n'a pas trouvé le stockage approprié, donc dans le cas d'une défaillance complète du site, ce serait une fois le basculement effectué.

Si le site A été basculé, la panne suivante de l'un des nœuds du site B survivant peut être gérée de manière transparente par le basculement vers le nœud survivant. Dans ce cas, le travail de quatre nœuds est désormais effectué par un seul nœud. Dans ce cas, la restauration consiste à effectuer un rétablissement vers le nœud local. Ensuite, lorsque le site A est restauré, une opération de rétablissement est effectuée pour restaurer le fonctionnement en état stable de la configuration.


Mgmt Network
Switch
Mgmt Network
Switch
vSphere Stretched HA Cluster
ESXi-B1
ESXi-B2
Storage Network
(FC/iSCSI/NFS)
Switch
Storage Network
(FC/iSCSI/NFS)
Switch
Long-haul ISLs
FC Switch
B1
FC Bridge B1
A1 Plex 1
A2 Plex 1
B1 Plex 0
B2 Plex 0
FC Bridge B2
10 GbE Cluster Interconnect
HA Interconnect
FC initiators and FC-VI connections
FC initiators and FC-VI connections
HA Interconnect
10 GbE Cluster Interconnect
Long-haul ISLs
FC Switch
B2
Cluster Peering Network