Cluster de armazenamento do vSphere Metro com o ONTAP

Enterprise applications

NetApp
February 11, 2026

Índice

Cluster de armazenamento do vSphere Metro com o ONTAP

Cluster de armazenamento do vSphere Metro com o ONTAP

O hypervisor vSphere líder do setor da VMware pode ser implantado como um cluster estendido chamado vSphere Metro Storage Cluster (vMSC).

As soluções vmsc são suportadas com o NetApp MetroCluster e o SnapMirror ative Sync (anteriormente conhecido como SnapMirror Business Continuity, ou SMBC) e fornecem continuidade de negócios avançada se um ou mais domínios de falha sofrerem uma interrupção total. A resiliência a diferentes modos de falha depende de quais opções de configuração você escolher.

Esta documentação substitui relatórios técnicos publicados anteriormente *TR-4128: VSphere no NetApp MetroCluster*

Soluções de disponibilidade contínua para ambientes vSphere

A arquitetura ONTAP é uma plataforma de armazenamento flexível e escalável que fornece serviços SAN (FCP, iSCSI e NVMe-oF) e NAS (NFS v3 e v4.1) para armazenamentos de dados. Os sistemas de armazenamento NetApp AFF, ASA e FAS usam o sistema operacional ONTAP para oferecer protocolos adicionais para acesso de armazenamento de convidados, como S3 e SMB/CIFS.

A NetApp MetroCluster usa a função de HA (failover de controladora ou CFO) da NetApp para proteger contra falhas de controladora. Ele também inclui tecnologia SyncMirror local, failover de cluster em desastre (failover de cluster em desastre ou CFOD), redundância de hardware e separação geográfica para alcançar altos níveis de disponibilidade. O SyncMirror espelha de forma síncrona os dados entre as duas metades da configuração do MetroCluster gravando dados em dois plexos: O Plex local (na gaveta local) fornecendo dados ativamente e o Plex remoto (na gaveta remota) normalmente não fornecendo dados. A redundância de hardware é implementada para todos os componentes MetroCluster, como controladores, armazenamento, cabos, switches (usados com Fabric MetroCluster) e adaptadores.

A sincronização ativa do NetApp SnapMirror, disponível em sistemas que não sejam MetroCluster e em sistemas ASA R2, oferece proteção granular do armazenamento de dados com protocolos FCP e SAN iSCSI. Ele permite que você proteja todo o vMSC ou proteja seletivamente cargas de trabalho de alta prioridade. Ele oferece acesso ativo-ativo a locais locais e remotos, ao contrário do NetApp MetroCluster, que é uma solução de espera ativa. A partir do ONTAP 9.15.1, o SnapMirror active Sync oferece suporte a uma funcionalidade ativo-ativo simétrica, permitindo operações de e/S de leitura e gravação de ambas as cópias de um LUN protegido com replicação síncrona bidirecional, permitindo que ambas as cópias do LUN atendam às operações de e/S localmente. Antes do ONTAP 9.15.1, a sincronização ativa do SnapMirror suporta apenas configurações ativas/ativas assimétricas, nas quais os dados no local secundário são proxy para a cópia primária de um LUN.

Para criar um cluster VMware HA/DRS em dois locais, os hosts ESXi são usados e gerenciados por um vCenter Server Appliance (VCSA). As redes de gerenciamento vSphere, vMotion e máquinas virtuais são conetadas por meio de uma rede redundante entre os dois sites. O vCenter Server que gerencia o cluster HA/DRS pode se conetar aos hosts ESXi em ambos os sites e deve ser configurado usando o vCenter HA.

```
https://docs.vmware.com/en/VMware-vSphere/8.0/vsphere-vcenter-esxi-management/GUID-F7818000-26E3-4E2A-93D2-FCDCE7114508.html["Como criar e configurar clusters no vSphere Client"]Consulte para configurar o vCenter HA.
```

Você também deve se referir "Práticas recomendadas do VMware vSphere Metro Storage Cluster" a .

O que é o vSphere Metro Storage Cluster?

O vSphere Metro Storage Cluster (vMSC) é uma configuração certificada que protege máquinas virtuais (VMs) e contêineres contra falhas. Isso é obtido usando conceitos de armazenamento estendido junto com clusters de hosts ESXi, que são distribuídos em diferentes domínios de falha, como racks, edifícios, campi ou até mesmo cidades. As tecnologias de armazenamento de sincronização ativa NetApp MetroCluster e SnapMirror são usadas para fornecer uma proteção de objetivo de ponto de recuperação zero (RPO=0) aos clusters de host. A configuração do vMSC foi projetada para garantir que os dados estejam sempre disponíveis, mesmo se um "site" físico ou lógico completo falhar. Um dispositivo de armazenamento que faz parte da configuração do vMSC deve ser certificado após passar por um processo de certificação do vMSC bem-sucedido. Todos os dispositivos de armazenamento suportados podem ser encontrados no "Guia de compatibilidade do VMware Storage".

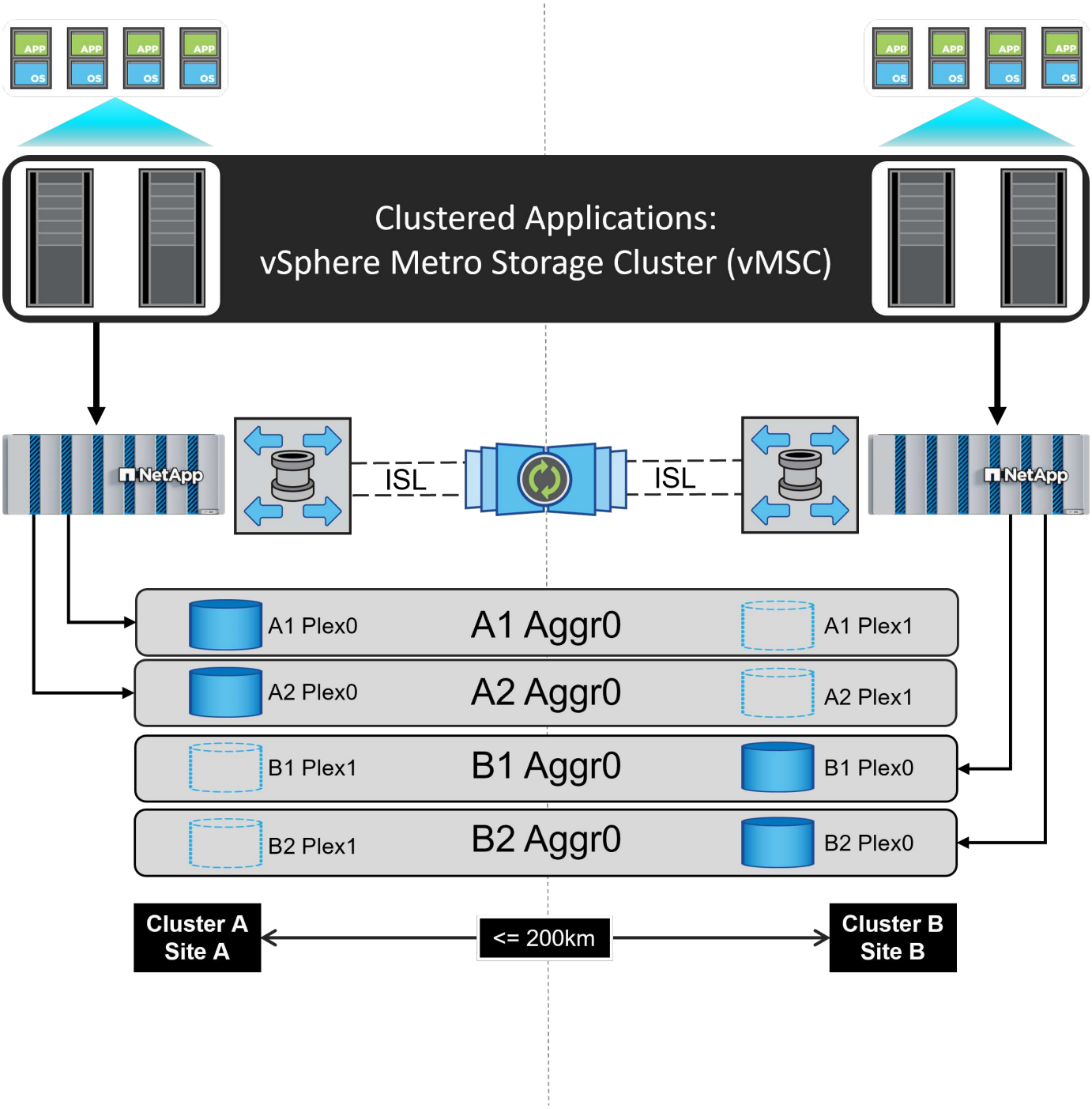
Se você quiser obter mais informações sobre as diretrizes de design do vSphere Metro Storage Cluster, consulte a seguinte documentação:

- "Suporte ao VMware vSphere com o NetApp MetroCluster"
- "Suporte ao VMware vSphere com o NetApp SnapMirror Business Continuity" (Agora conhecido como SnapMirror ative Sync)

O NetApp MetroCluster pode ser implantado em duas configurações diferentes para uso com o vSphere:

- Stretch MetroCluster
- Fabric MetroCluster

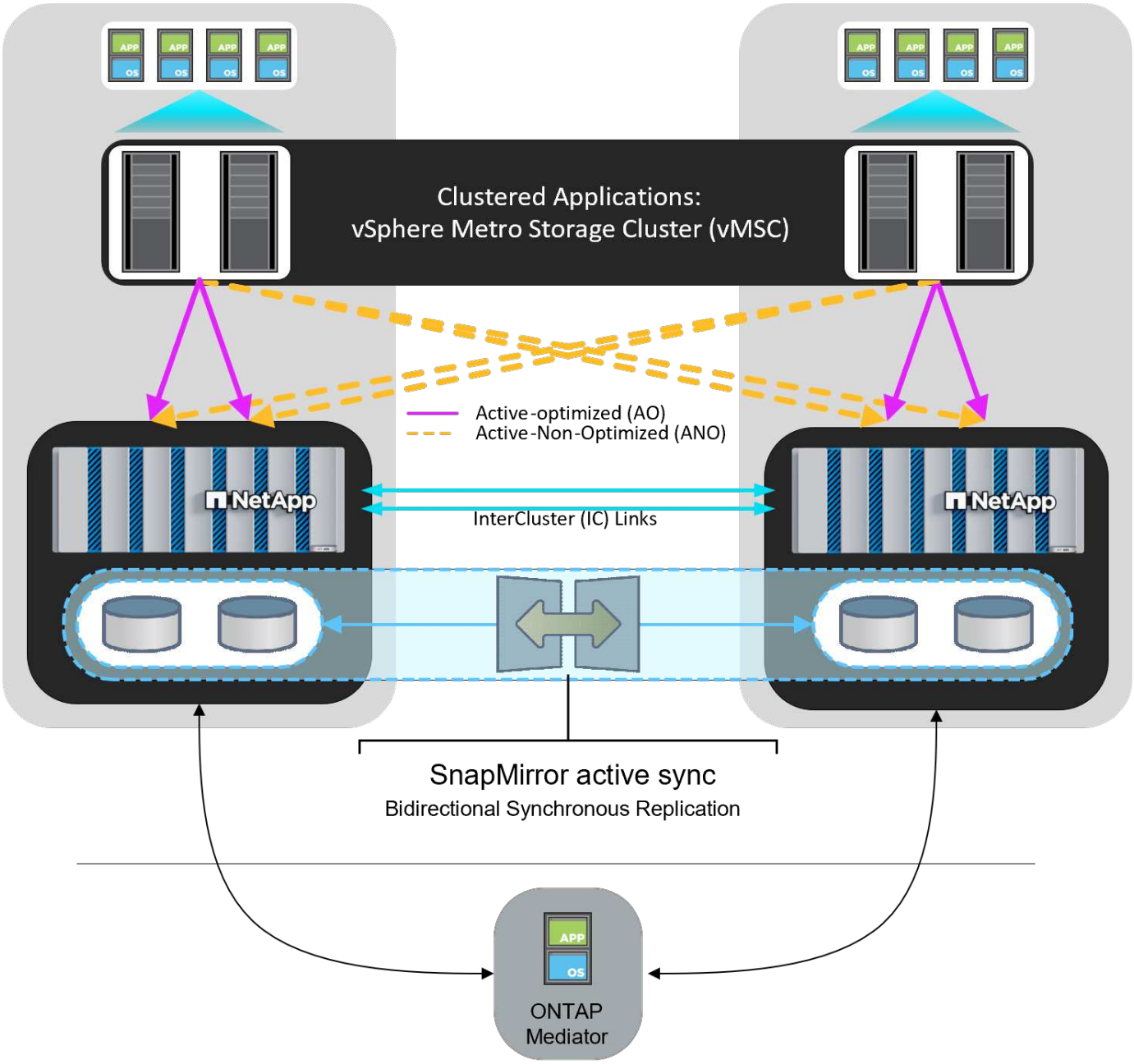
A seguir ilustra um diagrama de topologia de alto nível do Stretch MetroCluster.




```
https://www.netapp.com/support-and-
training/documentation/metrocluster/["Documentação do
MetroCluster"]Consulte para obter informações específicas sobre design e
implantação do MetroCluster.
```

O SnapMirror ative Sync também pode ser implantado de duas maneiras diferentes.

- Assimétrico
- Sincronização ativa simétrica (ONTAP 9.15.1)




```
https://docs.netapp.com/us-en/ontap/smbc/index.html["Documentos do
NetApp"]Consulte para obter informações específicas sobre design e
implementação para sincronização ativa do SnapMirror.
```

Visão geral da solução VMware vSphere

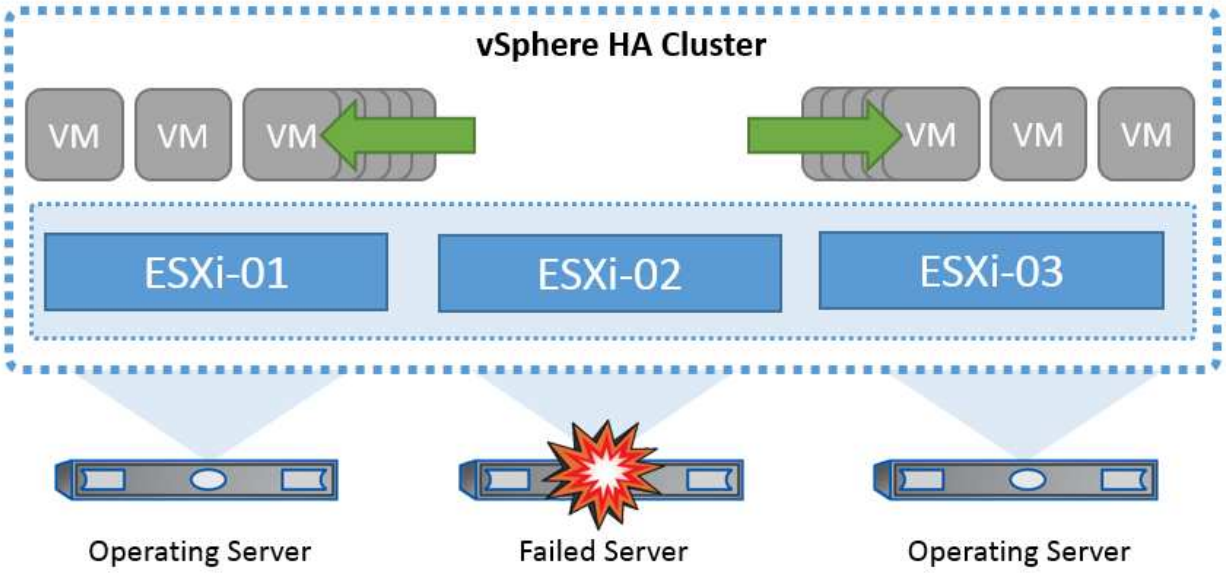
O vCenter Server Appliance (VCSA) é um poderoso sistema de gerenciamento centralizado e um painel de controle único para o vSphere que permite aos administradores operar clusters ESXi de forma eficaz. Ele facilita funções essenciais como provisionamento de máquinas virtuais, operação vMotion, alta disponibilidade (HA), Distributed Resource Scheduler (DRS), VMware vSphere Kubernetes Service (VKS) e muito mais. É um componente essencial em ambientes de nuvem VMware e

deve ser projetado levando em consideração a disponibilidade do serviço.

Alta disponibilidade do vSphere

A tecnologia de cluster da VMware agrupa os servidores ESXi em pools de recursos compartilhados para máquinas virtuais e fornece o vSphere High Availability (HA). O vSphere HA oferece alta disponibilidade e fácil de usar para aplicativos executados em máquinas virtuais. Quando o recurso HA está ativado no cluster, cada servidor ESXi mantém a comunicação com outros hosts para que, se qualquer host ESXi ficar sem resposta ou isolado, o cluster HA possa negociar a recuperação das máquinas virtuais que estavam sendo executadas nesse host ESXi entre os hosts sobreviventes no cluster. No caso de uma falha do sistema operacional convidado, o vSphere HA pode reiniciar a máquina virtual afetada no mesmo servidor físico. O vSphere HA possibilita reduzir o tempo de inatividade planejado, evitar tempo de inatividade não planejado e recuperar rapidamente de interrupções.

Cluster vSphere HA recuperando VMs de um servidor com falha.



É importante entender que o VMware vSphere não tem conhecimento da sincronização ativa do NetApp MetroCluster ou do SnapMirror e vê todos os hosts ESXi no cluster vSphere como hosts qualificados para operações de cluster de HA, dependendo das configurações de afinidade de host e grupo de VM.

Detecção de falha do host

Assim que o cluster HA é criado, todos os hosts do cluster participam da eleição, e um dos hosts se torna o mestre. Cada servidor escravo envia um sinal de pulsação à rede para o servidor mestre, e o servidor mestre, por sua vez, envia um sinal de pulsação a todos os servidores escravos. O host mestre de um cluster vSphere HA é responsável por detectar a falha dos hosts escravos.

Dependendo do tipo de falha detetada, as máquinas virtuais que estão sendo executadas nos hosts podem precisar ser reexecutadas.

Em um cluster do vSphere HA, três tipos de falha de host são detetados:

- Falha - Um host pára de funcionar.
- Isolamento - Um host se torna isolado na rede.
- Partição - Um host perde a conetividade de rede com o host mestre.

O host principal monitora os hosts escravos no cluster. Esta comunicação é feita através da troca de batimentos cardíacos de rede a cada segundo. Quando o host mestre deixa de receber esses batimentos cardíacos de um host escravo, ele verifica a presença de host antes de declarar que o host falhou. A verificação de vivacidade que o host mestre executa é determinar se o host escravo está trocando heartbeats com um dos datastores. Além disso, o host principal verifica se o host responde aos pings ICMP enviados para seus endereços IP de gerenciamento para detetar se ele é meramente isolado de seu nó mestre ou completamente isolado da rede. Ele faz isso fazendo ping no gateway padrão. Um ou mais endereços de isolamento podem ser especificados manualmente para melhorar a confiabilidade da validação de isolamento.

A NetApp recomenda especificar um mínimo de dois endereços de isolamento adicionais e que cada um desses endereços seja local. Isso aumentará a confiabilidade da validação de isolamento.

Resposta de isolamento do host

A Resposta de Isolamento é uma configuração no vSphere HA que determina a ação acionada nas máquinas virtuais quando um host em um cluster vSphere HA perde suas conexões de rede de gerenciamento, mas continua em execução. Existem três opções para esta configuração: "Desativado", "Desligar e reiniciar VMs" e "Desligar e reiniciar VMs".

"Desligar" é melhor do que "Desligar completamente", que não salva as alterações mais recentes no disco nem confirma as transações. Se as máquinas virtuais não forem desligadas em 300 segundos, elas serão desativadas. Para alterar o tempo de espera, use a opção avançada das.isolationshutdowntimeout.

Antes que o HA inicie a resposta de isolamento, ele primeiro verifica se o agente mestre do vSphere HA possui o datastore que contém os arquivos de configuração da VM. Caso contrário, o host não acionará a resposta de isolamento, porque não há mestre para reiniciar as VMs. O host verificará periodicamente o estado do datastore para determinar se ele é reivindicado por um agente do vSphere HA que detém a função mestre.

A NetApp recomenda definir a "resposta de isolamento do host" como Desativado.

Uma condição de split-brain pode ocorrer se um host ficar isolado ou particionado do host master do vSphere HA e o mestre não conseguir se comunicar por meio de datastores de heartbeat ou por ping. O mestre declara o host isolado morto e reinicia as VMs em outros hosts no cluster. Uma condição de split-brain agora existe porque existem duas instâncias da máquina virtual em execução, apenas uma das quais pode ler ou gravar os discos virtuais. As condições de split-brain agora podem ser evitadas configurando a VM Component Protection (VMCP).

Proteção de componentes VM (VMCP)

Um dos aprimoramentos de recursos do vSphere 6, relevantes para o HA, é o VMCP. O VMCP fornece proteção aprimorada contra todas as condições de APD (caminhos para baixo) e PDL (perda permanente de dispositivo) para bloco (FC, iSCSI, FCoE) e armazenamento de arquivos (NFS).

Perda permanente de dispositivo (PDL)

PDL é uma condição que ocorre quando um dispositivo de armazenamento falha permanentemente ou é removido administrativamente e não se espera que retorne. O array de armazenamento NetApp emite um código SCSI Sense para o ESXi, declarando que o dispositivo foi perdido permanentemente. Na seção Condições de Falha e Resposta da VM do vSphere HA, você pode configurar qual deve ser a resposta após a detecção de uma condição PDL.

A NetApp recomenda definir a "Resposta para armazenamento de dados com PDL" como "**Desligar e reiniciar as VMs**". Quando essa condição for detectada, uma máquina virtual será reiniciada instantaneamente em um host íntegro dentro do cluster vSphere HA.

Todos os caminhos para baixo (APD)

APD é uma condição que ocorre quando um dispositivo de armazenamento se torna inacessível ao host e não há caminhos disponíveis para o array. O ESXi considera este um problema temporário com o dispositivo e espera que ele volte a estar disponível em breve.

Quando uma condição APD é detetada, um temporizador é iniciado. Após 140 segundos, a condição APD é oficialmente declarada e o dispositivo é marcado como APD time out. Quando os 140 segundos tiverem passado, o HA começará a contar o número de minutos especificado no atraso para o APD de failover da VM. Quando o tempo especificado tiver passado, o HA reiniciará as máquinas virtuais afetadas. Você pode configurar o VMCP para responder de forma diferente, se desejado (Desativado, Eventos de problemas ou Desligar e reiniciar VMs).

- O NetApp recomenda configurar a "resposta para datastore com APD" para "**Desligar e reiniciar VMs (conservative)**".
- O termo "conservador" refere-se à probabilidade de o HA (alta disponibilidade) conseguir reiniciar as VMs (máquinas virtuais). Quando configurado para o modo Conservador, o HA reiniciará a VM afetada pelo APD somente se souber que outro host pode reiniciá-la. No caso do modo Agressivo, o HA tentará reiniciar a VM mesmo que não conheça o estado dos outros hosts. Isso pode resultar na não reinicialização das VMs caso não haja um host com acesso ao armazenamento de dados onde elas estão localizadas.
- Se o status do APD for resolvido e o acesso ao armazenamento for restaurado antes que o tempo limite tenha passado, o HA não reiniciará desnecessariamente a máquina virtual, a menos que você a configure explicitamente para fazê-lo. Se uma resposta for desejada, mesmo quando o ambiente foi recuperado da condição APD, então Response for APD Recovery After APD Timeout deve ser configurado para Reset VMs.
- O NetApp recomenda configurar a resposta para recuperação do APD após o tempo limite do APD para Desativado.

Implementação do VMware DRS para o NetApp SnapMirror ative Sync

O VMware DRS é um recurso que agrega os recursos de host em um cluster e é usado principalmente para o balanceamento de carga em um cluster em uma infraestrutura virtual. O VMware DRS calcula principalmente os recursos de CPU e memória para realizar o balanceamento de carga em um cluster. Como o vSphere não tem conhecimento do clustering estendido, ele considera todos os hosts em ambos os locais quando o balanceamento de carga.

Implementação do VMware DRS para NetApp MetroCluster

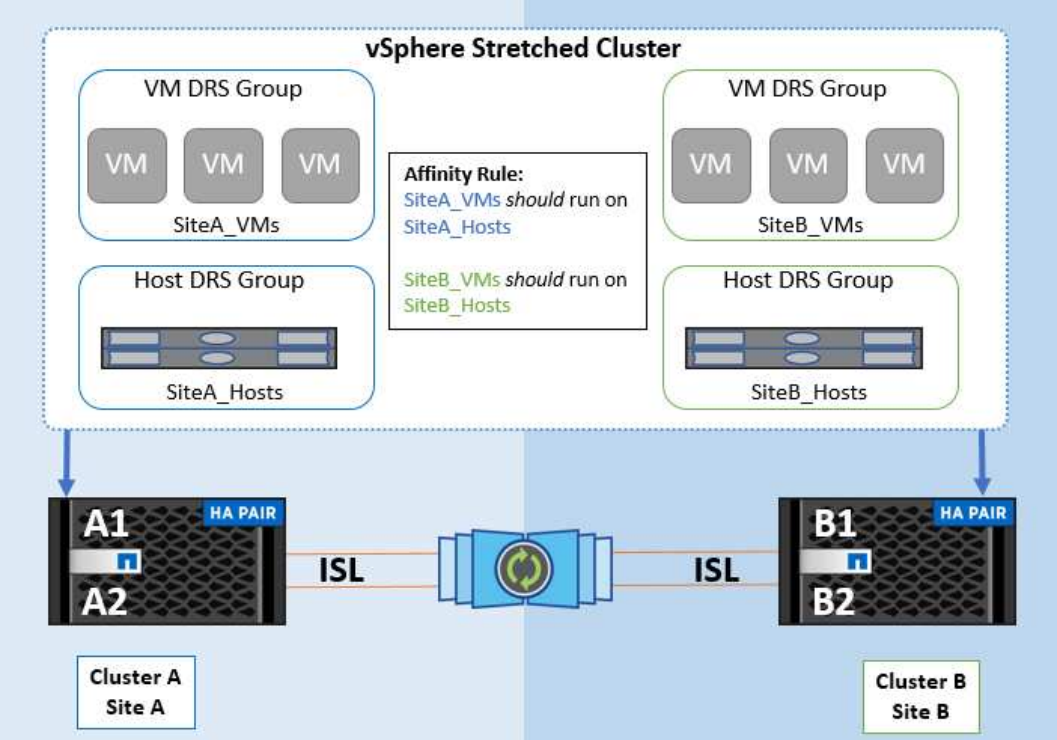
```
 To avoid cross-site traffic, NetApp recommends configuring DRS affinity
rules to manage a logical separation of VMs. This will ensure that, unless
there is a complete site failure, HA and DRS will only use local hosts.
Se você criar uma regra de afinidade DRS para o cluster, poderá
especificar como o vSphere aplica essa regra durante um failover de
máquina virtual.
```

Existem dois tipos de regras que você pode especificar para o comportamento de failover do vSphere HA:

- As regras de anti-afinidade da VM forçam as máquinas virtuais especificadas a permanecerem separadas durante as ações de failover.
- As regras de afinidade de host da VM colocam máquinas virtuais especificadas em um host específico ou em um membro de um grupo definido de hosts durante ações de failover.

Usando regras de afinidade de host de VM no VMware DRS, pode-se ter uma separação lógica entre o local A e o local B para que a VM seja executada no host no mesmo local do array configurado como o controlador de leitura/gravação primário para um determinado datastore. Além disso, as regras de afinidade de host da VM permitem que as máquinas virtuais permaneçam locais para o armazenamento, o que, por sua vez, verifica a conexão da máquina virtual em caso de falhas de rede entre os sites.

A seguir está um exemplo de grupos de hosts de VM e regras de afinidade.



Melhor prática

A NetApp recomenda a implementação de regras "devem" em vez de regras "obrigatórias" porque elas são violadas pelo vSphere HA em caso de falha. O uso de regras "obrigatórias" pode potencialmente levar a interrupções de serviço.

A disponibilidade dos serviços deve sempre prevalecer sobre o desempenho. No cenário em que um data center inteiro falha, as regras "obrigatórias" devem escolher hosts do grupo de afinidade de hosts de VMs e, quando o data center estiver indisponível, as máquinas virtuais não serão reiniciadas.

Implementação do VMware Storage DRS com o NetApp MetroCluster

O recurso VMware Storage DRS permite a agregação de armazenamentos de dados em uma única unidade e equilibra discos de máquina virtual quando os limites de controle de e/S de armazenamento (SIOC) são excedidos.

O controle de e/S de armazenamento é habilitado por padrão nos clusters DRS habilitados para Storage DRS.

O controle de e/S de armazenamento permite que um administrador controle a quantidade de e/S de armazenamento que é alocada a máquinas virtuais durante períodos de congestionamento de e/S, o que permite que máquinas virtuais mais importantes tenham preferência sobre máquinas virtuais menos importantes para alocação de recursos de e/S.

O Storage DRS usa o Storage vMotion para migrar as máquinas virtuais para diferentes datastores dentro de um cluster de datastore. Em um ambiente NetApp MetroCluster, a migração de uma máquina virtual precisa ser controlada nos datastores desse site. Por exemplo, a máquina virtual A, em execução em um host no local A, deve idealmente migrar dentro dos armazenamentos de dados do SVM no local A. se não o fizer, a máquina virtual continuará operando, mas com desempenho degradado, uma vez que a leitura/gravação do disco virtual será do local B através de links entre sites.

*Ao usar o armazenamento ONTAP, é recomendável desativar o DRS de armazenamento.

- O DRS de armazenamento geralmente não é necessário ou recomendado para uso com sistemas de armazenamento ONTAP.
- O ONTAP oferece seus próprios recursos de eficiência de storage, como deduplicação, compressão e compactação, que podem ser afetados pelo Storage DRS.
- Se você estiver usando snapshots do ONTAP , o Storage vMotion deixará para trás a cópia da VM presente no snapshot, o que pode aumentar a utilização do armazenamento e afetar aplicativos de backup como o NetApp SnapCenter, que rastreiam VMs e seus snapshots do ONTAP .

Diretrizes de projeto e implementação do vMSC

Este documento descreve as diretrizes de design e implementação do vMSC com sistemas de storage ONTAP.

Configuração de armazenamento NetApp

As instruções de configuração do NetApp MetroCluster estão disponíveis em "Documentação do MetroCluster". As instruções para a sincronização ativa do SnapMirror (SMA) também estão disponíveis em "Visão geral da continuidade dos negócios da SnapMirror".

Depois de configurar o MetroCluster, administrá-lo é como gerenciar um ambiente ONTAP tradicional. É possível configurar máquinas virtuais de storage (SVMs) usando várias ferramentas, como a interface de linha de comando (CLI), o System Manager ou o Ansible. Uma vez configurados os SVMs, crie interfaces lógicas (LIFs), volumes e números de unidades lógicas (LUNs) no cluster que será usado para operações normais. Esses objetos serão replicados automaticamente para o outro cluster usando a rede de peering de cluster.

Se não estiver usando o MetroCluster ou se você tiver sistemas ONTAP que não sejam compatíveis com o MetroCluster, como sistemas ASA R2, poderá usar o SnapMirror active Sync, que fornece proteção granular do armazenamento de dados e acesso ativo-ativo em vários clusters ONTAP em diferentes domínios de falha. O SMA usa grupos de consistência (CGS) para garantir a consistência da ordem de gravação entre um ou mais datastores e você pode criar vários CGS dependendo dos requisitos do aplicativo e do datastore. Os grupos de consistência são especialmente úteis para aplicativos que exigem sincronização de dados entre vários datastores. Por exemplo, LVMs convidadas distribuídas entre datastores. O SMA também suporta RDMs (Mapeamentos de dispositivo brutos) e armazenamento conetado a convidados com iniciadores iSCSI in-Guest. Você pode aprender mais sobre grupos de consistência em "Visão geral dos grupos de consistência".

Há alguma diferença no gerenciamento de uma configuração vmsc com a sincronização ativa do SnapMirror

quando comparada a um MetroCluster. Primeiro, SMA é uma configuração somente SAN, nenhum datastores NFS pode ser protegido com a sincronização ativa do SnapMirror. Em segundo lugar, você deve mapear ambas as cópias dos LUNs para os hosts ESXi para que eles acessem os datastores replicados em ambos os domínios de falha. Terceiro, você deve criar um ou mais grupos de consistência para os datastores que deseja proteger com a sincronização ativa do SnapMirror. Finalmente, você deve criar uma política do SnapMirror para os grupos de consistência criados. Tudo isso pode ser feito facilmente usando o assistente "proteger cluster" no plug-in do vCenter Tools do ONTAP, ou usando manualmente a CLI do ONTAP ou o Gerenciador do sistema.

Usando o plug-in do vCenter Tools do ONTAP para o SnapMirror ative Sync

O plug-in do ONTAP Tools vCenter oferece uma maneira simples e intuitiva de configurar a sincronização ativa do SnapMirror para vmsc. Você pode usar o plug-in do vCenter Tools do ONTAP para criar e gerenciar relações de sincronização ativa do SnapMirror entre dois clusters do ONTAP. Este plugin fornece uma interface fácil de usar para estabelecer e gerenciar esses relacionamentos de forma eficiente. Você pode saber mais sobre o plugin do vCenter Tools do ONTAP em "Ferramentas do ONTAP para VMware vSphere", ou ir direto para "Proteger usando a proteção do cluster de host"o .

Configuração do VMware vSphere

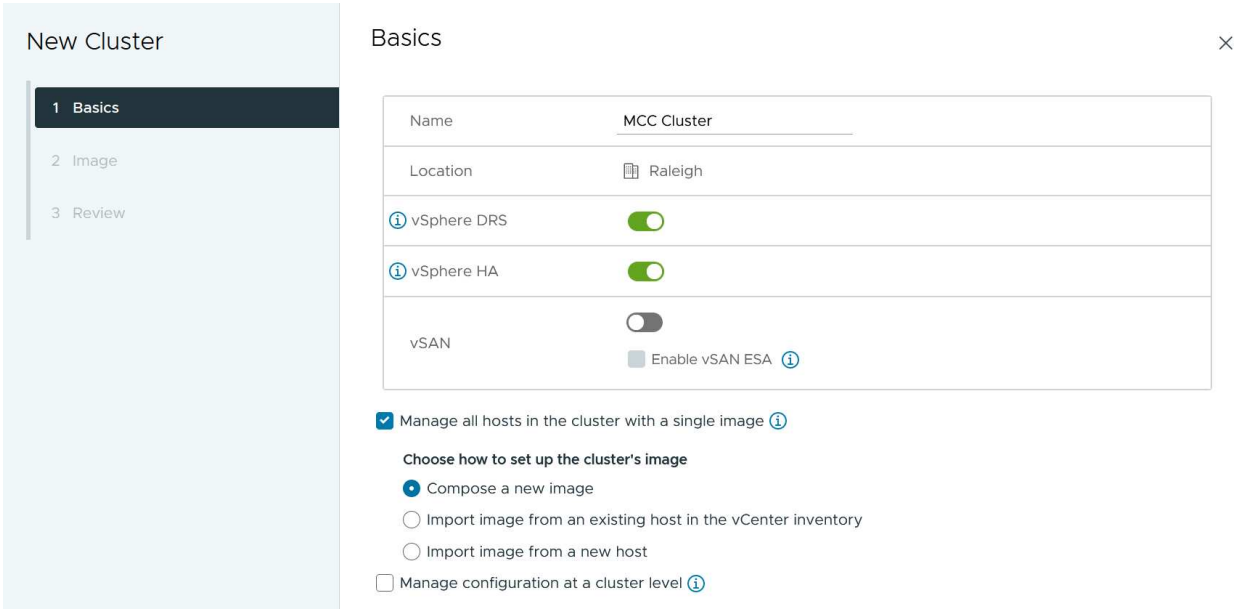
Crie um cluster do vSphere HA

A criação de um cluster do vSphere HA é um processo de várias etapas totalmente documentado em "Como criar e configurar clusters no vSphere Client em docs.vmware.com". Em suma, primeiro você deve criar um cluster vazio e, usando o vCenter, você deve adicionar hosts e especificar o vSphere HA e outras configurações do cluster.

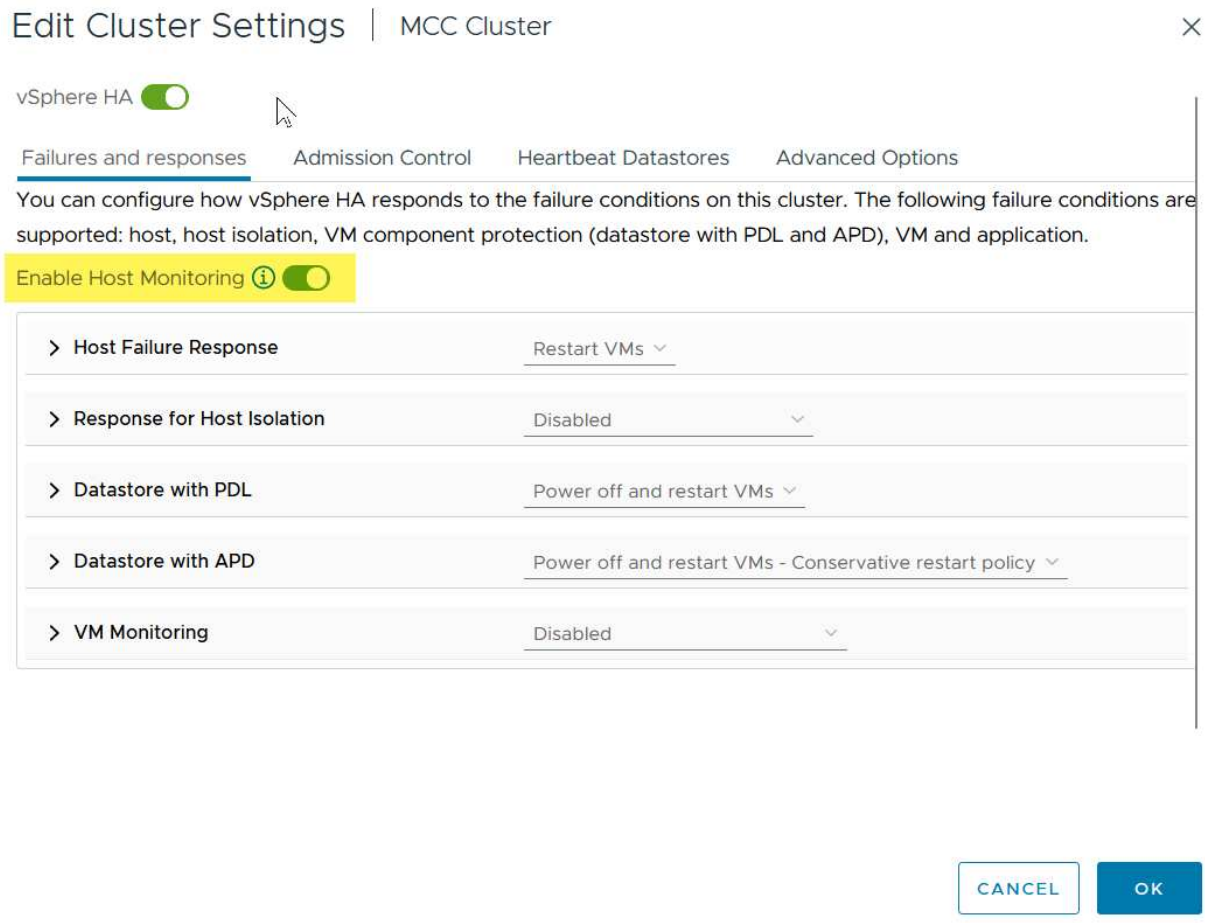
> ⓘ Nada neste documento substitui "Práticas recomendadas do VMware vSphere Metro Storage Cluster". Este conteúdo é fornecido para fácil referência e não substitui a documentação oficial da VMware.

Para configurar um cluster de HA, execute as seguintes etapas:

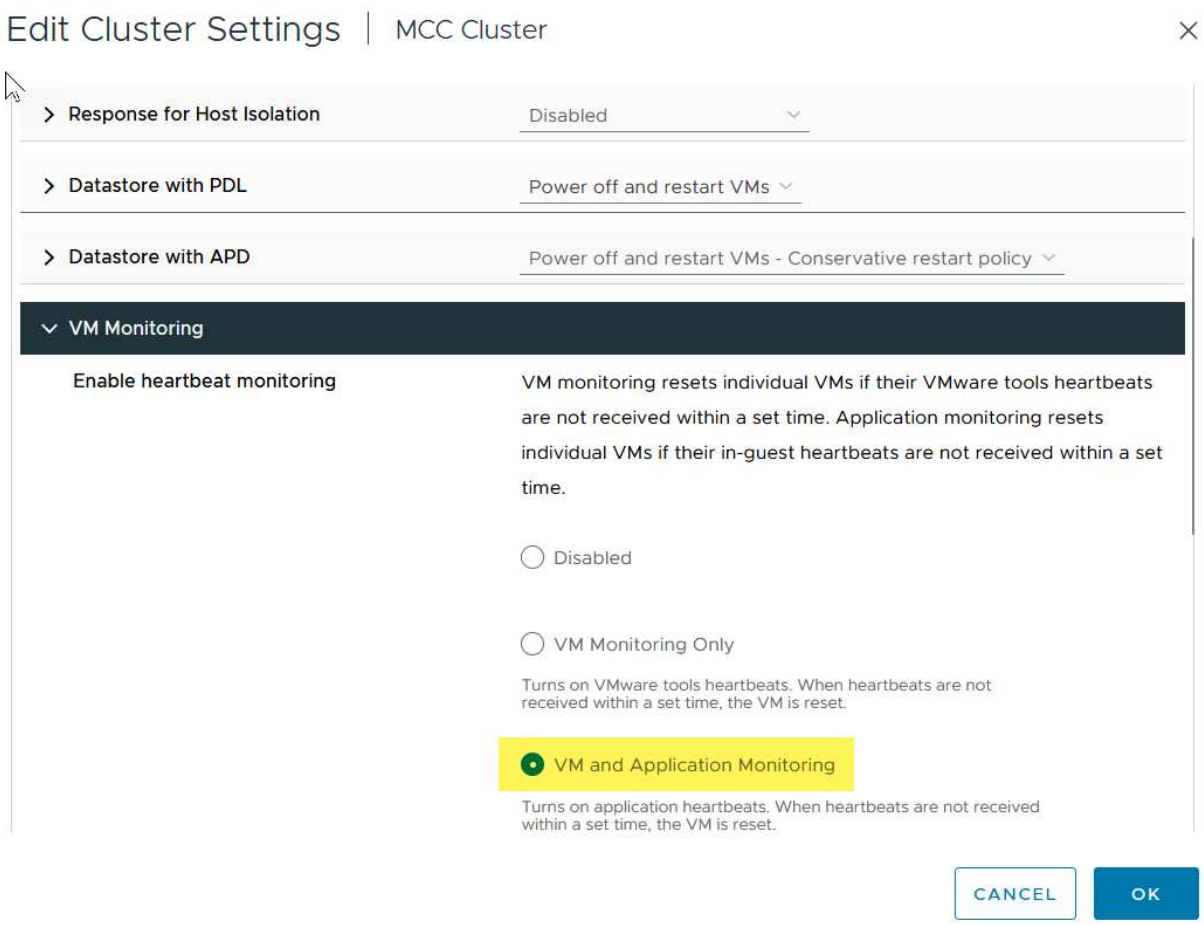
1. Conete-se à IU do vCenter.
2. Em hosts e clusters, navegue até o data center onde você deseja criar seu cluster de HA.
3. Clique com o botão direito do rato no objeto do data center e selecione novo cluster. Em noções básicas, certifique-se de que você ativou o vSphere DRS e o vSphere HA. Conclua o assistente.



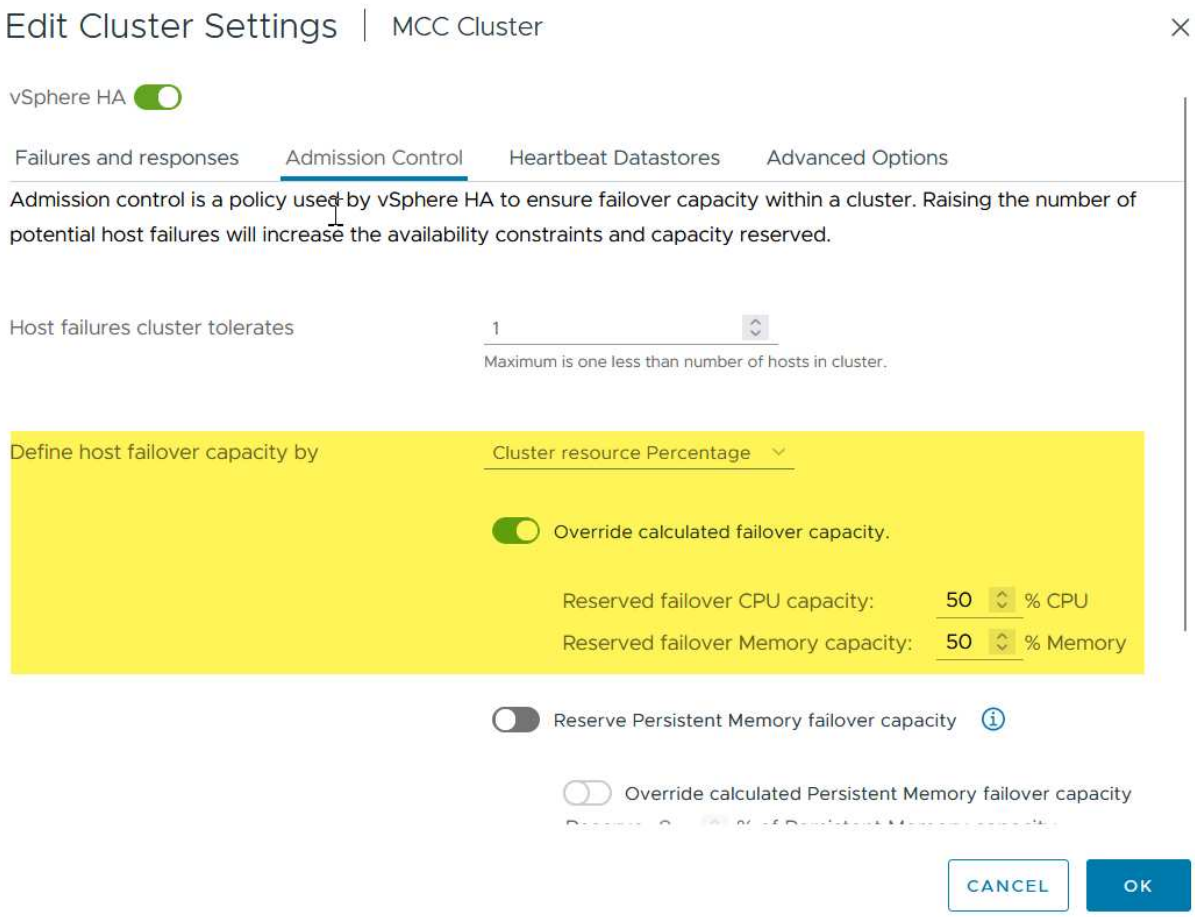
1. Selecione o cluster e vá para a guia configurar. Selecione vSphere HA e clique em Edit.
2. Em Monitoramento de host, selecione a opção Ativar monitoramento de host.



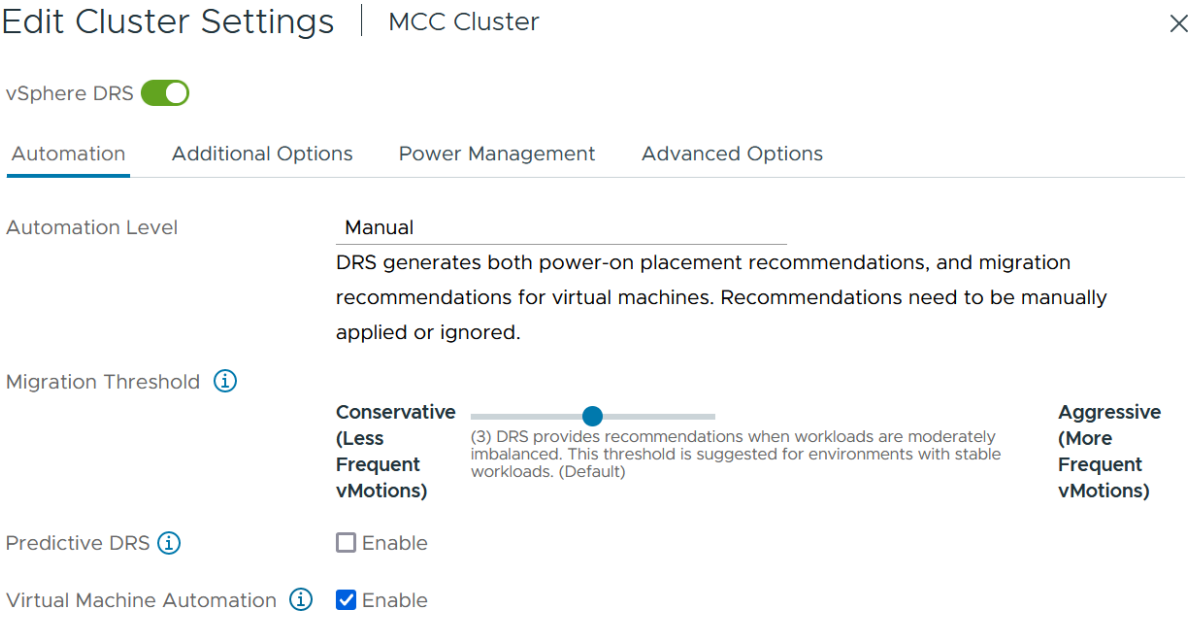
1. Enquanto ainda estiver na guia falhas e respostas, em Monitoramento de VM, selecione a opção somente Monitoramento de VM ou a opção Monitoramento de VM e aplicativo.



1. Em Controle de admissão, defina a opção de controle de admissão HA para reserva de recursos de cluster; use 50% CPU/MEM.



1. Clique em "OK".
2. Selecione DRS e clique EM editar.
3. Defina o nível de automação para manual, a menos que seja necessário pelas suas aplicações.



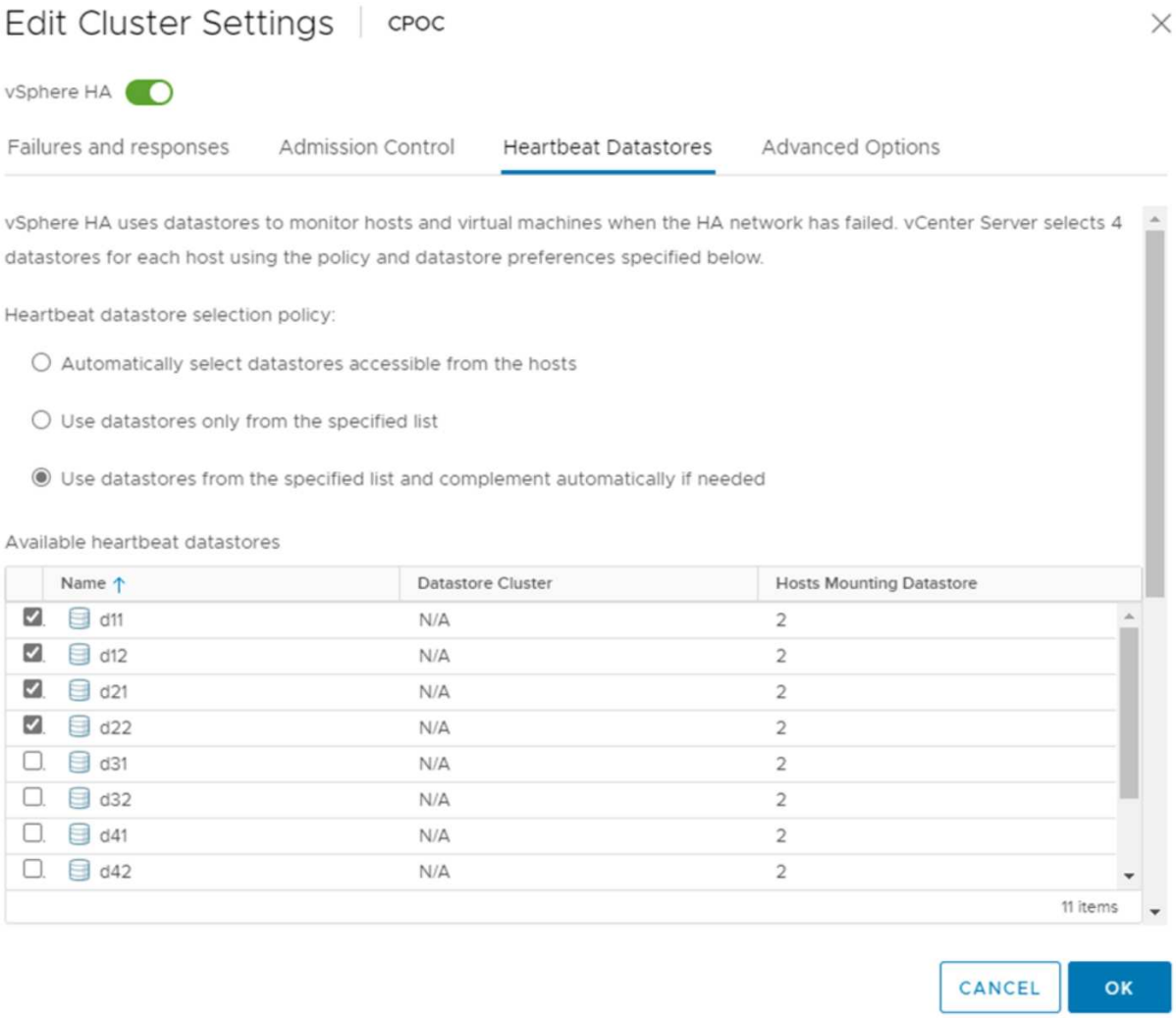
1. Ativar a proteção de componentes VM, "docs.vmware.com" consulte a .
2. As seguintes configurações adicionais do vSphere HA são recomendadas para vMSC com MetroCluster:

Falha	Resposta
Falha do host	Reinicie as VMs
Isolamento de host	Desativado
Armazenamento de dados com perda permanente de dispositivo (PDL)	Desligue e reinicie as VMs
Datastore com todos os caminhos para baixo (APD)	Desligue e reinicie as VMs
Hóspede não é coração batendo	Repor as VMs
Política de reinicialização da VM	Determinado pela importância da VM
Resposta para isolamento do host	Encerre e reinicie as VMs
Resposta para datastore com PDL	Desligue e reinicie as VMs
Resposta para datastore com APD	Desligar e reiniciar as VMs (conservadoras)
Atraso para failover de VM para APD	3 minutos
Resposta para recuperação APD com tempo limite APD	Desativado
Sensibilidade de monitoramento da VM	Predefinição alta

Configurar datastores para Heartbearing

O vSphere HA usa datastores para monitorar hosts e máquinas virtuais quando a rede de gerenciamento falhou. Você pode configurar como o vCenter seleciona armazenamentos de dados Heartbeat. Para configurar armazenamentos de dados para batimentos cardíacos, execute as seguintes etapas:

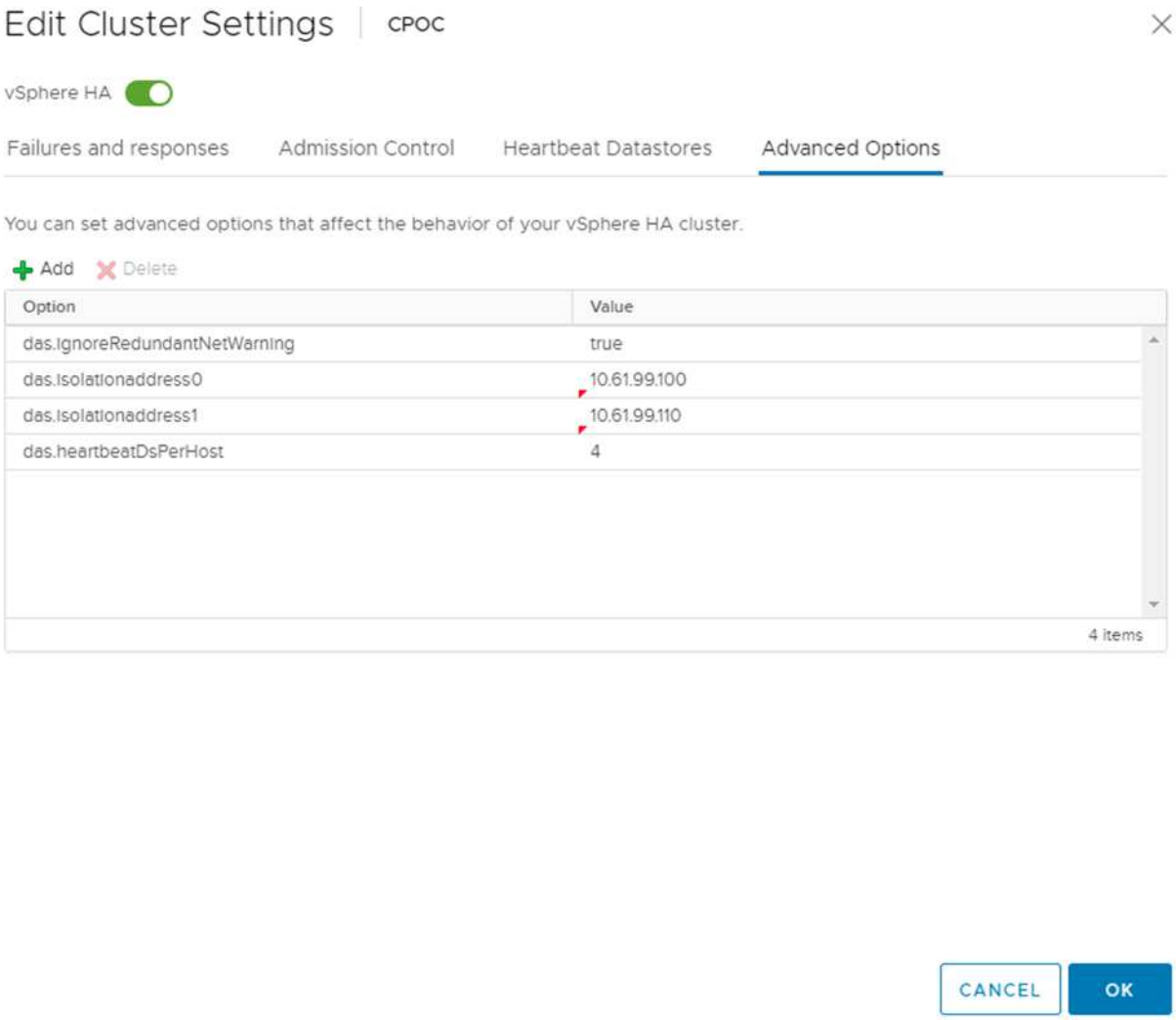
1. Na seção Heartbearing do datastore, selecione usar datastores na Lista especificada e elogiar automaticamente, se necessário.
2. Selecione os datastores que você deseja que o vCenter use em ambos os sites e pressione OK.



Configurar opções avançadas

Os eventos de isolamento ocorrem quando os hosts dentro de um cluster de HA perdem a conetividade com a rede ou com outros hosts no cluster. Por padrão, o vSphere HA usará o gateway padrão para sua rede de gerenciamento como endereço de isolamento padrão. No entanto, você pode especificar endereços de isolamento adicionais para o host fazer ping para determinar se uma resposta de isolamento deve ser acionada. Adicione dois IPs de isolamento que podem fazer ping, um por local. Não utilize o IP do gateway. A configuração avançada do vSphere HA usada é das.isolationaddress. Você pode usar endereços IP do ONTAP ou Mediator para esse fim.

```
https://www.vmware.com/docs/vmw-vmware-vsphere-metro-storage-cluster-
recommended-practices["Práticas recomendadas do VMware vSphere Metro
Storage Cluster"]Consulte para obter mais informações___.__
```



Adicionar uma configuração avançada chamada das.heartbeatDsPerHost pode aumentar o número de datastores de heartbeat. Use quatro datastores de heartbeat (HB DSS) - dois por local. Utilize a opção "Selecionar a partir da lista mas elogio". Isso é necessário porque, se um local falhar, você ainda precisará de dois DSS HB. No entanto, eles não precisam ser protegidos com a sincronização ativa do MetroCluster ou do SnapMirror.

```
https://www.vmware.com/docs/vmw-vmware-vsphere-metro-storage-cluster-
recommended-practices["Práticas recomendadas do VMware vSphere Metro
Storage Cluster"]Consulte para obter mais informações___.__
```

Afinidade do VMware DRS para NetApp MetroCluster

Nesta seção, criamos grupos DRS para VMs e hosts para cada site/cluster no ambiente MetroCluster. Em seguida, configuramos regras VM/Host para alinhar a afinidade do host da VM com os recursos de armazenamento local. Por exemplo, as VMs do Site A pertencem ao grupo VM sitea_vms e os hosts do Site A pertencem ao grupo de hosts sitea_hosts. Em seguida, nas regras VM/Host, declaramos que o sitea_vms deve ser executado em hosts no sitea_hosts.

- O NetApp recomenda altamente a especificação **deve ser executada em hosts no Grupo** em vez da especificação **deve ser executada em hosts no Grupo**. No caso de uma falha de host de um local, as VMs do local A precisam ser reiniciadas em hosts no local B por meio do vSphere HA, mas a última especificação não permite que o HA reinicie VMs no local B porque é uma regra geral. A especificação anterior é uma regra suave e será violada em caso de HA, permitindo assim disponibilidade em vez de desempenho.
- Você pode criar um alarme baseado em eventos que é acionado quando uma máquina virtual viola uma regra de afinidade VM-Host. No vSphere Client, adicione um novo alarme para a máquina virtual e selecione "VM is violing VM-Host Affinity Rule" como gatilho de evento. Para obter mais informações sobre como criar e editar alarmes, "Monitoramento e desempenho do vSphere"consulte a documentação.

Crie grupos de hosts DRS

Para criar grupos de hosts DRS específicos ao local A e local B, execute as seguintes etapas:

1. No cliente da Web vSphere, clique com o botão direito do Mouse no cluster no inventário e selecione Configurações.
2. Clique em VM/Host Groups.
3. Clique em Adicionar.
4. Digite o nome do grupo (por exemplo, sitea_hosts).
5. No menu tipo, selecione Grupo anfitrião.
6. Clique em Adicionar e selecione os hosts desejados no site A e clique em OK.
7. Repita estas etapas para adicionar outro grupo de hosts para o local B.
8. Clique em OK.

Crie grupos de VM DRS

Para criar grupos de VM DRS específicos para o local A e o local B, execute as seguintes etapas:

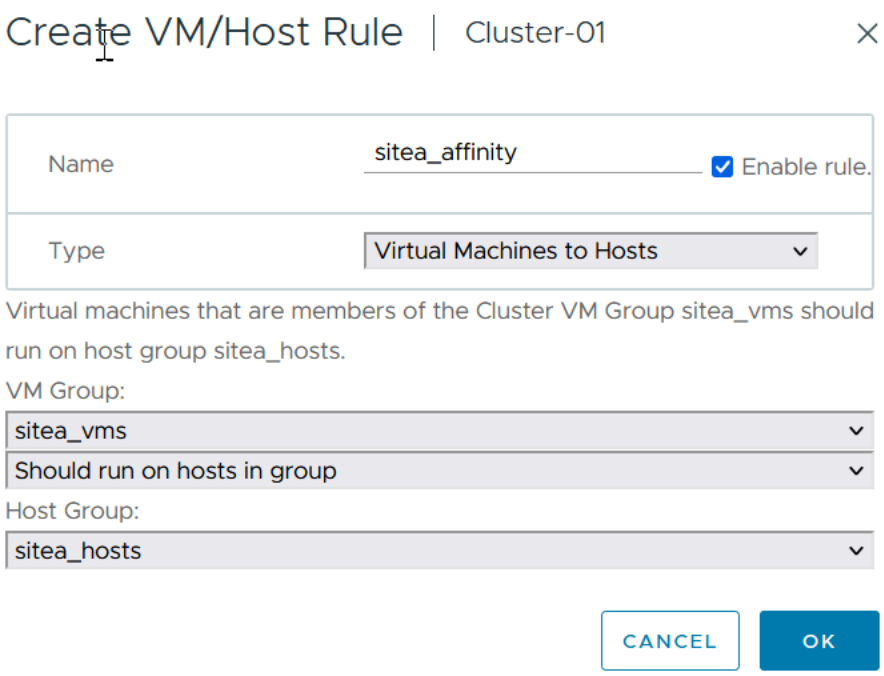
1. No cliente da Web vSphere, clique com o botão direito do Mouse no cluster no inventário e selecione Configurações.
2. Clique em VM/Host Groups.
3. Clique em Adicionar.
4. Digite o nome do grupo (por exemplo, sitea_vms).
5. No menu tipo, selecione Grupo VM.
6. Clique em Adicionar e selecione as VMs desejadas no local A e clique em OK.
7. Repita estas etapas para adicionar outro grupo de hosts para o local B.
8. Clique em OK.

Criar regras de host de VM

Para criar regras de afinidade do DRS específicas ao local A e ao local B, execute as seguintes etapas:

1. No cliente da Web vSphere, clique com o botão direito do Mouse no cluster no inventário e selecione Configurações.

2. Clique em VM/Host Rules.
3. Clique em Adicionar.
4. Digite o nome da regra (por exemplo, sitea_Affinity).
5. Verifique se a opção Ativar regra está marcada.
6. No menu tipo, selecione máquinas virtuais para hosts.
7. Selecione o grupo VM (por exemplo, sitea_vms).
8. Selecione o grupo Host (por exemplo, sitea_hosts).
9. Repita estas etapas para adicionar outra VM/regra de host para o local B.
10. Clique em OK.



Crie clusters de datastore, se necessário

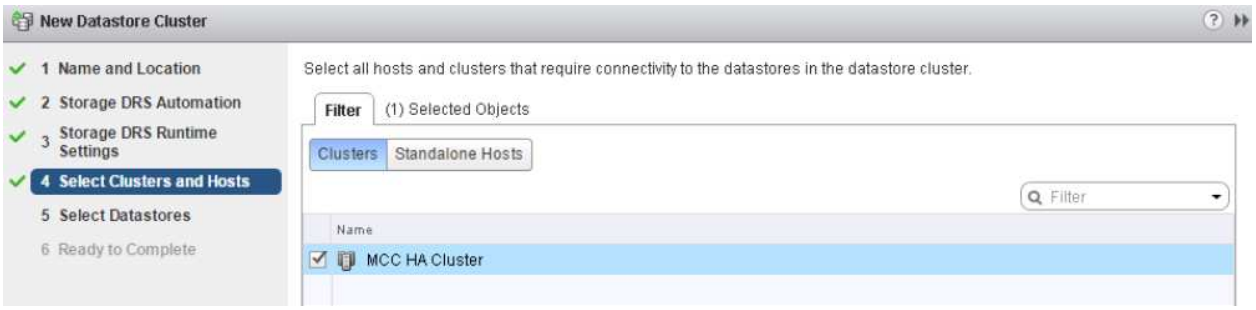
Para configurar um cluster de datastore para cada site, execute as seguintes etapas:

1. Usando o cliente da Web vSphere, navegue até o data center em que o cluster HA reside em Storage.
2. Clique com o botão direito do rato no objeto do data center e selecione armazenamento > novo cluster do datastore.

*Ao usar o armazenamento ONTAP, é recomendável desativar o DRS de armazenamento.

- O DRS de armazenamento geralmente não é necessário ou recomendado para uso com sistemas de armazenamento ONTAP.
- O ONTAP oferece seus próprios recursos de eficiência de storage, como deduplicação, compressão e compactação, que podem ser afetados pelo Storage DRS.
- Se você estiver usando snapshots do ONTAP, o storage vMotion deixaria para trás a cópia da VM no snapshot, aumentando potencialmente a utilização do storage e pode afetar aplicativos de backup, como o NetApp SnapCenter, que rastreiam VMs e seus snapshots do ONTAP.

1. Selecione o cluster HA e clique em Next (seguinte).



1. Selecione os datastores pertencentes ao site A e clique em Avançar.

1. Reveja as opções e clique em concluir.
2. Repita essas etapas para criar o cluster do datastore do site B e verifique se somente os datastores do site B estão selecionados.

Disponibilidade do vCenter Server

Os dispositivos do vCenter Server (VCSAs) devem ser protegidos com o vCenter HA. O vCenter HA permite implantar dois VCSAs em um par de HA ativo-passivo. Um em cada domínio de falha. Você pode ler mais sobre o vCenter HA no "docs.vmware.com".

Resiliência para eventos planejados e não planejados

O NetApp MetroCluster e o SnapMirror ative Sync são ferramentas poderosas que melhoram a alta disponibilidade e as operações ininterruptas do hardware NetApp e do software ONTAP.

Essas ferramentas fornecem proteção em todo o local para todo o ambiente de storage, garantindo que seus dados estejam sempre disponíveis. Quer você esteja usando servidores autônomos, clusters de servidores de alta disponibilidade, contêineres ou servidores virtualizados, a tecnologia NetApp mantém perfeitamente a disponibilidade do storage em caso de uma interrupção total devido à perda de energia, resfriamento ou conectividade de rede, desligamento do storage array ou erro operacional.

O MetroCluster e o SnapMirror ative Sync oferecem três métodos básicos para a continuidade dos dados em caso de eventos planejados ou não planejados:

- Componentes redundantes para proteção contra falha de componente único
- Takeover de HA local para eventos que afetam um único controlador
- Proteção completa do local – retomada rápida do serviço movendo o armazenamento e o acesso do cliente do cluster de origem para o cluster de destino

Isso significa que as operações continuam sem problemas em caso de falha de um único componente e retornam automaticamente à operação redundante quando o componente com falha é substituído.

Todos os clusters do ONTAP, exceto clusters de nó único (normalmente versões definidas por software, como o ONTAP Select, por exemplo), têm recursos de HA incorporados, chamados de takeover e giveback. Cada controlador no cluster é emparelhado com outro controlador, formando um par de HA. Esses pares garantem que cada nó esteja conetado localmente ao storage.

O takeover é um processo automatizado no qual um nó assume o storage do outro para manter os serviços de dados. Giveback é o processo reverso que restaura a operação normal. O takeover pode ser planejado, como ao executar a manutenção de hardware ou atualizações de ONTAP, ou não planejado, resultante de um pânico de nó ou falha de hardware.

Durante uma takeover, LIFs nas nas configurações do MetroCluster fazem failover automático. No entanto, os LIFs SAN não fazem failover; eles continuarão a usar o caminho direto para os LUNs (Logical Unit Numbers).

Para obter mais informações sobre a aquisição de HA e a giveback, consulte o "Visão geral do gerenciamento do par HA". Vale a pena notar que essa funcionalidade não é específica para a sincronização ativa do MetroCluster ou do SnapMirror.

O switchover do local com o MetroCluster ocorre quando um local está off-line ou como uma atividade planejada para manutenção em todo o local. O local restante assume a propriedade dos recursos de storage (discos e agregados) do cluster off-line, e os SVMs no site com falha são colocados on-line e reiniciados no local de desastre, preservando sua identidade completa para acesso ao cliente e ao host.

Com a sincronização ativa do SnapMirror, uma vez que ambas as cópias são usadas ativamente simultaneamente, seus hosts existentes continuarão operando. O Mediador ONTAP é necessário para garantir que o failover do site ocorra corretamente.

Cenários de falha para vMSC com MetroCluster

As seções a seguir descrevem os resultados esperados de vários cenários de falha com

sistemas vMSC e NetApp MetroCluster.

Falha de caminho de storage único

Nesse cenário, se componentes como a porta HBA, a porta de rede, a porta do switch de dados front-end ou um cabo FC ou Ethernet falharem, esse caminho específico para o dispositivo de armazenamento será marcado como morto pelo host ESXi. Se vários caminhos forem configurados para o dispositivo de storage fornecendo resiliência na porta HBA/rede/switch, o ESXi executará idealmente um switchover de caminho. Durante esse período, as máquinas virtuais permanecem em execução sem serem afetadas, pois a disponibilidade para o armazenamento é tratada fornecendo vários caminhos para o dispositivo de armazenamento.

Não há nenhuma mudança no comportamento do MetroCluster neste cenário, e todos os datastores continuam intactos de seus respetivos sites.

Melhor prática

Em ambientes em que os volumes NFS/iSCSI são usados, a NetApp recomenda ter pelo menos dois uplinks de rede configurados para a porta NFS vmkernel no vSwitch padrão e o mesmo no grupo de portas em que a interface NFS vmkernel é mapeada para o vSwitch distribuído. O agrupamento de NIC pode ser configurado em ativo-ativo ou ativo-standby.

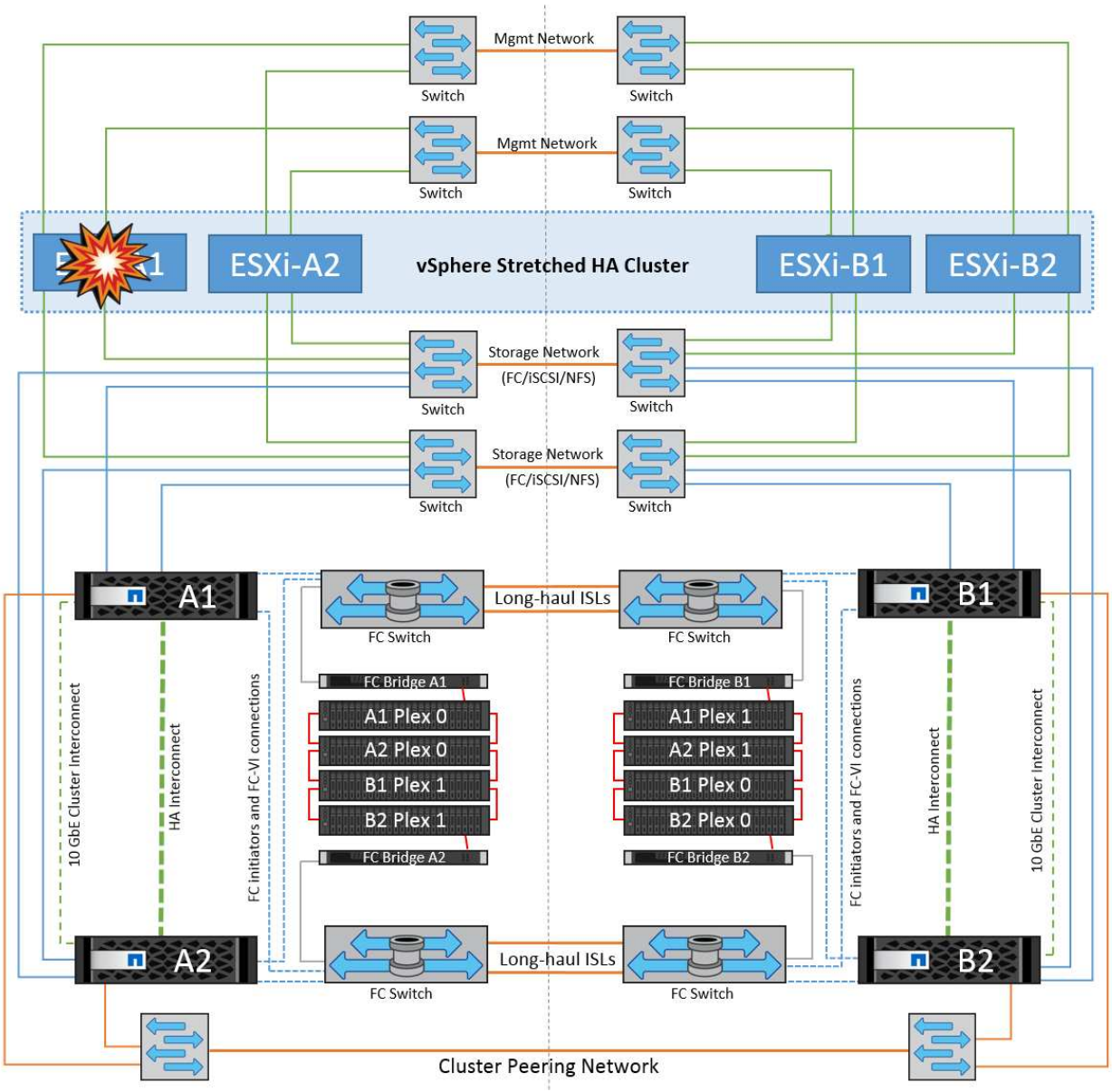
Além disso, para iSCSI LUNs, multipathing deve ser configurado vinculando as interfaces vmkernel aos adaptadores de rede iSCSI. Para obter mais informações, consulte a documentação de armazenamento do vSphere.

Melhor prática

Em ambientes em que LUNs de Fibre Channel são usados, a NetApp recomenda ter pelo menos dois HBAs, o que garante resiliência no nível de HBA/porta. O NetApp também recomenda um iniciador único para o zoneamento de destino único como a melhor prática para configurar o zoneamento.

O VSC (Virtual Storage Console) deve ser usado para definir políticas de multipathing porque define políticas para todos os dispositivos de armazenamento NetApp novos e existentes.

Falha única do host ESXi



Nesse cenário, se houver uma falha do host ESXi, o nó mestre no cluster do VMware HA deteta a falha do host, já que ele não recebe mais batimentos cardíacos da rede. Para determinar se o host está realmente inativo ou apenas uma partição de rede, o nó mestre monitora os batimentos cardíacos do datastore e, se eles estiverem ausentes, ele executa uma verificação final fazendo ping nos endereços IP de gerenciamento do host com falha. Se todas essas verificações forem negativas, o nó principal declara que este host é um host com falha e todas as máquinas virtuais que estavam sendo executadas nesse host com falha são reinicializadas no host sobrevivente no cluster.

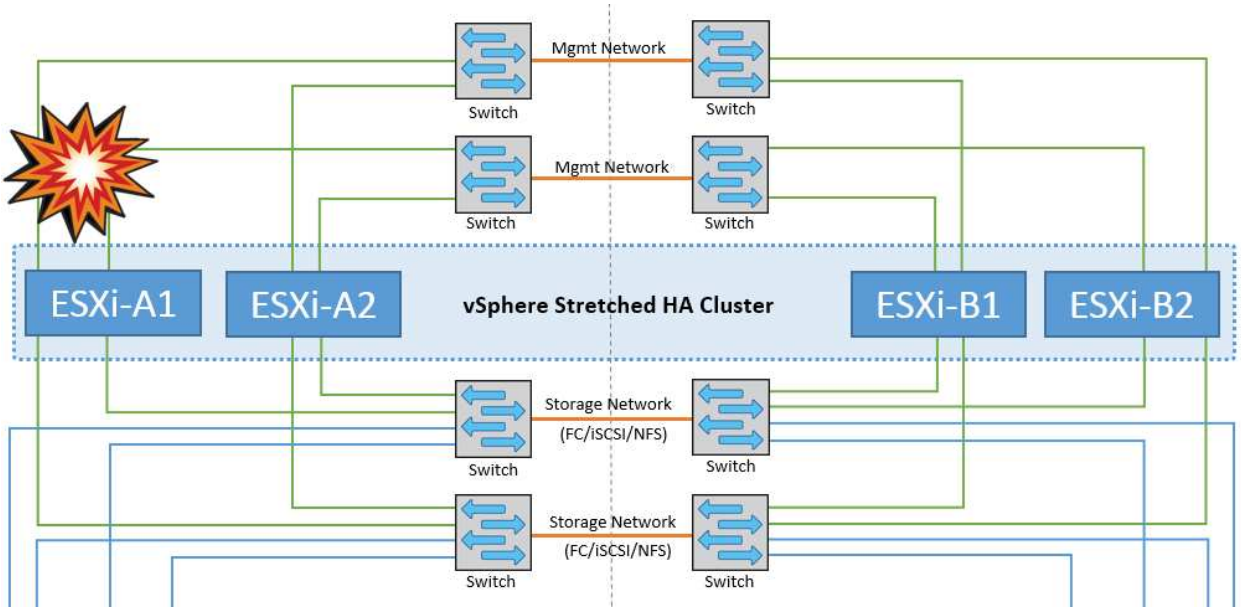
Se as regras de afinidade de host e VM DRS tiverem sido configuradas (as VMs no grupo VM sitea_vms devem executar hosts no grupo host sitea_hosts), o mestre HA primeiro verifica se há recursos disponíveis no local A. se não houver hosts disponíveis no local A, o mestre tentará reiniciar as VMs nos hosts no local B.

É possível que as máquinas virtuais sejam iniciadas nos hosts ESXi no outro site se houver uma restrição de recursos no site local. No entanto, as regras de afinidade de host e VM DRS definidas corrigirão se alguma regra for violada migrando as máquinas virtuais de volta para qualquer host ESXi sobrevivente no site local. Nos casos em que o DRS é definido como manual, o NetApp recomenda chamar o DRS e aplicar as recomendações para corrigir o posicionamento da máquina virtual.

Não há nenhuma mudança no comportamento do MetroCluster neste cenário e todos os datastores continuam

intactos de seus respetivos sites.

Isolamento do host ESXi

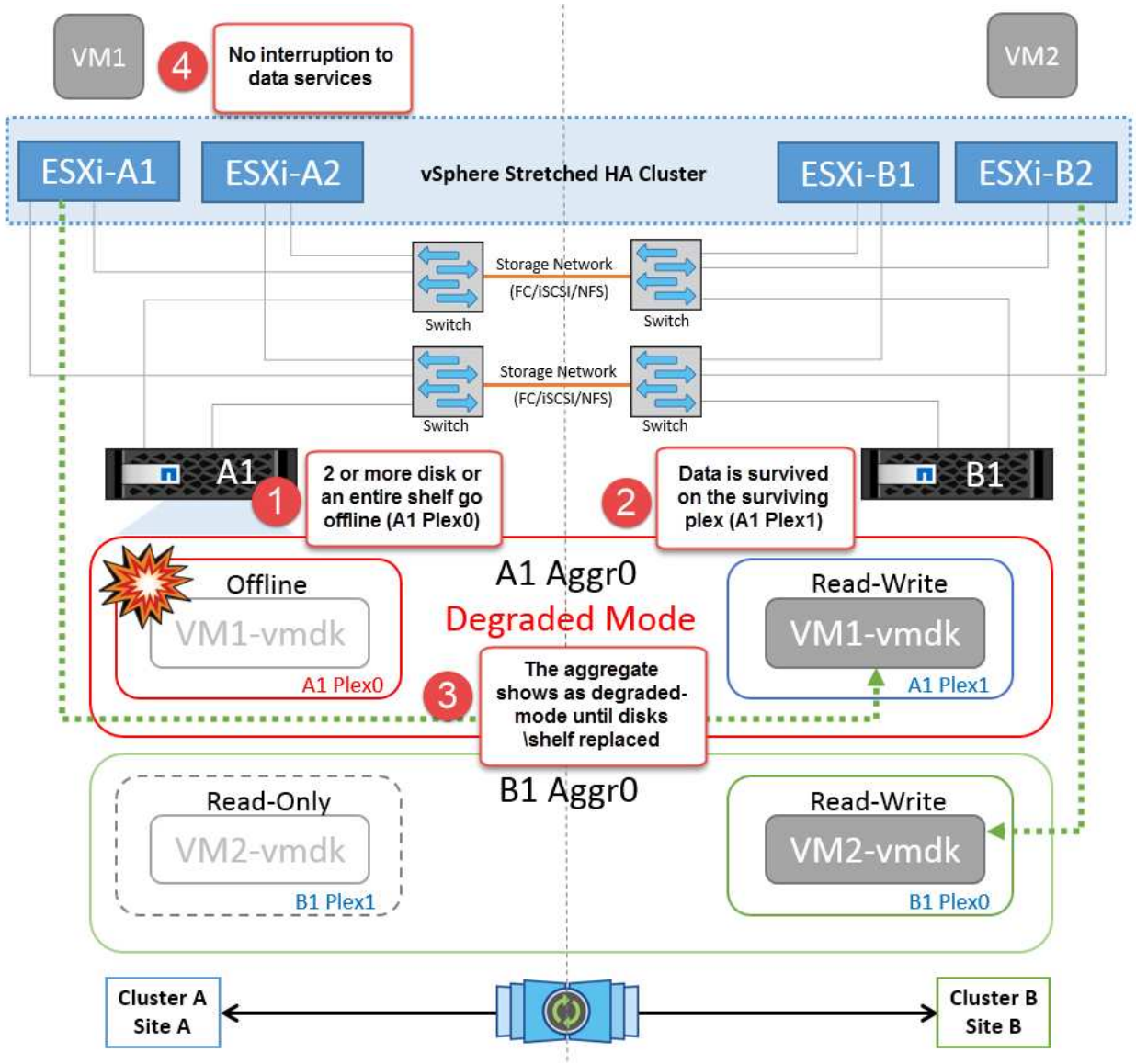


Nesse cenário, se a rede de gerenciamento do host ESXi estiver inativa, o nó mestre no cluster HA não receberá nenhum heartbeats e, portanto, esse host fica isolado na rede. Para determinar se ele falhou ou está isolado apenas, o nó principal começa a monitorar o batimento cardíaco do datastore. Se estiver presente, o host é declarado isolado pelo nó mestre. Dependendo da resposta de isolamento configurada, o host pode optar por desligar, desligar as máquinas virtuais ou até mesmo deixar as máquinas virtuais ligadas. O intervalo padrão para a resposta de isolamento é de 30 segundos.

Não há nenhuma mudança no comportamento do MetroCluster neste cenário e todos os datastores continuam intactos de seus respetivos sites.

Falha no compartimento de disco

Nesse cenário, há uma falha de mais de dois discos ou de uma gaveta inteira. Os dados são fornecidos do Plex sobrevivente sem interrupção para os serviços de dados. A falha do disco pode afetar um Plex local ou remoto. Os agregados serão apresentados como modo degradado porque apenas um Plex está ativo. Depois que os discos com falha forem substituídos, os agregados afetados serão ressincronizados automaticamente para reconstruir os dados. Após a ressincronização, os agregados retornarão automaticamente ao modo espelhado normal. Se mais de dois discos dentro de um único grupo RAID falharem, o Plex terá de ser reconstruído.

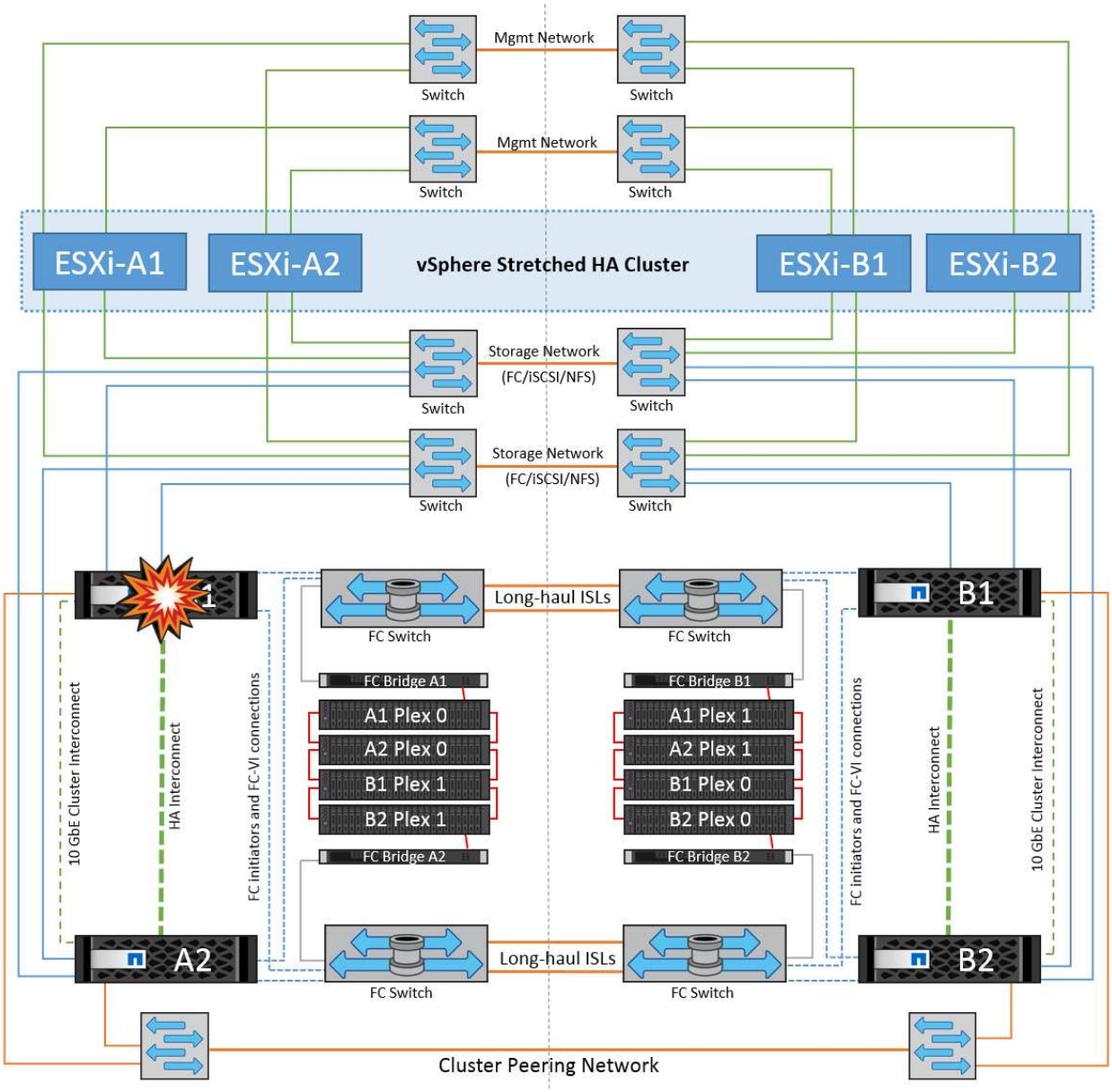


*[NOTA]

- Durante esse período, não há impactos nas operações de e/S da máquina virtual, mas há desempenho degradado porque os dados estão sendo acessados do compartimento de disco remoto por meio de links ISL.

Falha no controlador de storage único

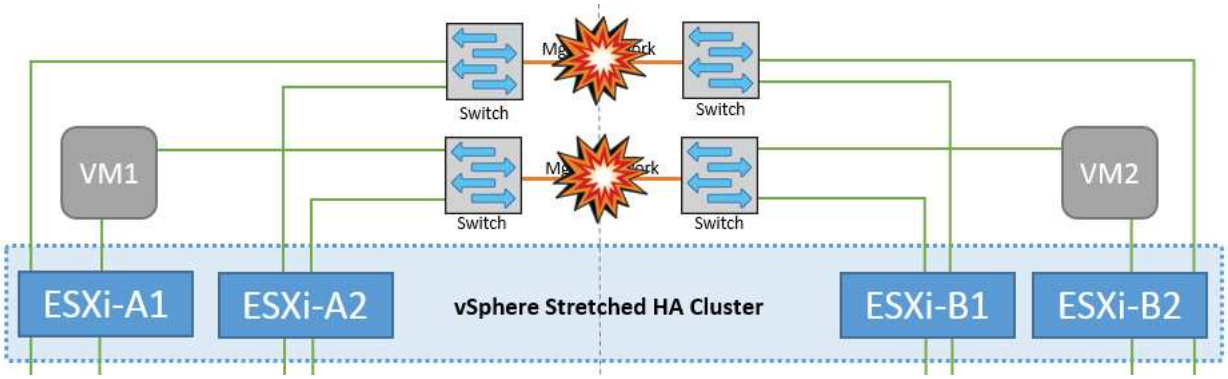
Nesse cenário, um dos dois controladores de storage falha em um local. Como há um par de HA em cada local, uma falha de um nó aciona o failover automaticamente para o outro nó. Por exemplo, se o nó A1 falhar, o storage e os workloads serão transferidos automaticamente para o nó A2. As máquinas virtuais não serão afetadas porque todos os plexos permanecem disponíveis. Os segundo nós do local (B1 e B2) não são afetados. Além disso, o vSphere HA não tomará nenhuma ação porque o nó mestre no cluster ainda estará recebendo os batimentos cardíacos da rede.



Se o failover fizer parte de um desastre contínuo (nó A1 faz failover para A2) e houver uma falha subsequente de A2 ou a falha completa do local A, o switchover após um desastre pode ocorrer no local B.

Avarias na ligação InterSwitch

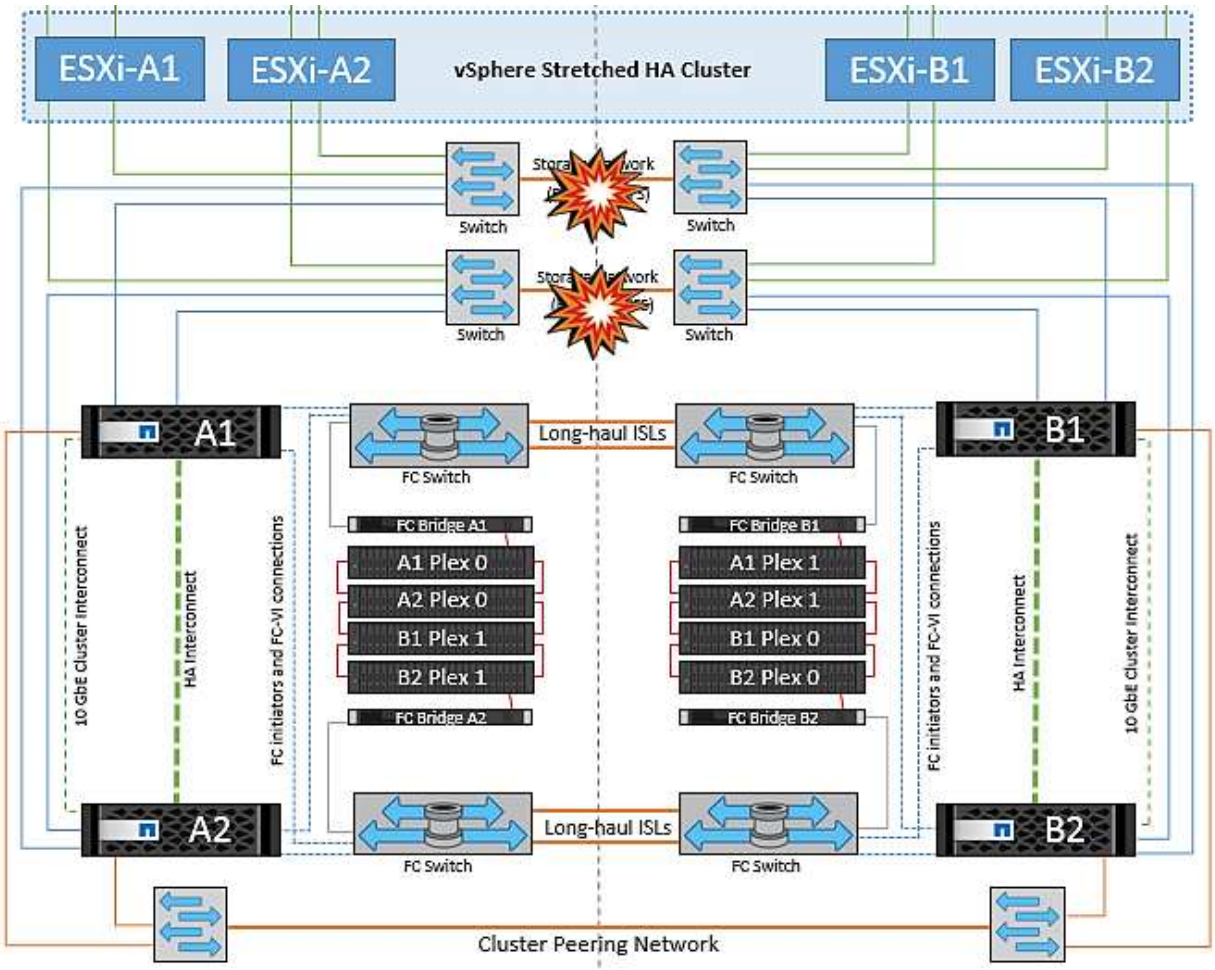
Falha de ligação InterSwitch na rede de gestão



Nesse cenário, se os links ISL na rede de gerenciamento de host front-end falharem, os hosts ESXi no local A não poderão se comunicar com hosts ESXi no local B. isso levará a uma partição de rede porque os hosts ESXi em um determinado local não poderão enviar os batimentos cardíacos da rede para o nó mestre no cluster HA. Como tal, haverá dois segmentos de rede por causa da partição e haverá um nó mestre em cada segmento que protegerá as VMs de falhas de host dentro do site específico.

Durante esse período, as máquinas virtuais permanecem em execução e não há alteração no comportamento do MetroCluster nesse cenário. Todos os armazenamentos de dados continuam intactos de seus respetivos sites.

Falha na ligação InterSwitch na rede de armazenamento

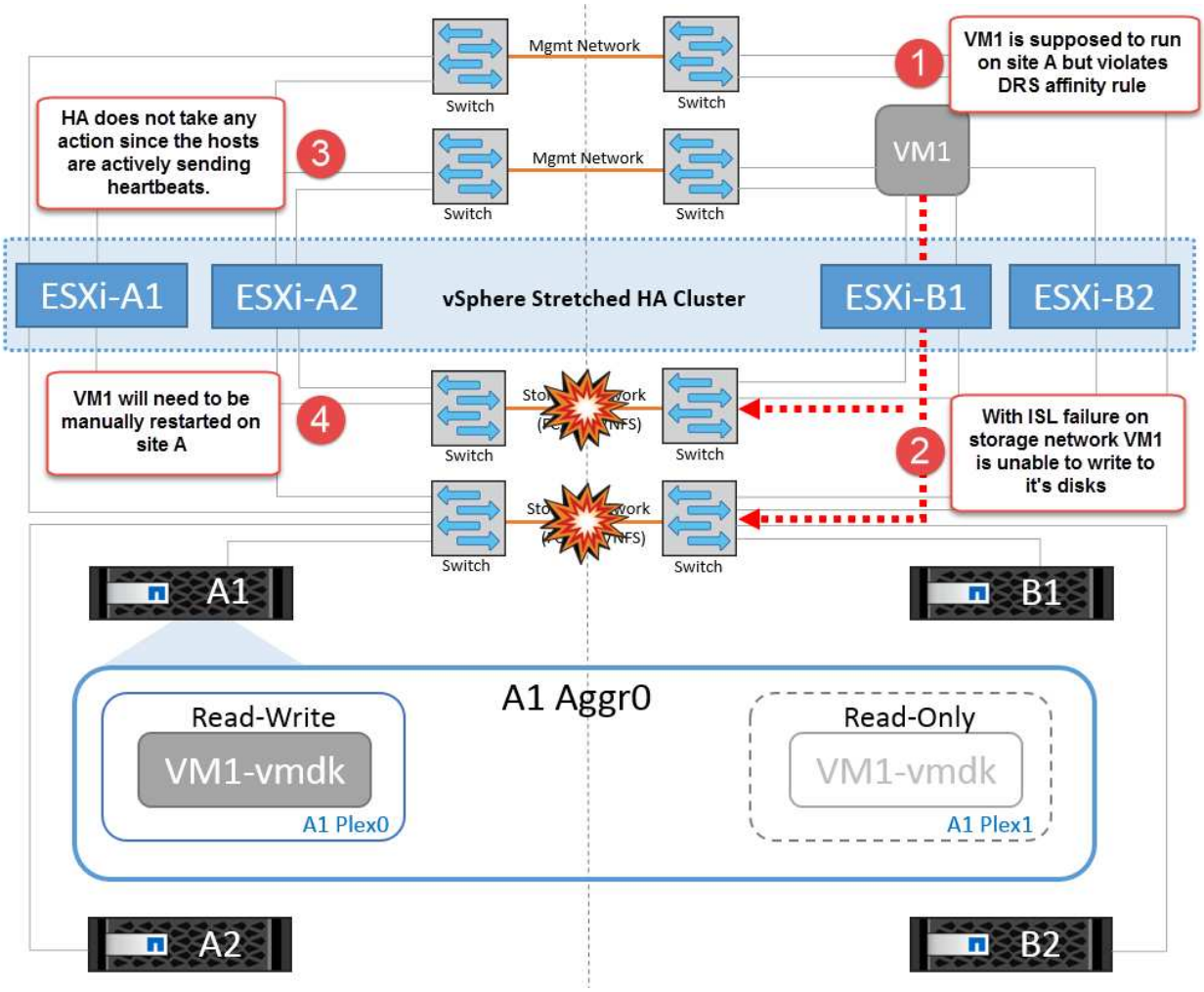


Nesse cenário, se os links ISL na rede de armazenamento de back-end falharem, os hosts no local A perderão acesso aos volumes de armazenamento ou LUNs do cluster B no local B e vice-versa. As regras do VMware DRS são definidas para que a afinidade do local de armazenamento de host facilite a execução das máquinas virtuais sem impactos no local.

Durante esse período, as máquinas virtuais permanecem em execução em seus respetivos sites e não há alteração no comportamento do MetroCluster nesse cenário. Todos os armazenamentos de dados continuam intactos de seus respetivos sites.

Se, por algum motivo, a regra de afinidade foi violada (por exemplo, VM1, que deveria ser executado a partir do site A, onde seus discos residem em nós de cluster local A, está sendo executado em um host no local B), o disco da máquina virtual será acessado remotamente por meio de links ISL. Devido à falha do link ISL, o VM1 em execução no local B não seria capaz de gravar em seus discos porque os caminhos para o volume

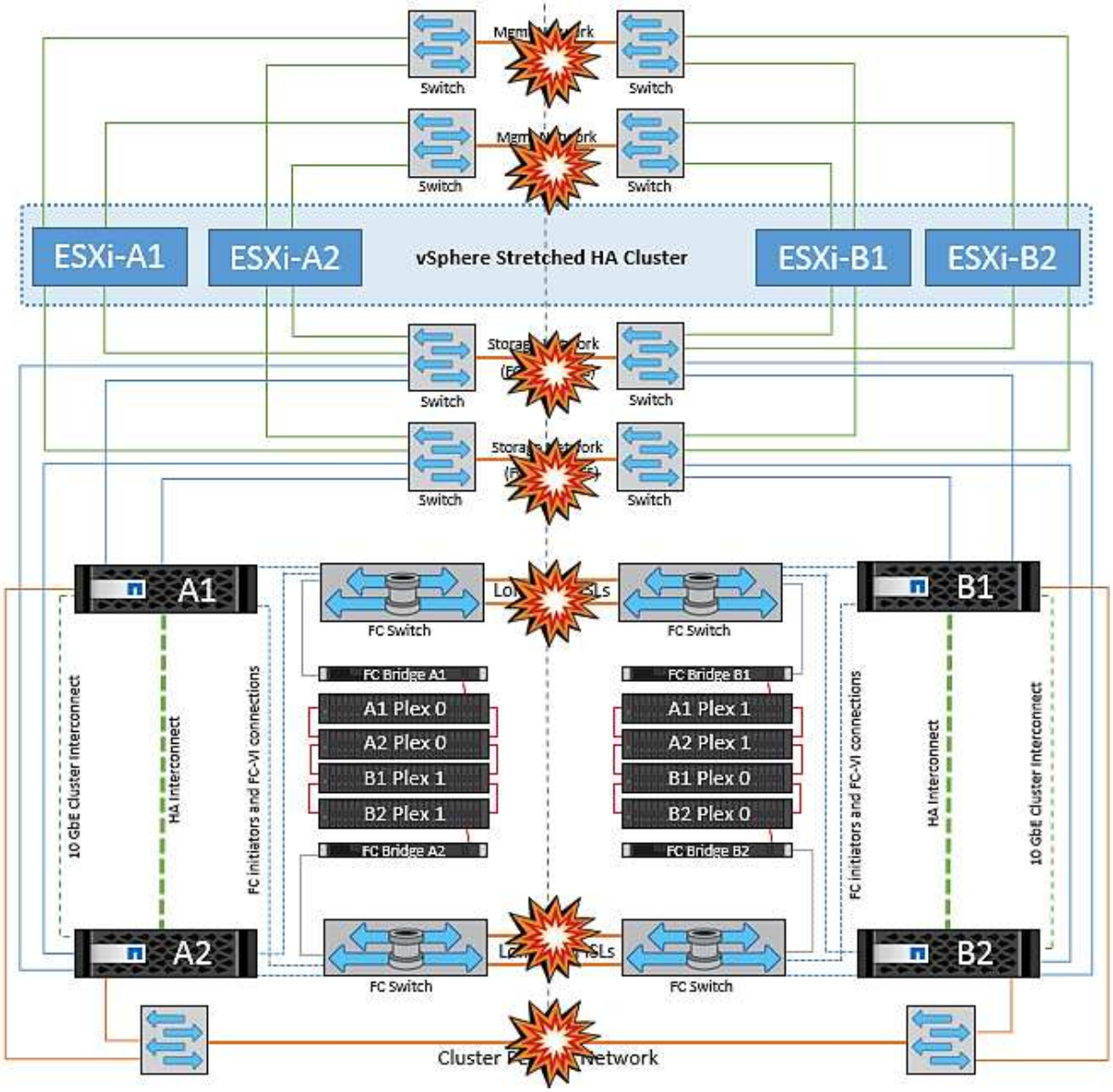
de armazenamento estão inativos e essa máquina virtual específica está inativa. Nessas situações, o VMware HA não toma nenhuma ação, uma vez que os hosts estão enviando batimentos cardíacos ativamente. Essas máquinas virtuais precisam ser manualmente desligadas e ligadas em seus respetivos sites. A figura a seguir ilustra uma VM que viola uma regra de afinidade DRS.



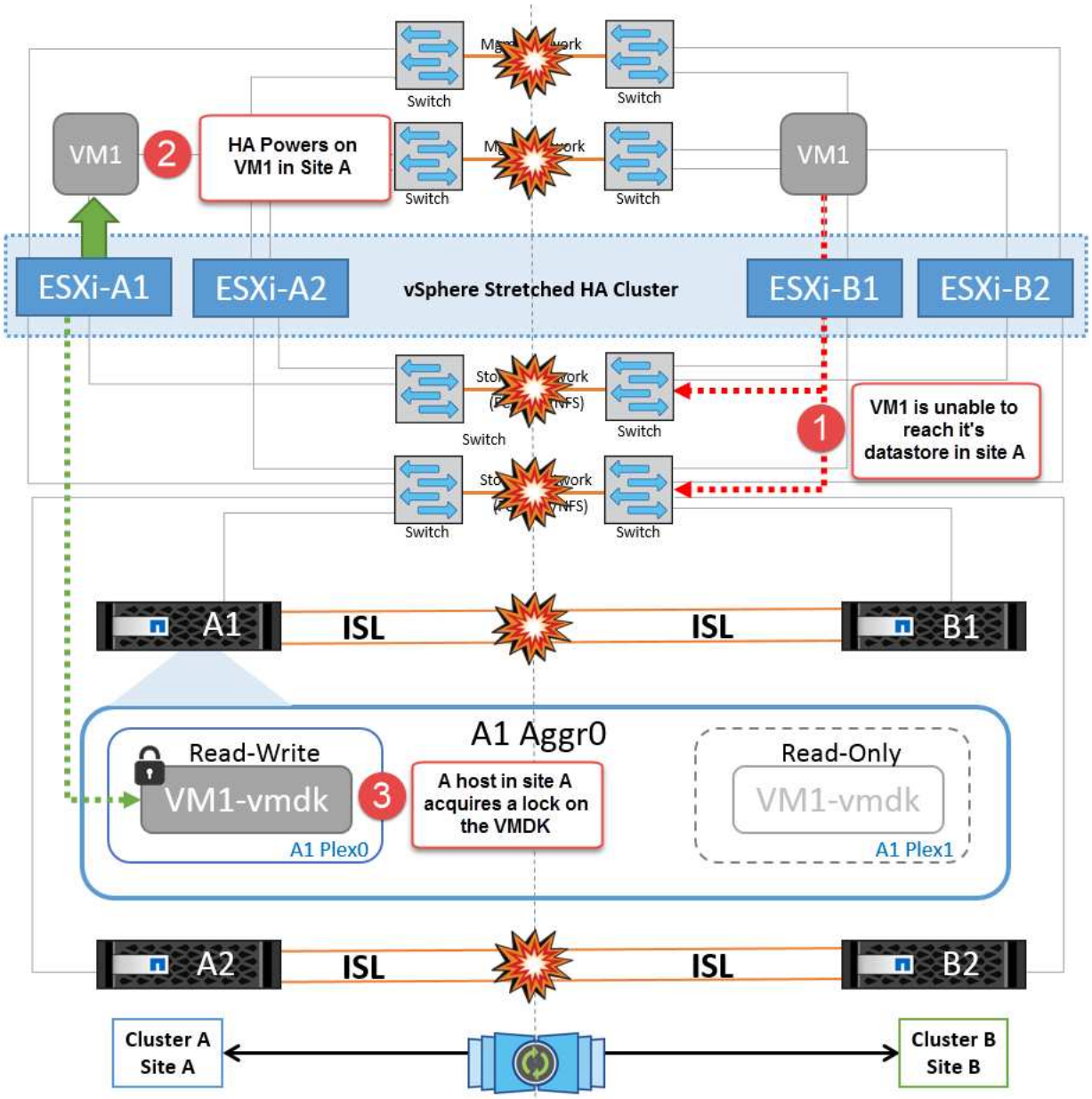
Todas as falhas do InterSwitch ou completa partição do data center

Neste cenário, todos os links ISL entre os sites estão inativos e ambos os sites são isolados uns dos outros. Como discutido em cenários anteriores, como falha de ISL na rede de gerenciamento e na rede de armazenamento, as máquinas virtuais não são afetadas em falha completa de ISL.

Depois que os hosts ESXi forem particionados entre sites, o agente do vSphere HA verificará os batimentos cardíacos do datastore e, em cada site, os hosts ESXi locais poderão atualizar os batimentos cardíacos do datastore para o respetivo volume/LUN de leitura/gravação. Os hosts no local A assumirão que os outros hosts ESXi no local B falharam porque não há heartbeats de rede/datastore. O vSphere HA no local A tentará reiniciar as máquinas virtuais do local B, o que acabará falhará porque os datastores do local B não estarão acessíveis devido a falha do ISL de armazenamento. Uma situação semelhante é repetida no local B..



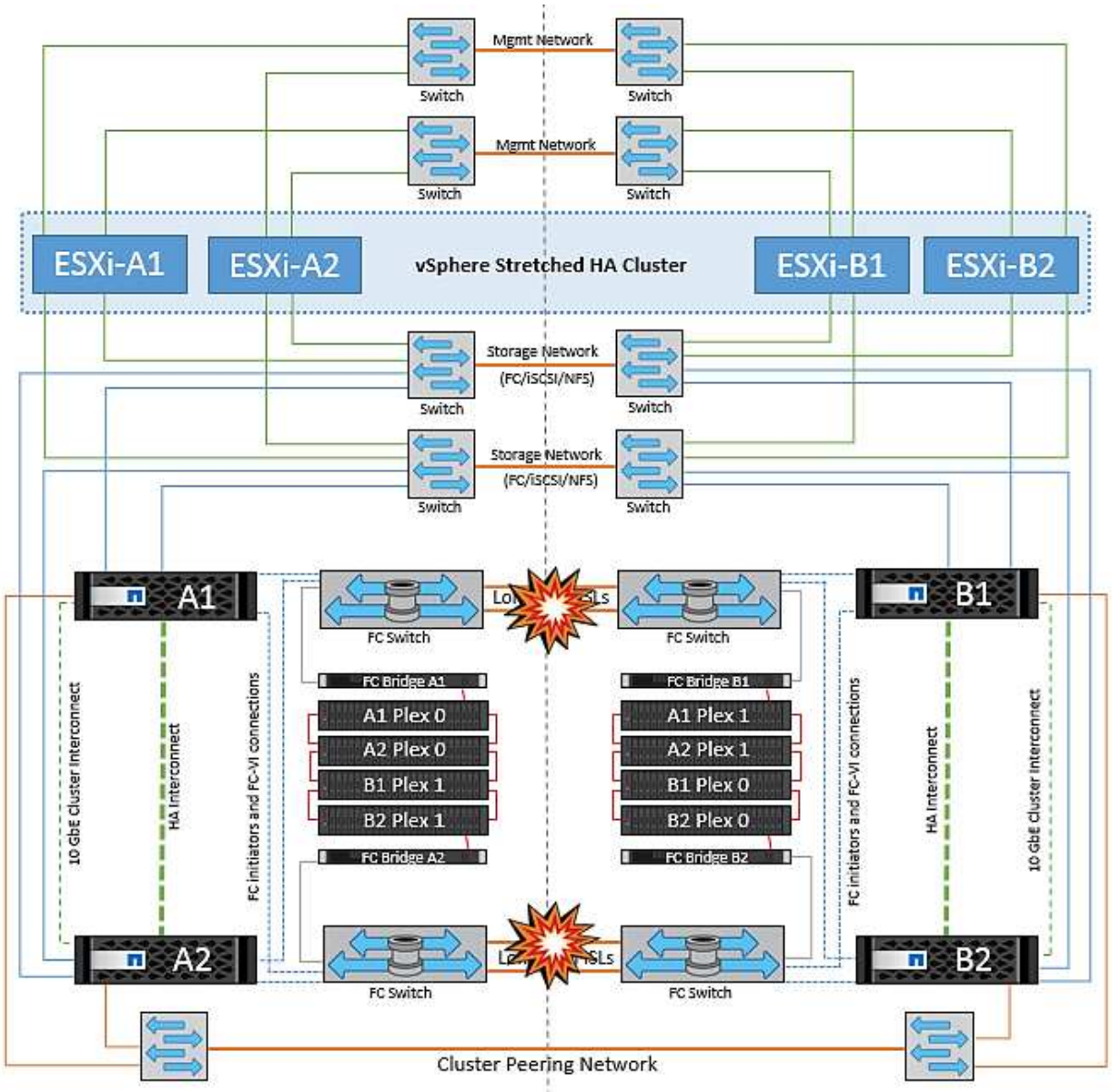
A NetApp recomenda determinar se alguma máquina virtual violou as regras do DRS. Todas as máquinas virtuais executadas a partir de um site remoto ficarão inativas, uma vez que não poderão acessar o datastore, e o vSphere HA reiniciará essa máquina virtual no site local. Depois que os links ISL estiverem novamente online, a máquina virtual que estava sendo executada no local remoto será morta, uma vez que não pode haver duas instâncias de máquinas virtuais executando com os mesmos endereços MAC.



Falha de ligação InterSwitch em ambas as malhas no NetApp MetroCluster

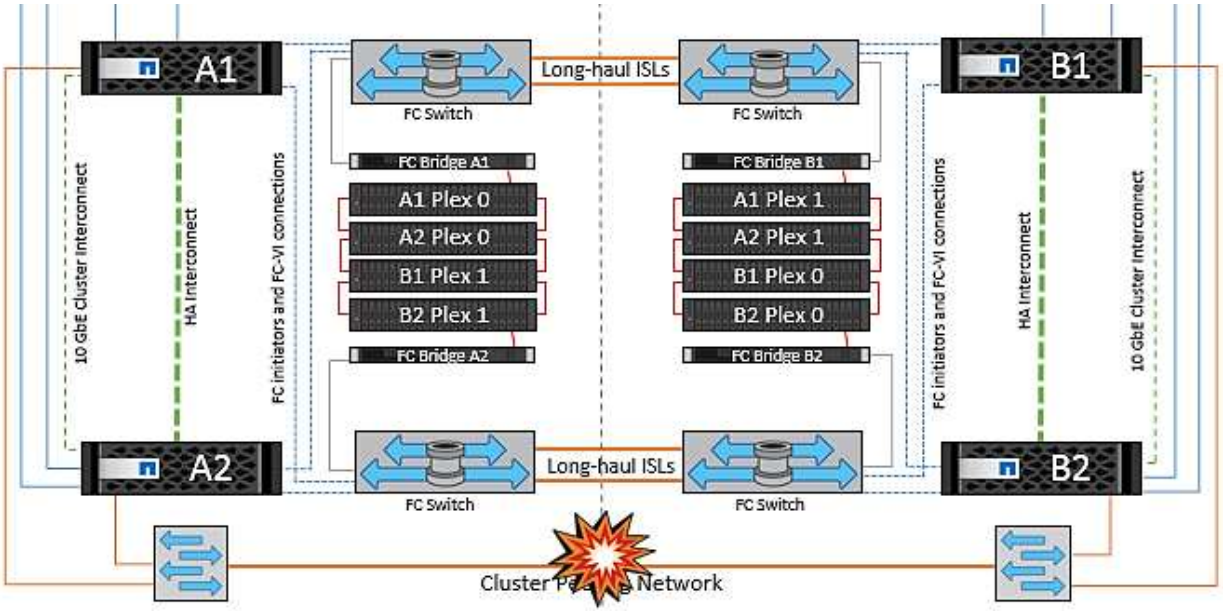
Em um cenário de falha de um ou mais ISLs, o tráfego continua através dos links restantes. Se todos os ISLs em ambas as malhas falharem, de modo que não haja nenhum link entre os locais para armazenamento e replicação do NVRAM, cada controladora continuará fornecendo seus dados locais. Em um mínimo de um ISL é restaurado, a ressincronização de todos os plexos acontecerá automaticamente.

Quaisquer gravações que ocorram depois de todos os ISLs estarem inativos não serão espelhadas para o outro site. Um switchover em caso de desastre, enquanto a configuração estiver nesse estado, incorreria, portanto, na perda dos dados que não haviam sido sincronizados. Neste caso, a intervenção manual é necessária para a recuperação após a mudança. Se for provável que nenhum ISLs esteja disponível por um período prolongado, um administrador pode optar por encerrar todos os serviços de dados para evitar o risco de perda de dados se for necessário um switchover em caso de desastre. A execução dessa ação deve ser ponderada contra a probabilidade de um desastre exigir mudança antes de pelo menos uma ISL ficar disponível. Alternativamente, se os ISLs estiverem falhando em um cenário em cascata, um administrador pode acionar um switchover planejado para um dos sites antes que todos os links tenham falhado.



Falha na ligação do cluster de peered

Em um cenário de falha de enlace de cluster com peered, como os ISLs de malha ainda estão ativos, os serviços de dados (leituras e gravações) continuam em ambos os locais para ambos os plexos. Qualquer alteração na configuração de cluster (por exemplo, adicionar um novo SVM, provisionar um volume ou LUN em um SVM existente) não pode ser propagado para o outro local. Estes são mantidos nos volumes de metadados locais do CRS e propagados automaticamente para o outro cluster após a recuperação do link do cluster peered. Se for necessário um switchover forçado antes que o link do cluster com peering possa ser restaurado, as alterações pendentes na configuração do cluster serão reproduzidas automaticamente da cópia replicada remota dos volumes de metadados no local que sobreviveu como parte do processo de switchover.



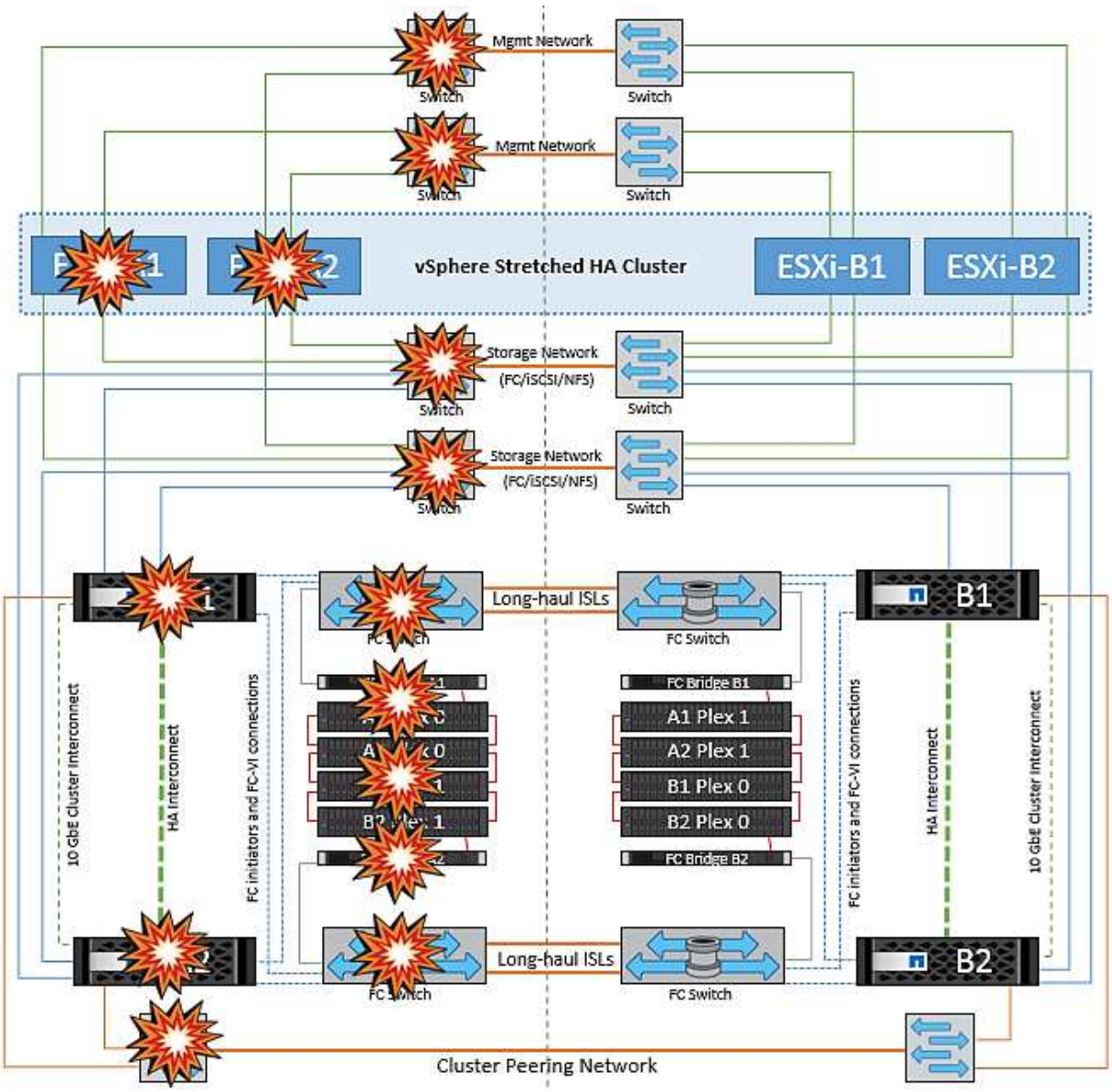
Falha no local completo

Em um local completo Um cenário de falha, os hosts ESXi no local B não receberão o heartbeat da rede dos hosts ESXi no local A porque estão inoperantes. O mestre de HA no local B verificará se os batimentos cardíacos do armazenamento de dados não estão presentes, declarará que os hosts no local A estão com falha e tentará reiniciar o local Uma máquina virtual no local B. durante esse período, o administrador de storage executa um switchover para retomar os serviços dos nós com falha no local sobrevivente, o que restaurará todos os serviços de armazenamento do Local A no local B. após o local A volumes ou LUNs estarem disponíveis no local B, o agente de HA tentará reiniciar o local B.

Se a tentativa do agente mestre do vSphere HA de reiniciar uma VM (que envolve registrá-la e ligá-la) falhar, a reinicialização será novamente tentada após um atraso. O atraso entre reinicializações pode ser configurado até um máximo de 30 minutos. O vSphere HA tenta reiniciar para um número máximo de tentativas (seis tentativas por padrão).

O mestre de HA não inicia as tentativas de reinicialização até que o gerente de colocação encontre um armazenamento adequado, portanto, no caso de uma falha completa no local, isso seria depois que o switchover foi executado.

Se o local A tiver sido substituído, uma falha subsequente de um dos nós do local B sobreviventes pode ser tratada de forma otimizada pelo failover para o nó sobrevivente. Neste caso, o trabalho de quatro nós agora está sendo executado por apenas um nó. A recuperação neste caso consistiria em realizar um giveback para o nó local. Em seguida, quando o local A é restaurado, uma operação de switchback é executada para restaurar a operação de estado estável da configuração.


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

Informações sobre direitos autorais

Informações sobre marcas comerciais