Disponibilité vSphere

VMware vSphere 6.5
VMware ESXi 6.5
vCenter Server 6.5

Ce document prend en charge la version de chacun des produits répertoriés, ainsi que toutes les versions publiées par la suite jusqu'au remplacement dudit document par une nouvelle édition. Pour rechercher des éditions plus récentes de ce document, rendez-vous sur : http://www.vmware.com/fr/support/pubs.

FR-002085-01

vmware®

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

http://www.vmware.com/fr/support/

Le site Web de VMware propose également les dernières mises à jour des produits.

N'hésitez pas à nous transmettre tous vos commentaires concernant cette documentation à l'adresse suivante :

docfeedback@vmware.com

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

VMware, Inc.
100-101 Quartier Boieldieu
92042 Paris La Défense
France
www.vmware.com/fr

Table des matières

À propos de Disponibilité de vSphere

Disponibilité vSphere présente les solutions permettant d'assurer la continuité d'activité, et explique notamment comment mettre en place vSphere® High Availability (HA) et vSphere Fault Tolerance.

Public cible

Ces informations sont destinées à tous ceux qui veulent assurer la continuité d'activité à l'aide des solutions vSphere HA et Fault Tolerance. Les informations fournies dans ce livre sont destinées aux administrateurs du système Windows ou Linux expérimentés qui connaissent le fonctionnement de la technologie des machines virtuelles et des centres de données.

Les instructions relatives aux tâches présentées dans ce guide se basent sur vSphere Web Client. Vous pouvez également exécuter la plupart des tâches de ce guide en utilisant la nouvelle version de vSphere Client. La terminologie, la topologie et le workflow de la nouvelle interface utilisateur de vSphere Clientcorrespondent fidèlement aux aspects et éléments de l'interface utilisateur de vSphere Web Client. Vous pouvez appliquer les instructions de vSphere Web Client à la nouvelle version de vSphere Client sauf mention du contraire.

Remarque Les fonctionnalités de vSphere Web Client n'ont pas toutes été mises en œuvre pour vSphere Client dans la version vSphere 6.5. Pour obtenir une liste actualisée des fonctionnalités non prises en charge, consultez le *Guide des mises à jour des fonctionnalités de vSphere Client* sur http://www.vmware.com/info?id=1413.

Informations mises à jour pour vSphere Availability

vSphere Availability est mis à jour avec chaque édition du produit ou lorsque cela est nécessaire.

Ce tableau indique l'historique des mises à jour de *vSphere Availability*.

Révision	Description
002085-01	■ Ajout d'informations sur la gestion des licences requises pour vCenter HA. Reportez-vous à la section « Configurations matérielle et logicielle requises de vCenter HA », page 69. ■ Exigence de mapper le nom de domaine complet supprimée des conditions préalables pour le réseau vCenter HA. Reportez-vous à la section « Configurer le réseau », page 73.
002085-00	Version initiale.

Continuité d'activité et minimisation des interruptions de service 1

Qu'elles soient prévues ou imprévues, les interruptions de service engendrent des coûts considérables. Cependant les solutions assurant des niveaux élevés de disponibilité sont généralement chères et difficiles à implémenter et à gérer.

Les logiciels de VMware assurent facilement et à moindre coût un niveau élevé de disponibilité pour les applications importantes. Avec vSphere, les entreprises peuvent augmenter facilement le niveau de disponibilité de base assuré pour toutes les applications et fournir des niveaux élevés de disponibilité plus facilement et à moindre frais. Avec vSphere, vous pouvez :

- Assurer une disponibilité élevée quels que soient les matériels, le système d'exploitation et les applications.
- Réduire les interruptions de service prévues pour les opérations de maintenance ordinaires.
- Assurer la restauration automatique en cas de dysfonctionnement.

vSphere permet de réduire les interruptions de service prévues, d'éviter des interruptions de service imprévues et de récupérer rapidement suite à des interruptions.

Ce chapitre aborde les rubriques suivantes :

Réduire les interruptions de service prévues

Les interruptions de service prévues représentent généralement plus de 80 % des interruptions de service d'un centre de données. La maintenance matérielle, la migration des serveurs et les mises à niveau des microprogramme imposent une interruption du service des serveurs physiques. Pour réduire les répercussions de ces interruptions de service, les entreprises doivent reporter la maintenance à des plages horaires peu pratiques et difficiles à planifier.

vSphere permet aux entreprises de réduire considérablement les interruptions de service prévues. Comme les charges de travail d'un environnement vSphere peuvent être déplacées dynamiquement sur différents serveurs physiques sans interruptions de service, la maintenance des serveurs peut être effectuée sans exiger une interruption des applications et du service. Avec vSphere, les entreprises peuvent :

- éliminer les interruptions de service pour les opérations de maintenance ordinaires.

- éliminer les plages de maintenance prévues.
- exécuter la maintenance à tout moment sans perturber les utilisateurs et les services.

vSphere vMotion® et la fonctionnalité Storage vMotion de vSphere permettent aux entreprises de réduire les interruptions de service prévues car les charges de travail d'un environnement VMware peuvent être déplacées dynamiquement sur d'autres serveurs physiques ou sur d'autres stockages sous-jacents sans interruption de service. Les administrateurs peuvent effectuer plus rapidement des opérations de maintenance entièrement transparentes, sans devoir planifier des plages de maintenance peu pratiques.

Prévenir les interruptions de service imprévues

Alors qu'un hôte ESXi offre une plate-forme stable pour exécuter des applications, les entreprises doivent aussi se protéger contre les interruptions de service imprévues provoquées par des pannes matérielles ou logicielles. vSphere renforce considérablement les capacités des infrastructures des centres de données, ce qui contribue à éviter les interruptions de service imprévues.

Ces capacités vSphere font partie d'une infrastructure virtuelle et sont transparentes pour le système d'exploitation et les applications exécutées sur les machines virtuelles. Ces fonctions peuvent être configurées et utilisées par toutes les machines virtuelles sur un système physique, ce qui réduit le coût et la complexité de la prévision d'une disponibilité supérieure. Des fonctions clés de disponibilité sont intégrées à vSphere :

- Stockage partagé. Élimine des points de panne isolés en stockant les fichiers des machines virtuelles dans des espaces de stockage partagés, comme Fibre Channel ou iSCSI SAN, ou encore NAS. Il est possible de faire appel aux fonctions de réplication et de mise en miroir SAN pour conserver les copies mises à niveau des disques virtuels dans des sites de reprise.
- Association d'interfaces réseau. Assure la tolérance aux défaillances des adaptateurs réseau individuelles.
- chemins multiples du stockage. Assure la tolérance aux défaillances des emplacements de stockage.

En outre, les fonctions vSphere HA et Fault Tolerance peuvent réduire ou éliminer les interruptions de service imprévues en assurant respectivement la reprise rapide de l'activité suite à une interruption et la continuité de la disponibilité.

vSphere HA assure une reprise d'activité rapide suite à une interruption

vSphere HA a recours à plusieurs hôtes ESXi configurés en cluster pour assurer une reprise d'activité rapide suite à une interruption et une haute disponibilité à moindres coûts pour les applications exécutées sur des machines virtuelles.

vSphere HA protège la disponibilité des applications de la manière suivante :

- Il protège contre une défaillance du serveur en redémarrant les machines virtuelles sur d'autres hôtes au sein du cluster.
- Il protège contre les défaillances des applications en surveillant en permanence une machine virtuelle et en la réinitialisant en cas de détection d'une défaillance.
- Il protège contre les erreurs d'accessibilité de la banque de données en redémarrant les machines virtuelles affectées sur d'autres hôtes ayant toujours accès à leurs banques de données.
- Il protège les machines virtuelles contre l'isolation réseau en les redémarrant si leurs hôtes se retrouvent isolés sur le réseau de gestion ou Virtual SAN. Cette protection est assurée même si le réseau s'est retrouvé partitionné.

Contrairement aux autres solutions de mise en cluster, vSphere HA fournit l'infrastructure nécessaire à la protection de toutes les charges de travail :

- Il n'est pas nécessaire d'installer des logiciels spéciaux dans l'application ou sur la machine virtuelle. Toutes les charges de travail sont protégées par vSphere HA. Une fois que vSphere HA est configuré, aucune action n'est requise pour protéger de nouvelles machines virtuelles. Elles sont protégées automatiquement.
- Vous pouvez associer vSphere HA à vSphere Distributed Resource Scheduler (DRS) pour assurer la protection contre les pannes, et pour répartir la charge entre tous les hôtes d'un cluster.

vSphere HA présente plusieurs avantages face aux solutions de basculement habituelles :

Configuration minimale	Quand un cluster vSphere HA a été configuré, toutes les machines virtuelles du cluster sont incluses dans le basculement sans configuration supplémentaire.
Coûts et configuration matérielle réduits	La machine virtuelle fait office de conteneur portable pour les applications et elle peut être déplacée parmi les hôtes. Les administrateurs évitent ainsi de reproduire les configurations sur plusieurs machines. Lorsque vous utilisez vSphere HA, vous devez disposer de suffisamment de ressources pour le basculement des hôtes que vous souhaitez protéger avec vSphere HA. Toutefois, le système vCenter Server® gère automatiquement les ressources et configure les clusters.
Disponibilité accrue des applications	Une application exécutée au sein d'une machine virtuelle a accès à une disponibilité accrue. Comme la machine virtuelle peut récupérer d'une défaillance matérielle, toutes les applications qui démarrent au moment de l'initialisation ont une disponibilité accrue sans accroître la charge de calcul, même si l'application n'est pas en cluster. En surveillant et en répondant aux signaux de pulsation de VMware Tools et en redémarrant les machines virtuelles qui ne répondent plus, elle assure également une protection contre les défaillances du système d'exploitation client.
Intégration DRS et vMotion	En cas de défaillance d'un hôte et du redémarrage des machines virtuelles sur d'autres hôtes, DRS peut fournir des recommandations de migration ou faire migrer les machines virtuelle en équilibrant les ressources allouées. Si l'hôte source et/ou l'hôte de destination d'une migration sont défaillants, vSphere HA peut faciliter la récupération suite à la défaillance.

vSphere Fault Tolerance assure la continuité de la disponibilité

vSphere HA assure un niveau de protection de base pour vos machines virtuelles en les redémarrant en cas de défaillance de l'hôte. vSphere Fault Tolerance assure un niveau de disponibilité supérieur en permettant aux utilisateurs de protéger les machines virtuelles contre une défaillance de l'hôte sans perte de données, de transactions ou de connexions.

Fault Tolerance assure la continuité de la disponibilité en vérifiant que les états des machines virtuelles principales et secondaires demeurent identiques tout au long de l'exécution des instructions de la machine virtuelle.

Si l'hôte faisant fonctionner la machine virtuelle principale ou l'hôte faisant fonctionner la machine virtuelle secondaire est défaillant, un basculement immédiat et transparent se produit. L'hôte ESXi en état de marche devient la machine virtuelle principale sans qu'il y ait perte des connexions réseau ou des transactions en cours. Le basculement transparent évite toute perte de données et assure le maintien des connexions réseau. En cas de basculement transparent, une nouvelle machine virtuelle est réaffectée et la redondance est rétablie. Le processus est entièrement transparent et automatisé et se produit même en cas d'indisponibilité du vCenter Server.

Protection de vCenter Server Appliance avec vCenter High Availability

vCenter High Availability (vCenter HA) protège non seulement contre les défaillances matérielles et de l'hôte mais également contre les défaillances de l'application vCenter Server. Grâce au basculement automatisé entre actif et passif, vCenter HA prend en charge la haute disponibilité avec un temps d'arrêt minimal.

Options de déploiement de vCenter HA

vCenter HA protège votre dispositif vCenter Server Appliance. Cependant, Platform Services Controller fournit l'authentification, la gestion des certificat et les licences pour vCenter Server Appliance. Par conséquent, vous devez garantir la haute disponibilité de Platform Services Controller. Vous avez plusieurs options.

- Déployez un nœud actif avec une instance intégrée de Platform Services Controller. Dans le cadre du processus de clonage, Platform Services Controller est cloné, ainsi que tous ses services. Dans le cadre de la synchronisation entre le nœud actif et le nœud passif, , Platform Services Controller est mis à jour sur le nœud passif.

  Lorsque le basculement se produit du mode actif au mode passif, les instances de Platform Services Controller sur le nœud passif sont disponibles, ainsi que l'environnement complet.

- Déployez au moins deux Platform Services Controller instances et placez-les derrière un équilibrage de charge.

  Lorsque le basculement se produit du nœud actif au nœud passif, le nœud passif continue de pointer vers l'équilibrage de charge. Lorsque l'une des instances de Platform Services Controller n'est plus disponible, l'équilibrage de charge dirige les requêtes vers la seconde instance de Platform Services Controller.

Reportez-vous à « Options de déploiement de vCenter HA », page 70.

Options de configuration de vCenter HA

Vous configurez vCenter HA à partir de vSphere Web Client. L'assistant de configuration offre les options suivantes.

Option	Description
De base	L'option De base clone le nœud actif en nœud passif et en nœud témoin, et configure le nœud pour vous. Vous pouvez utiliser cette option si votre environnement répond à l'une des conditions requises suivantes. ■ Soit l'instance de vCenter Server Appliance, qui deviendra le nœud actif, gère son propre hôte ESXi et sa propre machine virtuelle. Cette configuration de vCenter Server est parfois appelée gestion automatique. ■ Ou le dispositif vCenter Server Appliance est géré par une autre instance de vCenter Server (vCenter Server de gestion) et les deux instances de vCenter Server se trouvent dans le même domaine vCenter Single Sign-On. Cela implique que toutes deux utilisent un dispositif Platform Services Controller externe et qu'elles exécutent toutes deux vSphere 6.5. Reportez-vous à « Configurer vCenter HA avec l'option Basique », page 74.
Avancé	L'option Avancé offre davantage de flexibilité. Vous pouvez utiliser cette option tant que votre environnement satisfait les configurations matérielles et logicielles requises. Si vous sélectionnez cette option, vous devez cloner le nœud actif sur le nœud passif et sur le nœud témoin. Vous devez également effectuer une configuration de mise en réseau. Reportez-vous à « Configurer vCenter HA avec l'option Avancé », page 75.

Protection de vCenter Server avec VMware Service Lifecycle Manager

La disponibilité de vCenter Server est fournie par VMware Service Lifecycle Manager.

Si le service vCenter échoue, VMware Service Lifecycle Manager le redémarre. VMware Service Lifecycle Manager surveille la santé des services et exécute une action corrective préconfigurée en cas de détection de panne. Le service ne redémarre pas en cas de plusieurs tentatives de correction de panne.

Créer et utiliser des clusters vSphere HA

2

Les clusters vSphere HA permettent à un ensemble d'hôtes ESXi de travailler conjointement, de façon à fournir aux machines virtuelles, en tant que groupe, un niveau de disponibilité supérieur à celui d'un seul hôte ESXi. Si vous envisagez de créer et d'utiliser un nouveau cluster vSphere HA, les options choisies affectent la manière dont ce cluster réagit aux pannes des hôtes ou des machines virtuelles.

Avant de créer un cluster vSphere HA, vous devez savoir comment vSphere HA identifie les pannes et l'isolation de l'hôte et comment il réagit à ces situations. Vous devez aussi connaître le mode de fonctionnement du contrôle d'admission de façon à être capable de choisir les règles qui répondent à vos besoins de basculement. Après avoir créé un cluster, vous pouvez en personnaliser le comportement avec des options avancées et en optimiser les performances en suivant les recommandations.

Remarque Un message d'erreur peut apparaître lorsque vous essayez d'utiliser vSphere HA. Pour plus d'informations sur les messages d'erreur relatifs à vSphere HA, reportez-vous à l'article de la base de connaissances VMware sur http://kb.vmware.com/kb/1033634.

Ce chapitre aborde les rubriques suivantes :

Fonctionnement de vSphere HA

vSphere HA assure la disponibilité élevée des machines virtuelles en les rassemblant avec leurs hôtes respectifs dans un cluster. Les hôtes du cluster sont surveillés et, en cas de défaillance, les machines virtuelles d'un hôte défectueux sont redémarrées sur d'autres hôtes.

Lorsque vous créez un cluster vSphere HA , un seul hôte est automatiquement sélectionné en tant qu'hôte maître. L'hôte maître communique avec vCenter Server et surveille l' état de protection de toutes les machines virtuelles et des hôtes esclaves. Différents types de défaillances d'hôtes sont possibles, et l'hôte principal doit les détecter et les traiter de façon adaptée. L'hôte principal doit faire la différence entre un hôte défaillant et un hôte se trouvant dans une partition de réseau ou réseau isolé. L'hôte principal utilise le signal de pulsation de banques de données pour déterminer le type de panne.

Clusters vSphere HA (http://link.brightcove.com/services/player/bcpid2296383276001?bctid=ref:vSphereHAClusters)

Hôte maître et hôtes esclaves

Lorsque vous ajoutez un hôte à un cluster vSphere HA, un agent est transféré vers l'hôte et configuré pour communiquer avec les autres agents du cluster. Chaque hôte du cluster fonctionne en tant qu'hôte principal (maître) ou hôte secondaire (esclave).

Lorsque vSphere HA est activé pour un cluster, tous les hôtes actifs (ceux qui ne sont pas en mode standby ou maintenance, ou qui ne sont pas déconnectés) participent au choix de l'hôte principal du cluster. L'hôte contenant le plus grand nombre de banques de données a l'avantage pour être choisi. Habituellement, il n'existe qu'un hôte principal par cluster, tous les autres sont des hôtes secondaires. Si l'hôte principal est défaillant, fermé, mis en mode standby ou éliminé du cluster, un nouvel hôte principal doit être choisi.

L'hôte principal d'un cluster a un certain nombre de responsabilités :

- Surveiller l'état des hôtes secondaires. Si un hôte secondaire est défaillant ou devient inaccessible, l'hôte principal identifie les machines virtuelles qui doivent être redémarrées.
- Surveiller l'état d'alimentation de toutes les machines virtuelles protégées. Si une machine virtuelle est défaillante, l'hôte principal s'assure qu'elle est redémarrée. Grâce à un moteur de placement local, l'hôte principal détermine également où le redémarrage doit avoir lieu.
- Gérer les listes d'hôtes et de machines virtuelles protégées du cluster.
- Servir d'interface de gestion vCenter Server du cluster et rendre compte de l'état de santé du cluster.

Les hôtes secondaires apportent une contribution essentielle au cluster en exécutant des machines virtuelles localement, en surveillant leur état d'exécution et en communiquant les mises à jour d'état à l'hôte principal. Un hôte principal peut également exécuter et surveiller des machines virtuelles. Les hôtes principaux et les hôtes secondaires mettent en œuvre les fonctions de surveillance de VM et d'application.

Une des fonctions exercées par l'hôte maître est la coordination des redémarrages de machines virtuelles protégées. Une VM est protégée par un hôte maître après que vCenter Server observe que l'état d'alimentation de la VM est passé de hors tension à sous tension en réponse à une action de l'utilisateur. L'hôte maître conserve la liste des machines virtuelles protégées dans les banques de données du cluster. Un hôte maître nouvellement élu utilise ces informations pour déterminer quelles machines virtuelles doivent être protégées.

Remarque Si vous déconnectez un hôte d'un cluster, aucune des machines virtuelles enregistrées sur cet hôte n'est protégée par vSphere HA.

Types de pannes d'hôte

L'hôte principal d'un cluster vSphere HA est responsable de la détection des pannes des hôtes secondaires. Selon le type de panne détecté, les machines virtuelles exécutées sur les hôtes peuvent nécessiter un basculement.

Dans un cluster vSphere HA, trois types de pannes d'hôtes sont détectés :

- Panne : un hôte cesse de fonctionner.
- Isolation : un hôte se retrouve isolé sur le réseau.
- Partition : un hôte perd sa connectivité réseau avec l'hôte principal.

L'hôte principal surveille la réactivité des hôtes secondaires du cluster. Cette communication s'effectue par l'échange, toutes les secondes, de signaux de pulsation réseau. Lorsqu'un hôte principal cesse de recevoir des signaux de pulsation d'un hôte secondaire, il vérifie la réactivité de l'hôte avant de le déclarer défaillant. Le contrôle de réactivité effectué par l'hôte principal permet de déterminer si l'hôte secondaire échange des signaux de pulsation avec une des banques de données. Reportez-vous à « Signal de pulsation de banque de données », page 22. Par ailleurs, l'hôte principal vérifie si l'hôte répond aux pings ICMP envoyés à ses adresses IP de gestion.

Si un hôte principal ne peut pas communiquer directement avec l'agent sur un hôte secondaire, celui-ci ne répond pas aux commandes ping ICMP. Si l'agent n'émet pas de pulsations, il est considéré comme défaillant. Les machines virtuelles des hôtes sont redémarrées sur d'autres hôtes. Si cet hôte secondaire échange des pulsations avec une banque de données, l'hôte principal suppose que l'hôte secondaire se trouve dans une partition du réseau ou est isolé du réseau. L'hôte principal continue donc à surveiller l'hôte et ses machines virtuelles. Reportez-vous à « Partitions de réseau », page 22.

L'isolation du réseau de l'hôte survient lorsqu'un hôte, toujours en cours d'exécution, ne parvient plus à observer le trafic provenant des agents vSphere HA sur le réseau de gestion. Si un hôte cesse d'observer ce trafic, il tente d'envoyer un ping aux adresses d'isolation du cluster. Si cette commande ping échoue également, l'hôte déclare qu'il est isolé du réseau.

L'hôte principal surveille les machines virtuelles qui s'exécutent sur un hôte isolé. Si l'hôte principal remarque que les machines virtuelles se mettent hors tension et qu'il en est responsable, il les redémarre.

Remarque Si vous vous assurez que l'infrastructure réseau est suffisamment redondante et qu'au moins un chemin d'accès au réseau est disponible en permanence, l'isolation du réseau de l'hôte est moins susceptible de se produire.

Pannes de Proactive HA

Une panne de Proactive HA se produit lorsqu'un composant hôte est défaillant, ce qui entraîne une perte de redondance ou une panne non grave. Cependant, le comportement de fonctionnement des machines virtuelles qui résident sur l'hôte n'est pas affecté. Par exemple, si une alimentation électrique sur l'hôte tombe en panne, mais que les autres alimentations sont disponibles, il s'agit d'une panne de Proactive HA.

En cas de panne de Proactive HA, vous pouvez automatiser la mesure corrective prise dans la section Disponibilité vSphere de vSphere Web Client. Les machines virtuelles sur l'hôte concerné peuvent être évacuées vers d'autres hôtes et l'hôte est mis en mode de quarantaine ou de maintenance.

Remarque Pour que la surveillance des pannes de Proactive HA fonctionne, votre cluster doit utiliser vSphere DRS.

Déterminer les réponses aux problèmes de l'hôte

Si un hôte échoue et que ses machines virtuelles doivent être redémarrées, vous pouvez contrôler l'ordre dans lequel cela se fait avec le paramètre de priorité de redémarrage des machines virtuelles. De même, vous pouvez configurer la réponse de vSphere HA lorsque des hôtes perdent la connectivité au réseau de gestion à d'autres hôtes en utilisant les paramètres de réponse d'isolation. D'autres facteurs sont également pris en compte lorsque vSphere HA redémarre une machine virtuelle après un échec.

Les paramètres suivants s'appliquent à toutes les machines virtuelles du cluster en cas d'échec ou d'isolation d'un hôte. Vous pouvez configurer des exceptions pour des machines virtuelles spécifiques. Reportez-vous à « Personnaliser une machine virtuelle secondaire », page 44.

Réponse d'isolation de l'hôte

La réponse d'isolation d'hôte détermine les événements survenant lorsqu'un hôte d'un cluster vSphere HA perd ses connexions au réseau de gestion, mais continue à s'exécuter. Vous pouvez utiliser la réponse d'isolation afin que vSphere HA mette hors tension les machines virtuelles en cours d'exécution sur un hôte isolé et les redémarre sur un hôte non isolé. Les réponses d'isolation d'hôte exigent que l'état de surveillance de l'hôte soit activé. Si l'état de surveillance de l'hôte est désactivé, les réponses d'isolation d'hôte sont également suspendues. Un hôte détermine qu'il est isolé lorsqu'il est incapable de communiquer avec les

agents en cours d'exécution sur les autres hôtes et d'envoyer un ping à ses adresses d'isolation. L'hôte exécute ensuite sa réponse d'isolation. Les réponses sont Mettre hors tension et redémarrer les VM ou Arrêter et redémarrer les machines virtuelles. Vous pouvez personnaliser cette propriété pour des machines virtuelles individuelles.

Remarque Si le paramètre de priorité de redémarrage d'une machine virtuelle est défini sur Désactivée, aucune réponse d'isolation d'hôte n'est fournie.

Pour utiliser le paramètre Arrêter et redémarrer les machines virtuelles, vous devez installer VMware Tools dans le système d'exploitation invité de la machine virtuelle. L'arrêt de la machine virtuelle offre l'avantage de préserver son état. L'arrêt est préférable à la mise hors tension de la machine virtuelle qui ne prend pas en compte pas les dernières modifications apportées aux disques ni ne valide les transactions. Le basculement des machines virtuelles qui sont en train de s'arrêter est plus long car la fermeture doit aussi être effectuée. Les machines virtuelles qui n'ont pas été arrêtées au bout de 300 secondes ou du délai défini par l'option avancée `das.isolationshutdowntimeout` sont mises hors tension.

Lorsque vous avez créé un cluster vSphere HA, vous pouvez changer les paramètres par défaut du cluster relatifs à la priorité de redémarrage et à la réponse d'isolation de machines virtuelles spécifiques. Ces remplacements sont utiles pour les machines virtuelles qui sont utilisées pour des tâches spéciales. Par exemple, les machines virtuelles qui fournissent des services d'infrastructure, comme DNS ou DHCP, doivent éventuellement être mises sous tension avant d'autres machines virtuelles du cluster.

Une condition de split-brain peut se produire sur une machine virtuelle lorsqu'un hôte se retrouve isolé ou partitionné depuis un hôte principal qui ne peut pas communiquer avec lui à l'aide des banques de données des signaux de pulsation. Dans une telle situation, l'hôte principal n'est pas en mesure de déterminer si l'hôte est actif et le déclare inactif. L'hôte principal fait ensuite une tentative pour redémarrer les machines virtuelles qui s'exécutent sur l'hôte isolé ou partitionné. Cette tentative réussit si les machines virtuelles continuent de s'exécuter sur l'hôte isolé ou partitionné et celui-ci perd l'accès aux banques de données des machines virtuelles quand il s'est retrouvé isolé ou partitionné. Il existe alors une condition de split-brain, car la machine virtuelle se retrouve avec deux instances. Toutefois, seule une de ces instances est en mesure de lire ou d'écrire sur les disques virtuels de la machine virtuelle. VM Component Protection peut vous aider à empêcher cette condition de split-brain. Lorsque vous activez VMCP avec le paramètre intensif, il contrôle l'accessibilité de la banque de données sur les machines virtuelles sous tension et arrête celles qui perdent l'accès à leurs banques de données.

Pour résoudre ce problème, ESXi génère une question sur la machine virtuelle qui a perdu les verrouillages disque pour le moment où l'hôte quitte son état d'isolation et est dans l'impossibilité obtenir de nouveau les verrouillages disque. vSphere HA répond automatiquement à cette question ce qui permet à l'instance de la machine virtuelle qui a perdu les verrouillages disque de se mettre hors tension, laissant uniquement l'instance qui dispose des verrouillages disque.

Dépendances des machines virtuelles

Vous pouvez créer des dépendances entre les groupes de machines virtuelles. Pour cela, vous devez d'abord créer les groupes de VM dans vSphere Web Client en accédant à l'onglet **Configurer** du cluster et en sélectionnant **Groupes de VM/Hôte**. Une fois les groupes créés, vous pouvez créer des règles de redémarrage des dépendances entre les groupes en accédant à l'onglet **Règles de VM/Hôte** et en sélectionnant dans le menu déroulant **Machines virtuelles vers machines virtuelles**. Ces règles peuvent spécifier que certains groupes de VM ne peuvent pas être redémarrés tant que d'autres groupes spécifiés n'ont pas été démarrés en premier.

Facteurs pris en charge pour le redémarrage de la machine virtuelle

Après un échec, l'hôte principal du cluster fait une tentative de redémarrage des machines virtuelles concernées en identifiant un hôte susceptible de les mettre sous tension. Lors de la sélection de cet hôte, l'hôte principal tient compte d'un certain nombre de facteurs.

Accessibilité des fichiers	Avant de pouvoir redémarrer une machine virtuelle, ses fichiers doivent être accessibles depuis l'un des hôtes actif du cluster avec lequel l'hôte principal peut communiquer via le réseau.
Machine virtuelle et compatibilité de l'hôte	S'il existe des hôtes accessibles, la machine virtuelle doit être compatible avec au moins l'un d'entre eux. La compatibilité définie pour une machine virtuelle comprend l'effet de l'une des règles d'affinité machine virtuelle/hôte. Par exemple, si une règle permet à une machine virtuelle de s'exécuter sur deux hôtes, elle est prise en compte pour le placement sur ces deux hôtes.
Réservations de ressources	Parmi les hôtes sur lesquels la machine virtuelle peut s'exécuter, au moins un doit disposer d'une capacité non réservée suffisante pour satisfaire aux besoins de la mémoire de temps système de la machine virtuelle et aux réservations de ressources. Quatre types de réservations sont prises en compte : CPU, mémoire, vNIC et lecteur Flash virtuel. De plus, un nombre de ports réseau suffisant doit être disponible pour mettre sous tension la machine virtuelle.
Limites d'hôtes	En plus des réservations de ressources, une machine virtuelle ne peut être placée sur un hôte que si cela ne lui fait pas dépasser le nombre maximal de machines virtuelles autorisées ou de vCPU utilisés.
Contraintes de la fonctionnalité	Si l'option avancée qui a été définie nécessite que vSphere HA fasse respecter les règles d'affinité machine virtuelle/machine virtuelle, vSphere HA n'enfreint pas cette règle. De plus, vSphere HA n'enfreint pas les limites configurée pour chaque hôte pour les machines virtuelles Fault Tolerance.

Si aucun hôte ne répond aux considérations précédentes, l'hôte principal émet un événement indiquant qu'il ne dispose pas des ressources suffisantes pour que vSphere HA démarre la machine virtuelle et ressaiera une fois les conditions du cluster améliorées. Par exemple, si la machine virtuelle n'est pas accessible, l'hôte principal réessaie après une modification de l'accessibilité des fichiers.

Surveillance des VM et applications

Surveillance de VM redémarre les machines virtuelles si leurs signaux de pulsation de VMware Tools n'ont pas été reçus pendant un certain temps. De même, la Surveillance d'application peut redémarrer une machine virtuelle si les signaux de pulsation d'une application exécutée ne sont pas reçus. Il est possible d'activer ces fonctions et de configurer la sensibilité de la surveillance de l'absence de réaction par vSphere HA.

Lorsque vous activez la Surveillance de VM, le service Surveillance de VM (à l'aide de VMware Tools) vérifie si chaque machine virtuelle du cluster fonctionne en vérifiant la régularité des signaux de pulsations et l'activité des E/S à partir du processus VMware Tools exécuté sur le client. Si aucun signal de pulsation ou activité des E/S n'est reçu, cela est probablement dû à une défaillance du système d'exploitation client ou au fait que les VMware Tools n'ont pas eu le temps de terminer certaines tâches. Dans ce cas, le service Surveillance de VM détermine que la machine virtuelle est défectueuse et la machine virtuelle redémarre pour restaurer le service.

Il arrive qu'occasionnellement, les machines virtuelles ou les applications qui continuent à fonctionner correctement, cessent d'émettre des signaux de pulsation. Pour éviter les réinitialisations inutiles, le service Surveillance de VM surveille aussi l'activité des E/S d'une machine virtuelle. Si aucun signal de pulsation n'est reçu pendant la période de défaillance, la fréquence des statistiques des E/S (attribut défini au niveau du cluster) est vérifiée. La fréquence des statistiques des E/S détermine si un disque ou une activité réseau s'est produite sur la machine virtuelle au cours des deux minutes (120 secondes) précédentes. Si ce n'est pas le cas, la machine virtuelle est réinitialisée. Cette valeur par défaut (120 secondes) peut être modifiée à l'aide de l'option avancée `das.iostatsinterval`.

Pour activer la surveillance d'application, il faut d'abord obtenir le SDK approprié (ou utiliser une application qui prend en charge la surveillance de l'application VMware) et l'utiliser pour configurer des signaux de pulsation personnalisés pour les applications à surveiller. Après avoir fait cela, la surveillance d'application fonctionne de la même manière que la Surveillance de VM. Si les signaux de pulsation d'une application ne sont pas reçus pendant un certain temps, sa machine virtuelle est redémarrée.

Vous pouvez configurer le niveau de sensibilité de la surveillance. Une sensibilité de surveillance élevée permet de conclure plus rapidement à un dysfonctionnement. Même si cela est peu probable, une sensibilité de surveillance élevée peut entraîner l'identification erronée de dysfonctionnements alors que la machine virtuelle ou l'application en question fonctionne toujours mais les signaux de pulsation ne sont pas reçus du fait de certains facteurs tels que des contraintes de ressources. Une sensibilité de surveillance basse se traduit par des interruptions de service prolongées entre les défaillances avérées et le redémarrage des machines virtuelles. Sélectionnez l'option qui offre un compromis intéressant par rapport à vos besoins.

Les paramètres par défaut de la sensibilité de surveillance sont décrits dans Tableau 2-1. Vous pouvez aussi indiquer des valeurs personnalisées à la fois pour la sensibilité de la surveillance et les intervalles de statistiques d'E/S en cochant la case **Personnalisé**.

Tableau 2-1. Paramètres de surveillance des machines virtuelles

Paramètre	Intervalle de défaillance (en secondes)	Période de réinitialisation
Haut	30	1 heure
Moyen	60	24 heures
Faible	120	7 jours

Lorsque des dysfonctionnements sont détectés, vSphere HA réinitialise les machines virtuelles. La réinitialisation contribue à garantir que les services restent disponibles. Pour éviter de réinitialiser constamment des machines virtuelles en cas d'erreurs non transitoires, les machines virtuelles sont réinitialisées par défaut trois fois seulement au cours d'une période configurable. Après trois réinitialisations des machines virtuelles, vSphere HA n'effectue aucune tentative supplémentaire pour redémarrer les machines virtuelles en cas de nouvel échec et ce jusqu'à ce que la période définie ne soit écoulée. Vous pouvez configurer le nombre de réinitialisations à l'aide du paramètre personnalisé **Nbre maximum de réinitialisations par machine virtuelle**.

Remarque Les statistiques de réinitialisation sont effacées lorsque la machine virtuelle est mise hors tension puis sous tension, ou quand elle est migrée à un autre hôte en utilisant vMotion. Cela provoque le redémarrage du système d'exploitation d'hôte, mais de façon différente à un «redémarrage» dans lequel l'état d'alimentation de la VM est changé.

Si une machine virtuelle rencontre un problème d'accessibilité à la banque de données (Tous chemins hors service ou Perte de périphérique permanente), le service de surveillance de machine virtuelle interrompt sa réinitialisation jusqu'à ce que le problème ait été résolu.

VM Component Protection

Si VM Component Protection (VMCP) est activé, vSphere HA peut détecter les erreurs d'accessibilité à la banque de données et fournir une récupération automatisée pour les machines virtuelles concernées.

VMCP offre une protection contre les erreurs d'accessibilité à la banque de données qui affectent une machine virtuelle s'exécutant sur un hôte dans un cluster vSphere HA. En cas d'erreur d'accessibilité à une banque de données, l'hôte affecté ne peut plus accéder au chemin de stockage d'une banque de données spécifique. Vous pouvez déterminer la réaction de vSphere HA face à cette erreur, depuis la création d'alarmes d'événement jusqu'au redémarrage de la machine virtuelle sur d'autres hôtes.

Remarque Pour utiliser la fonctionnalité VM Component Protection, la version de vos hôtes ESXi doit être 6.0 ou une version ultérieure.

Types d'erreurs

Il existe deux types d'erreurs d'accessibilité à une banque de données :

PDL — PDL (perte de périphérique permanente) est une perte d'accessibilité irrécupérable qui se produit lorsqu'un périphérique de stockage signale que la banque de données n'est plus accessible à l'hôte. Cette condition ne peut pas être rétablie sans mettre hors tension les machines virtuelles.

APD — APD (Tous chemins hors service) représente une perte d'accessibilité temporaire ou inconnue, ou tout autre retard non identifié dans le traitement des E/S. Ce type d'erreur d'accessibilité est récupérable.

Configuration de VMCP

La fonctionnalité VM Component Protection est configurée dans vSphere Web Client. Accédez à l'onglet **Configurer** et cliquez sur **Disponibilité vSphere**, puis cliquez sur **Modifier**. Sous **Pannes et réponses**, vous pouvez sélectionnez l'option **Banque de données avec PDL** ou **Banque de données avec APD**. Les niveaux de protection du stockage que vous pouvez sélectionner et les actions de correction de la machine virtuelle disponibles varient selon le type d'erreur d'accessibilité à la base de données.

Erreurs PDL — Sous **Banque de données avec PDL**, vous pouvez sélectionner l'option **Émission d'événements** ou **Mettre hors tension et redémarrer les VM**.

Erreurs APD — La réponse aux événements APD est plus complexe et, en fonction de la configuration, est définie avec une plus grande précision. Vous pouvez sélectionner l'option **Émission d'événements**, **Mettre hors tension et redémarrer les VM : stratégie de redémarrage modérée** ou **Mettre hors tension et redémarrer les VM : stratégie de redémarrage agressive**.

Remarque Si les paramètres Surveillance VM ou Priorité redémarrage VM sont désactivés, VMCP ne peut pas redémarrer la machine virtuelle. Toutefois, la santé du stockage peut toujours être surveillée et les événements être émis.

Partitions de réseau

En cas de défaillance du réseau de gestion d'un cluster vSphere HA, un sous-ensemble d'hôtes du cluster risque d'être incapable de communiquer avec les autres hôtes sur le réseau de gestion. De multiples partitions peuvent se produire dans un cluster.

Un cluster partitionné entraîne une diminution de la protection des machines virtuelles et une altération des fonctions de gestion du cluster. Réparez le cluster partitionné dès que possible.

- Protection de VM. vCenter Server permet de mettre sous tension une VM, mais celle-ci n'est protégée que si elle s'exécute sur la même partition que l'hôte principal qui en est responsable. L'hôte principal doit communiquer avec vCenter Server. Un hôte principal est responsable d'une machine virtuelle s'il a bloqué exclusivement un fichier défini par le système sur la banque de données contenant le fichier de configuration de la machine virtuelle.
- Gestion de cluster. vCenter Server peut communiquer avec l'hôte principal, mais uniquement un sous-ensemble d'hôtes secondaires. Par conséquent, il se peut que les modifications de configuration relatives à vSphere HA ne prennent pas effet tant que le problème de partition n'est pas résolu. Suite à cette défaillance, une des partitions pourrait s'exécuter selon l'ancienne configuration, tandis qu'une autre utiliserait les nouveaux paramètres.

Signal de pulsation de banque de données

Lorsque l'hôte principal d'un cluster vSphere HA ne peut pas communiquer avec un hôte secondaire sur le réseau de gestion, l'hôte principal utilise le signal de pulsation de banque de données pour déterminer si l'hôte secondaire est défaillant, s'il se trouve dans une partition de réseau ou s'il est réseau isolé. Si l'hôte secondaire a arrêté le signal de pulsation de banque de données, il est considéré comme défaillant et ses machines virtuelles sont redémarrées ailleurs.

vCenter Server sélectionne un ensemble de banques de données préférées pour le signal de pulsation. Cette sélection a pour but d'optimiser le nombre d'hôtes ayant accès à une banque de données de signaux de pulsation et de minimiser le risque que les banques de données soient sauvegardées par le même LUN ou le même serveur NFS.

Vous pouvez utiliser l'option avancée `das.heartbeatdsperhost` pour modifier le nombre de banques de données de signaux de pulsation sélectionné par vCenter Server pour chaque hôte. La valeur par défaut est deux et la valeur maximale est cinq.

vSphere HA crée un répertoire à la racine de chaque banque de données qui sert à la fois au signal de pulsation de banques de données et à maintenir l'ensemble des machines virtuelles protégées. Le nom de ce répertoire est `.vSphere-HA`. Vous ne devez ni supprimer ni modifier les fichiers stockés dans ce répertoire car cela peut avoir des répercussions sur les opérations. Plusieurs clusters peuvent utiliser une banque de données. Des sous-répertoires sont donc créés dans ce répertoire pour chaque cluster. Ces répertoires et fichiers font partie de la racine, et seule celle-ci peut les lire et les modifier. L'espace disque utilisé par vSphere HA dépend de plusieurs facteurs, notamment la version de VMFS et le nombre d'hôtes qui utilisent la banque de données pour le signal de pulsation. Avec vmfs3, l'utilisation maximale est d'environ 2 Go et l'utilisation type est d'environ 3 Mo. Avec vmfs5, l'utilisation normale maximale est d'environ 3 Mo. L'utilisation vSphere HA de la banque de données ajoute une charge additionnelle négligeable et n'a pas d'impact sur la performance des autres opérations de la banque de données.

vSphere HA limite le nombre de machines virtuelles qui peuvent avoir des fichiers de configuration sur une banque de données unique. Consultez *Configurations Maximales* pour connaître les limites mises à jour. Si vous placez plus que ce nombre de machines virtuelles sur une banque de données et que vous les mettez sous tension, vSphere HA ne protège un certain nombre de machines virtuelles que jusqu'à cette limite.

Remarque Une banque de données de Virtual SAN ne peut pas être utilisée pour le signal de pulsation de banque de données. Par conséquent, si aucun autre stockage partagé n'est accessible à tous les hôtes du cluster, il se peut qu'aucune banque de données de signaux de pulsation ne soit utilisée. Toutefois, si vous disposez d'un stockage qui peut être atteint par un chemin réseau alternatif indépendant de Virtual SAN, vous pouvez l'utiliser pour configurer une banque de données de signaux de pulsation.

Sécurité vSphere HA

Plusieurs fonctions de sécurité permettent d'améliorer vSphere HA.

Sélectionner les ports de pare-feu ouverts

vSphere HA utilise les ports 8182 TCP et UDP pour la communication d'agent à agent. Les ports de pare-feu s'ouvrent et se ferment automatiquement pour assurer qu'ils sont ouverts uniquement lorsque cela est nécessaire.

Fichiers de configuration protégés par les autorisations du système de fichiers

vSphere HA stocke les informations de configuration sur le système de stockage local ou sur le ramdisk s'il n'existe aucune banque de données locale. Ces fichiers sont protégés par les autorisations du système de fichiers et sont accessibles uniquement par l'utilisateur racine. Les hôtes sans stockage local sont pris en charge uniquement si ils sont gérés par Auto Deploy.

Journalisation détaillée

L'emplacement des fichiers journaux choisi par vSphere HA dépend de la version de l'hôte.

- Pour les hôtes ESXi 5.x, vSphere HA écrit sur syslog uniquement par défaut. Les journaux sont donc placés à l'endroit indiqué dans la configuration de syslog. Les noms des fichiers journaux de vSphere HA sont précédés de `fdm`, fault domain manager (gestionnaire de domaine de pannes), qui est un service de vSphere HA.
- Pour les hôtes existants 4.x ESXi, vSphere HA écrit dans `/var/log/vmware/fdm` sur le disque local, ainsi que syslog si il est configuré.
- Pour les hôtes hérités ESX 4.x, vSphere HA écrit sur `/var/log/vmware/fdm`.

Connexions vSphere HA sécurisées

vSphere HA se connecte aux agents vSphere HA à l'aide d'un compte d'utilisateur, **vpxuser**, créé par vCenter Server. Ce compte est le même que celui utilisé par vCenter Server pour la gestion de l'hôte. vCenter Server crée un mot de passe aléatoire pour ce compte et le change régulièrement. La fréquence de renouvellement du mot de passe est définie par le paramètre `VirtualCenter.VimPasswordExpirationInDays` de vCenter Server. Les utilisateurs ayant des privilèges d'administration sur le dossier racine de l'hôte peut se connecter à l'agent.

Communication sécurisée	Toutes les communications entre vCenter Server et l'agent vSphere HA sont sécurisées par SSL. La communication d'agent à agent utilise également le protocole SSL sauf pour les messages d'élection, qui utilisent UDP. Les messages d'élection sont vérifiés via SSL de sorte qu'un agent non autorisé puisse empêcher uniquement l'hôte sur lequel l'agent s'exécute d'être choisi comme hôte principal. Dans ce cas, un problème de configuration du cluster est émis afin que l'utilisateur soit informé du problème.
Vérification du certificat SSL de l'hôte requise	vSphere HA exige que chaque hôte dispose d'un certificat SSL vérifié. Chaque hôte génère un certificat auto-signé lors de son premier démarrage. Ce certificat peut être généré une nouvelle fois ou remplacé par un certificat émis par une autorité. Si le certificat est remplacé, vSphere HA doit être reconfiguré sur l'hôte. Si un hôte se déconnecte de vCenter Server après la mise à jour de son certificat et si l'agent de l'hôte ESXi ou ESX est redémarré, vSphere HA est automatiquement reconfiguré au moment où l'hôte est reconnecté à vCenter Server. Si la déconnexion n'est pas due au fait que la vérification du certificat SSL de l'hôte de vCenter Server est désactivée à ce moment-là, vérifiez le nouveau certificat et reconfigurez vSphere HA sur l'hôte.

Contrôle d'admission vSphere HA

vSphere HA utilise le contrôle d'admission pour s'assurer que des ressources suffisantes sont réservées à la récupération des machine virtuelles en cas de défaillance d'un hôte.

Le contrôle d'admission impose des contraintes sur l'utilisation des ressources. Les actions qui risquent d'enfreindre ces contraintes ne sont pas autorisées. Les actions qui peuvent ne pas être autorisées incluent les exemples suivants :

- Mise sous tension d'une machine virtuelle
- Migration d'une machine virtuelle
- Augmentation de la réserve de CPU ou de mémoire d'une machine virtuelle

La base du contrôle d'admission vSphere HA est le nombre de défaillances d'hôte que le cluster est autorisé à tolérer et qui continue à garantir le basculement. La capacité de basculement des hôtes peut être définie de trois manières différentes :

- Pourcentage de ressources du cluster
- Stratégie d'emplacement
- Hôtes de basculement dédiés

Remarque Le contrôle d'admission vSphere HA peut être désactivé. Cependant, sans ce contrôle, il est impossible de garantir que le nombre de machines virtuelles attendu puisse être redémarré après une défaillance. Ne désactivez pas le contrôle d'admission de façon permanente.

Quelle que soit l'option de contrôle d'admission choisie, un seuil de réduction des ressources de VM existe également. Ce paramètre permet de spécifier le pourcentage de dégradation des ressources pouvant être toléré, mais il n'est pas disponible si vSphere DRS n'est pas activé.

Le calcul de la réduction des ressources est vérifié pour le CPU et la mémoire. Il prend en compte la mémoire réservée d'une machine virtuelle et la surcharge de la mémoire pour décider de l'autoriser ou non à être mise sous tension, migrée ou à modifier sa réservation. La mémoire réelle utilisée par la machine virtuelle n'est pas prise en compte dans le calcul car la réservation de mémoire ne correspond pas toujours à l'utilisation réelle de la mémoire de la machine virtuelle. Si l'utilisation réelle est supérieure à la mémoire réservée, une capacité de basculement insuffisante est disponible, ce qui entraîne la dégradation des performances lors du basculement.

Définir un seuil de réduction des performances vous permet de spécifier l'occurrence d'un problème de configuration. Par exemple :

- La valeur par défaut est 100 %, qui ne produit pas d'avertissements.
- Si vous réduisez le seuil à 0 %, un avertissement est généré dès que l'utilisation du cluster est supérieure à la capacité disponible.
- Si vous réduisez le seuil à 20 %, la réduction des performances pouvant être tolérée est calculée de la manière suivante : `performance reduction = current utilization * 20%`. Lorsque l'utilisation actuelle moins la réduction des performances dépasse la capacité disponible, une notification concernant la configuration est émise.

Contrôle d'admission Pourcentage de ressources de cluster

Il est possible de configurer vSphere HA pour effectuer le contrôle d'admission en réservant un pourcentage spécifique de ressources de CPU et de mémoire du cluster à la récupération en cas de pannes d'hôtes.

Avec ce type de contrôle d'admission, vSphere HA vérifie qu'un pourcentage spécifié de ressources cumulées de CPU et de mémoire est réservé au basculement.

Lorsque l'option de pourcentage de ressources de cluster est configurée, vSphere HA met en œuvre le contrôle d'admission de la manière suivante :

1 Calcule les besoins totaux en ressources pour toutes les machines virtuelles sous tension dans le cluster.

2 Calcule les ressources totales de l'hôte disponibles pour les machines virtuelles.

3 Calcule la Capacité CPU de basculement actuelle et la Capacité mémoire de basculement actuelle du cluster.

4 Détermine si la Capacité de basculement de CPU actuelle ou la Capacité de basculement mémoire actuelle sont inférieures ou non à la Capacité de basculement configurée correspondante (spécifiée par l'utilisateur).

   Si c'est le cas, le contrôle d'admission n'autorise pas l'opération.

vSphere HA utilise les réserves effectives des machines virtuelles. Si une machine virtuelle n'a pas de réserves, c'est-à-dire que la valeur de réserve est nulle, les valeurs utilisées par défaut sont 0 Mo de mémoire et 32 MHz de CPU.

Remarque L'option de pourcentage de ressources de cluster du contrôle d'admission vérifie également qu'il existe au moins deux hôtes compatibles vSphere HA dans le cluster (à l'exception des hôtes qui passent en mode maintenance). S'il n'y a qu'un hôte compatible vSphere HA, aucune opération n'est autorisée, même si le pourcentage de ressources disponibles est suffisant. Cette vérification supplémentaire s'explique par le fait que vSphere HA ne peut pas effectuer de basculement s'il n'y a qu'un seul hôte dans le cluster.

Calcul de la Capacité de basculement actuelle

Les ressources totales requises par les machines virtuelles sous tension incluent deux composants, CPU et mémoire. vSphere HA calcule ces valeurs.

- Le besoin en composant CPU est obtenu en additionnant le CPU réservé par les machines virtuelles sous tension. Si aucun CPU n'a été réservé pour une machine virtuelle, une valeur de 32 MHz est définie par défaut (cette valeur peut être modifiée par l'option avancée `das.vmcpuminmhz`).
- La taille du composant de mémoire est obtenue en additionnant la mémoire réservée (plus la capacité supplémentaire de mémoire) de chaque machine virtuelle sous tension.

Les ressources totales des hôtes disponibles pour les machines virtuelles sont calculées en additionnant les ressources de CPU et de mémoire des hôtes. Ces valeurs sont celles contenues dans le pool de ressources racine de l'hôte, et non dans les ressources physiques totales de l'hôte. Les ressources utilisées à des fins de virtualisation ne sont pas incluses. Seuls les hôtes qui sont connectés, qui ne sont pas en mode maintenance et qui ne présentent pas d'erreurs vSphere HA sont pris en compte.

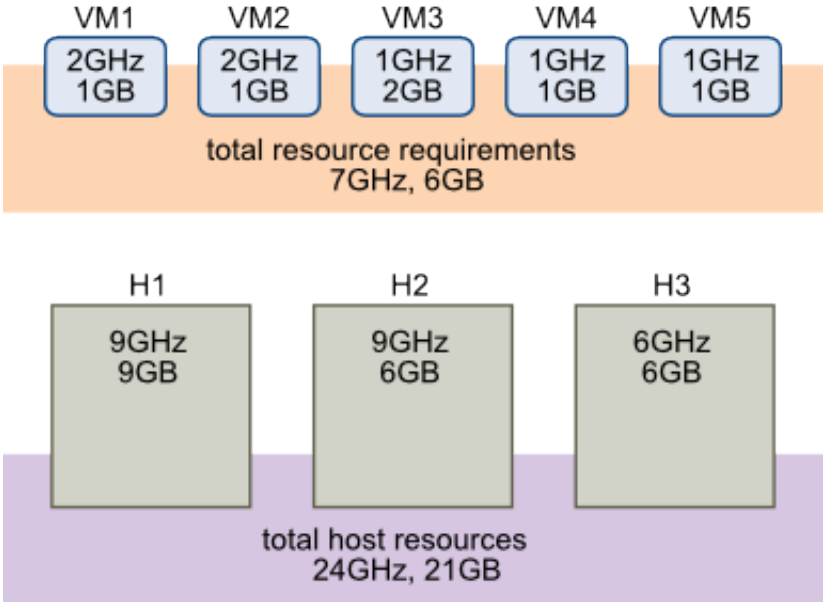
La Capacité CPU de basculement actuelle est calculée en soustrayant les besoins totaux en ressources CPU des ressources CPU totales des hôtes et en divisant le résultat par les ressources CPU totales des hôtes. La Capacité mémoire de basculement actuelle est calculée de la même manière.

Exemple : Contrôle d'admission en utilisant un pourcentage de ressources de cluster

Nous allons illustrer par un exemple le mode de calcul de la Capacité de basculement actuelle et son utilisation avec cette règle de contrôle d'admission. Prenons les hypothèses suivantes pour un cluster :

- Le cluster est composé de trois hôtes, ayant chacun des quantités différentes de CPU et de ressources mémoire disponibles. Le premier hôte (H1) a 9 Ghz de ressources CPU et 9 Go de mémoire disponibles. Le second (H2) a 9 Ghz de CPU et 6 Go de mémoire disponibles et le troisième (H3) a 6 Ghz de CPU et 6 Go de mémoire disponibles.
- Il y a cinq machines virtuelles sous tension dans le cluster avec des besoins en CPU et en mémoire différents. VM1 a besoin de 2 Ghz de ressources CPU et 1 Go de mémoire, tandis que VM2 a besoin de 2 Ghz et 1 Go, VM3 a besoin de 1 Ghz et de 2 Go, VM4 a besoin de 1 Ghz et 1 Go, VM5 a besoin de 1 Ghz et 1 Go.
- La capacité de basculement configurée pour le processeur et la mémoire est pour tous deux de 25 %.

Figure 2-1. Exemple de contrôle d'admission utilisant les règles de Pourcentage de ressources de cluster réservées



Les besoins totaux en ressources des machines virtuelles sous tension sont de 7 Ghz et 6 Go. Les ressources totales de l'hôte disponibles pour les machines virtuelles sont de 24 Ghz et 21 Go. Partant de là, la Capacité CPU de basculement actuelle s'élève à 70% ((24 Ghz - 7 Ghz)/24 Ghz). De même, la Capacité mémoire de basculement actuelle s'élève à 71% ((21 Go - -6 Go)/21 Go).

Comme la Capacité de basculement configurée pour le cluster est de 25 %, 45 % des ressources CPU totales du cluster et 46 % des ressources mémoire totales du cluster sont toujours disponibles pour les machines virtuelles supplémentaires.

Contrôle d'admission Stratégie d'emplacement

Lorsque l'option de stratégie d'emplacement est configurée, vSphere HA s'assure que même si un nombre d'hôtes spécifié est défaillant, les ressources demeurent en quantité suffisante sur le cluster pour permettre le basculement de toutes les machines virtuelles depuis ces hôtes.

Avec la stratégie d'emplacement, vSphere HA effectue le contrôle d'admission de la manière suivante :

1 Calcule la taille d'emplacement.

   Un emplacement est une représentation logique de la mémoire et des ressources CPU. Par défaut, il est dimensionné pour satisfaire aux exigences de chaque machine virtuelle sous tension dans le cluster.

2 Détermine le nombre d'emplacements pouvant se trouver sur chaque hôte du cluster.

3 Détermine la Capacité de basculement actuelle du cluster.

   Il s'agit du nombre d'hôtes défectueux permettant de conserver un nombre suffisant d'emplacements pour satisfaire toutes les machines virtuelles sous tension.

4 Détermine si la Capacité de basculement actuelle est inférieure ou non à la Capacité de basculement configurée (précisée par l'utilisateur).

   Si c'est le cas, le contrôle d'admission n'autorise pas l'opération.

Remarque Vous pouvez définir une taille d'emplacement spécifique pour les CPU et la mémoire dans la section de contrôle d'admission des paramètres vSphere HA dans vSphere Web Client

---

Calcul de la taille d'emplacement

Taille d'emplacement et contrôle d'admission de vSphere HA (http://link.brightcove.com/services/player/bcpid2296383276001?bctid=ref:video_vsphere_slot_admission_control)

La taille d'un emplacement est déterminée par deux composants, le CPU et la mémoire.

- vSphere HA calcule la taille de CPU à partir du CPU réservé par chaque machine virtuelle sous tension, en sélectionnant la valeur la plus élevée. Si aucun CPU n'a été réservé pour une machine virtuelle, une valeur de 32 MHz est définie par défaut. Cette valeur peut être modifiée par l'option avancée `das.vmcpuminmhz`.)
- vSphere HA calcule la taille de la mémoire à partir de la mémoire réservée (plus la capacité supplémentaire de mémoire) de chaque machine virtuelle sous tension, en sélectionnant la valeur la plus élevée. Il n'y a pas de valeur par défaut pour la mémoire réservée.

Si le cluster contient des machines virtuelles ayant des valeurs de réservation bien plus élevées que d'autres, celles-ci influeront sur le calcul de la taille d'emplacement. Pour éviter cela, vous pouvez préciser une limite supérieure pour le CPU ou le composant de mémoire de la taille d'emplacement en utilisant respectivement les options avancées `das.slotcpuinmhz` ou `das.slotmeminmb`. Reportez-vous à « Options avancées de vSphere HA », page 42.

Vous pouvez également déterminer le risque de fragmentation des ressources dans le cluster en regardant le nombre de machines virtuelles qui nécessitent plusieurs emplacements. Ceci peut être calculé dans la section de contrôle d'admission des paramètres vSphere HA dans vSphere Web Client. Les machines virtuelles peuvent nécessiter plusieurs emplacements si vous avez spécifié une taille fixe ou maximale d'emplacements dans les options avancées.

Utiliser les emplacements pour déterminer la capacité de basculement actuelle

Une fois la taille d'emplacement calculée, vSphere HA détermine les ressources de CPU et de mémoire disponibles sur chaque hôte pour les machines virtuelles. Ces valeurs sont celles contenues dans le pool de ressources racine de l'hôte, et non dans les ressources physiques totales de l'hôte. Vous trouverez les données sur les ressources d'un hôte utilisé par vSphere HA dans l'onglet **Résumé** de l'hôte, sur vSphere Web Client. Si tous les hôtes de votre cluster sont identiques, vous pouvez obtenir ces données en divisant les chiffres relatifs au cluster dans son ensemble par le nombre d'hôtes. Les ressources utilisées à des fins de virtualisation ne sont pas incluses. Seuls les hôtes qui sont connectés, qui ne sont pas en mode maintenance et qui ne présentent pas d'erreurs vSphere HA sont pris en compte.

Le nombre maximum d'emplacements pouvant être pris en charge par chaque hôte est alors déterminé. À cette fin, la quantité de ressources CPU de l'hôte est divisée par le composant de CPU de la taille d'emplacement et le résultat est arrondi. Le même calcul est fait pour la quantité de ressources de mémoire de l'hôte. Ces deux valeurs sont comparées et la plus basse équivaut au nombre d'emplacements pouvant être pris en charge par l'hôte.

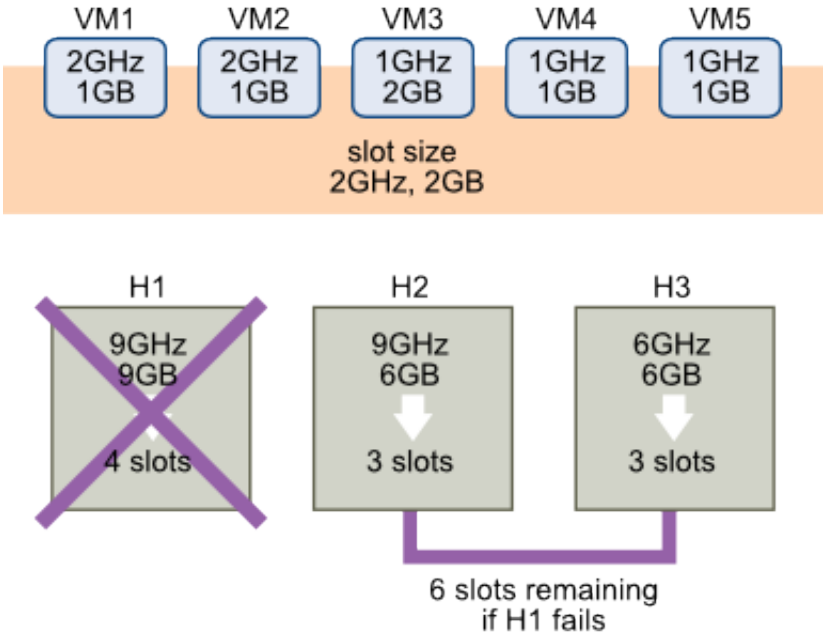
La Capacité de basculement actuelle est calculée en déterminant le nombre d'hôtes (en commençant par le plus gros) pouvant être défectueux tout en conservant un nombre suffisant d'emplacements pour satisfaire toutes les machines virtuelles sous tension.

Exemple : Contrôle d'admission en utilisant la stratégie d'emplacement

Nous allons illustrer par un exemple le mode de calcul de la taille d'emplacement et son utilisation avec cette stratégie de contrôle d'admission. Prenons les hypothèses suivantes pour un cluster :

- Le cluster est composé de trois hôtes, ayant chacun des quantités différentes de CPU et de ressources mémoire disponibles. Le premier hôte (H1) a 9 Ghz de ressources CPU et 9 Go de mémoire disponibles. Le second (H2) a 9 Ghz de CPU et 6 Go de mémoire disponibles et le troisième (H3) a 6 Ghz de CPU et 6 Go de mémoire disponibles.
- Il y a cinq machines virtuelles sous tension dans le cluster avec des besoins en CPU et en mémoire différents. VM1 a besoin de 2 Ghz de ressources CPU et 1 Go de mémoire, tandis que VM2 a besoin de 2 Ghz et 1 Go, VM3 a besoin de 1 Ghz et de 2 Go, VM4 a besoin de 1 Ghz et 1 Go, VM5 a besoin de 1 Ghz et 1 Go.
- Les défaillances d'hôte tolérées par le cluster sont définies sur la valeur 1.

Figure 2-2. Exemple de contrôle d'admission avec la stratégie Défaillances d'hôte tolérées par le cluster



1 La taille d'emplacement est calculée en comparant à la fois les exigences de CPU et de mémoire des machines virtuelles et en sélectionnant la plus élevée.

   Le besoin en CPU le plus élevé (partagé par VM1 et VM2) est de 2 GHz, tandis que le besoin en mémoire le plus élevé (VM3) est de 2 Go. Partant de là, la taille d'emplacement se compose d'un CPU de 2 GHz et d'une mémoire de 2 Go.

2 Le nombre maximum d'emplacements pouvant être pris en charge par chaque hôte est déterminé.

   H1 peut prendre en charge quatre emplacements. H2 peut prendre en charge trois emplacements (le plus bas de 9 GHz/2 GHz et 6 Go/2 Go) et H3 peut aussi en prendre en charge trois.

3 La Capacité de basculement actuelle est calculée.

   Le plus gros hôte est H1 et s'il est défectueux, le cluster contient toujours six slots, ce qui est suffisant pour les cinq machines virtuelles sous tension. Si H1 et H2 sont défectueux, il ne reste que trois emplacements, ce qui est insuffisant. Par conséquent, la Capacité de basculement actuelle est de 1.

Le cluster a un slot disponible (les six slots de H2 et H3 moins les cinq slots utilisés).

Contrôle d'admission sur des hôtes de basculement dédiés

Il est possible de configurer vSphere HA afin de désigner des hôtes spécifiques comme hôtes de basculement.

Avec le contrôle d'admission sur des hôtes de basculement dédiés, en cas de panne d'un hôte, vSphere HA tente de redémarrer ses machines virtuelles sur un des hôtes de basculement prédéfinis. Si ce n'est pas possible car les hôtes de basculement sont eux-même en panne ou leurs ressources sont insuffisantes, par exemple, vSphere HA tente de redémarrer ces machines virtuelles sur d'autres hôtes du cluster.

Pour que des capacités restent disponibles sur un hôte de basculement, vous ne pouvez pas mettre sous tension des machines virtuelles ni utiliser vMotion pour faire migrer des machines virtuelles vers un hôte de basculement. De plus, DRS n'utilise pas d'hôte de basculement pour la répartition de la charge.

Remarque Si vous utilisez le contrôle d'admission sur des hôtes de basculement dédiés et désignez plusieurs hôtes de basculement, DRS ne cherche pas à faire respecter les règles d'affinité VM-VM pour les machines virtuelles qui s'exécutent sur des hôtes de basculement.

Interopérabilité de vSphere HA

vSphere HA peut interagir avec de nombreuses autres fonctionnalités, comme DRS et Virtual SAN.

Avant de configurer vSphere HA, vous devez connaître les limitations de son interopérabilité avec ces autres fonctionnalités ou produits.

Utilisation de vSphere HA avec Virtual SAN

Vous pouvez utiliser Virtual SAN comme stockage partagé pour un cluster vSphere HA. Lorsqu'il est activé, Virtual SAN regroupe les disques de stockage locaux spécifiés qui sont disponibles sur les hôtes dans une banque de données unique partagée par tous les hôtes.

Avant d'utiliser vSphere HA avec Virtual SAN, vous devez connaître les exigences et les limitations liées à l'interopérabilité de ces deux fonctions.

Pour plus d'informations sur Virtual SAN, reportez-vous à la section *Administration de VMware Virtual SAN*.

Remarque Vous pouvez utiliser vSphere HA avec des clusters étendus Virtual SAN.

Conditions requises pour les hôtes ESXI

Pour utiliser Virtual SAN avec un cluster vSphere HA, les conditions suivantes doivent être remplies :

- Tous les hôtes ESXi du cluster doivent être de la version 5.5 ou ultérieure.
- Le cluster doit avoir au moins trois hôtes ESXi.

Différences de mise en réseau

Virtual SAN dispose de son propre réseau. Lorsque Virtual SAN et vSphere HA sont activés sur le même cluster, le trafic entre agents HA circule sur ce réseau de stockage et non pas sur le réseau de gestion. vSphere HA utilise le réseau de gestion uniquement si Virtual SAN est désactivé. Si vSphere HA est configuré sur un hôte, vCenter Server choisit le réseau approprié.

Remarque Vous ne pouvez activer Virtual SAN que si vSphere HA est désactivé.

Si vous modifiez la configuration de Virtual SAN, les agents vSphere HA ne choisissent pas automatiquement les nouveaux paramètres réseau. Pour modifier le réseau Virtual SAN, vous devez effectuer la procédure suivante dans vSphere Web Client :

1 Désactivez la surveillance de l'hôte pour le cluster vSphere HA.

2 Modifiez Virtual SAN.

3 Cliquez avec le bouton droit sur chacun des hôtes du cluster et sélectionnez **Reconfigurer pour vSphere HA**.

4 Réactivez la surveillance de l'hôte pour le cluster vSphere HA.

Tableau 2-2 montre les différences dans la mise en réseau vSphere HA, que Virtual SAN soit utilisé ou non.

Tableau 2-2. Différences de mise en réseau de vSphere HA

	Virtual SAN activé	Virtual SAN désactivé
Réseau utilisé par vSphere HA	Réseau de stockage de Virtual SAN	Réseau de gestion
Banques de données de signaux de pulsation	Toutes les banques de données montées sur plusieurs hôtes, sauf les banques de données Virtual SAN	Toutes les banques de données montées sur plusieurs hôtes
Hôte déclaré comme isolé	Adresses d'isolation ne répondant pas aux commandes ping et réseau de stockage de Virtual SAN inaccessible.	Adresses d'isolation ne répondant pas aux commandes ping et réseau de gestion inaccessible.

Paramètres de réservation de capacité

Lorsque vous réservez de la capacité pour votre cluster vSphere HA en utilisant une stratégie de contrôle d'admission, ce paramètre doit être cohérent avec le paramètre de Virtual SAN correspondant qui permet d'assurer l'accessibilité des données en cas de panne. Plus précisément, la valeur du paramètre définissant le nombre de pannes toléré dans l'ensemble des règles de Virtual SAN ne doit pas être inférieure à la capacité réservée par le paramètre de contrôle d'admission de vSphere HA.

Par exemple, si l'ensemble de règles de Virtual SAN n'autorise que deux pannes, la stratégie du contrôle d'admission de vSphere HA doit réserver une capacité équivalente à seulement une ou deux pannes d'hôte. Si vous utilisez la stratégie du pourcentage de ressources de cluster réservées sur un cluster disposant de huit hôtes, vous ne devez pas réserver plus de 25 % des ressources du cluster. Si vous utilisez la stratégie des pannes d'hôtes tolérées par le cluster sur ce même cluster, la valeur du paramètre ne doit pas dépasser deux hôtes. Si vSphere HA réserve moins de capacité, l'activité de basculement peut s'avérer imprévisible. La réservation d'une capacité trop grande impose une contrainte excessive à l'activation des machines virtuelles et aux migrations vSphere vMotion entre clusters.

Utilisation conjointe de vSphere HA et DRS

L'utilisation de vSphere HA avec Distributed Resource Scheduler (DRS) allie le basculement automatique à l'équilibrage de la charge. Cette association peut aboutir à un cluster mieux équilibré une fois que vSphere HA a déplacé les machines virtuelles sur d'autres hôtes.

Quand vSphere HA exécute le basculement et redémarre les machines virtuelles sur des hôtes différents, sa première priorité est la disponibilité immédiate de toutes les machines virtuelles. Après le redémarrage des VM, les hôtes sur lesquels elles sont mises sous tension peuvent se retrouver surchargés, tandis que la charge d'autres hôtes est, en comparaison, plus légère. vSphere HA utilise le CPU et la réservation de mémoire de la VM pour déterminer si un hôte dispose de suffisamment de capacité disponible pour prendre en charge la VM.

Dans un cluster utilisant DRS et vSphere HA avec le contrôle d'admission activé, les machines virtuelles ne sont pas nécessairement évacuées des hôtes passant en mode maintenance. Ce comportement intervient par suite des ressources réservées pour le redémarrage des machines virtuelles en cas de panne. Il faut migrer manuellement les machines virtuelles en dehors des hôtes avec vMotion.

Dans certains cas, vSphere HA ne parvient pas à basculer les machines virtuelles en raison de contraintes de ressources. Ceci peut se produire pour plusieurs raisons.

- Le contrôle d'admission HA est désactivé et Gestion de l'alimentation distribuée (DPM) est activé. Cela peut aboutir à la consolidation par DPM des machines virtuelles sur un nombre inférieur d'hôtes et à la mise en veille des hôtes vides, ce qui ne laisse pas suffisamment de réserve de capacité active pour effectuer un basculement.
- Les règles (requises) d'affinité de machine virtuelle/hôte peuvent limiter les hôtes sur lesquels certaines machines virtuelles peuvent être placées.
- Il peut y avoir suffisamment de ressources cumulées mais celles-ci sont fragmentées sur plusieurs hôtes de sorte qu'elles ne peuvent pas être utilisées par les machines virtuelles pour le basculement.

Dans ces cas-là, vSphere HA peut utiliser DRS pour essayer d'ajuster le cluster (par exemple, en sortant les hôtes du mode veille ou en migrant les machines virtuelles pour défragmenter les ressources du cluster) de sorte que HA puisse exécuter les basculements.

Si DPM est en mode manuel, vous devrez éventuellement confirmer les recommandations de mise sous tension des hôtes. De même, si DPM est en mode manuel, vous devrez éventuellement confirmer les recommandations de migration.

Si vous utilisez les règles d'affinité entre VM et hôte requises, sachez que ces règles doivent obligatoirement être respectées. vSphere HA n'effectue pas de basculement si cela risque d'enfreindre une règle.

Pour plus d'informations sur DRS, consultez la documentation *Gestion des ressources vSphere*.

Règles d'affinités de vSphere HA et DRS

Si vous créez une règle d'affinité DRS pour votre cluster, vous pouvez indiquer de quelle manière vSphere HA doit appliquer cette règle en cas de basculement d'une machine virtuelle.

Les deux types de règles pour lesquelles vous pouvez le comportement de vSphere HA en cas de basculement sont les suivants :

- Les règles d'anti-affinité de machine virtuelle contraignent les machines virtuelles spécifiées à rester séparées pendant les opérations de basculement.
- Les règles d'affinité machine virtuelle/hôte placent les machines virtuelles spécifiées sur un hôte particulier ou un membre d'un groupe d'hôtes défini pendant les opérations de basculement.

Lorsque vous modifiez une règle d'affinité DRS, cochez la ou les cases appliquant le comportement de basculement souhaité pour vSphere HA.

- **HA doit respecter les règles d'anti-affinité VM pendant le basculement** : si les machines virtuelles avec cette règle doivent être placées ensemble, le basculement est abandonné.
- **HA devrait respecter les règles d'anti-affinité VM pendant le basculement** : vSphere HA tente de placer les machines virtuelles soumises à cette règle sur les hôtes spécifiés le cas échéant.

Remarque vSphere HA peut redémarrer une machine virtuelle dans un cluster sur lequel DRS est désactivé, en remplaçant un mappage de règles d'affinité machine virtuelle/hôte si l'échec de l'hôte a lieu rapidement (par défaut en moins de 5 minutes) après avoir défini la règle.

Autres problèmes d'interopérabilité de vSphere HA

Pour utiliser vSphere HA, vous devez connaître les problèmes d'interopérabilité supplémentaires suivants.

VM Component Protection

VM Component Protection (VMCP) connaît les problèmes et limitations de l'interopérabilité suivants :

- VMCP ne prend pas en charge vSphere Fault Tolerance. Si VMCP est activé pour cluster utilisant Fault Tolerance, les machines virtuelles FT affectées recevront automatiquement des remplacements qui désactivent VMCP.
- VMCP ne détecte pas ni ne réagit aux problèmes d'accessibilité des fichiers situés sur les banques de données Virtual SAN. Si la configuration et les fichiers VMDK d'une machine virtuelle sont situés uniquement sur des banques de données Virtual SAN, ils ne sont pas protégés par VMCP.
- VMCP ne détecte pas ni ne réagit aux problèmes d'accessibilité des fichiers situés sur les banques de données de volume virtuel. Si la configuration et les fichiers VMDK d'une machine virtuelle sont situés uniquement sur des banques de données de volume virtuel, ils ne sont pas protégés par VMCP.
- VMCP ne protège pas contre le mappage de périphérique brut (Raw Device Mapping, RDM) inaccessible.

IPv6

vSphere HA peut être utilisé avec des configurations réseau IPv6, qui sont entièrement pris en charge si les considérations suivantes sont prises en compte :

- Le cluster contient uniquement des hôtes ESXi 6.0 ou version ultérieure.
- Le réseau de gestion de tous les hôtes dans le cluster doit être configuré avec la même version d'adresse IP, IPv6 ou IPv4. Les clusters vSphere HA ne peuvent pas contenir les deux types de configuration de la mise en réseau.
- Les adresses d'isolation réseau utilisées par vSphere HA doivent correspondre à la version de l'adresse IP utilisée par le cluster pour son réseau de gestion.
- IPv6 ne peut pas être utilisé dans les clusters vSphere HA qui utilisent également Virtual SAN.

En plus des restrictions précédentes, les types suivants d'adresses IPv6 ne sont pas pris en charge pour être utilisés avec l'adresse d'isolation ou le réseau de gestion vSphere HA : adresse de lien local, ORCHID et adresse de lien local avec indices de zone. De plus, le type d'adresse loopback ne peut pas être utilisé pour le réseau de gestion.

Remarque Pour mettre à jour le déploiement de l'IPv4 vers l'IPv6, vous devez d'abord désactiver vSphere HA.

Création d'un cluster vSphere HA

vSphere HA fonctionne dans le cadre d'un cluster d'hôtes ESXi (ou ESX hérités). Vous devez créer un cluster, le remplir d'hôtes et configurer les paramètres vSphere HA pour que la protection du basculement puisse être établie.

Lorsque vous créez un cluster vSphere HA, vous devez configurer divers paramètres qui déterminent le mode de fonctionnement de la fonction. Avant de commencer, identifiez les nœuds du cluster. Ces nœuds sont les hôtes ESXi qui fourniront les ressources pour la prise en charge des machines virtuelles et qui seront utilisés par vSphere HA pour la protection du basculement. Déterminez ensuite la manière dont ces nœuds doivent être reliés les uns aux autres et au stockage partagé où résident les données de la machine virtuelle. Lorsque l'architecture de mise en réseau est en place, vous pouvez ajouter les hôtes au cluster et terminer la configuration de vSphere HA.

Vous pouvez activer et configurer vSphere HA avant d'ajouter des nœuds d'hôtes au cluster. Toutefois, tant que les hôtes n'ont pas été ajoutés, le cluster n'est pas entièrement opérationnel et quelques paramètres du cluster ne sont pas disponibles. Par exemple, les règles de contrôle d'admission Spécifier un hôte de basculement ne sont pas disponibles tant qu'un hôte n'a pas été défini comme hôte de basculement.

Remarque La fonction de démarrage et d'arrêt de la machine virtuelle (démarrage automatique) est désactivée pour toutes les machines virtuelles résidant sur des hôtes qui se trouvent dans un cluster vSphere HA (ou qui y ont été déplacées). Le démarrage automatique n'est pas pris en charge avec vSphere HA.

Liste de contrôle de vSphere HA

La liste de contrôle de vSphere HA contient les conditions requises que vous devez connaître pour pouvoir créer et utiliser un cluster vSphere HA.

Consultez cette liste avant de configurer un cluster vSphere HA. Pour plus d'informations, suivez les références croisées appropriées.

- Tous les hôtes doivent disposer d'une licence pour vSphere HA.
- Un cluster doit contenir au moins deux hôtes.

- Tous les hôtes doivent être configurés avec des adresses IP statiques. Si vous utilisez DHCP, vérifiez que l'adresse de chaque hôte est conservée après les redémarrages.
- Tous les hôtes doivent avoir au moins un réseau de gestion en commun. Il est recommandé d'avoir au moins deux réseaux de gestion en commun. Vous devez utiliser le réseau VMkernel avec la case **Trafic de gestion** cochée. Les réseaux doivent être accessibles l'un à l'autre et vCenter Server et les hôtes doivent être accessibles les uns aux autres sur les réseaux de gestion. Reportez-vous à « Meilleures pratiques pour la mise en réseau », page 45.
- Pour vous assurer que toutes les machines virtuelles peuvent s'exécuter sur n'importe quel hôte du cluster, tous les hôtes doivent avoir accès aux même réseaux et banques de données de machines virtuelles. De même, les machines virtuelles doivent se trouver sur des stockages partagés, et non locaux, sinon il ne peut pas y avoir de basculement en cas de défaillance de l'hôte.

  Remarque vSphere HA utilise le signal de pulsation de banque de données pour différencier les hôtes partitionnés, isolés ou défaillants. Par conséquent, s'il y a des banques de données plus fiables dans votre environnement, configurez vSphere HA pour leur donner la préférence.

- Le fonctionnement de surveillance des machines virtuelles nécessite l'installation de VMware tools. Reportez-vous à « Surveillance des VM et applications », page 19.
- vSphere HA prend en charge IPv4 et IPv6. Voir « Autres problèmes d'interopérabilité de vSphere HA », page 32 pour consulter les considérations à prendre en compte lors de l'utilisation d'IPv6.
- Pour que VM Component Protection fonctionne, la fonctionnalité de délai d'expiration Tous les chemins hors service (All Paths Down, APD) doit être activée.
- Pour utiliser VM Component Protection, les clusters doivent comporter des hôtes ESXi 6.0 hosts ou version ultérieure.
- Seuls les clusters vSphere HA contenant des hôtes ESXi 6.0 ou version ultérieure peuvent être utilisés pour activer VMCP. Les clusters contenant des hôtes d'une version antérieure ne peuvent pas activer VMCP et ne peuvent pas être ajoutés à un cluster sur lequel VMCP est activé.
- Si votre cluster utilise des banques de données de volume virtuel, lorsque vSphere HA est activé, un configuration de volume virtuel est créée sur chaque banque de données par vCenter Server. Dans ces conteneurs, vSphere HA stocke les fichiers qu'il utilise pour protéger les machines virtuelles. vSphere HA ne fonctionne pas correctement si vous supprimez ces conteneurs. Un seul conteneur est créé par banque de données de volume virtuel.

Créer un cluster vSphere HA

Pour activer votre cluster pour vSphere HA, vous devez d'abord créer un cluster vide. Après avoir planifié les ressources et l'architecture de réseau de votre cluster, utiliser vSphere Web Client pour ajouter des hôtes au cluster et spécifier les paramètres du cluster vSphere HA.

Un cluster doit obligatoirement être compatible avec vSphere HA pour que vSphere Fault Tolerance fonctionne.

Prérequis

- Vérifiez que toutes les machines virtuelles et leurs fichiers de configuration résident sur des stockages partagés.
- Vérifiez que les hôtes sont configurés pour accéder au stockage partagé, afin de pouvoir mettre sous tension les machines virtuelles à l'aide des différents hôtes dans le cluster.
- Vérifiez que les hôtes sont configurés pour avoir accès au réseau de machines virtuelles.
- Vérifiez que vous utilisez des connexions réseau de gestion redondant pour vSphere HA. Pour plus d'informations sur la configuration d'un réseau redondant, consultez la rubrique« Meilleures pratiques pour la mise en réseau », page 45.

- Vérifiez que vous avez configuré les hôtes avec au moins deux banques de données afin de fournir de la redondance au signal de pulsation de la banque de données vSphere HA.
- Connectez vSphere Web Client à vCenter Server en utilisant un compte disposant des autorisations d'administrateur de cluster.

Procédure

1 Dans vSphere Web Client, accédez au centre de données où vous voulez que le cluster réside et cliquez sur **Créer un cluster**.

2 Complétez le paramètre de l'assistant Nouveau cluster.

Ne pas mettre sous tension vSphere HA (ou DRS).

3 Cliquez sur **OK** pour fermer l'assistant et créer un cluster vide.

4 Sur la base de votre plan pour les ressources et l'architecture de réseau du cluster, utiliser le vSphere Web Client pour ajouter des hôtes au cluster.

5 Accédez au cluster et activez vSphere HA.

a Cliquez sur l'onglet **Configurer**.

b Sélectionnez **Disponibilité vSphere** et cliquez sur **Modifier**.

c Sélectionnez **Activer vSphere HA**.

d Sélectionnez **Activer Proactive HA** pour autoriser les migrations proactives de machines virtuelles depuis les hôtes sur lesquels un fournisseur a signalé une dégradation de santé.

6 Sous **Pannes et réponses**, sélectionnez **Activer la surveillance d'hôte**

Lorsque l'option de surveillance d'hôte est activée, les hôtes du cluster peuvent échanger des signaux de pulsation réseau et vSphere HA peut agir lorsqu'il détecte des pannes. La surveillance d'hôte est aussi requise pour le bon fonctionnement du processus de récupération de vSphere Fault Tolerance.

7 Sélectionnez un paramètre de **Surveillance de VM**.

Sélectionnez **Surveillance de VM seulement** pour redémarrer des machines virtuelles individuelles si leurs signaux de pulsation ne sont pas reçus dans un délai déterminé. Vous pouvez également sélectionner **Surveillance de VM et d'application** pour activer la surveillance des applications.

8 Cliquez sur **OK**.

Vous disposez désormais d'un cluster vSphere HA rempli d'hôtes.

Suivant

Configurez les paramètres vSphere HA appropriés pour votre cluster.

- Pannes et réponses
- Pannes et réponses de Proactive HA
- Contrôle d'admission
- banques de données de signaux de pulsation
- Options avancées

Reportez-vous à « Configuration des paramètres de disponibilité vSphere », page 36.

Configuration des paramètres de disponibilité vSphere

Lorsque vous créez un cluster vSphere HA ou que vous configurez un cluster existant, vous devez configurer les paramètres qui déterminent le mode de fonctionnement de la fonction.

Dans vSphere Web Client vous pouvez configurer les paramètres vSphere HA suivants :

Pannes et réponses	Fournissez ici les paramètres de réponses aux pannes d'hôte, d'isolation des hôtes, de surveillance des VM et de protection des composants des machines virtuelles.
Pannes et réponses de Proactive HA	Fournissez des paramètres précisant comment Proactive HA répond lorsqu'un fournisseur a notifié sa dégradation de santé à vCenter, indiquant une panne partielle de cet hôte.
Contrôle d'admission	Activez ou désactivez le contrôle d'admission pour le cluster vSphere HA et choisissez une règle pour déterminer son application.
banques de données de signaux de pulsation	Indiquez vos préférences pour les banques de données que vSphere HA utilise pour le signal de pulsation des banques de données.
Options avancées	Personnalisez le comportement de vSphere HA en définissant les options avancées.

Configuration des réponses aux pannes

Le volet **Panne et réponses** des paramètres de vSphere HA vous permet de configurer le fonctionnement du cluster lorsque des problèmes se produisent.

Dans cette partie de vSphere Web Client, vous pouvez déterminer les réponses spécifiques du cluster vSphere HA en cas de pannes ou d'isolation d'un hôte. Vous pouvez également configurer les actions de VM Component Protection (VMCP) lorsque des situations de type PDL (perte de périphérique permanente) et APD (Tous chemins hors service) se produisent et vous pouvez activer la surveillance de VM.

Les tâches suivantes sont disponibles :

1 Répondre en cas de panne d'hôte page 36

Vous pouvez définir des réponses spécifiques en cas de pannes d'un hôte dans votre cluster vSphere HA.

2 Réponse en cas d'isolation d'hôte page 37

Vous pouvez définir des réponses spécifiques en cas d'isolation d'hôte dans votre cluster vSphere HA.

3 Configurer les réponses de VMCP page 37

Configurez la réponse de VMCP (VM Component Protection) en cas de défaillance de banque de données avec PDL ou APD.

4 Activer la surveillance de VM page 38

Vous pouvez activer la surveillance des VM et des applications, et également définir la sensibilité de surveillance de votre cluster vSphere HA.

Répondre en cas de panne d'hôte

Vous pouvez définir des réponses spécifiques en cas de pannes d'un hôte dans votre cluster vSphere HA.

Cette page est modifiable uniquement si vous avez activé vSphere HA.

Procédure

1 Dans vSphere Web Client, accédez au cluster vSphere HA .

2 Cliquez sur l'onglet **Configurer**.

3 Sélectionnez **Disponibilité vSphere** et cliquez sur **Modifier**.

4 Cliquez sur **Pannes et réponses** et développez l'option **Réponse en cas de panne de l'hôte**

5 Sélectionnez une des options de configuration suivantes.

Option	Description
Réponse en cas de panne	Si vous sélectionnez **Désactivé**, ce paramètre désactive la surveillance et les machines virtuelles ne sont pas redémarrées en cas de panne d'un hôte. Si l'option **Redémarrer les machines virtuelles** est sélectionnée, en cas de panne d'un hôte, les VM sont basculées en fonction de leur priorité de redémarrage.
Priorité de redémarrage des VM par défaut	La priorité de redémarrage détermine l'ordre de redémarrage des machines virtuelles en cas d'échec de l'hôte. Les machines virtuelles de plus haute priorité sont démarrées en premier. Si plusieurs hôtes échouent, toutes les machines virtuelles sont migrées du premier hôte par ordre de priorité, puis toutes les machines virtuelles du deuxième hôte par ordre de priorité, et ainsi de suite.
Condition de redémarrage des dépendances de VM	Une condition spécifique doit être sélectionnée ainsi que le délai après que cette condition a été remplie, avant que vSphere HA soit autorisé à passer à la priorité de redémarrage de la VM suivante.

6 Cliquez sur **OK**.

Vos paramètres de réponse en cas de panne d'hôte sont appliqués.

Réponse en cas d'isolation d'hôte

Vous pouvez définir des réponses spécifiques en cas d'isolation d'hôte dans votre cluster vSphere HA.

Cette page est modifiable uniquement si vous avez activé vSphere HA.

Procédure

1 Dans vSphere Web Client, accédez au cluster vSphere HA .

2 Cliquez sur l'onglet **Configurer**.

3 Sélectionnez **Disponibilité vSphere** et cliquez sur **Modifier**.

4 Cliquez sur **Pannes et réponses** et développez l'option **Réponse en cas d'isolation d'hôte**.

5 Pour configurer la réponse en cas d'isolation d'hôte, sélectionnez **Désactivé**, **Arrêter et redémarrer les machines virtuelles** ou **Mettre hors tension et redémarrer les VM**.

6 Cliquez sur **OK**.

Votre paramètre de réponse en cas d'isolation d'hôte est appliqué.

Configurer les réponses de VMCP

Configurez la réponse de VMCP (VM Component Protection) en cas de défaillance de banque de données avec PDL ou APD.

Cette page est modifiable uniquement si vous avez activé vSphere HA.

Procédure

1 Dans vSphere Web Client, accédez au cluster vSphere HA .

2 Cliquez sur l'onglet **Configurer**.

3 Sélectionnez **Disponibilité vSphere** et cliquez sur **Modifier**.

4 Cliquez sur **Pannes et réponses** et développez l'option **Banque de données avec PDL** ou **Banque de données avec APD** .

5 Si vous avez cliqué sur **Banque de données avec PDL**, vous pouvez définir la réponse de VMCP pour ce type de problème : **Désactivé**, **Émission d'événements** ou **Mettre hors tension et redémarrer les VM**.

6 Si vous avez cliqué sur **Banque de données avec APD**, vous pouvez définir la réponse de VMCP pour ce type de problème : **Désactivé**, **Émission d'événements**, **Mettre hors tension et redémarrer les VM : stratégie de redémarrage modérée** ou **Mettre hors tension et redémarrer les VM : stratégie de redémarrage agressive**. Vous pouvez également définir l'option **Récupération de réponse**, qui est le nombre de minutes pendant lesquelles VMCP attend avant d'exécuter une action.

7 Cliquez sur **OK**.

Vos paramètres de réponse aux défaillances de VMCP sont appliqués.

Activer la surveillance de VM

Vous pouvez activer la surveillance des VM et des applications, et également définir la sensibilité de surveillance de votre cluster vSphere HA.

Cette page est modifiable uniquement si vous avez activé vSphere HA.

Procédure

1 Dans vSphere Web Client, accédez au cluster vSphere HA .

2 Cliquez sur l'onglet **Configurer**.

3 Sélectionnez **Disponibilité vSphere** et cliquez sur **Modifier**.

4 Cliquez sur **Pannes et réponses** et développez l'option **Surveillance de VM**.

5 Sélectionnez **Surveillance de VM** puis **Surveillance d'application**.

Ces paramètres activent les signaux de pulsation de VMware Tools et des applications, respectivement.

6 Pour définir la sensibilité de surveillance des signaux de pulsation, déplacez le curseur entre **Basse** et **Élevée** ou sélectionnez **Personnalisée** pour fournir des paramètres personnalisés.

7 Cliquez sur **OK**.

Vos paramètres de surveillance sont appliqués.

Configurer Proactive HA

Vous pouvez configurer la manière dont Proactive HA répond lorsqu'un fournisseur a signalé la dégradation de sa santé à vCenter, ce qui est le signe d'une panne partielle de cet hôte.

Cette page est modifiable uniquement si vous avez activé vSphere DRS.

Procédure

1 Dans vSphere Web Client, accédez au cluster Proactive HA.

2 Cliquez sur l'onglet **Configurer**.

3 Sélectionnez **Disponibilité vSphere** et cliquez sur **Modifier**.

4 Sélectionnez **Activer Proactive HA**.

5 Cliquez sur **Pannes et réponses de Proactive HA**

6 Sélectionnez une des options de configuration suivantes.

Option	Description
Niveau d'automatisation	Déterminez si le mode de quarantaine ou de maintenance des hôtes et les migrations de VM sont des recommandations ou des réponses automatiques. ■ **Manuel**. vCenter Server suggérera des recommandations de migration pour les machines virtuelles. ■ **Automatisé**. Les machines virtuelles seront migrées vers des hôtes sains et les hôtes dégradés seront mis en quarantaine ou en mode de maintenance selon la configuration du niveau d'automatisation de Proactive HA.
Correction	Déterminez ce qui se produit pour les hôtes partiellement dégradés. ■ **Mode Quarantaine pour toutes les pannes**. Maintient un équilibre entre performances et disponibilité, en évitant l'utilisation d'hôtes partiellement dégradés tant que les performances des machines virtuelles ne sont pas affectées. ■ **Mode de quarantaine pour les pannes modérées et mode de maintenance pour les pannes graves (mixte)**. Maintient un équilibre entre performances et disponibilité, en évitant l'utilisation d'hôtes modérément dégradés tant que les performances des machines virtuelles ne sont pas affectées. Garantit que les machines virtuelles ne s'exécutent pas sur des hôtes subissant une grave panne. ■ **Mode Maintenance pour toutes les pannes**. Garantit que les machines virtuelles ne s'exécutent pas sur des hôtes subissant une panne partielle. Les privilèges `Host.Config.Quarantine` et `Host.Config.Maintenance` doivent mettre les hôtes respectivement en mode de quarantaine et en mode de maintenance.

Cochez les cases pour activer les fournisseurs Proactive HA pour ce cluster. Les fournisseurs s'affichent ci-dessous lorsque leur plug-in vSphere Web Client correspondant a été installé et que les fournisseurs surveillent chaque hôte du cluster. Cliquez sur le lien de modification pour afficher/modifier les conditions de panne prises en charge par le fournisseur.

7 Cliquez sur **OK**.

Vos paramètres pour la réponse de Proactive HA prennent effet.

Configurer le contrôle d'admission

Après avoir créé un cluster, vous pouvez configurer le contrôle d'admission afin de spécifier si les machines virtuelles peuvent être démarrées si elles ne respectent pas les contraintes de disponibilité. Le cluster réserve des ressources pour permettre le basculement de toutes les machines virtuelles en cours d'exécution sur le nombre d'hôtes spécifié.

La page Contrôle admission apparaît uniquement si vous avez activé vSphere HA.

Procédure

1 Dans vSphere Web Client, accédez au cluster vSphere HA .

2 Cliquez sur l'onglet **Configurer**.

3 Sélectionnez **Disponibilité vSphere** et cliquez sur **Modifier**.

4 Cliquez sur **Contrôle d'admission** pour afficher les options de configuration.

5 Sélectionnez un nombre dans **Pannes de l'hôte tolérées par le cluster**. Il s'agit du nombre maximal de pannes de l'hôte dont le cluster peut récupérer ou pour lesquelles il peut garantir le basculement.

6 Sélectionnez une option pour **Définir la capacité de basculement de l'hôte par**.

Option	Description
Pourcentage de ressources du cluster	Spécifiez un pourcentage des ressources CPU et de mémoire du cluster à réserver comme capacité disponible pour prendre en charge les basculements.
Stratégie d'emplacement (VM sous tension)	Sélectionnez une stratégie de taille d'emplacement qui couvre toutes les machines virtuelles sous tension ou qui correspond à une taille fixe. Vous pouvez également calculer le nombre de machines virtuelles qui ont besoin d'emplacements multiples.
Hôtes de basculement dédiés	Sélectionnez les hôtes à utiliser pour les actions de basculement. Les basculements peuvent toujours se produire sur d'autres hôtes du cluster si l'hôte de basculement par défaut ne dispose pas des ressources suffisantes.
Désactivé	Sélectionnez cette option pour désactiver le contrôle d'admission et autoriser la mise sous tension des machines virtuelles qui ne respectent pas les contraintes d'admission.

7 Définissez le pourcentage pour **Dégradation des performances tolérées par les VM**.

Ce paramètre détermine quel pourcentage de dégradation des performances les machines virtuelles du cluster sont autorisées à tolérer lors d'une panne.

8 Cliquez sur **OK**.

Vos paramètres de contrôle d'admission sont appliqués.

Configurer les banques de données de signal de pulsation

vSphere HA utilise le signal de pulsation de banque de données pour identifier les hôtes défaillants et les hôtes qui résident dans une partition réseau. Avec le signal de pulsation des banques de données, vSphere HA peut surveiller les hôtes en cas de partitionnement du réseau de gestion et continuer à répondre aux pannes.

Vous pouvez spécifier les banques de données que vous voulez utiliser pour le signal de pulsation des banques de données.

Procédure

1 Dans vSphere Web Client, accédez au cluster vSphere HA .

2 Cliquez sur l'onglet **Configurer**.

3 Sélectionnez **Disponibilité vSphere** et cliquez sur **Modifier**.

4 Cliquez sur **Banques de données de signal de pulsation** pour afficher les options de configuration de signal de pulsation des banques de données.

5 Pour indiquer à vSphere HA comment sélectionner les banques de données et comment traiter vos préférences, sélectionnez une des options suivantes.

Tableau 2-3.

Options de signal de pulsation de banque de données
Sélectionner automatiquement les banques de données accessibles depuis l'hôte
Utiliser uniquement les banques de données de la liste spécifiée
Utiliser les banques de données de la liste spécifiée, puis compléter automatiquement si nécessaire

6 Dans le volet Banques de données des signaux de pulsation disponibles, sélectionnez les banques de données que vous souhaitez utiliser pour le signal de pulsation.

   Les banques de données répertoriées sont partagées par plusieurs hôtes du cluster vSphere HA. Lorsque vous sélectionnez une banque de données, le volet inférieur affiche tous les hôtes du cluster vSphere HA qui peuvent y accéder.

7 Cliquez sur **OK**.

Définir les options avancées

Pour personnaliser le comportement de vSphere HA, définissez les options avancées de vSphere HA.

Prérequis

Vérifiez que vous possédez des privilèges d'administrateur sur les clusters.

Remarque Ces options affectent le fonctionnement de vSphere HA. Modifiez-les donc avec prudence.

Procédure

1 Dans vSphere Web Client, accédez au cluster vSphere HA .

2 Cliquez sur l'onglet **Configurer**.

3 Sélectionnez **Disponibilité vSphere** et cliquez sur **Modifier**.

4 Cliquez sur **Options avancées**.

5 Cliquez sur **Ajouter** et tapez le nom de l'option avancée dans la zone de texte.

   Vous pouvez définir la valeur de l'option dans la zone de texte dans la colonne Valeur.

6 Répétez l'étape 5 pour chaque nouvelle option que vous souhaitez ajouter et cliquez sur **OK**.

Le cluster utilise les options que vous avez ajoutées ou modifiées.

Suivant

Après avoir défini une option avancée vSphere HA, elle est conservée jusqu'à ce que vous procédiez à ce qui suit :

- À l'aide de vSphere Web Client, réinitialisez sa valeur à la valeur par défaut.
- Modifiez ou supprimez manuellement l'option depuis le fichier fdm.cfg sur tous les hôtes du cluster.

Options avancées de vSphere HA

Vous pouvez définir des options avancées qui affectent le comportement du cluster vSphere HA.

Tableau 2-4. Options avancées de vSphere HA

Option	Description
`das.isolationaddress[...]`	définit l'adresse pour exécuter un ping afin de déterminer si un hôte est isolé du réseau. Le ping est uniquement envoyé à cette adresse lorsqu'aucun autre hôte du cluster ne reçoit de signaux de pulsation. En l'absence de précision, la passerelle par défaut du réseau de gestion est utilisée. Cette passerelle par défaut doit être une adresse fiable et disponible, de sorte que l'hôte puisse déterminer s'il est isolé du réseau. Vous pouvez indiquer plusieurs adresses d'isolation (jusqu'à 10) pour le cluster : `das.isolationaddressX`, où X = 0-9. Vous devez généralement en indiquer une par réseau de gestion. L'indication d'un nombre excessif d'adresses ralentit la détection de l'isolement.
`das.usedefaultisolationaddress`	Par défaut, vSphere HA utilise la passerelle par défaut du réseau de console comme adresse d'isolement. Cette option indique l'utilisation ou non de ce paramètre par défaut (vrai|faux).
`das.isolationshutdowntimeout`	Période pendant laquelle le système attend que la machine virtuelle s'arrête avant de la mettre hors tension. Cela s'applique uniquement si la réponse à l'isolement de l'hôte est Arrêter la machine virtuelle. La valeur par défaut est de 300 secondes.
`das.slotmeminmb`	Définit la limite maximum de la taille d'un emplacement de mémoire. Si cette option est utilisée, la taille d'emplacement est la plus petite de cette valeur ou la réserve de mémoire maximale plus la capacité supplémentaire de n'importe quelle machine virtuelle sous tension dans le cluster.
`das.slotcpuinmhz`	Définit la limite maximale de la taille d'un emplacement de CPU. Si cette option est utilisée, la taille d'emplacement est la plus petite de cette valeur ou la réserve de CPU maximale de n'importe quelle machine virtuelle sous tension dans le cluster.
`das.vmmemoryminmb`	Définit la valeur de ressources de mémoire par défaut associée à une machine virtuelle si sa réserve de mémoire n'est pas précisée ou nulle. Celle-ci est utilisée pour la stratégie de contrôle d'admission Défaillances d'hôte tolérées par le cluster. Si aucune valeur n'est spécifiée, la valeur par défaut est de 0 Mo.
`das.vmcpuminmhz`	Définit la valeur des ressources CPU par défaut associée à une machine virtuelle si sa réserve de CPU n'est pas précisée ou nulle. Celle-ci est utilisée pour la stratégie de contrôle d'admission Défaillances d'hôte tolérées par le cluster. Si aucune valeur n'est spécifiée, la valeur par défaut est de 32 MHz.
`das.iostatsinterval`	Modifie l'intervalle de statistique des E/S par défaut pour la sensibilité de surveillance des machines virtuelles. La valeur par défaut est de 120 (secondes). Peut être définie sur une valeur supérieure ou égale à 0. Une valeur nulle désactive la vérification. **Remarque** Les valeurs inférieures à 50 ne sont pas recommandées, car elles peuvent entraîner la réinitialisation d'une machine virtuelle par vSphere HA de façon inattendue.

Tableau 2-4. Options avancées de vSphere HA (suite)

Option	Description
`das.ignoreinsufficienthbdatastore`	Désactive les problèmes de configuration créés si l'hôte n'a pas suffisamment de banques de données de signaux de pulsation pour vSphere HA. La valeur par défaut est "faux".
`das.heartbeatdsperhost`	Modifie le nombre de banques de données de signaux de pulsation nécessaire. Les valeurs peuvent s'étendre de 2 à 5 et la valeur par défaut est 2.
`fdm.isolationpolicydelaysec`	Le nombre de secondes pendant lesquelles le système attend avant d'exécuter la politique d'isolation une fois que l'isolation de l'hôte est déterminée. La valeur minimale est 30. S'il une valeur inférieure à 30 est définie, le délai sera de 30 secondes.
`das.respectvmvmantiaffinityrules`	Détermine si vSphere HA applique les règles d'anti-affinité VM-VM. Avec la valeur par défaut « false », les règles ne sont pas appliquées. Si la valeur « true » est choisie, les règles sont appliquées (même si vSphere DRS n'est pas activé). Dans ce cas, vSphere HA ne bascule pas sur une machine virtuelle s'il viole une règle en le faisant, mais émet un événement signalant que les ressources sont insuffisantes pour effectuer le basculement. Pour plus d'informations sur les règles d'anti-affinité, reportez-vous à *Gestion des ressources vSphere*.
`das.maxresets`	Nombre maximal de tentatives de réinitialisation par VMCP. En cas d'échec d'une opération de réinitialisation sur une machine virtuelle affectée par une situation d'APD, VMCP réessaie la réinitialisation plusieurs fois avant d'abandonner.
`das.maxterminates`	Nombre maximal de tentatives d'arrêt d'une machine virtuelle effectuées par VMCP.
`das.terminateretryintervalsec`	En cas d'échec de VMCP à arrêter une machine virtuelle, cette option correspond au nombre de secondes pendant lequel le système attend avant de refaire une tentative d'arrêt.
`das.config.fdm.reportfailoverfailevent`	Quand cette option est définie sur 1, elle permet de générer un événement par machine virtuelle lorsque vSphere HA échoue dans une tentative de redémarrage d'une machine virtuelle. La valeur par défaut est 0. Dans les versions antérieures à vSphere 6.0, cet événement est généré par défaut.
`vpxd.das.completemetadataupdateintervalsec`	Période (en secondes) après qu'une règle d'affinité machine virtuelle/hôte est définie pendant laquelle vSphere HA peut redémarrer une machine virtuelle dans un cluster sur lequel DRS est désactivé, remplaçant ainsi la règle. La valeur par défaut est de 300 secondes.

Tableau 2-4. Options avancées de vSphere HA (suite)

Option	Description
`das.config.fdm.memreservationmb`	Par défaut, les agents vSphere HA s'exécutent avec une limite de mémoire configurée de 250 Mo. Un hôte pourrait ne pas autoriser cette réservation si sa capacité réservable est épuisée. Vous pouvez utiliser cette option pour réduire la limite de mémoire et éviter ainsi ce problème. Seuls des nombres entiers supérieurs à 100, qui est la valeur minimale, peuvent être spécifiés. À l'inverse, pour prévenir tout problème lors des élections d'agents maîtres dans un cluster volumineux (contenant 6 000 à 8 000 machines virtuelles), cette limite doit être portée à 325 Mo. **Remarque** Une fois cette limite modifiée, vous devez exécuter une tâche Reconfigurer HA pour tous les hôtes dans le cluster. En outre, lorsqu'un nouvel hôte est ajouté au cluster ou qu'un hôte existant est redémarré, cette tâche doit être exécutée sur ces hôtes afin de mettre à jour ce paramètre de mémoire.
`das.reregisterrestartdisabledvms`	Lorsque vSphere HA est désactivé sur une VM spécifique, cette option garantit que la VM est enregistrée sur un autre hôte après une panne. Vous pouvez ainsi mettre cette machine virtuelle sous tension sans devoir la réenregistrer manuellement. **Remarque** Lorsque cette option est utilisée, vSphere HA ne met pas la machine virtuelle sous tension, mais l'enregistre uniquement.

Remarque Si vous modifiez la valeur de l'une des options avancées suivantes, vous devez désactiver, puis réactiver vSphere HA avant que les modifications ne s'appliquent.

- `das.isolationaddress[...]`
- `das.usedefaultisolationaddress`
- `das.isolationshutdowntimeout`

Personnaliser une machine virtuelle secondaire

Les paramètres par défaut du cluster relatifs à la priorité de redémarrage, à la réponse d'isolation de l'hôte, à la protection des composants des machines virtuelles et à la surveillance des machines virtuelles sont associés à chaque machine virtuelle d'un cluster vSphere HA. Vous pouvez préciser des comportements spécifiques pour chaque machine virtuelle en changeant ces valeurs par défaut. Si la machine virtuelle quitte le cluster, ces paramètres sont perdus.

Procédure

1 Dans vSphere Web Client, accédez au cluster vSphere HA .

2 Cliquez sur l'onglet **Configurer**.

3 Sous Configuration, sélectionnez **Remplacements de VM** et cliquez sur **Ajouter**.

4 Utilisez le bouton **+** pour sélectionner les machines virtuelles sur lesquelles appliquer les remplacements.

5 Cliquez sur **OK**.

6 (Facultatif) Vous pouvez modifier d'autres paramètres, comme **Niveau d'automatisation**, **Priorité redémarrage VM**, **Réponse d'isolement d'hôte**, VMCP, **Surveillance VM** ou **Sensibilité de surveillance VM**.

Remarque Vous pouvez afficher les paramètres par défaut du cluster pour ces paramètres en commençant par développer **Paramètres**, puis en développant **vSphere HA**.

7 Cliquez sur **OK**.

Le comportement de la VM est désormais différent des réglages par défaut du cluster pour chaque paramètre que avez modifié.

Meilleures pratiques pour les clusters vSphere HA

Pour garantir des performances optimales des clusters vSphere HA, vous devez suivre certaines meilleures pratiques. Cette rubrique met en évidence quelques-unes des recommandations essentielles concernant un cluster vSphere HA.

Vous pouvez également consulter la publication *Meilleures pratiques du déploiement vSphere High Availability* pour poursuivre la discussion.

Meilleures pratiques pour la mise en réseau

Suivez les meilleures pratiques pour la configuration des adaptateurs réseau hôtes et la topologie du réseau pour vSphere HA. Les pratiques d'excellence incluent des recommandations pour vos hôtes ESXi, et traitent aussi du câblage, des commutateurs, des routeurs et des pare-feu.

Configuration et maintenance du réseau

Les suggestions de maintenance du réseau suivantes contribuent à éviter une détection accidentelle d'hôtes défectueux et une isolation du réseau dues à la perte des signaux de pulsation vSphere HA.

- Lors d'une modification des réseaux sur lesquels se trouvent les hôtes ESXi en clusters, suspendez la fonction de surveillance d'hôte. Les changements de matériel ou de paramètres réseau peuvent interrompre les signaux de pulsation utilisés par vSphere HA pour détecter les défaillances d'hôtes, ce qui risque d'entraîner des tentatives intempestives de basculement des machines virtuelles.
- Lorsque, par exemple, vous modifiez la configuration du réseau sur les hôtes ESXi, l'ajout de groupes de ports, ou la suppression de vSwitches, suspendez la surveillance d'hôte. Après avoir effectué les modifications de configuration de réseau, vous devez reconfigurer vSphere HA sur tous les hôtes du cluster, ce qui provoque une nouvelle inspection des informations du réseau. Réactivez ensuite la Surveillance d'hôte.

Remarque La mise en réseau étant un aspect essentiel de vSphere HA, l'administrateur de vSphere HA doit être tenu informé de toute opération de maintenance du réseau.

Réseaux utilisés pour les communications vSphere HA

Pour identifier les opérations réseau qui risquent de perturber le bon fonctionnement de vSphere HA, il est nécessaire d'identifier les réseaux de gestion utilisés pour les signaux de pulsation et autres communications vSphere HA.

- Sur les hôtes hérités ESX du cluster, les communications vSphere HA sont acheminées via tous les réseaux qui sont identifiés comme réseaux de console de service. Les réseaux VMkernel ne sont pas utilisés par ces hôtes pour les communications vSphere HA. Pour contenir le trafic vSphere HA en un sous-ensemble de réseaux de la console ESX, utilisez l'option avancée `allowedNetworks`.

- Sur les hôtes ESXi du cluster, les communications vSphere HA, par défaut, sont acheminées via les réseaux VMkernel. Avec un hôte ESXi, si vous souhaitez utiliser un réseau autre que celui employé par vCenter Server pour communiquer avec l'hôte pour vSphere HA, vous devez cocher explicitement la case **Trafic de gestion**.

Pour garder le trafic de l'agent vSphere HA sur les réseaux que vous avez spécifiés, configurez des hôtes de façon à ce que les cartes vmkNICs utilisées par vSphere HA ne ne partagent pas les sous-réseaux avec les cartes vmkNIC utilisées à d'autres fins. vSphere HA envoient des paquets en utilisant une carte pNIC associée à un sous-réseau donné s'il y a aussi au moins une carte vmkNIC configurée pour le trafic de gestion vSphere HA. Par conséquent, pour assurer la séparation de flux réseau, les cartes vmkNIC utilisés par vSphere HA et par les autres fonctionnalités doivent être sur des sous-réseaux différents.

Adresses d'isolation réseau

Une adresse d'isolation réseau est une adresse IP qui reçoit une commande ping pour déterminer si un hôte est isolé du réseau. Le ping est uniquement envoyé à cette adresse lorsqu'un hôte a cessé de recevoir les signaux de pulsation de tous les autres hôtes du cluster. Si un hôte peut envoyer un ping à son adresse d'isolation réseau, l'hôte n'est pas isolé dans le réseau et soit les autres hôtes du cluster ont échoué, soit le réseau s'est partitionné. Mais si l'hôte ne peut pas envoyer de ping à son adresse d'isolation, il est probable que l'hôte ait été isolé du réseau et aucune action de basculement n'est entreprise.

L'adresse d'isolation réseau est la passerelle par défaut de l'hôte. Une seule passerelle est définie par défaut, quel que soit le nombre de réseaux de gestion définis. Vous devez utiliser l'option avancée `das.isolationaddress[...]` pour ajouter des adresses d'isolation à des réseaux supplémentaires. Reportez-vous à « Options avancées de vSphere HA », page 42.

Redondance des chemins de réseau

La redondance des chemins de réseau entre les nœuds de cluster est importante pour la fiabilité de vSphere HA. Un réseau de gestion isolé finit par être un point de panne isolé, ce qui aboutit à des basculements même si le réseau uniquement est défectueux. Si vous avez un seul réseau de gestion, toute défaillance entre l'hôte et le cluster peut provoquer une activité de basculement inutile (ou faux) si la connectivité du signal de pulsation des banques de données n'est pas conservé lors de la panne de réseau. Les défaillances possibles incluent les pannes de adaptateurs réseau, les pannes de câbles réseau, la suppression de câbles réseau et les réinitialisations de commutateurs. Examinez ces causes possibles de défaillances entre les hôtes et efforcez-vous de les minimiser en assurant une redondance du réseau.

Il vous est d'abord possible d'implémenter la redondance du réseau au niveau de l'association de cartes réseau. L'utilisation d'une association de deux adaptateurs réseau connectées pour séparer les commutateurs physiques améliore la fiabilité d'un réseau de gestion. Le cluster est plus résilient car les serveurs connectés par deux adaptateurs réseau (et par des commutateurs séparés) ont deux chemins indépendants pour la transmission et la réception de signaux de pulsation. Pour configurer une association de adaptateurs réseau pour réseau de gestion, configurez les vNIC de la configuration vSwitch pour la configuration Active ou Standby. Les réglages recommandés pour les paramètres des vNIC sont les suivants :

- Équilibrage de charge par défaut = Router en fonction de l'ID du port d'origine
- Retour arrière = Non

Lorsque vous avez ajouté une carte réseau à un hôte de votre cluster vSphere HA, vous devez reconfigurer vSphere HA sur cet hôte.

Dans la plupart des implémentations, l'association de cartes réseau offre une redondance suffisante, mais il vous est également possible de créer une connexion de réseau de gestion secondaire qui est liée à un commutateur virtuel distinct. La mise en réseau de gestion redondante garantit la fiabilité de la détection des pannes et évite la réalisation de conditions d'isolation ou de partition car les signaux de pulsation

peuvent être transmis via plusieurs réseaux. La connexion de réseau de gestion originelle est utilisée pour le réseau et à des fins de gestion. Lorsque la connexion de réseau de gestion secondaire est créée, vSphere HA transmet des signaux de pulsation sur les deux connexions de réseau de gestion à la fois. Si un chemin est défaillant, vSphere HA continue à transmettre et à recevoir des signaux de pulsation par l'autre chemin.

Remarque Configurez un nombre aussi réduit que possible de segments matériels entre les serveurs d'un cluster. L'objectif est de limiter les points de panne isolés. De plus, les chemins contenant trop de bonds peuvent provoquer des retards de paquets de signaux de pulsation et augmenter les points de panne éventuels.

Utilisation des configurations réseau IPv6

Une seule adresse IPv6 doit être attribuée à une interface réseau donnée utilisée par votre cluster vSphere HA. L'attribution de plusieurs adresses IP augmente le nombre de messages de signal de pulsation envoyés par l'hôte maître du cluster sans l'avantage correspondant.

Recommandations concernant l'interopérabilité

Suivez les recommandations suivantes pour permettre une interopérabilité adéquate entre vSphere HA et d'autres fonctionnalités.

Interopérabilité de vSphere HA et de Storage vMotion dans un cluster mixte

Dans les clusters où des hôtes ESXi 5.x et ESX/ESXi 4.1 ou des hôtes antérieurs sont présents et où Storage vMotion est largement utilisé ou Storage DRS est activé, ne déployez pas vSphere HA. vSphere HA pourrait répondre à une défaillance de l'hôte en redémarrant une VM sur un hôte avec une version ESXi différente de celle sur laquelle la VM a été lancée avant la défaillance. Un problème peut survenir si, au moment de la défaillance, la machine virtuelle participait à une action de Storage vMotion sur un hôte ESXi 5.x, et si vSphere HA redémarre la VM sur un hôte ayant une version antérieure à ESXi 5.0. Pendant l'allumage de la machine virtuelle, des tentatives ultérieures d'opérations de snapshot pourraient corrompre l'état du vdisk et rendre la machine virtuelle inutilisable.

Utiliser Auto Deploy avec vSphere HA

Vous pouvez utiliser simultanément vSphere HA et Auto Deploy pour améliorer la disponibilité de vos machines virtuelles. Auto Deploy approvisionne les hôtes lorsqu'ils s'allument. Vous pouvez également le configurer pour installer l'agent vSphere HA sur ces hôtes pendant le processus de démarrage. Pour plus de détails, consultez la documentation d'Auto Deploy incluse dans le guide Installation et configuration de vSphere.

Mise à niveau d'hôtes dans un cluster à l'aide de Virtual SAN

Si vous mettez à niveau les hôtes ESXi dans votre cluster vSphere HA vers la version 5.5 ou une version ultérieure, et que vous prévoyez également d'utiliser Virtual SAN, suivez ce processus.

1 Mettez à niveau tous les hôtes.

2 Désactivez vSphere HA.

3 Activez Virtual SAN.

4 Réactivez vSphere HA.

Recommandations concernant la surveillance d'un cluster

Suivez les recommandations suivantes lors de la surveillance de l'état et de la validité de votre cluster vSphere HA.

Définir des alarmes pour surveiller les changements des clusters

Quand vSphere HA ou Fault Tolerance interviennent pour préserver la disponibilité en effectuant un basculement de machine virtuelle, par exemple, vous avez la possibilité d'être averti de ces changements. Dans vCenter Server, configurez des alarmes qui seront déclenchées lorsque ces actions surviendront, et recevez des alertes, sous forme de messages électroniques, par exemple, envoyées à un groupe d'administrateurs prédéfini.

Plusieurs alarmes par défaut sont disponibles pour vSphere HA.

- Ressources de basculement insuffisantes (alarme de cluster)
- Impossible de trouver le cluster principal (alarme du cluster)
- Basculement en cours (alarme du cluster)
- Statut de l'hôte HA (alarme d'hôte)
- Erreur de surveillance de VM (alarme de machine virtuelle)
- Action de surveillance de VM (alarme de machine virtuelle)
- Échec du basculement (alarme de machine virtuelle)

Remarque Les alarmes par défaut contiennent le nom de la fonction, vSphere HA.

Surveillance de la validité du cluster

Un cluster valide est un cluster sur lequel il n'y eu aucune violation des stratégies de contrôle d'admission.

Un cluster sur lequel HA est activé devient invalide lorsque le nombre de machines virtuelles sous tension dépasse les exigences de basculement, ce qui signifie, que la capacité de basculement actuelle est inférieure à la capacité de basculement configurée. Si le contrôle d'admission est désactivé, les clusters ne deviennent pas non valides.

Dans vSphere Web Client, sélectionnez **vSphere HA** dans l'onglet **Moniteur** du cluster, puis sélectionnez **Problèmes de configuration**. La liste de problèmes actuels de vSphere HA apparaît.

Le comportement DRS n'est pas affecté par un cluster rouge à cause d'un problème lié à vSphere HA.

Assurer Fault Tolerance des machines virtuelles 3

Il est possible d'utiliser vSphere Fault Tolerance pour vos machines virtuelles afin d'assurer la continuité d'activité avec des niveaux de disponibilité et de protection des données supérieurs à ceux offerts par vSphere HA.

Fault Tolerance est basée sur la plate-forme hôte ESXi et elle fournit une disponibilité continue en exécutant des machines virtuelles identiques sur des hôtes distincts.

Pour obtenir des résultats optimaux de Fault Tolerance, il est nécessaire d'en comprendre le fonctionnement, de savoir comment l'activer sur un cluster et sur des machines virtuelles, et de connaître les meilleures pratiques pour son utilisation.

Protection Fault Tolerance pour machines virtuelles (http://link.brightcove.com/services/player/bcpid2296383276001?bctid=ref:video_fault_tolerance_protection_vms)

Ce chapitre aborde les rubriques suivantes :

Fonctionnement de Fault Tolerance

Il est possible d'utiliser vSphere Fault Tolerance (FT) sur la plupart des machines virtuelles cruciales pour une mission. FT assure la disponibilité continue d'une machine virtuelle de ce type en créant et en maintenant une autre machine virtuelle identique et disponible en permanence pour la remplacer en cas de situation de basculement.

La machine virtuelle protégée s'appelle la machine virtuelle principale. La copie de la machine virtuelle, la machine virtuelle secondaire, est créée et s'exécute sur un autre hôte. L'exécution de la machine virtuelle secondaire est identique à celle de la machine virtuelle principale et elle peut reprendre l'exécution à tout moment sans interruption, assurant ainsi une protection tolérante aux pannes.

Les machines virtuelles principale et secondaire surveillent continuellement l'état l'une de l'autre pour vérifier que la tolérance aux pannes est maintenue. Un basculement transparent se produit en cas de défaillance de l'hôte sur lequel la machine virtuelle principale est exécutée. Dans ce cas, la machine virtuelle secondaire est immédiatement activée pour remplacer la machine virtuelle principale. Une nouvelle machine virtuelle secondaire démarre et la redondance de Fault Tolerance est rétablie en quelques secondes. Si l'hôte de la machine virtuelle secondaire devient défectueux, il est aussi immédiatement remplacé. Dans l'un ou l'autre cas, les utilisateurs ne constatent aucune interruption de service ni perte de données.

Une machine virtuelle tolérante aux pannes et sa copie secondaire ne sont pas autorisées à fonctionner sur le même hôte. Cette restriction garantit qu'un échec de l'hôte ne peut pas entraîner la perte des deux machines virtuelles.

Remarque Vous pouvez aussi utiliser les règles d'affinité entre machine virtuelle et hôte pour préciser les hôtes sur lesquels certaines machines virtuelles peuvent être exécutées. Si vous utilisez ces règles, souvenez-vous que pour chaque machine virtuelle principale affectée par une règle précise, la machine virtuelle secondaire qui y est associée est aussi affectée par la même règle. Pour plus d'informations sur les règles d'affinité, reportez-vous à la documentation *Gestion des ressources vSphere*.

Fault Tolerance évite les situations de division qui peuvent se traduire par deux copies actives d'une machine virtuelle après la reprise suite à un dysfonctionnement. Le verrouillage atomique des fichiers sur les stockages partagés est utilisé pour coordonner le basculement de façon à ce qu'un côté seulement continue à exécuter la machine virtuelle principale et une nouvelle machine virtuelle secondaire est automatiquement réaffectée.

vSphere Fault Tolerance peut gérer les machines virtuelles à multiprocesseur symétrique (SMP) avec jusqu'à quatre vCPU. Les versions antérieures de vSphere utilisaient une technologie différente pour Fault Tolerance (connue sous le nom de FT héritée), avec différentes conditions requises et caractéristiques (notamment une limitation des vCPU uniques pour les machines virtuelles FT héritée). Si la compatibilité avec ces conditions antérieures est nécessaire, vous pouvez utiliser FT héritée à la place. Toutefois, cela implique la configuration d'une option avancée pour chaque machine virtuelle. Consultez « Fault Tolerance héritée », page 63 pour plus d'informations.

Cas d'utilisation de Fault Tolerance

Plusieurs situations types peuvent bénéficier de l'utilisation de vSphere Fault Tolerance.

Fault Tolerance assure un meilleur niveau de continuité d'activité que vSphere HA. Lorsqu'une machine virtuelle secondaire doit intervenir pour remplacer son homologue, la machine virtuelle principale, la machine virtuelle secondaire joue immédiatement le rôle de machine virtuelle principale, l'état de la machine virtuelle restant entièrement préservé. Les applications sont déjà en cours d'exécution et les données conservées en mémoire ne doivent pas être ressaisies ou rechargées. Ce n'est pas le cas du basculement assuré par vSphere HA qui redémarre les machines virtuelles affectées par un dysfonctionnement.

Ce haut niveau de continuité et la meilleure protection des informations d'états et des données informe les scénarios du déploiement possible de Fault Tolerance.

- Les applications qui doivent être disponibles en permanence, surtout celles présentant des connexions longues durées de clients que les utilisateurs veulent conserver pendant la défaillance matérielle.
- Applications personnalisées qui n'ont pas d'autres moyens de former un cluster.
- Cas où la grande disponibilité peut être assurée par des solutions de formation de cluster personnalisées qui sont très compliquées à configurer et à entretenir.

Un autre cas pratique de protection d'une machine virtuelle par Fault Tolerance s'intitule Fault Tolerance à la demande. Dans ce cas, une machine virtuelle est correctement protégée par vSphere HA pendant son fonctionnement normal. Pendant certaines périodes critiques, vous voudrez renforcer la protection de la machine virtuelle. Pendant la production d'un rapport trimestriel, par exemple, dont l'interruption pourrait

retarder la mise à disposition d'informations cruciales pour une mission. vSphere Fault Tolerance permet de protéger cette machine virtuelle avant la production du rapport, puis d'arrêter ou d'interrompre Fault Tolerance après la publication du rapport. Vous pouvez utiliser Fault Tolerance à la demande pour protéger la machine virtuelle pendant une période critique et revenir aux ressources normales pour les opérations non critiques.

Configuration requise, limites et licence de Fault Tolerance

Avant d'utiliser vSphere Fault Tolerance (FT), tenez compte des conditions requises de niveau supérieur, des limites et de l'attribution de licence qui s'appliquent à cette fonctionnalité.

Exigences

Les conditions de CPU et de mise en réseau requises suivantes s'appliquent à FT.

Les CPU qui sont utilisés sur les machines hôtes pour les machines virtuelles Fault Tolerance doivent être compatibles avec vSphere vMotion ou être améliorées par Enhanced vMotion Compatibility. De plus, les CPU qui prennent en charge la virtualisation du matériel MMU (Intel EPT ou AMD RVI) sont requis. Les CPU suivants sont pris en charge.

- Intel Sandy Bridge ou version ultérieure. Avoton n'est pas pris en charge.
- AMD Bulldozer ou version ultérieure.

Utilisez un réseau de journalisation de 10 Gbits pour FT et vérifiez que la latence du réseau est faible. Un réseau FT dédié est fortement recommandé.

Limites

Dans un cluster configuré pour utiliser Fault Tolerance, deux limites sont appliquées de manière distincte.

das.maxftvmsperhost	Le nombre maximal de machine virtuelles Fault Tolerance autorisées sur un hôte dans le cluster. Les machines virtuelles principale et secondaires sont prises compte vis-à-vis de cette limite. La valeur par défaut est 4.
das.maxftvcpusperhost	Le nombre maximal de vCPU regroupés dans toutes les machines virtuelles Fault Tolerance sur un hôte. Les vCPU des machines virtuelles principale et secondaires sont pris en compte vis-à-vis de cette limite. La valeur par défaut est 8.

Attribution de licences

Le nombre de vCPU pris en charge par une machine virtuelle unique est limité par le niveau d'attribution de licence acheté pour vSphere. Fault Tolerance est prise en charge comme suit :

- vSphere Standard et Enterprise. Autorise jusqu'à 2 vCPU
- vSphere Enterprise Plus. Autorise jusqu'à 4 vCPU

Remarque FT et FT héritée ne sont pas prises en charge dans vSphere Essentials et vSphere Essentials Plus.

Interopérabilité de Fault Tolerance

vSphere Fault Tolerance est soumise à certaines limitations concernant les fonctionnalités de vSphere, les périphériques et les autres fonctionnalités avec lesquelles elle peut interagir.

Avant de configurer vSphere Fault Tolerance, vous devez connaître les fonctions et produits incompatibles avec Fault Tolerance.

Fonctions vSphere non prises en charge par Fault Tolerance

Lors de la configuration de votre cluster, vous devez savoir que toutes les fonctionnalités de vSphere ne peuvent pas interagir avec Fault Tolerance.

Les fonctions vSphere suivantes ne sont pas prises en charge pour les machines virtuelles tolérantes aux pannes.

- Snapshots. Les snapshots doivent être supprimés ou engagés avant l'activation de Fault Tolerance sur une machine virtuelle. De plus, il n'est pas possible de prendre des snapshots de machines virtuelles sur lesquelles Fault Tolerance est activée.

  Remarque Les snapshots sur disque uniquement créés pour des sauvegardes de vStorage APIs - Data Protection (VADP) sont pris en charge avec l'option Fault Tolerance. Cependant, la protection FT héritée ne prend pas en charge VADP.

- Stockage vMotion Il n'est pas possible d'appeler le stockage vMotion pour les machines virtuelles pour lesquelles Fault Tolerance est activée. Pour migrer le stockage, il faut mettre hors tension temporairement Fault Tolerance et exécuter l'action de stockage vMotion. Une fois ceci fait, vous pouvez réactiver Fault Tolerance.
- Clones liés. Il n'est ni possible d'utiliser Fault Tolerance sur une machine virtuelle qui est un clone lié, ni de créer un clone lié à partir d'une machine virtuelle sur laquelle Fault Tolerance est activée.
- VM Component Protection (VMCP). Si VMCP est activé sur votre cluster, des remplacements sont créés pour les machines virtuelles Fault Tolerance qui désactivent cette fonctionnalité.
- Banques de données à volume virtuel.
- Gestion de stratégie basée sur le stockage.
- Filtres d'E/S.

Fonctions et périphériques incompatibles avec Fault Tolerance

Tous les périphériques, fonctionnalités ou produits tiers ne peuvent pas interagir avec Fault Tolerance.

Pour qu'une machine virtuelle soit compatible avec Fault Tolerance, celle-ci ne doit pas utiliser les fonctions ou périphériques suivants.

Tableau 3-1. Fonctions et périphériques incompatibles avec Fault Tolerance et les actions correctives

Fonction ou périphérique incompatible	Action corrective
Mappage disque brut physique (RDM).	Avec la fonctionnalité FT, vous pouvez reconfigurer les machines virtuelles avec des périphériques virtuels pris en charge par des RDM physiques de sorte qu'ils utilisent des RDM virtuels à la place.
Lecteur de CD-ROM ou de disquettes virtuels pris en charge par un périphérique physique ou distant.	Retirez le lecteur de CD-ROM ou de disquettes virtuels ou reconfigurez la sauvegarde avec une image ISO installée sur le stockage partagé.
Périphérique USB et audio.	Déconnectez ces périphériques de la machine virtuelle.
Virtualisation d'identification N-Port (NPIV).	Désactivez la configuration NPIV de la machine virtuelle
relais de adaptateurs réseau	Cette fonction n'est pas prise en charge par Fault Tolerance et doit donc être désactivée.

Tableau 3-1. Fonctions et périphériques incompatibles avec Fault Tolerance et les actions correctives (suite)

Fonction ou périphérique incompatible	Action corrective
Connexion de périphériques à chaud	La fonction de connexion à chaud est automatiquement désactivée pour les machines virtuelles tolérantes aux pannes. Pour la connexion des périphériques à chaud (ajout ou suppression), vous devez mettre hors tension temporairement Fault Tolerance, effectuer la connexion à chaud, puis réactiver Fault Tolerance. **Remarque** Lorsque vous utilisez Fault Tolerance, la modification des paramètres d'une carte réseau virtuelle pendant le fonctionnement d'une machine virtuelle constitue une connexion à chaud, car cela exige de « débrancher » la carte réseau, puis de la « rebrancher ». Prenons l'exemple d'une carte réseau virtuelle pour une machine virtuelle en cours d'exécution. Si vous modifiez le réseau auquel la carte réseau virtuelle est connectée, la tolérance aux pannes doit préalablement être arrêtée.
Ports série ou parallèles	Déconnectez ces périphériques de la machine virtuelle.
Périphériques vidéo dont la 3D est activée.	Fault Tolerance ne prend pas en charge les périphériques vidéo dont la 3D est activée.
Microprogramme EFI virtuel	Assurez-vous que la VM est configurée pour utiliser le firmware du BIOS avant d'installer le système d'exploitation d'hôte.
VMCI (Virtual machine communication interface)	Non prise en charge par Fault Tolerance.
Disque de machine virtuelle de plus de 2 To	Fault Tolerance n'est pas prise en charge sur les disques de machine virtuelle de plus de 2 To.

Utiliser Fault Tolerance avec DRS

Vous pouvez utiliser vSphere Fault Tolerance avec vSphere Distributed Resource Scheduler (DRS) quand la fonctionnalité EVC (Enhanced vMotion Compatibility) est activée. Ce processus permet aux machines Fault Tolerant de bénéficier d'un meilleur placement initial.

Quand la fonctionnalité EVC est activée pour un cluster, DRS émet les recommandations de placement initial pour les machines virtuelles Fault Tolerant et vous permet d'attribuer un niveau d'automatisation DRS aux machines virtuelles principales (la machine virtuelle secondaire adopte toujours le même paramètre que la machine virtuelle principale associée).

Quand vSphere Fault Tolerance est utilisé pour les machines virtuelles d'un cluster pour lequel EVC est désactivé, les machines virtuelles tolérantes aux pannes reçoivent des niveaux d'automatisation DRS "désactivés". Dans ce type de cluster, chaque machine virtuelle principale est uniquement mise sous tension sur son hôte enregistré et sa machine virtuelle secondaire est placée automatiquement.

Si vous utilisez des règles d'affinité avec deux machines virtuelles tolérantes aux pannes, une règle d'affinité VM-VM s'applique uniquement à la machine virtuelle principale, tandis qu'une règle d'affinité machine virtuelle-hôte s'applique à la fois à la machine virtuelle principale et à sa machine virtuelle secondaire. Si une règle d'affinité VM-VM est définie pour une machine virtuelle principale, DRS tente de corriger toutes les violations survenant après un basculement (c'est-à-dire, après le déplacement effectif de la machine virtuelle principale vers un nouvel hôte).

Préparer votre cluster et vos hôtes à Fault Tolerance

Pour activer vSphere Fault Tolerance pour votre cluster, les conditions préalables de la fonction doivent être remplies et il est nécessaire d'effectuer quelques étapes de configuration sur les hôtes. Une fois ces étapes accomplies et votre cluster créé, vous pouvez aussi vérifier que la configuration est conforme aux exigences requises pour l'activation de Fault Tolerance.

Les tâches devant être effectuées avant de tenter d'activer Fault Tolerance pour le cluster sont les suivantes :

- Vérifiez que vos cluster, vos hôtes et vos machines virtuelles satisfont les conditions requises par la liste de contrôle de Fault Tolerance.
- Configurer la mise en réseau de chaque hôte
- Créer un cluster vSphere HA, ajouter des hôtes et vérifier la conformité

Lorsque le cluster et les hôtes sont prêts, vous pouvez activer Fault Tolerance pour vos machines virtuelles. Reportez-vous à « Activer Fault Tolerance », page 58.

Liste de contrôle de Fault Tolerance

La liste de vérification suivante contient les spécifications en matière de cluster, d'hôte et de machine virtuelle que vous devez connaître avant d'utiliser vSphere Fault Tolerance.

Consultez cette liste avant de configurer Fault Tolerance.

Remarque Le basculement des machines virtuelles tolérantes aux pannes ne dépend pas de vCenter Server, mais vous devez utiliser vCenter Server pour configurer vos clusters de Fault Tolerance.

Spécifications des clusters pour Fault Tolerance

Les exigences suivantes aux clusters doivent être remplies avant d'utiliser Fault Tolerance.

- Journalisation de Fault Tolerance et réseau vMotion configuré. Reportez-vous à « Configurer la mise en réseau des machines hôtes », page 55.
- Cluster vSphere HA créé et activé. Reportez-vous à la section « Création d'un cluster vSphere HA », page 33. vSphere HA doit être activé avant la mise sous tension des machines virtuelles tolérantes aux pannes ou avant l'ajout d'un hôte dans un cluster qui prend déjà en charge des machines virtuelles tolérantes aux pannes.

Conditions requises pour les hôtes pour Fault Tolerance

Les conditions suivantes concernant les hôtes doivent être remplies avant d'utiliser Fault Tolerance.

- Les hôtes doivent utiliser des processeurs pris en charge.
- Les hôtes doivent avoir une licence pour Fault Tolerance.
- Les hôtes doivent être certifiés pour Fault Tolerance. Reportez-vous à la section http://www.vmware.com/resources/compatibility/search.php et sélectionnez **Recherche par ensembles compatibles Fault Tolerance** pour déterminer si vos hôtes sont certifiés.
- La configuration de chaque hôte implique l'activation de la virtualisation matérielle (HV) dans le BIOS.

Remarque VMware recommande que les paramètres de gestion de l'alimentation BIOS des hôtes que vous utilisez pour prendre en charge les machines virtuelles Fault Tolerant soient définis sur « Performances maximales » ou « Performances gérées par le système d'exploitation ».

Pour confirmer la compatibilité des hôtes dans le cluster pour la prise en charge de la tolérance aux pannes, vous pouvez aussi effectuer des vérifications de conformité de profils comme décrit dans « Créer un cluster et vérifier la conformité », page 56.

Conditions des machines virtuelles pour Fault Tolerance

Les conditions des machines virtuelles suivantes doivent être remplies avant d'utiliser Fault Tolerance.

- Aucun périphérique non pris en charge n'est attaché à la machine virtuelle. Reportez-vous à « Interopérabilité de Fault Tolerance », page 51.
- Les fonctions incompatibles ne doivent pas être exécutées avec les machines virtuelles tolérantes aux pannes. Reportez-vous à « Interopérabilité de Fault Tolerance », page 51.
- Les fichiers des machines virtuelles (sauf les fichiers VMDK) doivent être stockés sur le stockage partagé. Les solutions de stockage partagé approuvées comprennent Fibre Channel, iSCSI (matériel et logiciel), NFS et NAS.

Autres recommandations de configuration

Vous devez respecter les directives suivantes lors de la configuration de Fault Tolerance.

- Si vous accédez au stockage partagé par NFS, utilisez du matériel NAS dédié avec au moins une carte réseau 1 Gbit pour atteindre les performances réseaux requises pour le bon fonctionnement de Fault Tolerance.
- La réservation de mémoire d'une machine virtuelle Fault Tolerant est définie par la taille de la mémoire de la machine virtuelle lorsque Fault Tolerance est activée. Veillez à ce qu'un pool de ressources contenant des machines virtuelles Fault Tolerance dispose de réserves de mémoire dépassant la capacité de mémoire des machines virtuelles. Sans cet excédent de pool de ressources, il risque de ne pas y avoir de mémoire disponible comme capacité supplémentaire.
- Utilisez 16 disques virtuels au maximum par machine virtuelle tolérante aux pannes.
- Pour assurer la redondance et une protection maximale de Fault Tolerance, il est recommandé d'avoir au minimum trois hôtes par cluster. Dans une situation de basculement, on dispose ainsi d'un hôte capable de gérer la nouvelle machine virtuelle secondaire qui est créée.

Configurer la mise en réseau des machines hôtes

Vous devez configurer deux commutateurs de mise en réseau distincts (vMotion et journalisation de FT) sur chacun des hôtes que vous souhaitez ajouter à un cluster vSphere HA, de sorte que l'hôte puisse prendre en charge vSphere Fault Tolerance.

Pour configurer Fault Tolerance sur un hôte, vous devez exécuter cette procédure pour chaque option de groupe de ports (vMotion et journalisation de FT) afin de vous assurer qu'il y a suffisamment de bande passante disponible pour la journalisation de Fault Tolerance. Sélectionnez une option, terminez la procédure, et recommencez-la une seconde fois en sélectionnant l'autre option de groupes de port.

Prérequis

Des adaptateurs réseau (NIC) de plusieurs giga-octets sont nécessaires. Pour chaque hôte compatible avec Fault Tolerance, il faut au minimum deux cartes réseau physiques. par exemple, l'une dédiée à la journalisation de Fault Tolerance et l'autre dédiée à vMotion. Utilisation de trois adaptateurs réseau ou plus pour assurer la disponibilité.

Remarque Les cartes réseau de journalisation vMotion et de tolérance aux pannes doivent être sur des sous-réseaux différents. Si vous utilisez la fonctionnalité FT héritée, IPv6 n'est pas pris en charge sur la carte réseau de journalisation de FT.

Procédure

1 Dans vSphere Web Client, accédez à l'hôte

2 Cliquez sur l'onglet **Configurer**, puis sur **Mise en réseau**.

3 Sélectionnez l'option **Adaptateur réseau VMkernel**.

4 Cliquez sur l'icône **Ajouter mise en réseau d'hôte**.

5 Fournissez les informations appropriées pour votre type de connexion.

6 Cliquez sur **Terminer** .

Lorsque vous avez créé à la fois un commutateur virtuel de journalisation vMotion et de Fault Tolerance, vous pouvez créer d'autres commutateurs virtuels en cas de besoin. Ajoutez ensuite l'hôte au cluster et terminez toutes les étapes nécessaires à l'activation de Fault Tolerance.

Suivant

Remarque Si vous configurez la mise en réseau pour la prise en charge de FT mais que par la suite vous interrompez le port de journalisation de Fault Tolerance, les paires de machines virtuelles Fault Tolerance qui sont déjà sous tension le resteront. Mais dans le cas de situation de basculement, une nouvelle VM secondaire n'est pas démarrée après le remplacement de la VM principale par sa VM secondaire. Par conséquent, la nouvelle VM principale fonctionne en état non protégé.

Créer un cluster et vérifier la conformité

vSphere Fault Tolerance est utilisé dans le cadre d'un cluster vSphere HA. Après avoir configuré la mise en réseau de chaque hôte, créez le cluster vSphere HA et ajoutez-y les hôtes. Vous pouvez vérifier que le cluster est configuré correctement et qu'il est conforme aux exigences pour l'activation de Fault Tolerance.

Procédure

1 Dans vSphere Web Client, accédez au cluster.

2 Cliquez sur l'onglet **Surveiller** puis sur **Conformité de profil**.

3 Cliquez sur **Vérifier la conformité maintenant** pour exécuter les tests de conformité.

Les résultats des tests de conformité apparaissent et la conformité ou non de chaque hôte s'affiche.

Utilisation de la tolérance aux pannes

Après avoir suivi toutes les étapes nécessaires à l'activation de vSphere Fault Tolerance pour votre cluster, vous pouvez utiliser cette fonction en l'activant sur des machines virtuelles individuelles.

Avant de pouvoir activer Fault Tolerance, plusieurs vérifications de validation sont exécutés sur une machine virtuelle.

Après le passage de ces vérifications et après avoir activé vSphere Fault Tolerance pour une machine virtuelle, de nouvelles options sont ajoutées à la section Fault Tolerance de son menu contextuel. Elles comprennent notamment la mise hors tension ou la désactivation de Fault Tolerance, la migration de la machine virtuelle secondaire, le test du basculement et le test du redémarrage de la machine virtuelle secondaire.

Contrôles de validation pour l'activation de Fault Tolerance

Si l'option pour activer Fault Tolerance est disponible, cette tâche doit encore être validée et peut échouer si certaines conditions n'est pas remplies.

Plusieurs contrôles de validation sont exécutés sur une machine virtuelle avant de pouvoir activer Fault Tolerance.

- Le contrôle de certificat SSL doit être activé dans les paramètres de vCenter Server.
- L'hôte doit se trouver dans un cluster vSphere HA ou un cluster mixte vSphere HA et DRS.
- L'hôte doit disposer de ESXi 6.x ou version ultérieure (ESX/ESXi 4.x ou version ultérieure pour FT héritée).
- La machine virtuelle ne doit pas avoir de snapshots.
- La machine virtuelle ne doit pas être un modèle.
- La machine virtuelle ne doit pas avoir vSphere HA désactivé.
- Aucun périphérique vidéo dont la 3D est activée ne doit être présent sur la machine virtuelle.

Vérifications des machines virtuelles activées

Plusieurs vérifications de validation supplémentaires sont effectuées pour les machines virtuelles sous tension (ou celles qui sont en cours de mise sous tension).

- Le BIOS des hôtes où résident les machines virtuelles tolérantes aux pannes doit avoir la virtualisation matérielle (HV, Hardware Virtualization) activée.
- L'hôte qui prend en charge la machine virtuelle principale doit avoir un processeur qui prend en charge Fault Tolerance.
- Les composants matériels doivent être certifiés compatibles avec Fault Tolerance. Pour en avoir la confirmation, consultez le Guide de compatibilité VMware sur http://www.vmware.com/resources/compatibility/search.php et sélectionnez **Recherche par ensembles compatibles Fault Tolerance**.
- La configuration de la machine virtuelle doit être valide pour être utilisée avec une Fault Tolerance (par exemple, la configuration ne peut comporter aucun périphérique non pris en charge.).

Placement de la machine virtuelle secondaire

Quand votre effort d'activation de Fault Tolerance pour une machine virtuelle réussit aux contrôles de validation, la machine virtuelle secondaire est créée. Le placement et le statut immédiat de la machine virtuelle secondaire dépendent de l'état sous tension ou hors tension de la machine virtuelle principale quand vous avez activé Fault Tolerance.

Si la machine virtuelle principale est sous tension :

- L'état complet de la machine virtuelle principale est copié et la machine virtuelle secondaire est créée, placée sur un hôte compatible distinct et mise sous tension si elle passe le contrôle d'admission.
- Le statut de tolérance aux pannes affiché pour la machine virtuelle est **protégée**.

Si la machine virtuelle principale est hors tension :

- La machine virtuelle secondaire est créée immédiatement et enregistrée dans le cluster d'un hôte (Il doit être enregistré sur un hôte plus approprié lorsqu'il est mis sous tension.)
- La machine virtuelle secondaire est mise sous tension seulement après la mise sous tension de la machine virtuelle principale.

- Le statut de tolérance aux pannes affiché pour la machine virtuelle est **Non protégée, VM pas en exécution**.
- Quand vous essayez de mettre sous tension la machine virtuelle primaire après l'activation de Fault Tolerance, les contrôles supplémentaires de validation sont exécutés.

  Après le passage de ces contrôles, les machines virtuelles principales et secondaires sont mises sous tension et placées sur les hôtes distincts et compatibles. Le statut de tolérance aux pannes de la machine virtuelle est marqué comme **Protégée**.

Activer Fault Tolerance

Vous pouvez activer vSphere Fault Tolerance via vSphere Web Client.

Quand Fault Tolerance est activée, vCenter Server réinitialise la limite de mémoire de la VM et définit la réservation de mémoire en fonction de la taille de la mémoire de la VM. Si Fault Tolerance reste activée, il n'est pas possible de modifier la réservation de mémoire, sa taille, la limite, le nombre de vCPU ou les partages. Il est également impossible d'ajouter ou de supprimer des disques pour la machine virtuelle. Quand Fault Tolerance est désactivée, les valeurs d'origine de tous les paramètres qui ont été modifiés ne sont pas restaurées.

Connectez vSphere Web Client à vCenter Server en utilisant un compte ayant des droits d'accès administrateur au cluster.

Prérequis

L'option permettant d'activer Fault Tolerance n'est pas disponible (grisée) si l'une de ces conditions s'applique :

- La machine virtuelle réside sur un hôte qui n'a pas de licence pour la fonction.
- La machine virtuelle réside sur un hôte qui est en mode maintenance ou standby.
- La machine virtuelle est déconnectée ou orpheline (son fichier .vmx n'est pas accessible).
- L'utilisateur n'a pas l'autorisation d'activer la fonction.

Procédure

1. Dans vSphere Web Client, accédez à la VM pour laquelle vous souhaitez activer Fault Tolerance
2. Cliquez avec le bouton droit sur la machine virtuelle et sélectionnez **Fault Tolerance > Activer Fault Tolerance**.
3. Cliquez sur **Oui**.
4. Choisissez une banque de données sur laquelle placer les fichiers de configuration de la machine virtuelle secondaire. Puis cliquez sur **Suivant**.
5. Choisissez un hôte sur lequel placer la machine virtuelle secondaire. Puis cliquez sur **Suivant**.
6. Passez vos sélections en revue et cliquez sur **Terminer**.

La VM spécifiée est désignée comme VM principale et une VM secondaire est établie sur un autre hôte. La machine virtuelle principale est désormais tolérante aux pannes.

Désactiver la Fault Tolerance

La désactivation de vSphere Fault Tolerance supprime la machine virtuelle secondaire, sa configuration et l'ensemble de son historique.

Utilisez l'option **Désactiver la tolérance aux pannes** si vous n'avez pas prévu de réactiver la fonction. Dans le cas contraire, utilisez l'option **Interrompre Fault Tolerance**.

Remarque Si la VM secondaire réside sur un hôte en mode maintenance, déconnecté ou qui ne répond pas, vous ne pouvez pas utiliser l'option **Arrêter tolérance aux pannes**. Dans ce cas, interrompez, puis reprenez Fault Tolerance.

Procédure

1 Dans vSphere Web Client, accédez à la VM pour laquelle vous souhaitez arrêter la tolérance aux pannes.

2 Cliquez avec le bouton droit sur la machine virtuelle et sélectionnez **Fault Tolerance > Désactiver la Fault Tolerance**.

3 Cliquez sur **Oui**.

La tolérance aux pannes est arrêtée pour la machine virtuelle sélectionnée. L'historique, ainsi que la VM secondaire de la VM sélectionnée sont supprimés.

Interrompre Fault Tolerance

L'interruption de vSphere Fault Tolerance pour une machine virtuelle interrompt sa protection Fault Tolerance, mais conserve la machine virtuelle secondaire, sa configuration et l'ensemble de l'historique. Utilisez cette option pour reprendre la protection de Fault Tolerance à l'avenir.

Procédure

1 Dans vSphere Web Client, accédez à la machine virtuelle pour laquelle vous souhaitez interrompre Fault Tolerance.

2 Cliquez avec le bouton droit sur la machine virtuelle et sélectionnez **Fault Tolerance > Interrompre Fault Tolerance**.

3 Cliquez sur **Oui**.

Fault Tolerance est interrompue pour la machine virtuelle sélectionnée. L'historique et la machine virtuelle secondaire de la machine virtuelle sélectionnée sont préservés et seront utilisés si la fonctionnalité est reprise.

Suivant

Pour reprendre la fonctionnalité après avoir interrompu Fault Tolerance, sélectionnez **Relancer Fault Tolerance**.

Migration secondaire

Une fois que vSphere Fault Tolerance est activé pour une VM principale, vous pouvez migrer sa VM secondaire associée.

Procédure

1 Dans vSphere Web Client, accédez à la VM primaire pour laquelle vous souhaitez migrer sa VM secondaire.

2 Cliquez avec le bouton droit sur la machine virtuelle et sélectionnez **Fault Tolerance > Migration secondaire**.

3 Remplissez les options de la boîte de dialogue Migrer et validez les changements que vous faites.

4 Cliquez sur **Terminer** pour appliquer les modifications.

La VM secondaire associée à la machine virtuelle insensible aux défaillances sélectionnée est migrée vers l'hôte spécifié.

Tester le basculement

Vous pouvez provoquer une situation de basculement pour une VM principale sélectionnée afin de tester la protection de tolérance aux pannes.

Cette option est indisponible (grisée) si la VM est mise sous tension.

Procédure

1 Dans vSphere Web Client accédez à la VM primaire pour laquelle vous souhaitez tester le basculement.

2 Cliquez avec le bouton droit sur la machine virtuelle et sélectionnez **Fault Tolerance > Tester le basculement**.

3 Consultez les détails sur le basculement dans la console de travail.

Cette tâche provoque la défaillance de la VM principale afin de s'assurer que la VM secondaire la remplace. Une nouvelle VM secondaire est également démarrée, pour replacer la VM principale dans un état protégé.

Tester le redémarrage secondaire

Vous pouvez provoquer la défaillance d'une VM secondaire afin de tester la protection Tolérance aux pannes fournie pour une VM principale sélectionnée.

Cette option est indisponible (grisée) si la VM est mise sous tension.

Procédure

1 Dans vSphere Web Client, accédez à la VM primaire pour laquelle vous souhaitez effectuer le test.

2 Cliquez avec le bouton droit sur la machine virtuelle et sélectionnez **Fault Tolerance > Tester le redémarrage secondaire**.

3 Consultez les détails du test dans la Console des tâches

Cette tâche a pour conséquence l'arrêt de la VM secondaire qui assurait la protection Tolérance aux pannes pour la VM principale sélectionnée. Une nouvelle VM secondaire est alors démarrée, replaçant la VM principale dans un état protégé.

Mettre à niveau les hôtes utilisés pour Fault Tolerance

Procédez comme suit pour mettre à niveau les hôtes utilisés pour Fault Tolerance.

Prérequis

Vérifiez que vous possédez des privilèges d'administrateur sur les clusters.

Vérifiez que vous possédez des ensembles d'au moins quatre hôtes ESXi hébergeant des machines virtuelles tolérantes aux pannes qui sont sous tension. Si les machines virtuelles sont hors tension, les machines virtuelles principales et secondaires tolérantes aux pannes peuvent être déplacées sur des hôtes de versions différentes.

Remarque Cette procédure de mise à niveau est adaptée aux clusters de quatre nœuds au minimum. Les mêmes instructions peuvent être suivies avec un plus petit cluster, mais les intervalles sans protection seront légèrement plus longs.

Procédure

1 Avec vMotion, migrez les machines virtuelles tolérantes aux pannes à partir des deux hôtes.

2 Mettez à niveau les deux hôtes évacués de façon à ce qu'ils aient la même version d'ESXi.

3 Interrompez Fault Tolerance sur la machine virtuelle principale.

4 Avec vMotion, déplacez la machine virtuelle principale pour laquelle Fault Tolerance a été interrompue vers l'un des hôtes mis à niveau.

5 Reprenez Fault Tolerance sur la machine virtuelle principale qui a été déplacée.

6 RépétezÉtape 1 à Étape 5 pour autant de paires de machines virtuelles tolérantes aux pannes que les hôtes mis à niveau peuvent en accueillir.

7 Avec vMotion, répartissez les machines virtuelles tolérantes aux pannes.

Tous les hôtes ESXi d'un cluster sont mis à niveau.

Pratiques d'excellence pour Fault Tolerance

Pour garantir des résultats Fault Tolerance optimaux, vous devez respectez certaines meilleures pratiques.

Les recommandations suivantes concernant la configuration de l'hôte et de la mise en réseau peut améliorer la stabilité et les performances de votre cluster.

Configuration d'hôte

Les hôtes exécutant les machines virtuelles principales et secondaires doivent fonctionner à des fréquences de processeur assez proches sinon la machine virtuelle secondaire risque de redémarrer plus souvent. Les fonctions de gestion de l'alimentation de la plate-forme qui ne sont pas réglées selon la charge de travail (modes de limitation de puissance et de basse fréquence pour économiser de l'énergie, par exemple) peuvent entraîner de fortes variations des fréquences du processeur. Si des machines virtuelles secondaires sont redémarrées régulièrement, désactivez tous les modes de gestion de l'alimentation sur les hôtes exécutant des machines virtuelles tolérantes aux pannes ou veillez à ce que tous les hôtes soient exécutés avec les même modes de gestion de l'alimentation.

Configuration de la mise en réseau des hôtes

Les directives suivantes vous permettent de configurer la mise en réseau des hôtes pour la prise en charge de Fault Tolerance avec différentes combinaisons de types de trafic (par exemple, NFS) et plusieurs adaptateurs réseau physiques.

- Répartissez chaque association de adaptateurs réseau sur deux commutateurs physiques assurant la continuité des domaines L2 pour chaque VLAN entre les deux commutateurs physiques.
- Utilisez des règles d'association déterministe pour vous assurer que des types de trafic particuliers présentent une affinité avec une carte réseau particulière (active/veille) ou un ensemble de adaptateurs réseau (par exemple, ID port virtuel d'origine).

- Quand des règles active/veille sont utilisées, associez les types de trafic pour réduire les répercussions dans le cas de basculement où les deux types de trafic partagent un vmnic.
- Quand des règles active/veille sont utilisées, configurez tous les adaptateurs actifs pour un type de trafic particulier (par exemple, journalisation de la tolérance aux pannes) sur le même commutateur physique. Cela réduit le nombre de bonds réseau et diminue les possibilités de surabonner le commutateur à des liaisons de commutateurs.

Remarque Le trafic de la journalisation de la tolérance aux pannes entre les machines virtuelles primaires et secondaires est chiffré et contient un réseau client et des données E/S de stockage, ainsi que le contenu de la mémoire du système d'exploitation client. Ce trafic peut inclure des données sensibles telles que des mots de passe en texte brut. Pour éviter que ces données ne soient divulguées, assurez-vous que ce réseau est sécurisé, notamment pour éviter les « attaques de l'intercepteur ». Par exemple, vous pourriez utiliser un réseau privé pour le trafic de la journalisation de la tolérance aux pannes.

Clusters homogènes

vSphere Fault Tolerance peut fonctionner dans des clusters contenant des hôtes non uniformes, mais il est préférable que les clusters aient des nœuds compatibles. Au moment de la construction du cluster, tous les hôtes doivent être configurés comme suit :

- Accès commun aux banques de données utilisées par les machines virtuelles.
- La même configuration réseau de machines virtuelles.
- Les mêmes paramètres de BIOS (gestion de l'alimentation et hyperthreading) pour tous les hôtes.

Exécutez **Vérifier la conformité** pour identifier les incompatibilités et les corriger.

Performances

Pour accroître la bande passante disponible pour le trafic de journalisation entre les machines virtuelles principales et secondaires, utilisez une carte réseau de 10 Gbit et activez l'utilisation des Trames jumbo.

Vous pouvez sélectionner plusieurs cartes réseau pour le réseau de journalisation FT. En sélectionnant plusieurs cartes réseau, vous pouvez tirer parti de la bande passante de plusieurs cartes réseau, même si toutes les cartes réseau ne sont pas dédiées à l'exécution de FT.

Stocker les images ISO sur des stockages partagés pour un accès permanent

Les images ISO auxquelles accèdent les machines virtuelles dont Fault Tolerance est activée doivent être conservées sur des stockages partagés accessibles aux deux instances de la machine virtuelle tolérante aux pannes. Si vous utilisez cette configuration, le CD-ROM présent dans la machine virtuelle continue de fonctionner correctement, même en cas de basculement.

Pour les machines virtuelles dont Fault Tolerance est activée, il est possible d'utiliser les images ISO qui sont uniquement accessibles par la machine virtuelle principale. Dans ce cas, la machine virtuelle principale peut accéder à l'image ISO, mais en cas de basculement, le CD-ROM signale les erreurs comme s'il n'y avait pas de support. Cette situation peut être tolérée si le CD-ROM est utilisé pour une opération provisoire et non critique comme un correctif.

Éviter les partitions de réseau

Une partition de réseau survient quand un cluster vSphere HA connaît une défaillance du réseau de gestion qui isole certains hôtes de vCenter Server et les isole les uns des autres. Reportez-vous à « Partitions de réseau », page 22. En cas de partition, la protection de Fault Tolerance peut être réduite.

Dans un cluster vSphere HA partitionné utilisant Fault Tolerance, la machine virtuelle principale (ou sa machine virtuelle secondaire) pourrait se retrouver dans une partition gérée par un hôte principal qui n'est pas responsable de cette machine virtuelle. Si un basculement est nécessaire, une machine virtuelle secondaire est redémarrée uniquement si la machine virtuelle principale se trouvait dans une partition gérée par un hôte principal qui en était responsable.

Pour réduire les risques de panne de votre réseau de gestion entraînant une partition du réseau, suivez les recommandations figurant dans « Meilleures pratiques pour la mise en réseau », page 45.

Utilisation de banques de données Virtual SAN

vSphere Fault Tolerance peut utiliser des banques de données Virtual SAN, mais vous devez observer les restrictions suivantes :

- Un mélange de Virtual SAN et d'autres types de banques de données n'est pas pris en charge pour les VM principales et les VM secondaires.
- Les metro-clusters Virtual SAN ne sont pas pris en charge avec FT.

Pour augmenter les performances et la fiabilité lors de l'utilisation de FT avec Virtual SAN, les conditions suivantes sont également recommandées.

- Virtual SAN et FT doivent utiliser des réseaux distincts.
- Maintenez les VM principales et secondaires dans des domaines de pannes Virtual SAN distincts.

Fault Tolerance héritée

Par défaut, vSphere Fault Tolerance peut gérer les machines virtuelles à multiprocesseur symétrique (SMP) avec jusqu'à quatre vCPU. Si votre machine virtuelle dispose d'un vCPU unique, il vous est toutefois possible d'utiliser la fonctionnalité FT héritée à la place de la compatibilité descendante. Si la fonctionnalité FT héritée n'est pas nécessaire d'un point de vue technique, évitez de l'utiliser.

Pour utiliser la fonctionnalité Fault Tolerance héritée, vous devez configurer une option avancée pour la machine virtuelle. Une fois cette configuration terminée, la machine virtuelle avec FT héritée est quelque peu différente des autres machines virtuelles vSphere FT.

Différences des machines virtuelles avec FT héritée

Les machines virtuelles qui utilisent vSphere FT et celles qui utilisent FT héritée diffèrent de plusieurs manières.

Tableau 3-2. Différences entre FT héritée et vSphere FT

	FT héritée	vSphere FT
Extended Page Tables/Rapid Virtualization Indexing (EPT/RVI).	Non pris en charge	Requis
IPv6	Non pris en charge pour les cartes réseau de journalisation avec FT héritée.	Pris en charge pour les cartes réseau de journalisation avec vSphere FT.
DRS	Pris entièrement en charge pour le placement initial, l'équilibrage de charge et la prise en charge du mode de maintenance.	Prise en charge uniquement de la mise sous tension lors du placement de la machine virtuelle secondaire et du mode de maintenance.
API vStorage - sauvegarde de la protection des données	Non pris en charge	Pris en charge
Fichiers des disques .vmdk au format Thick eager-zeroed	Requis	Non requis, car vSphere FT prend en charge tous les types de disques, y compris les disques dynamiques et statiques.

Tableau 3-2. Différences entre FT héritée et vSphere FT (suite)

	FT héritée	vSphere FT
Redondance des fichiers .vmdk	Seulement une copie unique	Les machines virtuelles principales et les machines virtuelles secondaires conservent toujours des copies distinctes qui peuvent être placées dans des différentes banques de données afin d'augmenter la redondance.
Bande passante de la carte réseau	Carte réseau dédiée de 1 Go recommandée	Carte réseau dédiée de 10 Go recommandée
Compatibilité du CPU et de l'hôte	Nécessite un modèle et une famille de CPU identiques et des versions presque identiques de vSphere sur les hôtes.	Les CPU doivent être compatibles avec vSphere vMotion ou EVC. Les versions de vSphere sur les hôtes doivent être compatibles avec vSphere vMotion.
Activer FT sur les machines virtuelles en cours d'exécution	Pas toujours pris en charge. Il se peut que vous deviez d'abord mettre hors tension la machine virtuelle.	Pris en charge
Storage vMotion	Pris en charge uniquement sur les machines virtuelles hors tension. vCenter Server met automatiquement FT hors tension avant d'exécuter une action de Storage vMotion, puis remet FT sous tension une fois l'action de Storage vMotion terminée.	Non pris en charge. L'utilisateur doit mettre vSphere FT hors tension sur la machine virtuelle avant d'exécuter l'action de Storage vMotion, puis remettre vSphere FT sous tension.
Pilotes de mise en réseau vlance	Non pris en charge	Pris en charge

Conditions requises supplémentaires pour la fonctionnalité FT héritée

Outre les différences répertoriées concernant la fonctionnalité FT héritée, celle-ci est soumise aux conditions uniques suivantes.

- Les hôtes ESXi doivent avoir accès aux mêmes banques de données et réseaux des machines virtuelles.
- Les machines virtuelles doivent être conservées dans des fichiers de RDM virtuel ou de disque de machine virtuelle (VMDK) qui sont approvisionnés en lourd. Lorsqu'une machine virtuelle est stockée dans un fichier VMDK qui est provisionné dynamiquement et que vous tentez d'utiliser la fonctionnalité FT, un message s'affiche. Il indique que le fichier VMDK doit être converti. Vous devez mettre hors tension la machine virtuelle pour exécuter la conversion.
- Les hôtes doivent avoir des processeurs appartenant au groupe de processeurs compatibles avec vSphere FT. Vérifiez que les processeurs des hôtes sont compatibles les uns avec les autres.
- L'hôte qui prend en charge la machine virtuelle secondaire doit avoir un processeur qui prend en charge la tolérance de panne et dont la famille ou le modèle de CPU est le même que l'hôte qui prend en charge la machine virtuelle principale.
- Lorsque vous mettez à niveau des hôtes qui contiennent des machines virtuelles avec tolérance de panne, vérifiez que les machines virtuelles principales et secondaires continuent de s'exécuter sur des hôtes ayant le même numéro de version FT ou de build d'hôte. Cette exigence s'applique aux hôtes antérieurs à ESX/ESXi 4.1.

Remarque Si vous avez désigné une machine virtuelle devant utiliser FT avant de mettre à niveau les hôtes dans le cluster, celle-ci continuera d'utiliser la fonctionnalité FT héritée après la mise à niveau de l'hôte.

Mise à niveau des hôtes impliqués avec la fonctionnalité FT héritée

Pour mettre à niveau vos hôtes vers vSphere 6.5 ou une version ultérieure, vous devez désactiver FT héritée sur toutes les machines virtuelles concernées ou déplacer ces machines virtuelles vers d'autres hôtes. Si vous ne préparez pas la mise à niveau de cette manière, VMware vSphere Update Manager bloque la mise à niveau.

vCenter Server version 6.5 ou ultérieure peut gérer les machines virtuelles avec FT héritée, mais ne peut pas les créer, même sur les hôtes avec une version antérieure à la version 6.5. Les opérations vSphere FT suivantes peuvent être effectuées avec ce scénario :

- Interrompre ou relancer FT
- Basculement test
- Redémarrer la VM secondaire
- Migrer la VM secondaire
- Désactiver FT

Remarque Les machines virtuelles avec FT héritée existent uniquement sur les hôtes ESXi qui sont exécutés sur des versions de vSphere antérieures à la version 6.5.

vCenter High Availability 4

vCenter High Availability (vCenter HA) protège vCenter Server Appliance contre les défaillances matérielles et de l'hôte. L'architecture active-passive de la solution peut également vous aider à réduire considérablement les temps d'arrêt lorsque vous appliquez un correctif à vCenter Server Appliance.

Après avoir procédé à la configuration du réseau, vous créez un cluster à trois nœuds qui contient les nœuds actif, passif et témoin. Différents chemins de configuration sont possibles. Ce que vous sélectionnez dépend de votre configuration existante.

1 Planifier le déploiement de vCenter HA page 68

Avant de configurer vCenter HA, vous devez prendre en compte plusieurs facteurs. Un déploiement vCenter Server Appliance peut recourir à une instance interne ou externe de Platform Services Controller. Un déploiement dans un environnement établi, avec des composants utilisant différentes versions de vSphere, nécessite une préparation différente d'un déploiement dans un environnement utilisant exclusivement des composants vSphere 6.5. Les ressources et les logiciels nécessaires, ainsi que la configuration du réseau, doivent également faire l'objet d'une préparation soigneuse.

2 Configurer le réseau page 73

Quelles que soient l'option de déploiement et la hiérarchie d'inventaire sélectionnée, vous devez définir votre réseau avant de commencer la configuration. Pour définir les fondations du réseau vCenter HA, vous devez ajouter un groupe de ports à chaque hôte ESXi, puis ajouter une carte réseau virtuelle à vCenter Server Appliance, ce dernier devenant le nœud actif.

3 Configurer vCenter HA avec l'option Basique page 74

Lorsque vous utilisez l'option Basique, l'assistant vCenter HA crée et configure un deuxième adaptateur réseau sur vCenter Server Appliance, clone le nœud Actif et configure le réseau vCenter HA.

4 Configurer vCenter HA avec l'option Avancé page 75

La configuration du cluster vCenter HA avec l'option Avancé offre un contrôle accru de l'environnement et permet de contourner les conditions requises par la configuration Basique. Cependant, vous devez ajouter une deuxième carte réseau dans vCenter Server Appliance, cloner les nœuds actif, passif et témoin, et configurer les clones.

5 Gérer la configuration vCenter HA page 78

Après avoir configuré votre cluster vCenter HA, vous pouvez effectuer les tâches de gestion. Ces tâches incluent le remplacement de certificats, le remplacement des clés SSH et la configuration SNMP. Vous pouvez également modifier la configuration du cluster pour désactiver ou activer vCenter HA, passer en mode de maintenance et supprimer la configuration du cluster.

6 Corriger votre environnement vCenter HA page 84

En cas de problème, vous pouvez corriger votre environnement. La tâche que vous devez effectuer dépend des symptômes de la défaillance. Pour en savoir plus sur le dépannage, reportez-vous au système de la base de connaissances VMware.

7 Application de correctifs à un environnement vCenter High Availability page 87

Vous pouvez appliquer un correctif à un dispositif vCenter Server Appliance situé dans un cluster vCenter High Availability à l'aide de l'utilitaire <codeph>software-packages</codeph> disponible dans l'interpréteur du dispositif vCenter Server Appliance. Pour plus d'informations, reportez-vous à *Mise à niveau de vSphere*.

Planifier le déploiement de vCenter HA

Avant de configurer vCenter HA, vous devez prendre en compte plusieurs facteurs. Un déploiement vCenter Server Appliance peut recourir à une instance interne ou externe de Platform Services Controller. Un déploiement dans un environnement établi, avec des composants utilisant différentes versions de vSphere, nécessite une préparation différente d'un déploiement dans un environnement utilisant exclusivement des composants vSphere 6.5. Les ressources et les logiciels nécessaires, ainsi que la configuration du réseau, doivent également faire l'objet d'une préparation soigneuse.

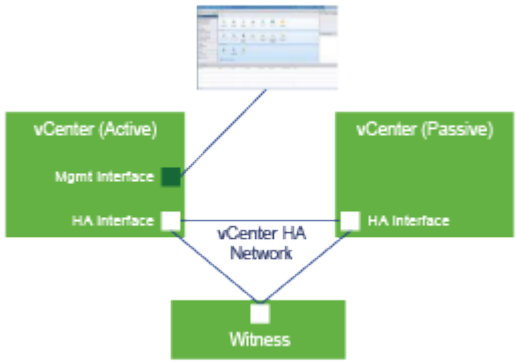
Vue d'ensemble de l'architecture de vCenter

Un cluster vCenter HA comprend trois instances de vCenter Server Appliance. La première instance, utilisée à l'origine comme un nœud actif, est clonée deux fois sur un nœud passif et un nœud témoin. Ensemble, ces trois nœuds forment une solution de basculement active-passive.

Le déploiement de chacun des nœuds sur une instance ESXi différente permet de se protéger contre les pannes matérielles. L'ajout des trois hôtes ESXi à un cluster DRS permet de mieux protéger votre environnement.

Une fois la configuration de vCenter HA terminée, seul le nœud actif dispose d'une interface de gestion active (IP public). Les trois nœuds communiquent sur un réseau privé appelé vCenter HA et qui a été défini lors de la configuration. Les nœuds actif et passif répliquent les données continuellement.

Figure 4-1. Cluster vCenter à trois nœuds



Les trois nœuds sont nécessaires pour que cette fonctionnalité fonctionne. Comparez les responsabilités des nœuds.

Tableau 4-1. Nœuds vCenter HA

Nœud	Description
Active	■ Exécute l'instance vCenter Server Appliance active ■ Utilise une adresse IP publique pour l'interface de gestion ■ Utilise le réseau vCenter HA pour la réplication des données sur le nœud passif. ■ Utilise le réseau vCenter HA pour communiquer avec le nœud témoin.
Passif	■ Est initialement un clone du nœud actif ■ Il reçoit constamment des mises à jour du nœud actif et synchronise son état avec ce dernier sur le réseau vCenter HA ■ Prend automatiquement le relais en tant que nœud actif en cas de panne
Témoin	■ Clone léger du nœud actif ■ Fournit un quorum afin d'assurer une protection en cas de division

Configurations matérielle et logicielle requises de vCenter HA

Avant de configurer vCenter HA, assurez-vous que vous disposez de suffisamment de ressources de mémoire, de CPU et de banques de données. Assurez-vous également que vous utilisez des versions de vCenter Server et ESXi qui prennent en charge vCenter HA.

Votre environnement doit répondre aux exigences suivantes.

Tableau 4-2. Configuration requise de vCenter HA

Composant	Configuration requise
ESXi	■ ESXi 5.5 ou version plus récente est requis. ■ Trois hôtes sont fortement recommandés. Chaque nœud vCenter HA peut ensuite s'exécuter sur un hôte différent pour une meilleure protection. ■ Il est recommandé d'utiliser VMware DRS pour protéger cet ensemble d'hôtes. Dans ce cas, un minimum de trois hôtes ESXi est requis.
vCenter Serverde gestion (si utilisé)	Votre environnement peut inclure un système de gestion de vCenter Server ou vous pouvez configurer votre instance de vCenter Server Appliance pour gérer l'hôte de ESXi sur lequel elle s'exécute (vCenter Server à gestion automatique) ■ vCenter Server 5.5 ou version plus récente est requis.
vCenter Server Appliance	■ vCenter Server 6.5 est requis. ■ Une taille de déploiement Petite (4 CPU et 16 Go de RAM) ou plus importante est requise pour atteindre l'objectif de temps de récupération. N'utilisez pas la version Minuscule dans les environnements de production. ■ vCenter HA est pris en charge et testé avec les banques de données VMFS, NFS et Virtual SAN. ■ Assurez-vous que vous disposez d'un espace disque suffisant pour collecter et stocker les bundles de support de ces trois nœuds sur le nœud actif. Reportez-vous à « Collecter des bundles de support pour un nœud vCenter HA », page 84.
Connexion réseau	■ La latence réseau de vCenter HA entre les nœuds actif, passif et témoin doit être inférieure à 10 ms. ■ Le réseau vCenter HA doit être sur un sous-réseau différent du réseau de gestion.
Gestion des licences requises pour vCenter HA	■ vCenter HA requiert une licence unique vCenter Server. ■ vCenter HA requiert une licence Standard.

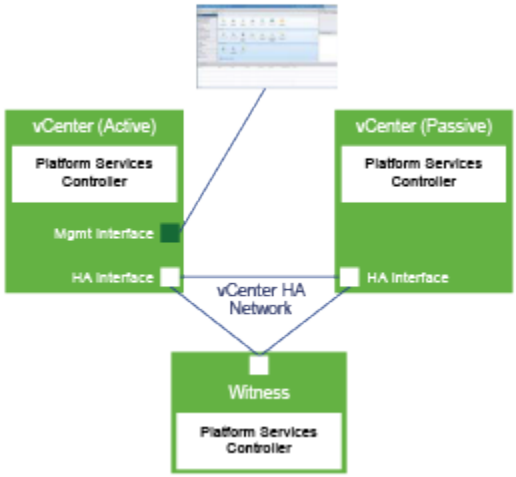
Options de déploiement de vCenter HA

Vous pouvez configurer votre environnement vCenter HA avec une instance intégrée de Platform Services Controller ou une instance externe de Platform Services Controller. Si vous décidez d'utiliser une instance externe de Platform Services Controller, vous pouvez la placer derrière un équilibreur de charge comme mesure de protection en cas de panne de Platform Services Controller.

vCenter HA avec une instance intégrée de Platform Services Controller

Lorsque vous utilisez vCenter HA avec une instance intégrée de Platform Services Controller, la configuration de l'environnement se présente comme suit.

Figure 4-2. vCenter HA avec une instance intégrée de Platform Services Controller



1 L'utilisateur provisionne vCenter Server Appliance avec une instance intégrée de Platform Services Controller.

2 Le clonage de vCenter Server Appliance vers un nœud passif et un nœud témoin s'effectue.

   - Dans une configuration de base, les clones sont créés et configurés automatiquement.
   - Dans une configuration avancée, les clones sont créés et configurés par l'utilisateur.

3 Dans le cadre du processus de clonage, Platform Services Controller et tous ses services sont également clonés.

4 Une fois la configuration terminée, vCenter HA procède à la réplication pour assurer la synchronisation du nœud passif avec le nœud actif. La réplication du nœud actif vers le nœud passif comprend les données de Platform Services Controller.

5 Une fois la configuration terminée, vCenter Server Appliance est protégé par vCenter HA. En cas de basculement, Platform Services Controller et tous ses services sont disponibles sur le nœud passif.

vCenter HA avec une instance externe de Platform Services Controller

Lorsque vous utilisez vCenter HA avec une instance externe de Platform Services Controller, vous devez configurer un équilibreur de charge externe pour protéger Platform Services Controller. Si une instance de Platform Services Controller devient indisponible, l'équilibreur de charge redirige vCenter Server Appliance vers une autre instance de Platform Services Controller.

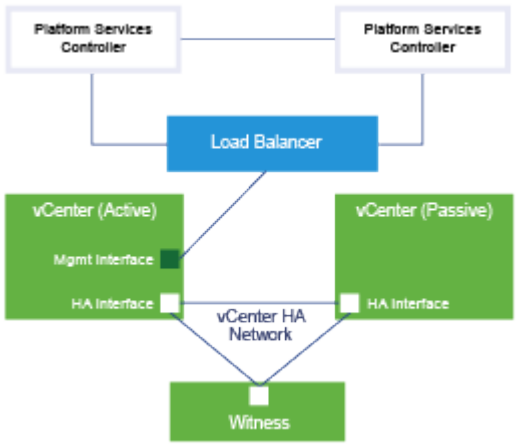
La configuration d'une instance externe de Platform Services Controller est traitée dans les articles de la base de connaissances de VMware suivants.

- 2147014: Configuring Netscaler Load Balancer for use with vSphere Platform Services Controller (PSC) 6.5

- 2147038 Configuring F5 BIG-IP Load Balancer for use with vSphere Platform Services Controller (PSC) 6.5
- 2147046 Configuring NSX Edge Load Balancer for use with vSphere Platform Services Controller (PSC) 6.5

La configuration de l'environnement se présente comme suit.

Figure 4-3. vCenter HA avec une instance externe de Platform Services Controller



1 L'utilisateur configure au moins deux instances de Platform Services Controller. Ces instances répliquent les informations de vCenter Single Sign-On et d'autres informations de Platform Services Controller, par exemple, les licences.

2 Pendant le provisionnement de vCenter Server Appliance, l'utilisateur sélectionne une instance externe de Platform Services Controller.

3 L'utilisateur configure vCenter Server Appliance de façon à pointer vers un équilibreur de charge qui offre la haute disponibilité pour Platform Services Controller.

4 L'utilisateur ou la configuration de base clone la première instance de vCenter Server Appliance pour créer un nœud passif et un nœud témoin.

5 Dans le cadre du processus de clonage, les informations sur l'instance externe de Platform Services Controller et l'équilibreur de charge sont également clonées.

6 Une fois la configuration terminée, vCenter Server Appliance est protégé par vCenter HA.

7 Si l'instance de Platform Services Controller devient indisponible, l'équilibreur de charge redirige les demandes d'authentification ou d'autres services à la deuxième instance de Platform Services Controller.

Présentation du workflow de configuration

Vous avez le choix entre une configuration de base ou avancée. L'option De base crée automatiquement les nœuds passifs et témoins dans le cadre de la configuration de vCenter HA. Avec l'option Avancée, il vous incombe de cloner manuellement le nœud actif afin de créer les nœuds passifs et témoins.

Votre environnement vous dicte l'option de configuration que vous choisissez. Les exigences de la configuration de base sont plus strictes, mais une plus grande partie de la configuration est automatique. Vous pouvez choisir une configuration avancée si votre environnement respecte les exigences matérielles et logicielles. Cette configuration offre plus de flexibilité. Cependant, la configuration avancée vous impose de créer et de configurer les clones du nœud actif.

Workflow de la configuration de base

La configuration de base clone automatiquement le nœud actif. Pour effectuer une configuration de base, vous devez respecter l'une des exigences suivantes.

- Soit l'instance de vCenter Server Appliance, qui deviendra le nœud actif, gère son propre hôte ESXi et sa propre machine virtuelle. Cette configuration de vCenter Server est parfois appelée gestion automatique.
- Ou le dispositif vCenter Server Appliance est géré par une autre instance de vCenter Server (vCenter Server de gestion) et les deux instances de vCenter Server se trouvent dans le même domaine vCenter Single Sign-On. Cela implique que toutes deux utilisent un dispositif Platform Services Controller externe et qu'elles exécutent toutes deux vSphere 6.5.

Si vous respectez les exigences, le workflow de base est le suivant.

1 L'utilisateur déploie le premier dispositif vCenter Server Appliance qui deviendra le nœud actif.

2 L'utilisateur ajoute un second réseau (groupe de ports) pour le trafic vCenter HA sur chaque hôte ESXi.

3 L'utilisateur commence à configurer vCenter HA, sélectionne l'option De base et fournit les adresses IP, l'hôte ESXi cible ou le cluster, et la banque de données pour chaque clone.

4 Le système clone le nœud actif et crée un nœud passif ayant exactement les mêmes paramètres, y compris le même nom d'hôte.

5 Le système clone à nouveau le nœud actif et crée un nœud témoin plus léger.

6 Le système configure le réseau vCenter HA sur lequel les trois nœuds communiquent, par exemple en échangeant des signaux de pulsation et d'autres informations.

Pour obtenir des informations étape par étape, reportez-vous à la section « Configurer vCenter HA avec l'option Basique », page 74.

Workflow de la configuration avancée

Si vous ne pouvez pas sélectionner l'option De base, ou si vous souhaitez disposer d'un contrôle plus étendu sur votre déploiement, vous pouvez effectuer une configuration avancée. Avec cette option, il vous incombe de cloner vous-même le nœud actif dans le cadre de la configuration de vCenter HA. Si vous sélectionnez cette option et supprimez ultérieurement la configuration de vCenter HA, il vous incombe de supprimer les nœuds que vous avez créés.

Le workflow de l'option Avancée est le suivant.

1 L'utilisateur déploie le premier dispositif vCenter Server Appliance qui deviendra le nœud actif.

2 L'utilisateur ajoute un second réseau (groupe de ports) pour le trafic vCenter HA sur chaque hôte ESXi.

3 L'utilisateur ajoute un second adaptateur réseau (NIC) au nœud actif.

4 L'utilisateur se connecte au dispositif vCenter Server Appliance (nœud actif) à l'aide de vSphere Web Client.

5 L'utilisateur démarre la configuration de vCenter HA, sélectionne l'option Avancée et fournit les adresses IP et les informations de sous-réseau pour les nœuds passifs et témoins. Éventuellement, l'utilisateur peut remplacer les adresses IP de gestion du basculement.

6 L'utilisateur se connecte à l'instance vCenter Server de gestion et crée deux clones du dispositif vCenter Server Appliance (nœud actif).

7 L'utilisateur revient à l'assistant de configuration du dispositif vCenter Server Appliance et achève le processus de configuration.

8 Le système configure le réseau vCenter HA sur lequel les trois nœuds échangent des signaux de pulsation et des informations de réplication.

9 Le dispositif vCenter Server Appliance est protégé par vCenter HA.

Reportez-vous à « Configurer vCenter HA avec l'option Avancé », page 75 pour plus de détails.

Configurer le réseau

Quelles que soient l'option de déploiement et la hiérarchie d'inventaire sélectionnée, vous devez définir votre réseau avant de commencer la configuration. Pour définir les fondations du réseau vCenter HA, vous devez ajouter un groupe de ports à chaque hôte ESXi, puis ajouter une carte réseau virtuelle à vCenter Server Appliance, ce dernier devenant le nœud actif.

Une fois la configuration terminée, le cluster vCenter HA a deux réseaux, le réseau de gestion sur la première carte réseau virtuelle et le réseau vCenter HA sur la deuxième carte réseau virtuelle.

Réseau de gestion Le réseau de gestion sert les demandes des clients (IP public). Les adresses IP du réseau de gestion doivent être statiques.

Réseau vCenter HA Le réseau vCenter HA connecte les nœuds actif, passif et témoin et réplique l'état du dispositif. Il surveille également les pulsations.

- Les adresses IP du réseau vCenter HA pour les nœuds actif, passif et témoin doivent être statiques.
- Le réseau vCenter HA doit être sur un sous-réseau différent du réseau de gestion. Les trois nœuds peuvent être sur le même sous-réseau ou sur des sous-réseaux différents.
- La latence du réseau entre les nœuds actif, passif et témoin doit être inférieure à 10 millisecondes.
- Vous ne devez pas ajouter d'entrée de passerelle par défaut pour le réseau du cluster.

Prérequis

- vCenter Server Appliance qui devient ensuite le nœud actif, est déployé.
- Vous pouvez y accéder et disposez des privilèges de modification de vCenter Server Appliance et de l'hôte ESXi sur lequel il s'exécute.
- Pendant la configuration du réseau, vous devez utiliser des adresses IP statiques pour le réseau de gestion. Les adresses des réseaux de gestion et du cluster doivent être de type IPv4 ou IPv6. Vous ne pouvez pas les mélanger.

Procédure

1 Connectez-vous au vCenter Server de gestion et trouvez l'hôte ESXi sur lequel le nœud actif s'exécute.

2 Ajoutez un groupe de ports à l'hôte ESXi.

Ce groupe de ports peut figurer sur un commutateur virtuel existant ou, pour une meilleure isolation réseau, vous pouvez créer un nouveau commutateur virtuel. Il doit figurer sur un autre sous-réseau que le réseau de gestion sur Eth0.

3 Si votre environnement inclut les trois hôtes ESXi recommandés, ajoutez le groupe de ports à chacun des hôtes.

Suivant

Les actions à réaliser ensuite dépendent du type de configuration sélectionné.

- Avec une configuration Basique, l'assistant crée la carte réseau virtuelle vCenter HA sur chaque clone et configure le réseau vCenter HA. Lorsque la configuration est terminée, le réseau vCenter HA devient disponible pour la réplication et le trafic des pulsations.

- Avec une configuration Avancée.
  - Vous devez commencer par créer et configurer une deuxième carte réseau sur le nœud actif. Reportez-vous à « Créer et configurer une deuxième carte réseau sur vCenter Server Appliance », page 76.
  - Lorsque vous exécutez la configuration, l'assistant demande les adresses IP des nœuds passif et témoin.
  - L'assistant vous demande de cloner le nœud actif. Dans le cadre du processus de clonage, vous exécutez des configurations réseau supplémentaires.

  Reportez-vous à « Configurer vCenter HA avec l'option Avancé », page 75.

Configurer vCenter HA avec l'option Basique

Lorsque vous utilisez l'option Basique, l'assistant vCenter HA crée et configure un deuxième adaptateur réseau sur vCenter Server Appliance, clone le nœud Actif et configure le réseau vCenter HA.

Prérequis

- Déployez l'instance de vCenter Server Appliance à utiliser en tant que nœud Actif initial.
  - Une adresse IP statique doit être mappée vers un nom de domaine complet pour l'instance de vCenter Server Appliance.
  - Le mode SSH doit être activé sur vCenter Server Appliance.
- Assurez-vous que votre environnement répond à au moins l'un des critères suivants.
  - Soit l'instance de vCenter Server Appliance, qui deviendra le nœud actif, gère son propre hôte ESXi et sa propre machine virtuelle. Cette configuration de vCenter Server est parfois appelée gestion automatique.
  - Ou vCenter Server Appliance est géré par une autre instance de vCenter Server (vCenter Server de gestion) et les deux dispositifs appartiennent au même domaine vCenter Single Sign-On. Cela implique que toutes deux utilisent un dispositif Platform Services Controller externe et qu'elles exécutent toutes deux vSphere 6.5.

  Si votre environnement ne correspond à aucun de ces deux critères, exécutez une configuration avancée. Reportez-vous à « Configurer vCenter HA avec l'option Avancé », page 75.
- Configurez l'infrastructure du réseau vCenter HA. Reportez-vous à « Configurer le réseau », page 73.
- Identifiez les adresses IP statiques à utiliser pour les deux nœuds vCenter Server Appliance qui deviendront les nœuds passif et témoin.

Procédure

1. Connectez-vous au nœud actif à l'aide de vSphere Web Client.
2. Cliquez avec le bouton droit sur l'objet vCenter Server dans l'inventaire et sélectionnez **Paramètres vCenter HA**.
3. Cliquez sur **Configurer**.
4. Sélectionnez l'option de configuration **Basique** et cliquez sur **Suivant**.

   Cette option n'est disponible que si votre environnement dispose des conditions requises pour l'option Basique.
5. Saisissez l'adresse IP, le masque de sous-réseau du nœud actif et le groupe de ports nécessaires pour la connexion au réseau vCenter HA, puis cliquez sur **Suivant**.

6 Entrez l'adresse IP du réseau vCenter HA, ainsi que le masque de sous-réseau des nœuds passif et témoin, puis cliquez sur **Suivant**.

L'assistant de configuration nécessite ces adresses pour créer le réseau vCenter HA et pour connecter les trois nœuds.

7 (Facultatif) Cliquez sur **Avancé** si vous souhaitez remplacer l'adresse IP de gestion du basculement du nœud passif.

8 Vérifiez les informations des nœuds passif et témoin, cliquez sur **Modifier** pour apporter des modifications, puis cliquez sur **Suivant**.

Si vous n'utilisez pas un cluster DRS, sélectionnez des hôtes et des banques de données différentes pour les nœuds passif et témoin, si possible.

9 Cliquez sur **Terminer** .

Les nœuds passif et témoin sont créés. Une fois la configuration de vCenter HA terminée, vCenter Server Appliance dispose de la protection haute disponibilité.

Suivant

Reportez-vous à « Gérer la configuration vCenter HA », page 78 pour consulter la liste des tâches de gestion de cluster.

Configurer vCenter HA avec l'option Avancé

La configuration du cluster vCenter HA avec l'option Avancé offre un contrôle accru de l'environnement et permet de contourner les conditions requises par la configuration Basique. Cependant, vous devez ajouter une deuxième carte réseau dans vCenter Server Appliance, cloner les nœuds actif, passif et témoin, et configurer les clones.

Procédure

1 Créer et configurer une deuxième carte réseau sur vCenter Server Appliance page 76

Avant de pouvoir lancer la configuration avancée, vous devez créer et configurer une deuxième carte réseau sur l'instance de vCenter Server Appliance qui deviendra le nouveau nœud actif. Cette carte réseau sera utilisée pour le trafic vCenter HA. Cette tâche doit être réalisée après la configuration du réseau, mais avant le début du processus de configuration.

2 Lancer le processus de configuration avancée page 76

Après avoir configuré le réseau et ajouté une deuxième carte réseau à vCenter Server Appliance, vous pouvez lancer le processus de configuration de vCenter HA.

3 Créer et configurer les clones du nœud actif page 77

Lors de la configuration avancée, vous devez cloner le nœud actif pour créer les nœuds passif et témoin. Ne quittez pas l'assistant Configurer vCenter HA pendant l'exécution des tâches de clonage.

4 Terminer la configuration avancée de vCenter HA page 78

Après avoir créé les nœuds passif et témoin, revenez à l'assistant de configuration du nœud actif afin de terminer la configuration.

Créer et configurer une deuxième carte réseau sur vCenter Server Appliance

Avant de pouvoir lancer la configuration avancée, vous devez créer et configurer une deuxième carte réseau sur l'instance de vCenter Server Appliance qui deviendra le nouveau nœud actif. Cette carte réseau sera utilisée pour le trafic vCenter HA. Cette tâche doit être réalisée après la configuration du réseau, mais avant le début du processus de configuration.

Prérequis

- Configurez l'infrastructure du réseau vCenter HA. Reportez-vous à « Configurer le réseau », page 73.
- Déployez l'instance de vCenter Server Appliance à utiliser en tant que nœud actif initial.
  - Une adresse IP statique doit être mappée vers un nom de domaine complet pour l'instance de vCenter Server Appliance.
  - Le mode SSH doit être activé sur vCenter Server Appliance.

Procédure

1 Connectez-vous au serveur de gestion de vCenter Server avec vSphere Web Client.

2 Sélectionnez la machine virtuelle vCenter Server Appliance (nœud actif), ajoutez un deuxième adaptateur réseau, puis attachez-le au groupe de ports vCenter HA que vous avez créé.

3 Connectez-vous directement à l'instance vCenter Server Appliance qui deviendra initialement le nœud actif.

Interface	Action
vCenter Server Appliance	Allez à https://*appliance-IP-address-or-FQDN*:5480
vSphere Web Client	a Allez à https://*appliance-IP-address-or-FQDN*/vsphere-client b Sélectionnez **Administration > Configuration système**

4 Configurez les paramètres IP du deuxième adaptateur réseau.

Lancer le processus de configuration avancée

Après avoir configuré le réseau et ajouté une deuxième carte réseau à vCenter Server Appliance, vous pouvez lancer le processus de configuration de vCenter HA.

Prérequis

- Déployez l'instance de vCenter Server Appliance à utiliser en tant que nœud actif initial.
  - Une adresse IP statique doit être mappée vers un nom de domaine complet pour l'instance de vCenter Server Appliance.
  - Le mode SSH doit être activé sur vCenter Server Appliance.
- Configurez le réseau. Reportez-vous à « Configurer le réseau », page 73.
- Identifiez les adresses IP statiques à utiliser pour les deux nœuds vCenter Server Appliance qui deviendront les nœuds passif et témoin.

Procédure

1 Connectez-vous au nœud actif à l'aide de vSphere Web Client.

2 Cliquez avec le bouton droit sur l'objet vCenter Server dans l'inventaire et sélectionnez **Paramètres vCenter HA**.

3 Cliquez sur **Configurer**.

4 Sélectionnez l'option de configuration **Avancé** et cliquez sur **Suivant**.

5 Fournissez l'adresse IP et le masque de sous-réseau pour les nœuds passif et témoin et cliquez sur **Suivant**.

Vous devez spécifier ces adresses IP maintenant même si les nœuds n'existent pas encore. Vous ne pouvez plus modifier ces adresses IP après avoir cliqué sur **Suivant**.

6 (Facultatif) Cliquez sur **Avancé** si vous souhaitez remplacer l'adresse IP de gestion du basculement du nœud passif.

7 Conservez la fenêtre de l'assistant ouverte et procédez aux tâches de clonage.

Suivant

« Créer et configurer les clones du nœud actif », page 77.

Créer et configurer les clones du nœud actif

Lors de la configuration avancée, vous devez cloner le nœud actif pour créer les nœuds passif et témoin. Ne quittez pas l'assistant Configurer vCenter HA pendant l'exécution des tâches de clonage.

Procédure

1 Connectez-vous à l'instance de gestion de vCenter Server, cliquez avec le bouton droit sur la machine virtuelle vCenter Server Appliance (nœud actif) et sélectionnez **Cloner > Cloner vers une machine virtuelle**.

2 Pour le premier clone, qui deviendra le nœud passif, entrez les valeurs suivantes.

Option	Valeur
Nouveau nom de machine virtuelle	Nom du nœud passif. Par exemple, utilisez vcsa-peer.
Sélectionner une ressource de calcul **Sélectionner le stockage**	Utilisez un hôte cible et une banque de données différents de ceux du nœud actif si possible.
Options de clonage	Sélectionnez les cases à cocher **Personnaliser le système d'exploitation** et **Mettre sous tension la machine virtuelle après la création**, puis cliquez sur l'icône Nouvelle spécification de personnalisation de la page suivante. Dans l'assistant Nouvelle spécification de personnalisation qui s'affiche, sélectionnez ce qui suit : a Utilisez le nom de l'hôte pour nommer le nœud actif. b Assurez-vous que le fuseau horaire est identique à celui du nœud actif. c Sur la page de configuration du réseau, spécifiez les paramètres IP de NIC1 et NIC2. Ces paramètres sont mappés avec l'interface de gestion et l'interface de vCenter HA. N'entrez rien dans le champ de passerelle par défaut de NIC2.

3 Une fois le premier clone créé, clonez à nouveau le nœud actif pour le nœud témoin.

Option	Valeur
Nouveau nom de machine virtuelle	Nom du nœud témoin. Par exemple, utilisez vcsa-witness.
Sélectionner une ressource de calcul **Sélectionner le stockage**	Utilisez un hôte cible et une banque de données différents de ceux des nœuds actif et passif, si possible.
Options de clonage	Sélectionnez les cases à cocher Personnaliser le système d'exploitation et Mettre sous tension la machine virtuelle après la création, puis cliquez sur l'icône Nouvelle spécification de personnalisation de la page suivante. Dans l'assistant Nouvelle spécification de personnalisation qui s'affiche, sélectionnez ce qui suit : a Utilisez le nom d'hôte de votre choix. b Assurez-vous que le fuseau horaire est identique à celui du nœud actif. c Sur la page de configuration du réseau, spécifiez les paramètres IP de NIC2. Ces paramètres sont mappés avec l'interface de vCenter HA. N'entrez rien dans le champ de passerelle par défaut de NIC2.

4 Assurez-vous que le processus de clonage se termine et que les machines virtuelles sont sous tension.

Suivant

Revenez à l'assistant vCenter HA du nœud actif pour terminer la configuration. Reportez-vous à « Terminer la configuration avancée de vCenter HA », page 78.

Terminer la configuration avancée de vCenter HA

Après avoir créé les nœuds passif et témoin, revenez à l'assistant de configuration du nœud actif afin de terminer la configuration.

Prérequis

Terminez le processus de clonage du nœud actif en nœud passif et en nœud témoin.

Procédure

1 Revenez à l'assistant de configuration et cliquez sur **Terminer**.

2 Attendez que la configuration de vCenter HA s'achève.

Gérer la configuration vCenter HA

Après avoir configuré votre cluster vCenter HA, vous pouvez effectuer les tâches de gestion. Ces tâches incluent le remplacement de certificats, le remplacement des clés SSH et la configuration SNMP. Vous pouvez également modifier la configuration du cluster pour désactiver ou activer vCenter HA, passer en mode de maintenance et supprimer la configuration du cluster.

- Configurer des interruptions SNMP page 79

  Vous pouvez configurer des interruptions SNMP (Simple Network Management Protocol) de manière à recevoir des notifications SNMP pour votre cluster vCenter HA.

- Configurer votre environnement pour utiliser des certificats personnalisés page 80

  Le certificat SSL de la machine sur chaque nœud est utilisé pour la communication en matière de gestion du cluster et pour le chiffrement du trafic de réplication. Si vous souhaitez utiliser des certificats personnalisés, vous devez supprimer la configuration vCenter HA, supprimer les nœuds passif et témoin, provisionner le nœud actif avec le certificat personnalisé et reconfigurer le cluster.

- Gérer les clés SSH de vCenter HA page 80

  vCenter HA utilise les clés SSH pour l'authentification sans mot de passe sur les nœuds actif, passif et témoin. L'authentification s'applique pour l'échange des signaux de pulsation et la réplication de fichiers et de données. Pour remplacer les clés SSH des nœuds d'un cluster vCenter HA, vous devez désactiver le cluster, générer de nouvelles clés SSH sur le nœud actif, transférer ces clés au nœud passif et activer le cluster.

- Initier le basculement de vCenter HA page 81

  Vous pouvez initier manuellement un basculement et faire en sorte que le nœud passif devienne le nœud actif.

- Modifier la configuration d'un cluster vCenter HA page 81

  Lorsque vous modifiez la configuration du cluster vCenter HA, vous pouvez désactiver ou activer le cluster, le placer en mode de maintenance ou le supprimer.

- Effectuer des opérations de sauvegarde et de restauration page 82

  Pour une sécurité supplémentaire, vous pouvez sauvegarder le nœud actif dans le cluster vCenter HA. Vous pouvez ensuite restaurer le nœud en cas de panne catastrophique.

- Supprimer une configuration de vCenter HA page 83

  Vous pouvez supprimer une configuration de vCenter HA de vSphere Web Client. Si vous utilisez une configuration avancée, ou si l'un des nœuds n'est pas détectable, vous devrez peut-être suivre des étapes de nettoyage supplémentaires.

- Redémarrer tous les nœuds vCenter HA page 83

  Si vous devez arrêter et redémarrer tous les nœuds dans le cluster, vous devez suivre un ordre spécifique pour l'arrêt afin d'empêcher le nœud passif d'assumer le rôle du nœud actif.

- Modifier l'environnement du dispositif page 83

  Lorsque vous déployez un dispositif vCenter Server Appliance, vous sélectionnez un environnement. Pour vCenter HA, Petit, Moyen, Grand et Très grand sont pris en charge pour les environnements de production. Si vous avez besoin de plus d'espace et souhaitez modifier l'environnement, vous devez supprimer la machine virtuelle du nœud passif avant de modifier la configuration.

- Collecter des bundles de support pour un nœud vCenter HA page 84

  Collecter un bundle de support à partir de tous les nœuds d'un cluster vCenter HA aide à dépanner les problèmes.

Configurer des interruptions SNMP

Vous pouvez configurer des interruptions SNMP (Simple Network Management Protocol) de manière à recevoir des notifications SNMP pour votre cluster vCenter HA.

Les interruptions sont définies par défaut sur SNMP version 1.

Configurez les interruptions SNMP pour le nœud actif et le nœud passif. Vous indiquez à l'agent où envoyer les interruptions associées en ajoutant une entrée cible à la configuration snmpd.

Procédure

1 Connectez-vous au nœud actif en utilisant la console de machine virtuelle ou les clés SSH.

2 Exécutez la commande `vicfg-snmp`, par exemple :

```
vicfg-snmp -t 10.160.1.1@1166/public
```

Dans cet exemple, `10.160.1.1` est l'adresse d'écoute du client, `1166` est le port d'écoute du client et `public` est la chaîne de la communauté.

3 Activez l'agent SNMP (snmpd) en exécutant la commande suivante.

```
vicfg-snmp -e
```

Suivant

Les commandes suivantes peuvent également être utiles.

- Pour accéder à l'aide complète concernant cette commande, exécutez **`vicfg-snmp -h`**.
- Pour désactiver l'agent SNMP, exécutez **`vicfg-snmp -D`**.
- Pour afficher la configuration de l'agent SNMP, exécutez **`vicfg-snmp -s`**.
- Pour rétablir la configuration par défaut, exécutez **`vicfg-snmp -r`**.

Configurer votre environnement pour utiliser des certificats personnalisés

Le certificat SSL de la machine sur chaque nœud est utilisé pour la communication en matière de gestion du cluster et pour le chiffrement du trafic de réplication. Si vous souhaitez utiliser des certificats personnalisés, vous devez supprimer la configuration vCenter HA, supprimer les nœuds passif et témoin, provisionner le nœud actif avec le certificat personnalisé et reconfigurer le cluster.

Si possible, remplacez les certificats dans vCenter Server Appliance qui deviendra le nœud actif, avant de procéder au clonage du nœud.

Procédure

1 Modifiez la configuration du cluster et sélectionnez **Supprimer**.

2 Supprimez le nœud passif et le nœud témoin.

3 Sur le nœud actif, qui est désormais un système vCenter Server Appliance autonome, remplacez le certificat SSL de la machine par un certificat personnalisé.

Consultez la documentation de *Administration de Platform Services Controller* .

4 Reconfigurez le cluster.

Gérer les clés SSH de vCenter HA

vCenter HA utilise les clés SSH pour l'authentification sans mot de passe sur les nœuds actif, passif et témoin. L'authentification s'applique pour l'échange des signaux de pulsation et la réplication de fichiers et de données. Pour remplacer les clés SSH des nœuds d'un cluster vCenter HA, vous devez désactiver le cluster, générer de nouvelles clés SSH sur le nœud actif, transférer ces clés au nœud passif et activer le cluster.

Procédure

1 Éditez le cluster et définissez son mode sur **Désactivé**.

2 Connectez-vous au nœud actif en utilisant la console de machine virtuelle ou les clés SSH.

3 Activez l'interpréteur de commandes de dépistage.

```
bash
```

4 Exécutez la commande suivante pour générer de nouvelles clés SSH sur le nœud actif.

```
/usr/lib/vmware-vcha/scripts/resetSshKeys.py
```

5 Utilisez SCP pour copier les clés sur les nœuds passif et témoin.

```
scp /vcha/.ssh/*
```

6 Modifiez la configuration du cluster et définissez le cluster vCenter HA sur **Activé**.

Initier le basculement de vCenter HA

Vous pouvez initier manuellement un basculement et faire en sorte que le nœud passif devienne le nœud actif.

Un cluster vCenter HA prend en charge deux types de basculement.

Basculement automatique	Le nœud passif tente de prendre le relais du nœud actif en cas de panne de celui-ci.
Basculement manuel	L'utilisateur peut forcer un nœud passif à prendre le relais du nœud actif en utilisant l'action Initier le basculement.

Initiez un basculement manuel à des fins de dépannage et de test.

Procédure

1 Connectez-vous au nœud actif vCenter Server Appliance à l'aide de vSphere Web Client et cliquez sur **Configurer**.

2 Dans **Paramètres**, sélectionnez **vCenter HA** et cliquez sur **Initier le basculement**.

3 Cliquez sur **Oui** pour déclencher le basculement.

Une boîte de dialogue qui s'ouvre vous permet de forcer un basculement sans synchronisation. Dans la plupart des cas, il est recommandé d'effectuer tout d'abord la synchronisation.

4 Après le basculement, vous pouvez vérifier que le nœud passif joue le rôle du nœud actif dans vSphere Web Client.

Modifier la configuration d'un cluster vCenter HA

Lorsque vous modifiez la configuration du cluster vCenter HA, vous pouvez désactiver ou activer le cluster, le placer en mode de maintenance ou le supprimer.

Le mode de fonctionnement d'un dispositif vCenter Server Appliance contrôle les capacités de basculement et la réplication de l'état dans un cluster vCenter HA.

Un cluster vCenter HA peut fonctionner dans l'un des modes suivants.

Tableau 4-3. Modes de fonctionnement d'un cluster vCenter HA

Mode	Basculement automatique	Basculement manuel	Réplication	
Activé	Oui	Oui	Oui	Ce mode de fonctionnement par défaut protège le dispositif vCenter Server Appliance des défaillances matérielles et logicielles en effectuant un basculement automatique.
Maintenance	Non	Oui	Oui	Utilisé pour les tâches de maintenance. Pour les autres tâches, vous devez désactiver vCenter HA.
Désactivé	Non	Non	Non	Si les nœuds passif et témoin sont perdus ou restaurés suite à une panne, la configuration de vCenter HA peut être désactivée. Le nœud actif continue à s'exécuter en tant que nœud vCenter Server Appliance autonome.

Remarque Si le cluster fonctionne en mode de maintenance ou Désactivé, un nœud actif peut continuer à desservir les requêtes du client même si les nœuds passif et témoin sont perdus ou inaccessibles.

Prérequis

Vérifiez que le cluster vCenter HA est déployé et contient les nœuds actif, passif et témoin.

Procédure

1 Connectez-vous au nœud actif vCenter Server Appliance à l'aide de vSphere Web Client et cliquez sur **Configurer**.

2 Sous **Paramètres**, sélectionnez **vCenter HA** et cliquez sur **Modifier**.

3 Sélectionnez l'une des options.

Option	Résultat
Activer vCenter HA	Autorise la réplication entre les nœuds actif et passif. Si l'état du cluster est sain, votre nœud actif est protégé par le basculement automatique du nœud passif.
Mode maintenance	En mode de maintenance, la réplication est toujours effectuée entre les nœuds actif et passif. Cependant, le basculement automatique est désactivé.
Désactiver vCenter HA	Désactive la réplication et le basculement. Conserve la configuration du cluster. Vous pourrez ensuite activer vCenter HA à nouveau.
Supprimer le cluster vCenter HA	Supprime le cluster. La réplication et le basculement ne sont plus fournis. Le nœud actif continue à s'exécuter en tant que nœud vCenter Server Appliance autonome. Reportez-vous à « Supprimer une configuration de vCenter HA », page 83 pour plus de détails.

4 Cliquez sur **OK**.

Effectuer des opérations de sauvegarde et de restauration

Pour une sécurité supplémentaire, vous pouvez sauvegarder le nœud actif dans le cluster vCenter HA. Vous pouvez ensuite restaurer le nœud en cas de panne catastrophique.

Remarque Supprimez la configuration du cluster avant de restaurer le nœud actif. Les résultats sont imprévisibles si vous restaurez le nœud actif alors que le nœud passif est toujours en cours d'exécution ou qu'une autre configuration de cluster est toujours en place.

Prérequis

Vérifiez l'interopérabilité de vCenter HA et de la solution de sauvegarde et de restauration. Une solution est la restauration de vCenter Server Appliance basée sur un fichier.

Procédure

1 Sauvegardez le nœud actif.

Ne sauvegardez pas le nœud passif et le nœud témoin.

2 Avant de restaurer le cluster, mettez hors tension et supprimez tous les nœuds vCenter HA.

3 Restaurez le nœud actif.

Le nœud actif est restauré en tant que dispositif vCenter Server Appliance autonome.

4 Reconfigurez vCenter HA.

Supprimer une configuration de vCenter HA

Vous pouvez supprimer une configuration de vCenter HA de vSphere Web Client. Si vous utilisez une configuration avancée, ou si l'un des nœuds n'est pas détectable, vous devrez peut-être suivre des étapes de nettoyage supplémentaires.

Procédure

1 Connectez-vous au nœud actif vCenter Server Appliance et cliquez sur **Configurer**.

2 Sous **Paramètres**, sélectionnez **vCenter HA** et cliquez sur **Modifier**.

3 Sélectionnez **Supprimer le cluster vCenter HA**.

- La configuration du cluster vCenter HA est supprimée des nœuds actif, passif et témoin.
- Le nœud actif continue à s'exécuter en tant que nœud vCenter Server Appliance autonome.
- Vous ne pouvez pas réutiliser les nœuds passif et témoin dans une nouvelle configuration de vCenter HA.
- Si vous procédez à la configuration en utilisant les options avancées, ou si les nœuds passif et témoin ne sont pas détectables, vous devez supprimer ces nœuds explicitement.
- Même si la seconde carte réseau virtuelle a été ajoutée par le processus de configuration, le processus de suppression ne supprime pas la carte réseau virtuelle.

Redémarrer tous les nœuds vCenter HA

Si vous devez arrêter et redémarrer tous les nœuds dans le cluster, vous devez suivre un ordre spécifique pour l'arrêt afin d'empêcher le nœud passif d'assumer le rôle du nœud actif.

Procédure

1 Arrêtez les nœuds dans cet ordre.

- Nœud passif
- Nœud actif
- Nœud témoin

2 Redémarrez chaque nœud.

Vous pouvez redémarrer les nœuds dans n'importe quel ordre.

3 Vérifiez que tous les nœuds joignent le cluster correctement et que le nœud actif précédent reprend ce rôle.

Modifier l'environnement du dispositif

Lorsque vous déployez un dispositif vCenter Server Appliance, vous sélectionnez un environnement. Pour vCenter HA, Petit, Moyen, Grand et Très grand sont pris en charge pour les environnements de production. Si vous avez besoin de plus d'espace et souhaitez modifier l'environnement, vous devez supprimer la machine virtuelle du nœud passif avant de modifier la configuration.

Procédure

1 Connectez-vous au nœud actif avec vSphere Web Client, modifiez la configuration du cluster, puis sélectionnez **Désactiver**.

2 Supprimez la machine virtuelle du nœud passif.

3 Modifiez la configuration vCenter Server Appliance du nœud actif (par exemple, d'un environnement petit à un environnement moyen).

4 Reconfigurez vCenter HA.

Collecter des bundles de support pour un nœud vCenter HA

Collecter un bundle de support à partir de tous les nœuds d'un cluster vCenter HA aide à dépanner les problèmes.

Lorsque vous collectez un bundle de support sur le nœud actif d'un cluster vCenter HA, le système procède de la façon suivante.

- Il collecte les informations du bundle de support directement sur le nœud actif.
- Il collecte des bundles de support depuis les nœuds passif et témoin, et les place dans le répertoire `commands` du bundle de support du nœud actif.

Remarque La collecte des bundles de support sur les nœuds passif et témoin se fait dans la mesure du possible et ne peut avoir lieu que si les nœuds sont atteignables.

Corriger votre environnement vCenter HA

En cas de problème, vous pouvez corriger votre environnement. La tâche que vous devez effectuer dépend des symptômes de la défaillance. Pour en savoir plus sur le dépannage, reportez-vous au système de la base de connaissances VMware.

- L'opération de clonage de vCenter HA échoue lors du déploiement page 84

  Si le processus de configuration de vCenter HA ne parvient pas à créer le clone, vous devez résoudre cette erreur de clonage.

- Le déploiement de vCenter HA échoue avec une erreur page 85

  L'échec du déploiement peut s'expliquer par des problèmes de configuration et en particulier par des incidents lors de la configuration de la mise en réseau.

- Dépannage d'un cluster vCenter HA dégradé page 85

  Pour qu'un cluster vCenter HA soit sain, chacun des nœuds actif, passif et témoin doivent être entièrement opérationnels et accessibles sur le réseau du cluster vCenter HA. En cas de panne d'un nœud, le cluster est considéré comme étant dans un état dégradé.

- Restauration de nœuds vCenter HA isolés page 86

  Si tous les nœuds d'un cluster vCenter HA ne peuvent pas communiquer les uns avec les autres, le nœud actif cesse de desservir les requêtes du client.

- Résolution des défaillances suite à un basculement page 87

  Lorsqu'un nœud passif ne devient pas un nœud actif lors d'un basculement, vous pouvez forcer le passage du nœud passif au nœud actif.

L'opération de clonage de vCenter HA échoue lors du déploiement

Si le processus de configuration de vCenter HA ne parvient pas à créer le clone, vous devez résoudre cette erreur de clonage.

Problème

L'opération de clonage échoue.

Cause

Recherchez l'exception de clone. Celle-ci peut indiquer l'un des problèmes suivants.

- Vous avez un cluster sur lequel DRS est activé, mais vous ne disposez pas de trois hôtes.
- La connexion à l'hôte ou à la base de données est perdue.
- L'espace disque est insuffisant.
- Autres erreurs **Cloner une machine virtuelle**

Solution

1 Corrigez l'erreur à l'origine de ce problème.

2 Supprimez le cluster, puis recommencez la configuration.

Le déploiement de vCenter HA échoue avec une erreur

L'échec du déploiement peut s'expliquer par des problèmes de configuration et en particulier par des incidents lors de la configuration de la mise en réseau.

Problème

Vous commencez la configuration d'un cluster vCenter HA, mais celle-ci échoue avec une erreur. L'erreur peut indiquer la cause du problème, par exemple, un message Échec de la connexion SSH peut s'afficher.

Solution

Si le déploiement échoue, prenez les mesures nécessaires pour résoudre les problèmes de réseau.

1 Vérifiez que les nœuds passif et témoin sont accessibles depuis le nœud actif.

2 Vérifiez que le routage entre les nœuds est configuré correctement.

3 Vérifiez le temps de réponse du réseau.

Dépannage d'un cluster vCenter HA dégradé

Pour qu'un cluster vCenter HA soit sain, chacun des nœuds actif, passif et témoin doivent être entièrement opérationnels et accessibles sur le réseau du cluster vCenter HA. En cas de panne d'un nœud, le cluster est considéré comme étant dans un état dégradé.

Problème

Si le cluster est dans un état dégradé, le basculement ne peut pas être effectué. Pour obtenir des informations sur les scénarios de panne lorsque le cluster est dans un état dégradé, reportez-vous à la section « Résolution des défaillances suite à un basculement », page 87.

Cause

Le cluster peut être dans un état dégradé pour plusieurs raisons.

L'un des nœuds est défaillant

- En cas de panne du nœud actif, un basculement du nœud actif vers le nœud passif est effectué automatiquement. Une fois le basculement effectué, le nœud passif devient le nœud actif.

  Le cluster est alors dans un état dégradé car le nœud actif d'origine n'est pas disponible.

  Après que le nœud défaillant a été rétabli ou mis en ligne, il devient le nouveau nœud passif et le cluster redevient sain une fois que les nœuds actif et passif ont été synchronisés.

- En cas de panne du nœud passif, le nœud actif continue à fonctionner, mais aucun basculement n'est possible et le cluster est dans un état dégradé.

  Si le nœud passif est rétabli ou mis en ligne, il rejoint automatiquement le cluster et l'état de celui-ci est sain une fois que les nœuds actif et passif ont été synchronisés.

- En cas de panne du nœud témoin, le nœud actif continue à fonctionner et la réplication entre les nœuds actif et passif continue, mais aucun basculement ne peut être effectué.

  Si le nœud témoin est rétabli ou mis en ligne, il rejoint automatiquement le cluster et l'état de celui-ci est sain.

La réplication de la base de données échoue

En cas d'échec de la réplication entre les nœuds actif et passif, le cluster est considéré comme étant dégradé. Le nœud actif continue de se synchroniser avec le nœud passif. S'il réussit, le cluster redevient sain. Cet état peut être du à des problèmes au niveau de la bande passante du réseau ou à d'autres manques de ressources.

Problèmes de réplication du fichier de configuration

Si les fichiers de configuration ne sont pas correctement répliqués entre le nœud actif et le nœud passif, le cluster est dans un état dégradé. Le nœud actif continue de tenter de se synchroniser avec le nœud passif. Cet état peut être du à des problèmes au niveau de la bande passante du réseau ou à d'autres manques de ressources.

Solution

La manière de résoudre ce problème dépend de l'origine de l'état de dégradation du cluster. Si le cluster se trouve dans un état dégradé, des événements, des alarmes et des interruptions SNMP affichent des erreurs.

Si l'un des nœuds est en panne, recherchez une éventuelle défaillance matérielle ou une isolation de réseau. Vérifiez si le nœud défaillant est mis sous tension.

En cas de défaillance de la réplication, vérifiez si le réseau vCenter HA dispose d'une bande passante suffisante et assurez-vous que la latence réseau est de 10 ms ou moins.

Restauration de nœuds vCenter HA isolés

Si tous les nœuds d'un cluster vCenter HA ne peuvent pas communiquer les uns avec les autres, le nœud actif cesse de desservir les requêtes du client.

Problème

L'isolation des nœuds est un problème de connectivité réseau.

Solution

1. Essayez de résoudre le problème de connectivité. Si vous parvenez à restaurer la connectivité, les nœuds isolés rejoignent le cluster automatiquement et le nœud actif commence à desservir les requêtes du client.
2. Si vous ne parvenez pas à résoudre le problème de connectivité, vous devez vous connecter directement à la console du nœud actif.
   a. Mettez les machines virtuelles des nœuds passif et actif hors tension, puis supprimez-les.
   b. Connectez-vous au nœud actif en utilisant SSH ou par l'intermédiaire de la console de la machine virtuelle.
   c. Pour activer le shell Bash, entrez **`shell`** à l'invite `appliancesh`.

d Exécutez la commande suivante pour supprimer la configuration de vCenter HA.

```
destroy-vcha -f
```

e Redémarrez le nœud actif.

Le nœud actif est désormais une instance autonome de vCenter Server Appliance.

f Procédez de nouveau à la configuration du cluster vCenter HA.

Résolution des défaillances suite à un basculement

Lorsqu'un nœud passif ne devient pas un nœud actif lors d'un basculement, vous pouvez forcer le passage du nœud passif au nœud actif.

Problème

Le nœud passif échoue lorsqu'il tente d'assurer le rôle du nœud actif.

Cause

Un basculement vCenter HA peut échouer pour les raisons suivantes.

- Le nœud témoin devient indisponible alors que le nœud passif tente d'assurer le rôle du nœud actif.
- Il existe un problème de synchronisation de l'état du dispositif entre les nœuds.

Solution

Vous pouvez résoudre ce problème de la manière suivante.

1 Si le nœud actif récupère de la défaillance, il redevient le nœud actif.

2 Si le nœud témoin récupère de la défaillance, suivez les étapes ci-dessous.

a Connectez-vous au nœud passif via la console de la machine virtuelle.

b Pour activer le shell Bash, entrez **`shell`** à l'invite `appliancesh`.

c Exécutez la commande suivante.

```
vcha-reset-primary
```

d Redémarrez le nœud passif.

3 Si le nœud actif et le nœud témoin ne peuvent pas récupérer de la défaillance, vous pouvez forcer le passage du nœud passif à une instance autonome de vCenter Server Appliance.

a Supprimez les machines virtuelles du nœud actif et du nœud témoin.

b Connectez-vous au nœud passif via la console de la machine virtuelle.

c Pour activer le shell Bash, entrez **`shell`** à l'invite `appliancesh`.

d Exécutez la commande suivante.

```
destroy-vcha
```

e Redémarrez le nœud passif.

Application de correctifs à un environnement vCenter High Availability

Vous pouvez appliquer un correctif à un dispositif vCenter Server Appliance situé dans un cluster vCenter High Availability à l'aide de l'utilitaire <codeph>software-packages</codeph> disponible dans l'interpréteur du dispositif vCenter Server Appliance. Pour plus d'informations, reportez-vous à *Mise à niveau de vSphere*.

Utilisation de Microsoft Clustering Service pour vCenter Server sur un cluster Windows haute disponibilité

5

Lorsque vous déployez vCenter Server, vous devez créer une architecture hautement disponible, capable de gérer des charges de travail de toutes tailles.

La disponibilité est essentielle pour les solutions nécessitant une connectivité continue à vCenter Server. Pour éviter des temps d'arrêt importants, vous pouvez obtenir une connectivité continue pour vCenter Server en utilisant un cluster Microsoft Cluster Service (MSCS).

Ce chapitre aborde les rubriques suivantes :

Avantages et limitations de l'utilisation de MSCS

vCenter Server 5.5 mise à jour 3.x prend en charge Microsoft Cluster Service (MSCS) comme option pour la disponibilité de vCenter Server.

Plusieurs instances de vCenter Server se trouvent dans un cluster MSCS, mais une seule est active à la fois. Utilisez cette solution pour effectuer la maintenance, comme les correctifs ou mises à niveau de systèmes d'exploitation, à l'exception de correctifs et mises à niveau de vCenter Server. Vous effectuez la maintenance sur un nœud du cluster sans arrêter la base de données vCenter Server.

Autre avantage éventuel de cette approche, MSCS utilise une architecture de cluster de type « sans partage ». Le cluster n'implique pas d'accès disque simultanés à partir de plusieurs nœuds. En d'autres termes, le cluster n'a pas besoin d'un gestionnaire de verrouillage distribué. Les clusters MSCS n'incluent généralement que deux nœuds et utilisent une connexion SCSI partagée entre ces nœuds. Comme un seul serveur a besoin des disques à un moment précis, il n'y a jamais d'accès simultané aux données. Ce partage minimise l'impact d'une panne de nœud.

Contrairement à l'option de cluster vSphere HA, l'option MSCS ne fonctionne que pour les machines virtuelles Windows. L'option MSCS ne prend pas en charge vCenter Server Appliance.

Remarque Cette configuration est prise en charge uniquement lorsque vCenter Server est exécuté en tant que VM et non sur un hôte physique.

Mettre à niveau vCenter Server dans un environnement MSCS

Si vous exécutez vCenter Server 6.0, vous devez mettre à niveau vCenter Server 6.5 pour configurer un environnement MSCS à haute disponibilité.

vCenter Server 6.0.x a 18 services, en supposant que le serveur PSC s'exécute sur un hôte différent. vCenter Server 6.5 a 3 services et les noms ont été modifiés. Le paramétrage d'un cluster MSCS créé pour configurer la haute disponibilité pour vCenter Server 6.0 devient non valide suite à une mise à niveau vers vCenter Server 6.5.

Le processus pour la haute disponibilité de vCenter Server dans un environnement MSCS est le suivant.

1 Supprimez la configuration MSCS pour vCenter Server.

2 Mettez à niveau vCenter Server de la version 6.0 à la version 6.5.

3 Configurez MSCS pour que vCenter Server devienne hautement disponible.

Prérequis

- Vérifiez que vous n'êtes pas en train de supprimer la machine virtuelle du nœud principal.
- Vérifiez que le nœud principal est le nœud actuellement actif.
- Vérifiez que l'ensemble des services de vCenter Server 6.0 s'exécutent sur le nœud principal.
- Vérifiez que la mise à niveau du nœud Platform Services Controller est terminée et qu'elle exécute vCenter Server 6.5.
- Collectez la sauvegarde de la base de données de l'inventaire.

Procédure

1 Mettez le nœud secondaire hors tension et attendez que tous les services vCenter Server soient démarrés sur le nœud principal.

2 Supprimez le nom du rôle.

3 Détruisez le cluster MSCS. Remettez les disques RDM en ligne avant de modifier le type de démarrage.

4 Ouvrez la vue Gestion des services et passez le type de démarrage pour les services vCenter Server de manuel à automatique.

5 Avant de mettre à niveau vers vCenter Server 6.5, remplacez l'adresse IP et le nom d'hôte par l'adresse IP et le nom d'hôte utilisés pour le rôle.

Vous devez redémarrer l'hôte et vous assurer que vCenter Server est accessible.

6 Montez l'image ISO vCenter Server 6.5 et démarrez l'installation.

7 Une fois l'installation terminée, ouvrez la vue Gestion des services et vérifiez que les nouveaux services sont installés et qu'ils s'exécutent.

8 Paramétrez de nouveau la configuration du cluster MSCS et définissez le type de démarrage de tous les services vCenter Server sur Manuel.

9 Arrêtez le nœud principal et détachez les disques RDM, mais ne les supprimez pas de la banque de données.

10 Une fois la reconfiguration terminée, sélectionnez **VM > Cloner > Cloner dans le modèle**. Clonez le nœud secondaire et modifiez ses adresse IP et nom d'hôte.

11 Gardez le nœud secondaire hors tension et ajoutez les disques RDM au nœud principal. Mettez ensuite le nœud principal sous tension et modifiez ses adresse IP et nom d'hôte.

12 Ajoutez les deux disques RDM au nœud secondaire. Puis mettez sous tension le nœud secondaire.

13 Ouvrez Failover Cluster Manager et configurez le cluster MSCS.

Vous devez utiliser l'adresse IP et le nom d'hôte du rôle de cluster.

Suivant

Lors de la configuration du cluster MSCS, vous devez ajouter des services vCenter Server comme le service VMware AFD et le service Configuration de VMware vCenter au rôle en tant que ressources.

Configurer MSCS pour la haute disponibilité

Utilisez la procédure suivante pour configurer Microsoft Cluster Service (MSCS) comme solution de disponibilité pour vCenter Server.

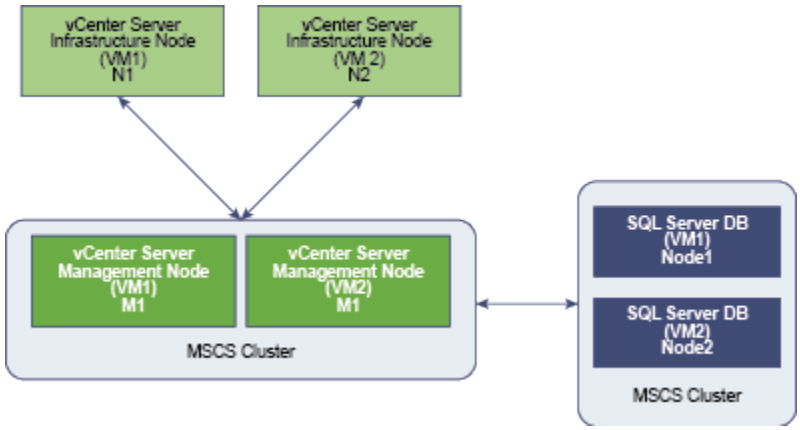
Prérequis

- Créer une machine virtuelle (VM) avec l'un des systèmes d'exploitation invités suivants :
- - Windows 2008 R2 Datacenter
  - Windows 2012 R2 Datacenter
- Ajoutez deux disques à mappage de périphériques bruts (RDM) à cette machine virtuelle. Ces disques doivent être montés lorsqu'ils sont ajoutés et les disques RDM doivent également être indépendants et permanents.
- Créez une contrôleur SCSI distinct avec l'option de partage de bus définie sur Physique.

Remarque Du fait que cette configuration utilise un contrôleur SCSI avec l'option de partage de bus définie sur Physique, la sauvegarde et la restauration ne sont pas prises en charge. Vous devez utiliser un agent basé sur l'hôte pour la sauvegarde ou la restauration.

- Définissez le nœud du périphérique virtuel sur le contrôleur SCSI 1 nouvellement créé.
- Ouvrez le lecteur MSCS et créez deux dossiers : un pour les données VC et un autre pour l'installation VC.
- Installez une instance de Platform Services Controller avant d'installer vCenter Server et fournissez son nom de domaine complet lors de l'installation.

Figure 5-1. Cluster MSCS pour la haute disponibilité de vCenter Server



Remarque MSCS en tant qu'option de disponibilité de vCenter Server est fourni uniquement pour les nœuds de gestion de vCenter Server (nœud M). Pour les nœuds d'infrastructure, les clients doivent déployer plusieurs nœuds N à des fins de haute disponibilité. Des nœuds M et N ne peuvent pas coexister sur la même machine virtuelle pour la protection MSCS.

Procédure

1 Mettez sous-tension la machine virtuelle.

2 Formatez les deux disques RDM, attribuez-leur des lettres de lecteur, puis convertissez-les en MBR.

3 À l'aide des options **Windows > Server Manager > Fonctionnalités**, installez .net.

4 Installez vCenter Server sur l'un des disques RDM et définissez l'option de démarrage sur Manuelle.

5 Mettez hors tension la machine virtuelle.

6 Détachez les disques RDM.

Le détachement des disques RDM n'est pas une suppression permanente. Ne sélectionnez pas **Supprimer du disque** et ne supprimez pas les fichiers vmdk.

7 Clonez la machine virtuelle et sélectionnez l'option **Personnaliser le système d'exploitation**, pour que le clone ait une identité unique.

Créez une identité unique au moyen du fichier sysrep par défaut ou du fichier sysrep personnalisé.

8 Attachez les RDM partagés aux deux machines virtuelles et mettez-les sous tension.

9 Modifiez le nom d'hôte et l'adresse IP sur la première machine virtuelle (VM1).

Notez l'adresse IP et le nom d'hôte qui ont été initialement utilisés au moment de l'installation de vCenter Server sur VM1. Ces informations sont utilisées pour attribuer une adresse IP de rôle de cluster.

10 Installez un clustering de basculement sur les deux nœuds.

11 Pour créer un cluster MSCS sur VM1, incluez les deux nœuds dans le cluster. Sélectionnez également l'option de validation sur le nouveau cluster.

12 Pour commencer à configurer les rôles, sélectionnez **Service générique** et cliquez sur **Suivant**.

13 Sélectionnez **VMware Service Lifecycle Manager** dans la liste des services et cliquez sur **Suivant**.

14 Entrez le nom de l'hôte et l'adresse IP utilisés pour VM1. Attribuez ensuite le disque RDM au rôle.

15 Dans l'assistant Répliquer les paramètres du registre, ajoutez la clé de registre `SYSTEM\CurrentControlSet\Services\VMwareDirectoryService` et cliquez sur **Suivant**.

16 À l'aide de l'option Ajouter une ressource, ajoutez les services VMware AFD et Configuration de VMware vCenter au rôle.

17 Arrêtez, puis redémarrez le rôle.

Vous avez créé un cluster MSCS pouvant prendre en charge la disponibilité de vCenter Server.

Suivant

Après avoir créée le cluster MSCS, vérifiez que le basculement s'effectue en mettant hors tension la machine virtuelle hébergeant vCenter Server (VM1). Quelques minutes après, vérifiez que les services s'exécutent sur l'autre VM (VM2).

Index