VSphere Metro Storage-Cluster mit ONTAP

Enterprise applications

NetApp
May 03, 2024

Inhalt

VSphere Metro Storage-Cluster mit ONTAP

VSphere Metro Storage-Cluster mit ONTAP

Der branchenführende vSphere-Hypervisor von VMware kann als erweiterbares Cluster, auch als vSphere Metro Storage-Cluster (vMSC) bezeichnet, implementiert werden.

VMSC Lösungen werden sowohl mit NetApp® MetroCluster™ als auch mit SnapMirror Active Sync (früher als SnapMirror Business Continuity oder SMBC bekannt) unterstützt und bieten erweiterte Business Continuity, wenn eine oder mehrere Ausfall-Domains ausfallen. Die Widerstandsfähigkeit gegenüber verschiedenen Fehlermodi hängt davon ab, welche Konfigurationsoptionen Sie wählen.

Lösungen für kontinuierliche Verfügbarkeit für vSphere Umgebungen

Die ONTAP-Architektur ist eine flexible und skalierbare Storage-Plattform, die SAN- (FCP, iSCSI und NVMe-of) und NAS-Services (NFS v3 und v4.1) für Datastores bietet. Die Storage-Systeme NetApp AFF, ASA und FAS nutzen das ONTAP Betriebssystem, um zusätzliche Protokolle für Gast-Storage wie S3 und SMB/CIFS anzubieten.

NetApp MetroCluster nutzt die NetApp HA-Funktion (Controller Failover oder CFO) zum Schutz vor Controller-Ausfällen. Außerdem beinhaltet es lokale SyncMirror Technologie, Cluster Failover bei Disaster (Controller Failover on Demand oder CFOD), Hardware-Redundanz und geografische Trennung, um ein hohes Maß an Verfügbarkeit zu erreichen. SyncMirror spiegelt Daten synchron auf die beiden Hälften der MetroCluster Konfiguration und schreibt sie in zwei Plexe: Der lokale Plex (auf dem lokalen Shelf), der aktiv Daten bereitstellt, und der Remote-Plex (auf dem Remote-Shelf), der normalerweise keine Daten bereitstellt. Für alle MetroCluster Komponenten wie Controller, Storage, Kabel, Switches (zur Verwendung mit Fabric MetroCluster) und Adapter besteht Hardwareredundanz.

NetApp SnapMirror Active Sync bietet granularen Schutz von Datenspeichern mit FCP- und iSCSI SAN-Protokollen, sodass Sie nur Workloads mit hoher Priorität selektiv schützen können. Es bietet aktiv/aktiv-Zugriff auf lokale und Remote-Standorte, im Gegensatz zu NetApp MetroCluster, die eine aktiv/Standby-Lösung ist. Derzeit ist Active Sync eine asymmetrische Lösung, bei der eine Seite der anderen bevorzugt wird, um eine bessere Performance zu bieten. Dies wird durch die ALUA-Funktionalität (Asymmetric Logical Unit Access) erreicht, die den ESXi Host automatisch informiert, welche Controller bevorzugt werden sollen. NetApp hat jedoch angekündigt, dass die aktive Synchronisierung in Kürze einen vollständig symmetrischen Zugriff ermöglichen wird.

Um einen VMware HA/DRS Cluster über zwei Standorte zu erstellen, werden ESXi-Hosts von einer vCenter Server Appliance (VCSA) verwendet und gemanagt. Die vSphere-Management-, vMotion®- und Virtual Machine-Netzwerke sind über ein redundantes Netzwerk zwischen den beiden Standorten verbunden. Der vCenter Server, der den HA/DRS Cluster verwaltet, kann eine Verbindung zu den ESXi-Hosts an beiden Standorten herstellen und sollte über vCenter HA konfiguriert werden.

Siehe "Wie erstellen und konfigurieren Sie Cluster im vSphere Client" Um vCenter HA zu konfigurieren.

Weitere Informationen finden Sie unter "Empfohlene Practices für VMware vSphere Metro Storage-Cluster".

Was ist vSphere Metro Storage-Cluster?

VSphere Metro Storage Cluster (vMSC) ist eine zertifizierte Konfiguration, die Virtual Machines (VMs) und Container vor Ausfällen schützt. Dies wird erreicht, indem erweiterte Storage-Konzepte zusammen mit Clustern von ESXi Hosts genutzt werden, die auf verschiedene Ausfall-Domains verteilt sind, zum Beispiel

Racks, Gebäude, Campus oder sogar Städte. Die NetApp MetroCluster und die Active Sync Storage-Technologien von SnapMirror sorgen für eine Sicherung von RPO=0 bzw. nahezu RPO=0 für die Host-Cluster. Mit der vMSC-Konfiguration ist sichergestellt, dass die Daten immer verfügbar sind, selbst wenn ein vollständiger physischer oder logischer „Standort“ ausfällt. Ein Speichergerät, das Teil der vMSC-Konfiguration ist, muss nach einer erfolgreichen vMSC-Zertifizierung zertifiziert werden. Alle unterstützten Speichergeräte finden Sie im "VMware Storage Compatibility Guide".

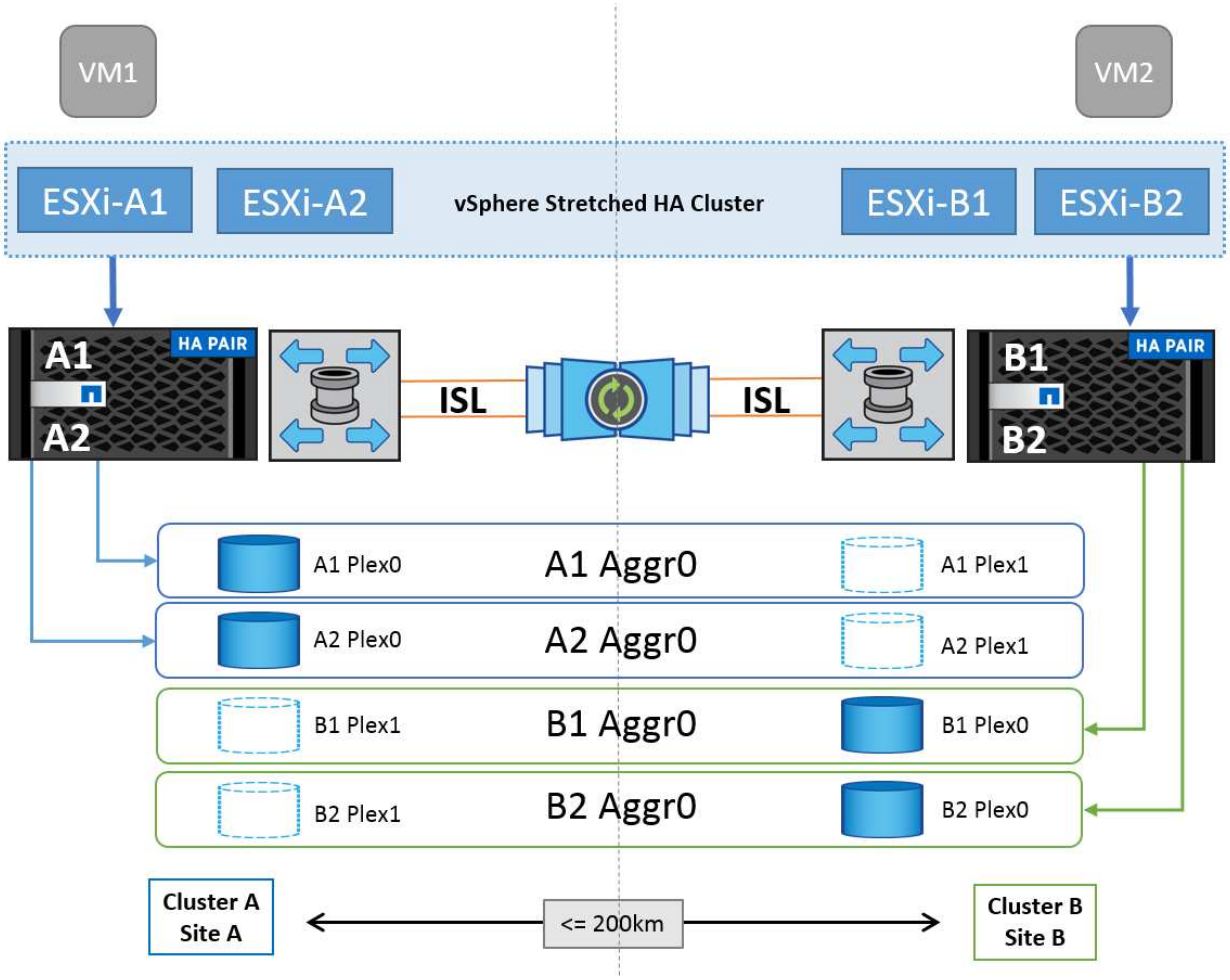
Weitere Informationen zu den Designrichtlinien für vSphere Metro Storage Cluster finden Sie in der folgenden Dokumentation:

- "Unterstützung von VMware vSphere für NetApp MetroCluster"
- "Unterstützung von VMware vSphere mit NetApp SnapMirror Business Continuity" (Jetzt bekannt als SnapMirror Active Sync)

Abhängig von den Überlegungen zur Latenz kann NetApp MetroCluster in zwei unterschiedlichen Konfigurationen für die Verwendung mit vSphere implementiert werden:

- Stretch-MetroCluster
- Fabric MetroCluster

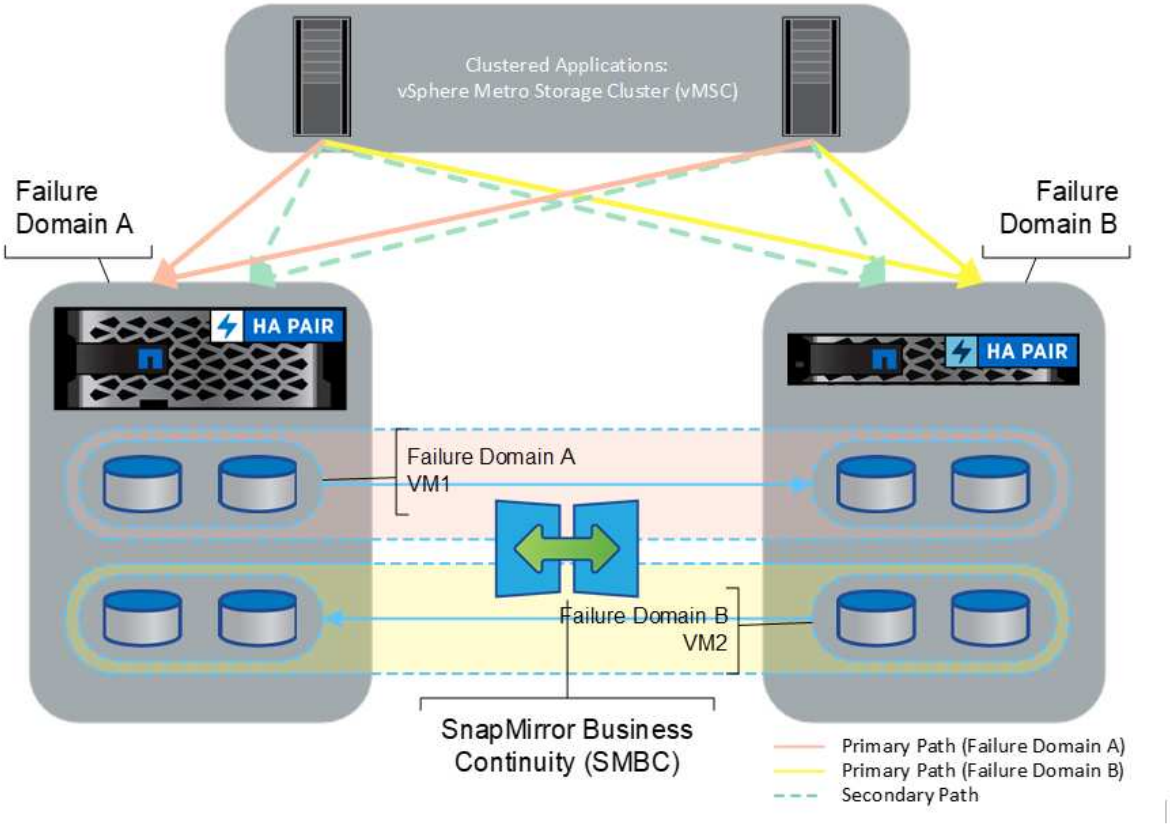
Im Folgenden finden Sie ein High-Level-Topologiediagramm von Stretch MetroCluster.



Siehe "MetroCluster-Dokumentation" Finden Sie spezifische Design- und Implementierungsinformationen für MetroCluster.

SnapMirror Active Sync kann darüber hinaus auf zwei verschiedene Arten implementiert werden.

- Asymmetrisch
- Symmetrisch (private Vorschau in ONTAP 9.14.1)



Siehe "NetApp Dokumente" Für spezifische Design- und Implementierungsinformationen zu SnapMirror Active Sync

VMware vSphere Lösungsübersicht

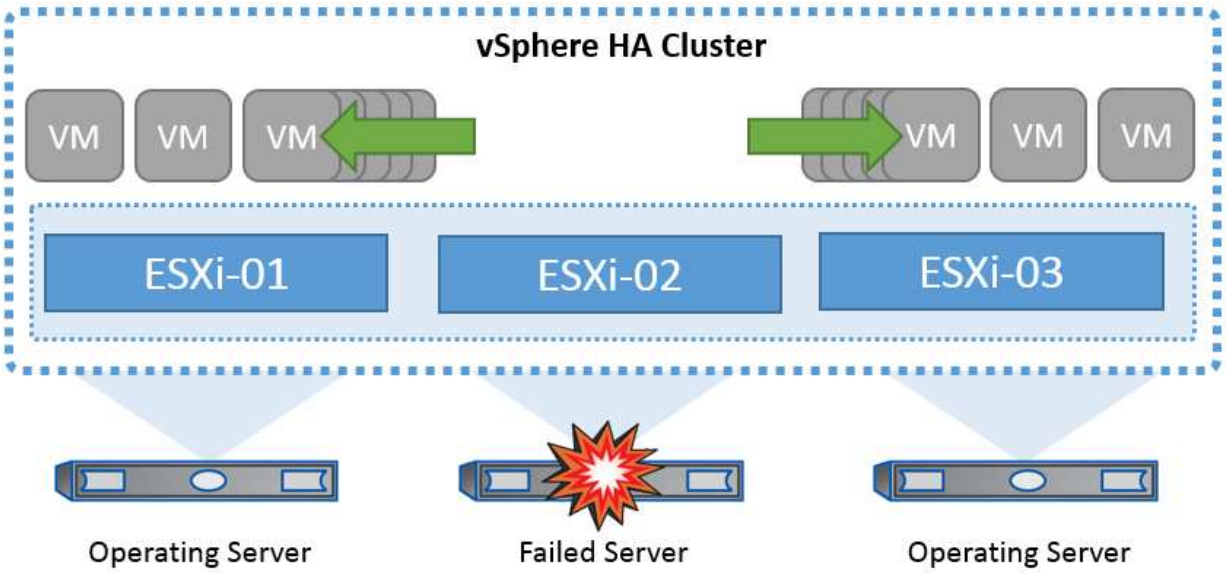
VCenter Server Appliance (VCSA) ist das leistungsstarke, zentralisierte Managementsystem und eine zentrale Konsole für vSphere, mit der Administratoren ESXi Cluster effizient betreiben können. Sie unterstützt wichtige Funktionen wie VM-Bereitstellung, vMotion Betrieb, Hochverfügbarkeit (HA), Distributed Resource Scheduler (DRS), Tanzu Kubernetes Grid und mehr. Sie ist eine wesentliche Komponente in VMware Cloud-Umgebungen und sollte im Hinblick auf die Service-Verfügbarkeit konzipiert werden.

VSphere High Availability

Die Cluster-Technologie von VMware gruppiert ESXi Server in Pools mit gemeinsam genutzten Ressourcen für Virtual Machines und stellt vSphere High Availability (HA) bereit. VSphere HA bietet benutzerfreundliche, hohe Verfügbarkeit für Anwendungen, die auf virtuellen Maschinen ausgeführt werden. Wenn die HA-Funktion auf dem Cluster aktiviert ist, hält jeder ESXi-Server die Kommunikation mit anderen Hosts aufrecht, sodass ein ESXi-Host nicht mehr reagiert oder isoliert wird. der HA-Cluster kann die Wiederherstellung der Virtual

Machines, die auf diesem ESXi-Host ausgeführt wurden, zwischen den noch intakten Hosts im Cluster aushandeln. Bei einem Ausfall eines Gastbetriebssystems startet vSphere HA die betroffene virtuelle Maschine auf demselben physischen Server neu. Mit vSphere HA werden geplante Ausfallzeiten reduziert, ungeplante Ausfallzeiten vermieden und eine schnelle Wiederherstellung nach Ausfällen ermöglicht.

VSphere HA Cluster Wiederherstellung von VMs aus ausgefallener Server.



Es ist wichtig zu wissen, dass VMware vSphere nicht über NetApp MetroCluster oder SnapMirror Active Sync verfügt und alle ESXi Hosts im vSphere Cluster als berechtigte Hosts für HA-Clustervorgänge erkennt, je nach Host- und VM-Gruppenaffinitätskonfigurationen.

Erkennung Von Host-Ausfällen

Sobald der HA-Cluster erstellt ist, nehmen alle Hosts im Cluster an der Auswahl Teil, und einer der Hosts wird zum Master. Jeder Slave führt den Netzwerk-Heartbeat zum Master aus, und der Master wiederum führt den Netzwerk-Heartbeat auf allen Slave-Hosts aus. Der Master-Host eines vSphere HA-Clusters ist für die Erkennung des Ausfalls von Slave-Hosts verantwortlich.

Je nach Art des erkannten Fehlers müssen die auf den Hosts ausgeführten virtuellen Maschinen möglicherweise ein Failover durchführen.

In einem vSphere HA-Cluster werden drei Arten von Host-Ausfällen erkannt:

- Fehler: Ein Host funktioniert nicht mehr.
- Isolierung: Ein Host wird zu einem isolierten Netzwerk.
- Partition: Ein Host verliert die Netzwerkverbindung mit dem Master-Host.

Der Master-Host überwacht die Slave-Hosts im Cluster. Diese Kommunikation erfolgt durch den Austausch von Netzwerk-Heartbeats jede Sekunde. Wenn der Master-Host diese Heartbeats nicht mehr von einem Slave-Host empfängt, prüft er die Host-Lebendigkeit, bevor er den Host für fehlgeschlagen erklärt. Die Liveness-Prüfung, die der Master-Host durchführt, besteht darin festzustellen, ob der Slave-Host Heartbeats mit einem der Datastores austauscht. Außerdem prüft der Master-Host, ob der Host auf ICMP-Pings reagiert, die an seine Management-IP-Adressen gesendet werden, um festzustellen, ob er lediglich von seinem Master-Knoten isoliert oder vollständig vom Netzwerk isoliert ist. Dies erfolgt durch Ping an das Standard-Gateway. Eine oder mehrere Isolationsadressen können manuell angegeben werden, um die Zuverlässigkeit der

Isolationsvalidierung zu erhöhen.

Best Practice

NetApp empfiehlt, mindestens zwei zusätzliche Isolationsadressen anzugeben, und jede dieser Adressen sollte standortlokal sein. Dies erhöht die Zuverlässigkeit der Isolationsvalidierung.

Antwort Der Hostisolation

Die Isolationsantwort ist eine Einstellung in vSphere HA, die die Aktion bestimmt, die auf virtuellen Maschinen ausgelöst wird, wenn ein Host in einem vSphere HA-Cluster seine Verwaltungsnetzwerkverbindungen verliert, aber weiterhin ausgeführt wird. Für diese Einstellung gibt es drei Optionen: „Disabled“, „Shut Down and Restart VMs“ und „Power Off and Restart VMs“.

„Herunterfahren“ ist besser als „Ausschalten“, das nicht die neuesten Änderungen auf Festplatte bereinigt oder Transaktionen festschreibt. Wenn virtuelle Maschinen nicht in 300 Sekunden heruntergefahren wurden, werden sie ausgeschaltet. Um die Wartezeit zu ändern, verwenden Sie die erweiterte Option das.isolationshutdowntimeout.

Bevor HA die Isolationsantwort initiiert, prüft es zunächst, ob der vSphere HA-Master-Agent den Datenspeicher besitzt, der die VM-Konfigurationsdateien enthält. Wenn dies nicht der Fall ist, löst der Host die Isolationsantwort nicht aus, da kein Master zum Neustart der VMs vorhanden ist. Der Host überprüft regelmäßig den Datastore-Status, um festzustellen, ob er von einem vSphere HA-Agent beansprucht wird, der die Master-Rolle besitzt.

Best Practice

NetApp empfiehlt, die „Host-Isolationsantwort“ auf deaktiviert zu setzen.

Ein Split-Brain-Zustand kann auftreten, wenn ein Host vom vSphere HA-Master-Host isoliert oder partitioniert wird und der Master nicht über Heartbeat Datastores oder Ping kommunizieren kann. Der Master erklärt den isolierten Host für tot und startet die VMs auf anderen Hosts im Cluster neu. Eine Split-Brain-Bedingung besteht jetzt, weil zwei Instanzen der virtuellen Maschine ausgeführt werden, von denen nur eine die virtuellen Laufwerke lesen oder schreiben kann. Split-Brain-Bedingungen können jetzt durch die Konfiguration von VM Component Protection (VMCP) vermieden werden.

Schutz von VM-Komponenten (VMCP)

Eine der Funktionsverbesserungen bei vSphere 6, relevant für HA, ist VMCP. VMCP bietet erweiterten Schutz vor All Paths Down (APD) und Permanent Device Loss (PDL) für Block (FC, iSCSI, FCoE) und File Storage (NFS).

Permanenter Geräteverlust (PDL)

PDL ist ein Zustand, der auftritt, wenn ein Speichergerät dauerhaft ausfällt oder administrativ entfernt wird und nicht zurückgegeben werden soll. Das NetApp-Speicher-Array gibt ESXi einen SCSI-Sense-Code aus, der erklärt, dass das Gerät dauerhaft verloren geht. Im Abschnitt Fehlerbedingungen und VM-Reaktion von vSphere HA können Sie konfigurieren, wie die Antwort nach dem Erkennen einer PDL-Bedingung aussehen soll.

Best Practice

NetApp empfiehlt, die „Antwort für Datastore mit PDL“ auf „**Ausschalten und Neustart von VMs**“ zu setzen. Wenn dieser Zustand erkannt wird, wird eine VM sofort auf einem funktionierenden Host im vSphere HA-

Cluster neu gestartet.

Alle Pfade nach unten (APD)

APD ist ein Zustand, der auftritt, wenn ein Speichergerät für den Host nicht mehr zugänglich ist und keine Pfade zum Array verfügbar sind. ESXi betrachtet dies als ein vorübergehendes Problem mit dem Gerät und erwartet, dass es wieder verfügbar wird.

Wenn eine APD-Bedingung erkannt wird, wird ein Timer gestartet. Nach 140 Sekunden wird der APD-Zustand offiziell deklariert und das Gerät als APD-Zeitabmeldung markiert. Nach Ablauf der 140 Sekunden zählt HA die Anzahl der Minuten, die in der Verzögerung für VM-Failover-APD angegeben sind. Wenn die angegebene Zeit verstrichen ist, startet HA die betroffenen virtuellen Maschinen neu. Sie können VMCP so konfigurieren, dass es bei Bedarf anders reagiert (deaktiviert, Ereignisse ausstellen oder VMs aus- und neu starten).

Best Practice

NetApp empfiehlt, die „Antwort für Datastore mit APD“ auf „**Ausschalten und Neustart von VMs (konservativ)**“ zu konfigurieren.

Konservativ bezieht sich auf die Wahrscheinlichkeit, dass HA die VMs neu starten kann. Wenn sie auf Conservative gesetzt ist, startet HA nur die VM neu, die vom APD betroffen ist, wenn sie weiß, dass ein anderer Host sie neu starten kann. Im Fall von aggressive, versucht HA, die VM neu zu starten, selbst wenn sie den Status anderer Hosts nicht kennt. Dies kann dazu führen, dass VMs nicht neu gestartet werden, wenn kein Host mit Zugriff auf den Datenspeicher vorhanden ist, auf dem sich dieser befindet.

Wenn der APD-Status aufgelöst ist und der Zugriff auf den Speicher wiederhergestellt wird, bevor die Zeiteinstellung überschritten wurde, startet HA die virtuelle Maschine nicht unnötig neu, es sei denn, Sie konfigurieren sie ausdrücklich dafür. Wenn eine Antwort gewünscht wird, selbst wenn sich die Umgebung von der APD-Bedingung erholt hat, sollte die Antwort für APD-Wiederherstellung nach APD-Timeout so konfiguriert werden, dass die VMs zurückgesetzt werden.

Best Practice

NetApp empfiehlt, die Antwort für die APD-Wiederherstellung nach der APD-Zeitüberschreitung auf deaktiviert zu konfigurieren.

VMware DRS Implementierung für NetApp MetroCluster

VMware DRS ist eine Funktion, die die Host-Ressourcen in einem Cluster aggregiert und hauptsächlich zum Lastausgleich innerhalb eines Clusters in einer virtuellen Infrastruktur verwendet wird. VMware DRS berechnet in erster Linie die CPU- und Arbeitsspeicherressourcen für den Lastausgleich in einem Cluster. Da vSphere das erweiterte Clustering nicht kennt, werden beim Lastausgleich alle Hosts an beiden Standorten berücksichtigt. Um standortübergreifenden Datenverkehr zu vermeiden, empfiehlt NetApp die Konfiguration der DRS Affinitätsregeln, um eine logische Trennung der VMs zu managen. So stellen Sie sicher, dass HA und DRS nur lokale Hosts verwenden, sofern es keinen vollständigen Standortausfall gibt.

Wenn Sie eine DRS-Affinitätsregel für Ihr Cluster erstellen, können Sie festlegen, wie vSphere diese Regel während eines Failover einer virtuellen Maschine anwendet.

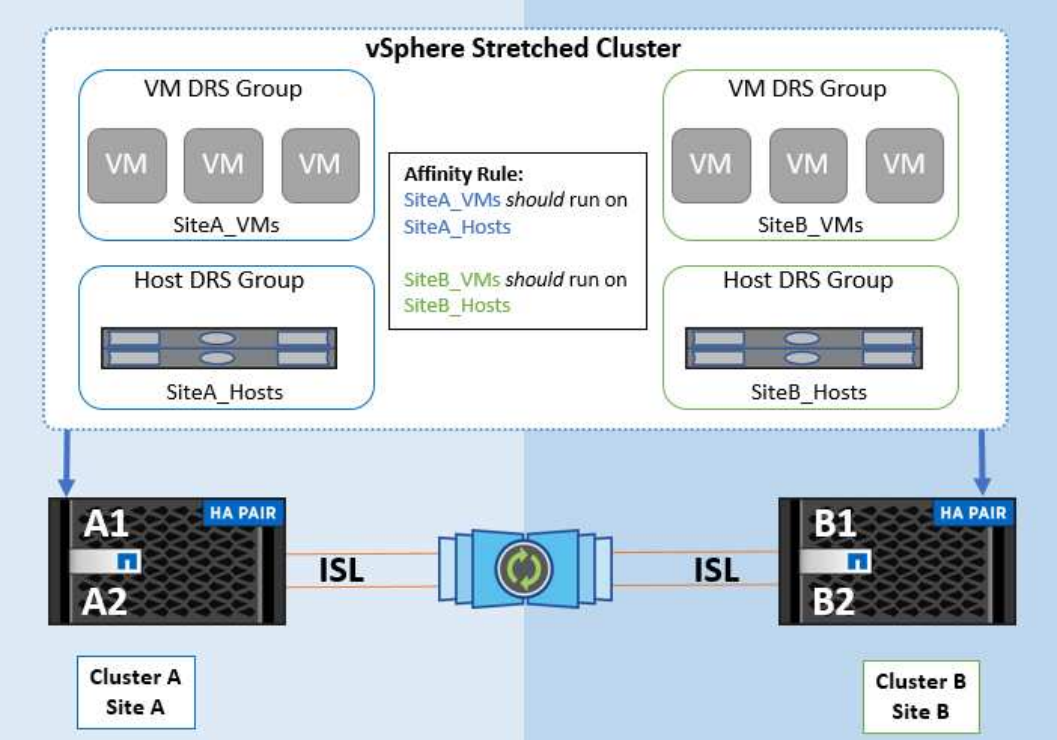
Es gibt zwei Arten von Regeln, die Sie vSphere HA-Failover-Verhalten angeben können:

- VM-Anti-Affinitätsregeln zwingen bestimmte Virtual Machines dazu, bei Failover-Aktionen getrennt zu bleiben.
- VM-Host-Affinitätsregeln platzieren angegebene Virtual Machines während Failover-Aktionen auf einem

bestimmten Host oder einem Mitglied einer definierten Gruppe von Hosts.

Mithilfe der VM Host-Affinitätsregeln in VMware DRS lässt sich eine logische Trennung zwischen Standort A und Standort B erreichen, sodass die VM auf dem Host am selben Standort ausgeführt wird wie das Array, das als primärer Lese-/Schreib-Controller für einen bestimmten Datenspeicher konfiguriert ist. Zudem bleiben Virtual Machines gemäß den Regeln zur VM Host-Affinität lokal im Storage, wodurch wiederum die Virtual Machine-Verbindung im Falle von Netzwerkausfällen zwischen den Standorten hergestellt wird.

Nachfolgend finden Sie ein Beispiel für VM-Hostgruppen und Affinitätsregeln.



Best Practice

NetApp empfiehlt die Implementierung der „sollte“-Regeln statt der „müssen“-Regeln, da im Falle eines Ausfalls von vSphere HA gegen diese verstoßen wird. Der Einsatz von „Must“-Regeln kann potenziell zu Serviceausfällen führen.

Die Verfügbarkeit von Services sollte immer Vorrang vor der Leistung haben. Wenn ein vollständiges Datacenter ausfällt, müssen die „Must“-Regeln Hosts aus der VM-Host-Affinitätsgruppe auswählen. Wenn das Datacenter nicht verfügbar ist, werden die Virtual Machines nicht neu gestartet.

VMware Storage DRS Implementierung mit NetApp MetroCluster

Die VMware Storage DRS-Funktion ermöglicht die Aggregation von Datastores in eine einzige Einheit und gleicht Festplatten der virtuellen Maschine aus, wenn die Storage-I/O-Kontrollschwellenwerte überschritten werden.

Die Storage-I/O-Steuerung ist bei DRS-Clustern mit Storage DRS standardmäßig aktiviert. Mit der Storage-I/O-Kontrolle kann ein Administrator die Menge an Storage-I/O steuern, die Virtual Machines bei I/O-Engpässen zugewiesen wird. Dadurch können wichtigeren Virtual Machines bei der I/O-Ressourcenzuweisung Vorrang vor weniger wichtigen Virtual Machines geben.

Storage DRS verwendet Storage vMotion, um die virtuellen Maschinen auf verschiedene Datastores innerhalb

eines Datastore-Clusters zu migrieren. In einer NetApp MetroCluster Umgebung muss eine Migration von Virtual Machines innerhalb der Datenspeicher dieses Standorts gesteuert werden. Eine Virtual Machine A, die auf einem Host an Standort A ausgeführt wird, sollte idealerweise innerhalb der Datenspeicher der SVM an Standort A migriert werden Wenn dies nicht der Fall ist, wird die virtuelle Maschine weiterhin betrieben, jedoch mit verminderter Leistung, da das Lesen/Schreiben der virtuellen Festplatte von Standort B über standortübergreifende Links erfolgt.

Best Practice

NetApp empfiehlt das Erstellen von Datastore-Clustern im Hinblick auf die Storage-Standortorientierung. Das heißt, Datastores mit Standortaffinität zu Standort A sollten nicht mit Datastore-Clustern mit Datastores mit Standortaffinität zu Standort B gemischt werden

Wenn eine virtuelle Maschine neu bereitgestellt oder mithilfe von Storage vMotion migriert wird, empfiehlt NetApp, alle für diese virtuellen Maschinen spezifischen VMware DRS-Regeln entsprechend manuell zu aktualisieren. Dadurch wird die Virtual Machine-Affinität auf Standortebene sowohl für Host als auch für Datenspeicher ermittelt und somit der Netzwerk- und Storage Overhead reduziert.

VMSC Design- und Implementierungsrichtlinien

Dieses Dokument enthält Design- und Implementierungsrichtlinien für vMSC mit ONTAP Storage-Systemen.

NetApp-Speicherkonfiguration

Setup-Anweisungen für NetApp MetroCluster (als MCC-Konfiguration bezeichnet) finden Sie unter "MetroCluster-Dokumentation". Anweisungen für SnapMirror Active Sync finden Sie auch unter "Überblick über die Business Continuity in SnapMirror".

Sobald Sie MetroCluster konfiguriert haben, ist die Verwaltung wie das Management einer herkömmlichen ONTAP-Umgebung. Sie können Storage Virtual Machines (SVMs) mithilfe verschiedener Tools wie Command Line Interface (CLI), System Manager oder Ansible einrichten. Sobald die SVMs konfiguriert sind, erstellen Sie logische Schnittstellen (LIFs), Volumes und LUNs (Logical Unit Numbers) auf dem Cluster, die für den normalen Betrieb verwendet werden. Diese Objekte werden automatisch über das Cluster-Peering-Netzwerk auf den anderen Cluster repliziert.

Wenn Sie MetroCluster nicht nutzen, können Sie SnapMirror Active Sync verwenden, was Datastores-granulare Sicherung und aktiv/aktiv-Zugriff über diverse ONTAP Cluster in verschiedenen Ausfall-Domains hinweg bietet. SnapMirror Active Sync verwendet Konsistenzgruppen, um die Konsistenz der Schreibreihenfolge zwischen einem oder mehreren Datastores sicherzustellen. Sie können je nach Applikations- und Datastore-Anforderungen mehrere Konsistenzgruppen erstellen. Konsistenzgruppen sind insbesondere für Applikationen nützlich, die eine Datensynchronisierung zwischen mehreren Datastores erfordern. SnapMirror Active Sync unterstützt außerdem Raw Device Mapping (RDMs) und über das Gastsystem verbundenen Storage mit iSCSI-Initiatoren in-Guest. Weitere Informationen zu Konsistenzgruppen finden Sie unter "Übersicht über Konsistenzgruppen".

Es gibt einen Unterschied beim Management einer vMSC Konfiguration mit aktiver SnapMirror Synchronisierung im Vergleich zu einer MetroCluster. Zunächst handelt es sich um eine reine SAN-Konfiguration. Es können keine NFS-Datenspeicher durch SnapMirror Active Sync geschützt werden. Als zweites müssen Sie Ihren ESXi-Hosts beide Kopien der LUNs zuordnen, damit sie auf die replizierten Datastores in beiden Ausfall-Domains zugreifen können.

VMware vSphere HA

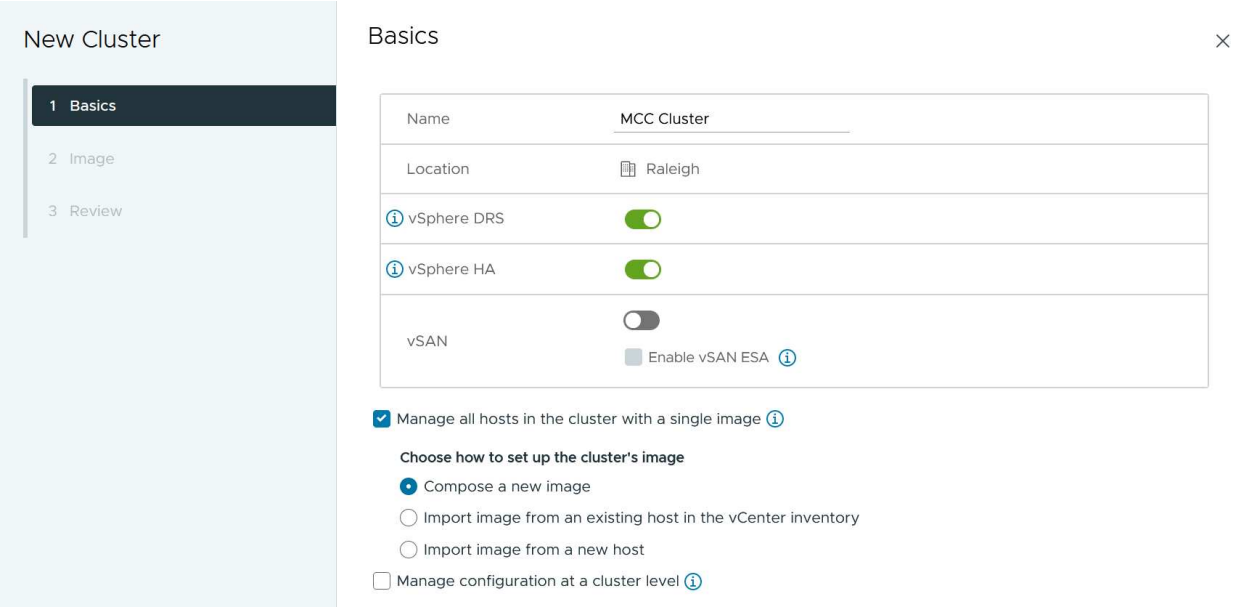
Erstellen Sie einen vSphere HA-Cluster

Die Erstellung eines vSphere HA-Clusters ist ein mehrstufiger Prozess, der in vollständig dokumentiert ist "Wie erstellen und konfigurieren Sie Cluster im vSphere Client auf docs.vmware.com". Kurz gesagt: Sie müssen zuerst einen leeren Cluster erstellen, dann mit vCenter Hosts hinzufügen und vSphere HA und andere Einstellungen des Clusters angeben.

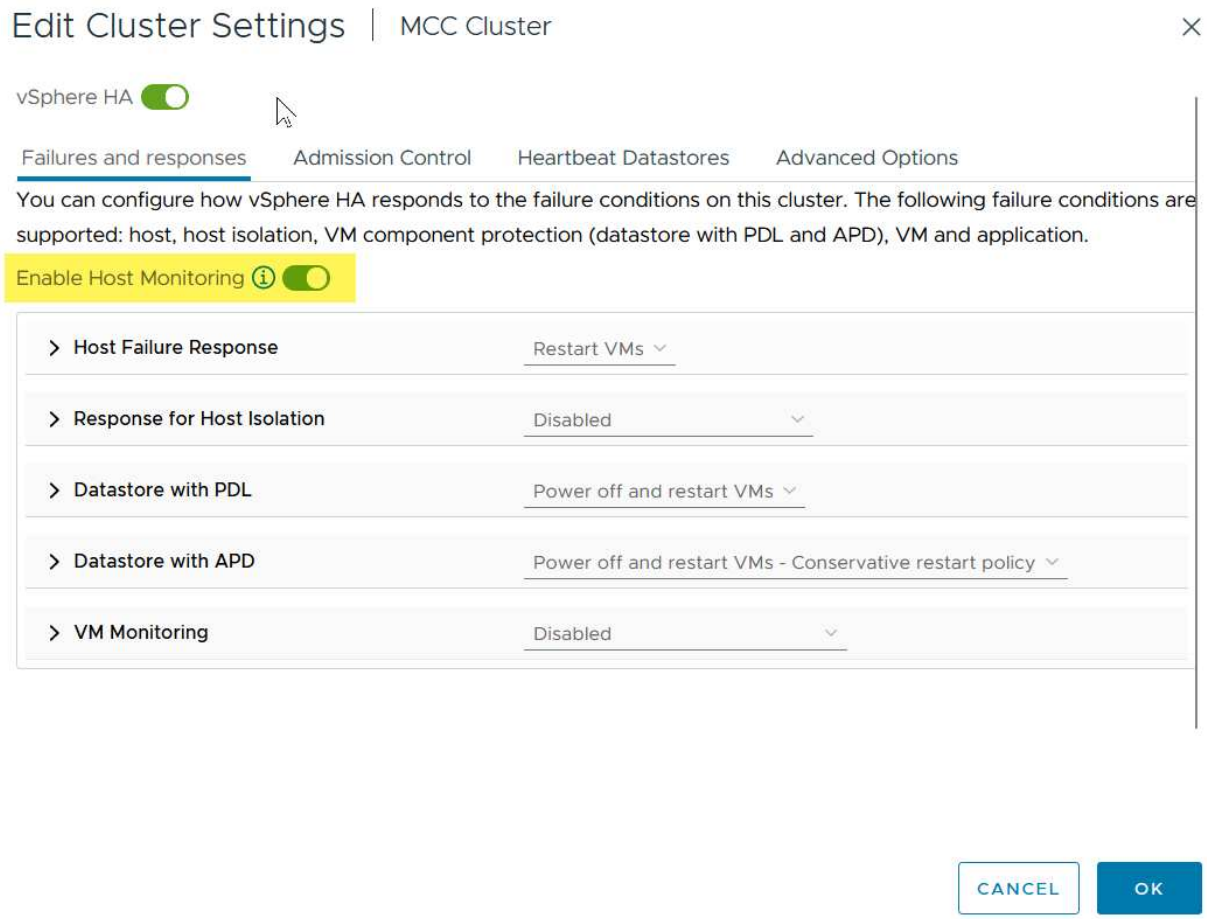
Hinweis: nichts in diesem Dokument ersetzt "Empfohlene Practices für VMware vSphere Metro Storage-Cluster"

Führen Sie zum Konfigurieren eines HA-Clusters die folgenden Schritte aus:

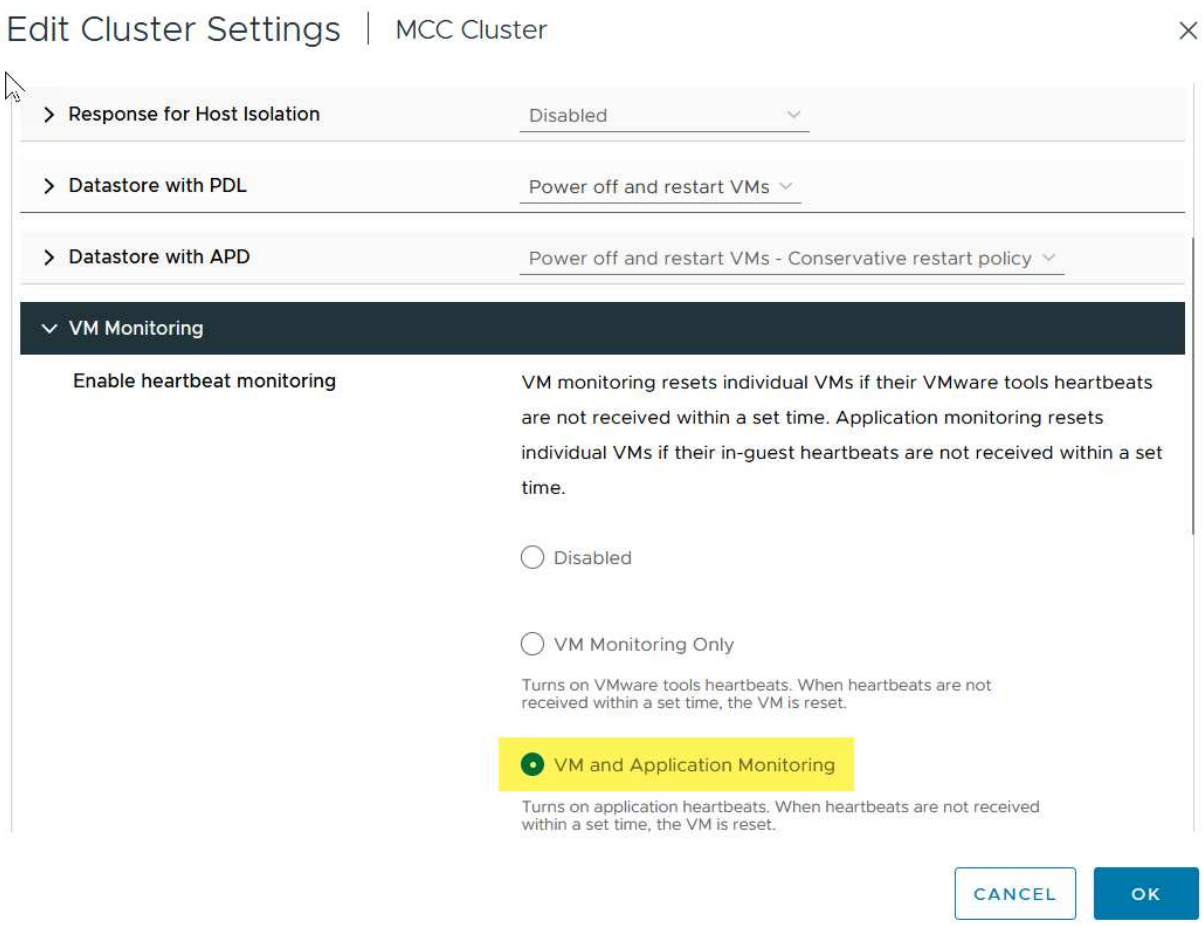
1. Stellen Sie eine Verbindung zur vCenter-Benutzeroberfläche her.
2. Navigieren Sie unter Hosts und Cluster zum Rechenzentrum, in dem Sie Ihr HA-Cluster erstellen möchten.
3. Klicken Sie mit der rechten Maustaste auf das Datacenter-Objekt, und wählen Sie Neuer Cluster aus. Unter Grundlagen stellen Sie sicher, dass Sie vSphere DRS und vSphere HA aktiviert haben. Schließen Sie den Assistenten ab.



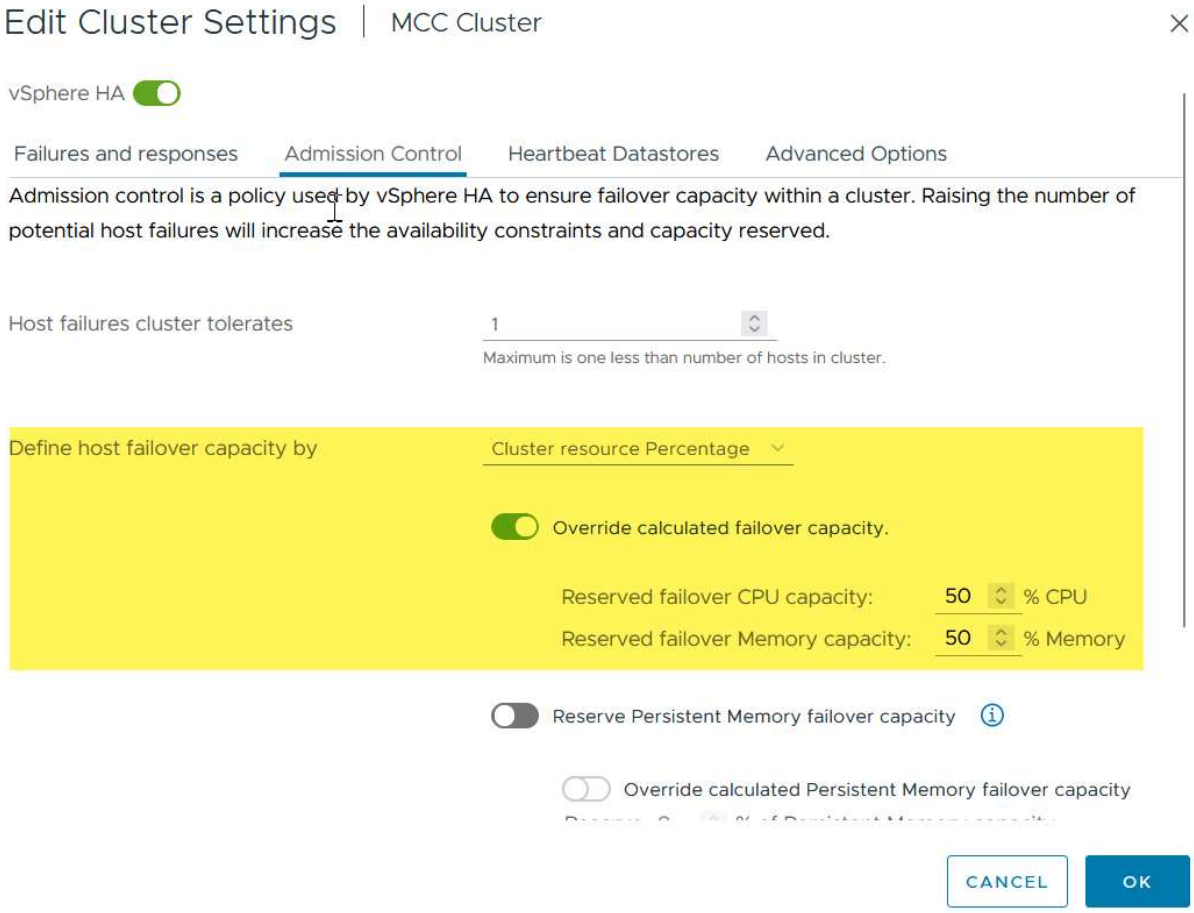
1. Wählen Sie den Cluster aus, und wechseln Sie zur Registerkarte Konfigurieren. Wählen Sie vSphere HA aus, und klicken Sie auf Bearbeiten.
2. Wählen Sie unter Host-Überwachung die Option Host-Überwachung aktivieren aus.



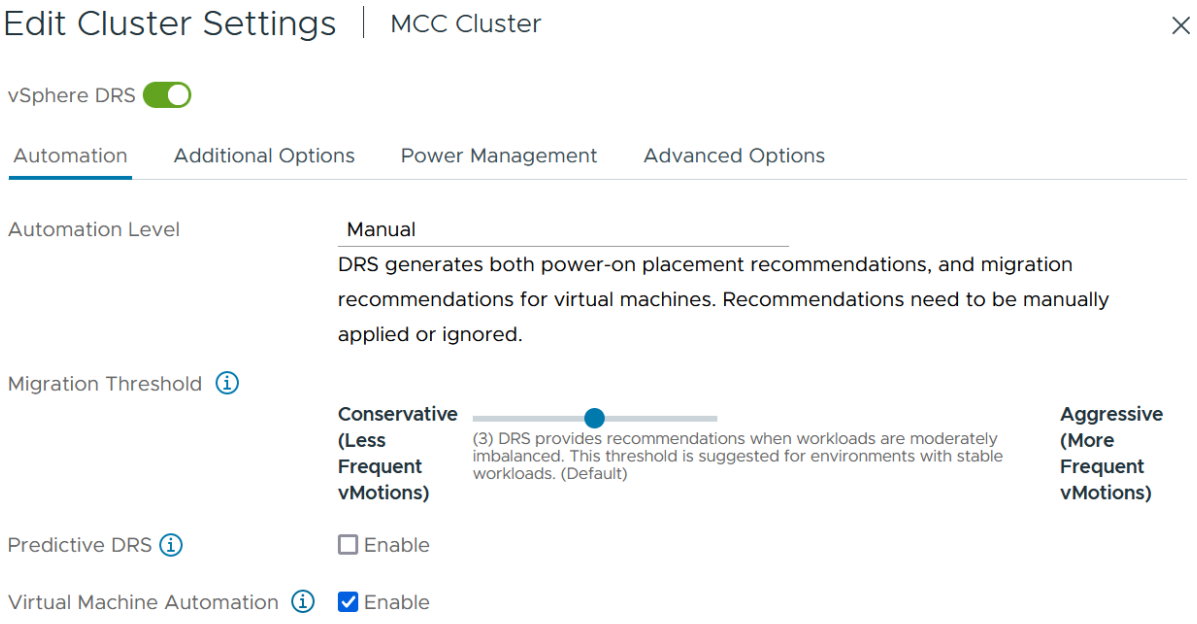
1. Wählen Sie auf der Registerkarte „Fehler und Antworten“ unter „VM-Überwachung“ die Option „nur VM-Überwachung“ oder „VM- und Anwendungsüberwachung“ aus.



1. Legen Sie unter Admission Control die Option HA-Eintrittskontrolle auf Cluster-Ressourcenreserve fest. Verwenden Sie 50 % CPU/MEM.



1. Klicken Sie auf „OK“.
2. Wählen Sie DRS und klicken Sie auf BEARBEITEN.
3. Setzen Sie den Automatisierungsgrad auf manuell, sofern dies nicht von Ihren Anwendungen erforderlich ist.



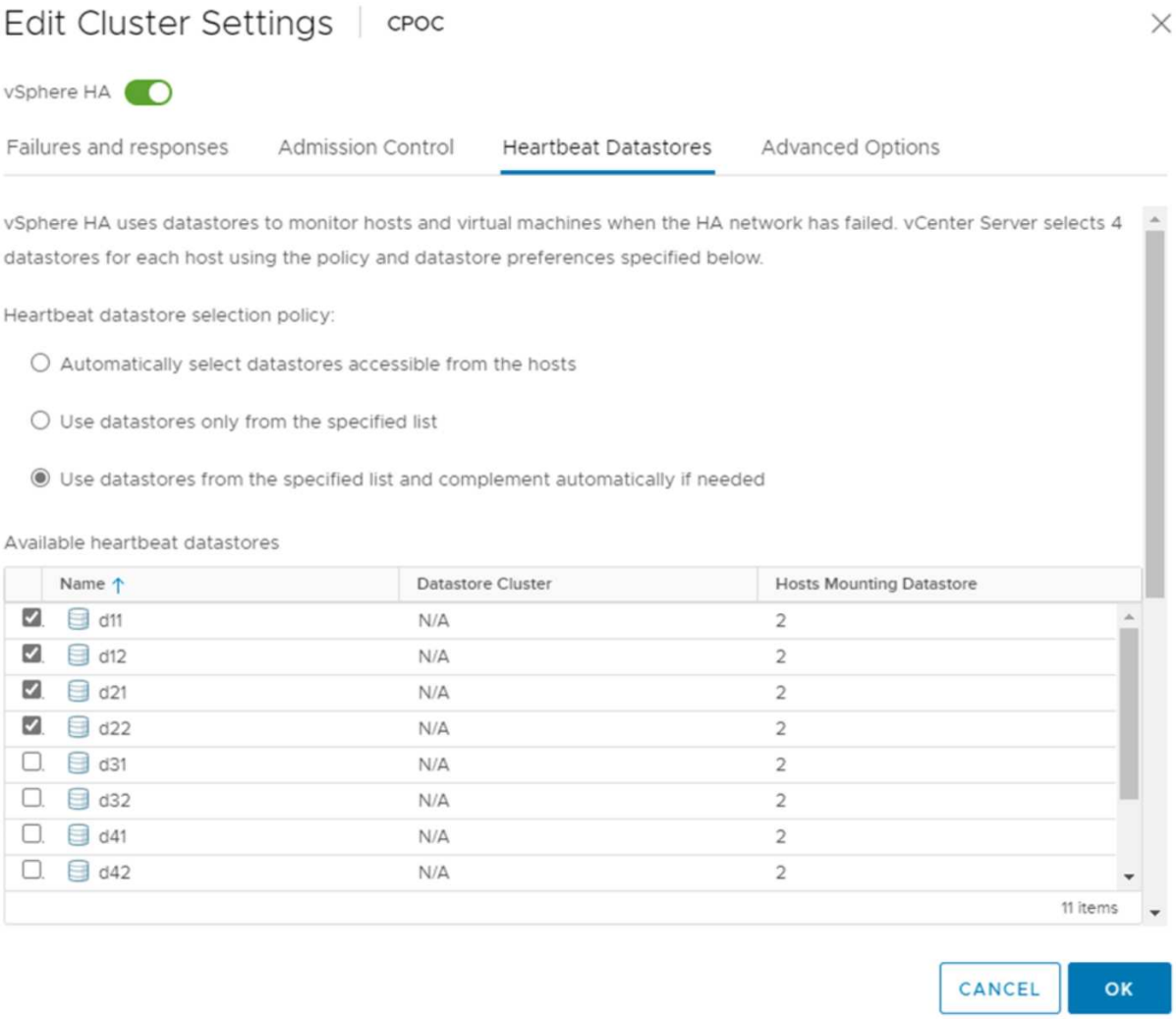
1. Aktivieren Sie den Schutz von VM-Komponenten, siehe "docs.vmware.com".
2. Die folgenden zusätzlichen vSphere HA-Einstellungen werden für vMSC mit MCC empfohlen:

Ausfall	Antwort
Host-Ausfall	Starten Sie die VMs neu
Host-Isolierung	Deaktiviert
Datenspeicher mit Permanent Device Loss (PDL)	Schalten Sie die VMs aus und starten Sie sie neu
Datastore mit All Paths Down (APD)	Schalten Sie die VMs aus und starten Sie sie neu
Der Gast ist nicht herzschlagend	Setzt die VMs zurück
Richtlinie für den Neustart der VM	Bestimmt durch die Bedeutung der VM
Antwort für Host-Isolation	Fahren Sie die VMs herunter, und starten Sie sie neu
Antwort für Datastore mit PDL	Schalten Sie die VMs aus und starten Sie sie neu
Antwort für Datenspeicher mit APD	VMs ausschalten und neu starten (konservativ)
Verzögerung bei VM-Failover für APD	3 Minuten
Antwort für APD-Wiederherstellung mit APD-Timeout	Deaktiviert
Sensitivität für VM-Monitoring	Voreinstellung hoch

Konfigurieren Sie Datastores für Heartbeating

VSphere HA verwendet Datastores, um Hosts und virtuelle Maschinen zu überwachen, wenn das Managementnetzwerk ausgefallen ist. Sie können konfigurieren, wie vCenter Heartbeat-Datenspeicher auswählt. Gehen Sie wie folgt vor, um Datastores für Heartbeating zu konfigurieren:

1. Wählen Sie im Abschnitt Datastore Heartbeating die Option Datastores aus der angegebenen Liste verwenden aus und ergänzen Sie bei Bedarf automatisch.
2. Wählen Sie die Datastores aus, die vCenter von beiden Standorten verwenden soll, und drücken Sie OK.

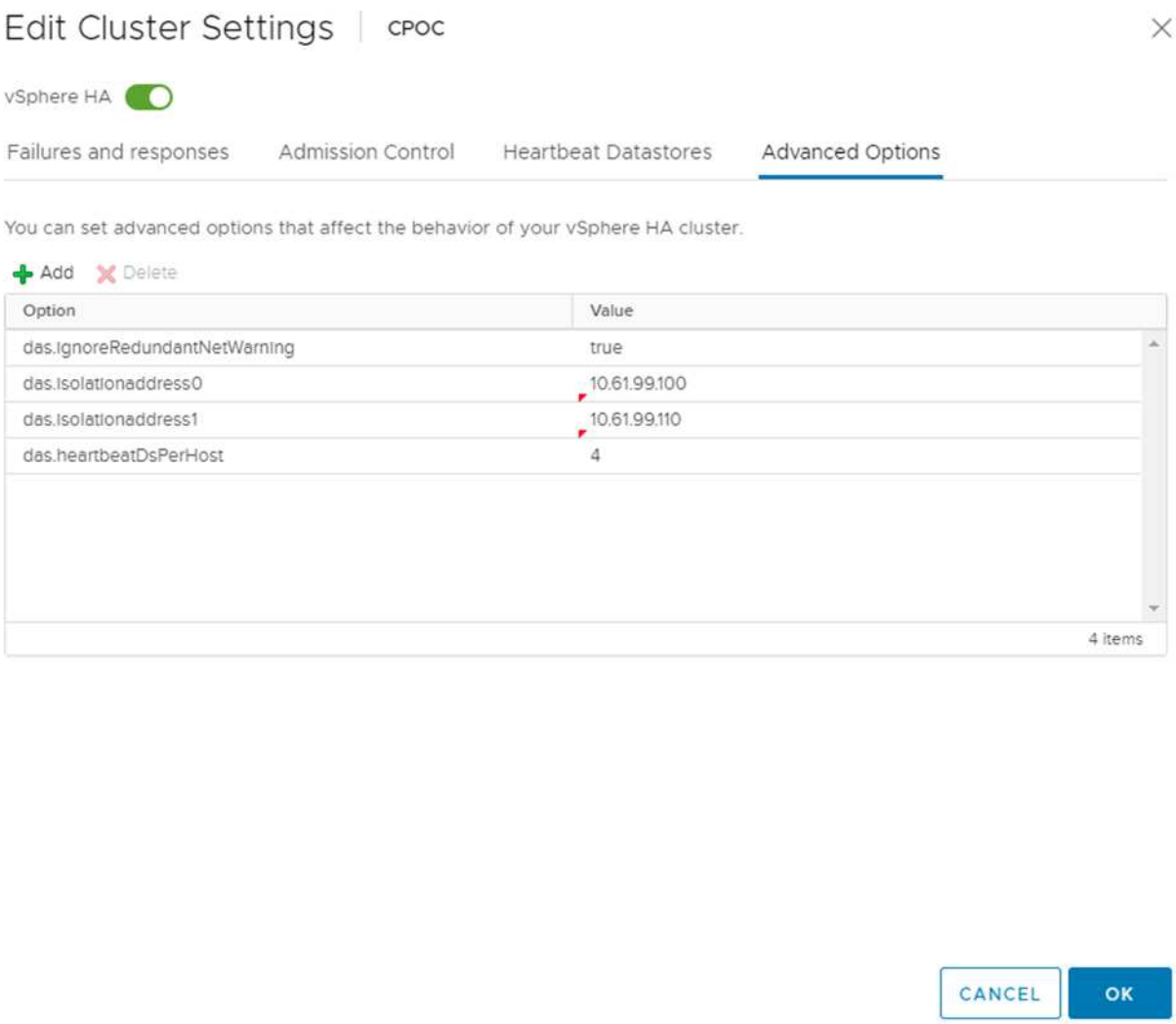


Konfigurieren Sie Die Erweiterten Optionen

Host-Fehlererkennung

Isolierungsereignisse treten auf, wenn Hosts innerhalb eines HA-Clusters die Verbindung zum Netzwerk oder zu anderen Hosts im Cluster verlieren. Standardmäßig verwendet vSphere HA das Standard-Gateway für sein Managementnetzwerk als Standard-Isolationsadresse. Sie können jedoch zusätzliche Isolationsadressen für den Host angeben, um zu bestimmen, ob eine Isolationsantwort ausgelöst werden soll. Fügen Sie zwei isolierte IPs hinzu, die Ping-Daten senden können, eine pro Standort. Verwenden Sie nicht die Gateway-IP. Die erweiterte vSphere HA-Einstellung ist das.isolationaddress. Dazu können Sie ONTAP- oder Mediator-IP-Adressen verwenden.

Siehe "core.vmware.com" Weitere Informationen.



Das Hinzufügen einer erweiterten Einstellung namens das.heartbeatDsPerHost kann die Anzahl der Heartbeat-Datenspeicher erhöhen. Verwenden Sie vier Heartbeat Datastores (HB DSS) – zwei pro Standort. Verwenden Sie die Option „aus Liste auswählen, aber Kompliment“. Dies wird benötigt, da Sie bei Ausfall eines Standorts immer noch zwei HB DSS benötigen. Diese müssen jedoch nicht durch MCC oder SnapMirror Active Sync geschützt werden.

Siehe "core.vmware.com" Weitere Informationen.

VMware DRS Affinity zu NetApp MetroCluster

In diesem Abschnitt erstellen wir DRS Gruppen für VMs und Hosts für jeden Standort\Cluster in der MetroCluster Umgebung. Anschließend konfigurieren wir VM\Host-Regeln, um die VM Host-Affinität mit lokalen Storage-Ressourcen auszurichten. Beispielsweise gehören Standort A VMs zur VM-Gruppe sitea_vms und Standort A Hosts zur Host-Gruppe sitea_Hosts. Als nächstes geben wir in VM\Host Rules an, dass sitea_vms auf Hosts in sitea_Hosts ausgeführt werden sollen.

Best Practice

- NetApp empfiehlt dringend die Spezifikation **sollte auf Hosts in Gruppe** laufen anstatt der Spezifikation **muss auf Hosts in Gruppe** ausgeführt werden. Im Falle eines Host-Ausfalls von Standort A müssen die VMs von Standort A über vSphere HA auf Hosts an Standort B neu gestartet werden. Bei der letzteren Spezifikation ist jedoch nicht möglich, dass HA die VMs auf Standort B neu starten, da es die harte Regel

ist. Die frühere Spezifikation ist eine weiche Regel und wird im Falle von HA verletzt, wodurch die Verfügbarkeit anstatt die Leistung ermöglicht wird.

Hinweis: Sie können einen ereignisbasierten Alarm erstellen, der ausgelöst wird, wenn eine virtuelle Maschine gegen eine VM-Host-Affinitätsregel verstößt. Fügen Sie im vSphere Client einen neuen Alarm für die virtuelle Maschine hinzu und wählen Sie als Ereignisauslöser „VM verletzt VM-Host Affinity Rule“ aus. Weitere Informationen zum Erstellen und Bearbeiten von Alarmen finden Sie unter "VSphere Monitoring und Performance" Dokumentation.

DRS-Host-Gruppen erstellen

So erstellen Sie DRS Host-Gruppen speziell für Standort A und Standort B:

1. Klicken Sie im vSphere-Webclient mit der rechten Maustaste auf den Cluster in der Bestandsaufnahme, und wählen Sie Einstellungen aus.
2. Klicken Sie auf VM\Host Groups.
3. Klicken Sie Auf Hinzufügen.
4. Geben Sie den Namen der Gruppe ein (z. B. sitea_Hosts).
5. Wählen Sie im Menü Typ die Option Host-Gruppe aus.
6. Klicken Sie auf Hinzufügen, wählen Sie die gewünschten Hosts von Standort A aus, und klicken Sie auf OK.
7. Wiederholen Sie diese Schritte, um eine weitere Host-Gruppe für Standort B hinzuzufügen
8. Klicken Sie auf OK.

DRS VM-Gruppen erstellen

So erstellen Sie DRS VM-Gruppen speziell für Standort A und Standort B:

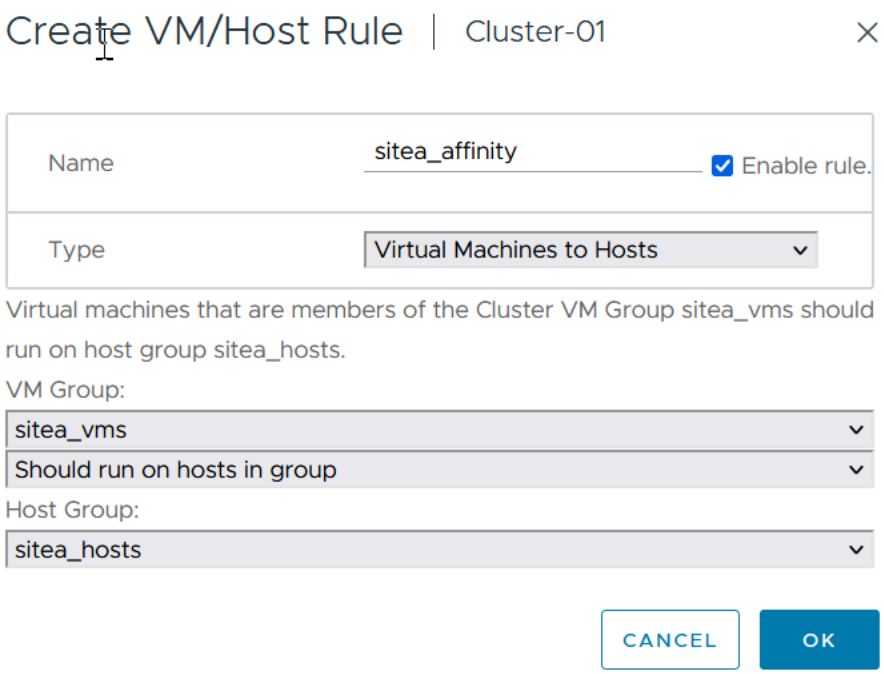
1. Klicken Sie im vSphere-Webclient mit der rechten Maustaste auf den Cluster in der Bestandsaufnahme, und wählen Sie Einstellungen aus.
2. Klicken Sie auf VM\Host Groups.
3. Klicken Sie Auf Hinzufügen.
4. Geben Sie den Namen der Gruppe ein (z. B. sitea_vms).
5. Wählen Sie im Menü Typ die Option VM-Gruppe aus.
6. Klicken Sie auf Hinzufügen, wählen Sie die gewünschten VMs von Standort A aus, und klicken Sie auf OK.
7. Wiederholen Sie diese Schritte, um eine weitere Host-Gruppe für Standort B hinzuzufügen
8. Klicken Sie auf OK.

Erstellen Sie VM-Hostregeln

Gehen Sie wie folgt vor, um DRS-Affinitätsregeln für Standort A und Standort B zu erstellen:

1. Klicken Sie im vSphere-Webclient mit der rechten Maustaste auf den Cluster in der Bestandsaufnahme, und wählen Sie Einstellungen aus.
2. Klicken Sie auf VM\Hostregeln.
3. Klicken Sie Auf Hinzufügen.
4. Geben Sie den Namen der Regel ein (z. B. sitea_Affinity).

5. Überprüfen Sie, ob die Option Regel aktivieren aktiviert ist.
6. Wählen Sie im Menü Typ die Option Virtuelle Maschinen zu Hosts aus.
7. Wählen Sie die VM-Gruppe aus (z.B. sitea_vms).
8. Wählen Sie die Host-Gruppe aus (z. B. sitea_Hosts).
9. Wiederholen Sie diese Schritte, um eine weitere VM\Host-Regel für Standort B hinzuzufügen
10. Klicken Sie auf OK.



VMware vSphere Storage DRS für NetApp MetroCluster

Datastore-Cluster Erstellen

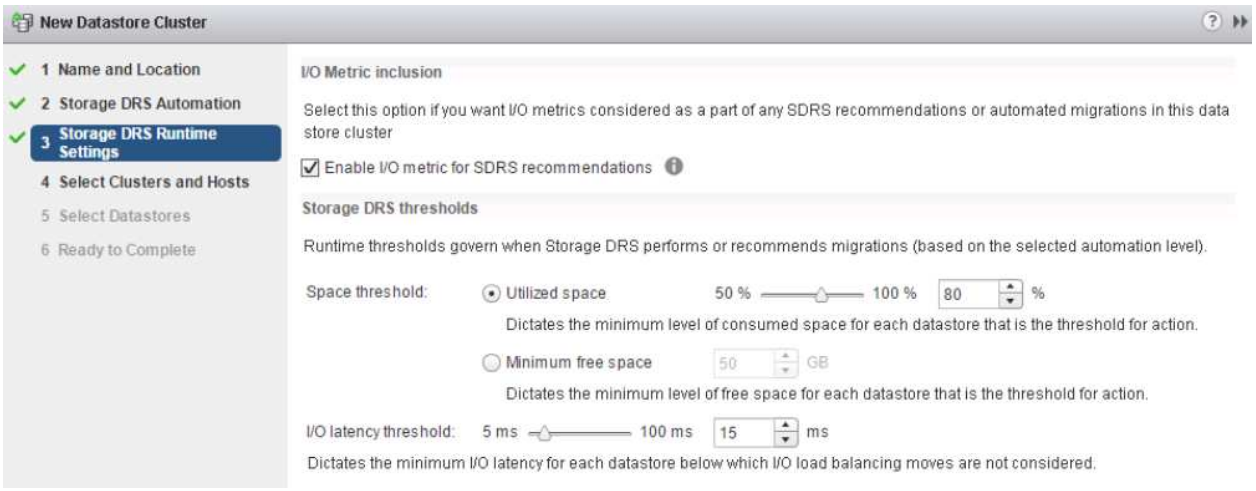
Führen Sie die folgenden Schritte aus, um ein Datastore-Cluster für jeden Standort zu konfigurieren:

1. Navigieren Sie mithilfe des vSphere-Webclients zum Rechenzentrum, in dem sich der HA-Cluster unter Speicher befindet.
2. Klicken Sie mit der rechten Maustaste auf das Datacenter-Objekt, und wählen Sie Storage > New Datastore Cluster aus.
3. Wählen Sie die Option Storage DRS aktivieren aus, und klicken Sie auf Weiter.
4. Stellen Sie alle Optionen auf Keine Automatisierung (manueller Modus) ein, und klicken Sie auf Weiter.

Best Practice

- NetApp empfiehlt, Storage DRS im manuellen Modus zu konfigurieren, sodass der Administrator entscheiden und kontrollieren kann, wann Migrationen stattfinden.

1. Vergewissern Sie sich, dass das Kontrollkästchen E/A-Metrik für SDRS-Empfehlungen aktivieren aktiviert ist. Die metrischen Einstellungen können mit Standardwerten belassen werden.



1. Wählen Sie das HA-Cluster aus, und klicken Sie auf Weiter.

1. Wählen Sie die Datastores aus, die zu Standort A gehören, und klicken Sie auf Weiter.

1. Überprüfen Sie die Optionen, und klicken Sie auf Fertig stellen.
2. Wiederholen Sie diese Schritte, um das Datastore-Cluster an Standort B zu erstellen und sicherzustellen, dass nur Datastores von Standort B ausgewählt sind.

VCenter Server-Verfügbarkeit

Ihre vCenter Server Appliances (VCSAs) sollten durch vCenter HA geschützt werden. Mit vCenter HA können

Sie zwei VCSAs in einem aktiv/Passiv-HA-Paar implementieren. Einer in jeder Ausfall-Domäne. Weitere Informationen zu vCenter HA finden Sie im "docs.vmware.com".

Ausfallsicherheit bei geplanten und ungeplanten Ereignissen

NetApp MetroCluster und SnapMirror Active Sync sind leistungsstarke Tools, die die Hochverfügbarkeit und den unterbrechungsfreien Betrieb von NetApp Hardware und ONTAP Software verbessern.

Diese Tools bieten standortweiten Schutz für die gesamte Storage-Umgebung und stellen sicher, dass Ihre Daten immer verfügbar sind. Ob Sie Standalone-Server, hochverfügbare Server-Cluster, Docker Container oder virtualisierte Server verwenden: Die NetApp-Technologie sorgt nahtlos für die Storage-Verfügbarkeit im Falle eines totalen Ausfalls aufgrund von Strom-, Kühlungs- oder Netzwerkkonnektivität, Herunterfahren des Storage-Arrays oder Bedienungsfehlern.

MetroCluster und SnapMirror Active Sync bieten drei grundlegende Methoden für die Datenverfügbarkeit bei geplanten und ungeplanten Ereignissen:

- Redundante Komponenten zum Schutz vor dem Ausfall einer einzelnen Komponente
- Lokaler HA-Takeover für Ereignisse, die sich auf einen einzelnen Controller auswirken
- Vollständiger Standortschutz: Schnelle Wiederaufnahme des Service durch Verschieben des Speicher- und Client-Zugriffs vom Quell-Cluster auf den Ziel-Cluster

Das bedeutet, dass die Abläufe bei Ausfall einer einzelnen Komponente reibungslos fortgesetzt werden und beim Austausch der ausgefallenen Komponente automatisch in den redundanten Betrieb zurückkehren.

Alle ONTAP Cluster außer Cluster mit einem Node (in der Regel softwaredefinierte Versionen, wie beispielsweise ONTAP Select) verfügen über integrierte HA-Funktionen für Takeover und Giveback. Jeder Controller im Cluster wird mit einem anderen Controller gepaart, wodurch ein HA-Paar entsteht. Diese Paare stellen sicher, dass jeder Node lokal mit dem Speicher verbunden ist.

Die Übernahme ist ein automatisierter Prozess, bei dem ein Node den Storage des anderen zur Aufrechterhaltung der Datenservices übernimmt. GiveBack bedeutet umgekehrter Prozess, der den normalen Betrieb wiederherstellt. Takeover können geplant werden, beispielsweise bei Hardware-Wartungsarbeiten oder ONTAP Upgrades oder aufgrund von Node-Panic- oder Hardware-Ausfällen.

Während einer Übernahme führen NAS-LIFs (Network Attached Storage Logical Interfaces) in MetroCluster-Konfigurationen automatisch ein Failover durch. Storage Area Network LIFs (SAN LIFs) führen jedoch keinen Failover durch. Sie verwenden weiterhin den direkten Pfad zu den Logical Unit Numbers (LUNs).

Weitere Informationen zu HA-Takeover und Giveback finden Sie im "HA-Paar-Management – Übersicht". Erwähnenswert ist, dass diese Funktion nicht spezifisch für MetroCluster oder SnapMirror für die aktive Synchronisierung ist.

Eine Standortumschaltung mit MetroCluster erfolgt, wenn ein Standort offline ist oder als geplante Aktivität für die standortweite Wartung vorgesehen ist. Der verbleibende Standort übernimmt die Eigentümerschaft der Storage-Ressourcen (Festplatten und Aggregate) des Offline-Clusters, und die SVMs am ausgefallenen Standort werden online geschaltet und am Disaster-Standort neugestartet. Dabei bleibt die volle Identität für den Client- und Host-Zugriff erhalten.

Da beide Kopien gleichzeitig aktiv verwendet werden, arbeiten Ihre vorhandenen Hosts mit der aktiven

SnapMirror Synchronisierung weiter. Der NetApp-Mediator ist erforderlich, um sicherzustellen, dass ein Standort-Failover korrekt ausgeführt wird.

Ausfallszenarien für vMSC mit MCC

In den folgenden Abschnitten werden die erwarteten Ergebnisse verschiedener Ausfallszenarien mit vMSC- und NetApp MetroCluster-Systemen beschrieben.

Ausfall Eines Einzelnen Storage-Pfads

Wenn in diesem Szenario Komponenten wie der HBA-Port, der Netzwerkport, der Front-End-Datenschalterport oder ein FC- oder Ethernet-Kabel ausfallen, wird dieser bestimmte Pfad zum Speichergerät vom ESXi-Host als „tot“ markiert. Wenn mehrere Pfade durch Ausfallsicherheit am HBA/Netzwerk/Switch Port für das Storage-Gerät konfiguriert sind, führt ESXi idealerweise eine Pfadumschaltung durch. Während dieser Zeit bleiben Virtual Machines ohne Beeinträchtigungen verfügbar, da für die Storage-Verfügbarkeit mehrere Pfade zum Storage-Gerät bereitgestellt werden.

Hinweis: Es gibt keine Änderung im MetroCluster-Verhalten in diesem Szenario, und alle Datastores bleiben weiterhin von ihren jeweiligen Seiten intakt.

Best Practice

In Umgebungen mit NFS/iSCSI-Volumes empfiehlt NetApp, mindestens zwei Netzwerk-Uplinks für den NFS-VMkernel-Port im Standard-vSwitch und dieselbe Port-Gruppe zu konfigurieren, bei der die NFS-VMkernel-Schnittstelle für den verteilten vSwitch zugeordnet ist. NIC-Teaming kann entweder aktiv-aktiv oder aktiv-Standby konfiguriert werden.

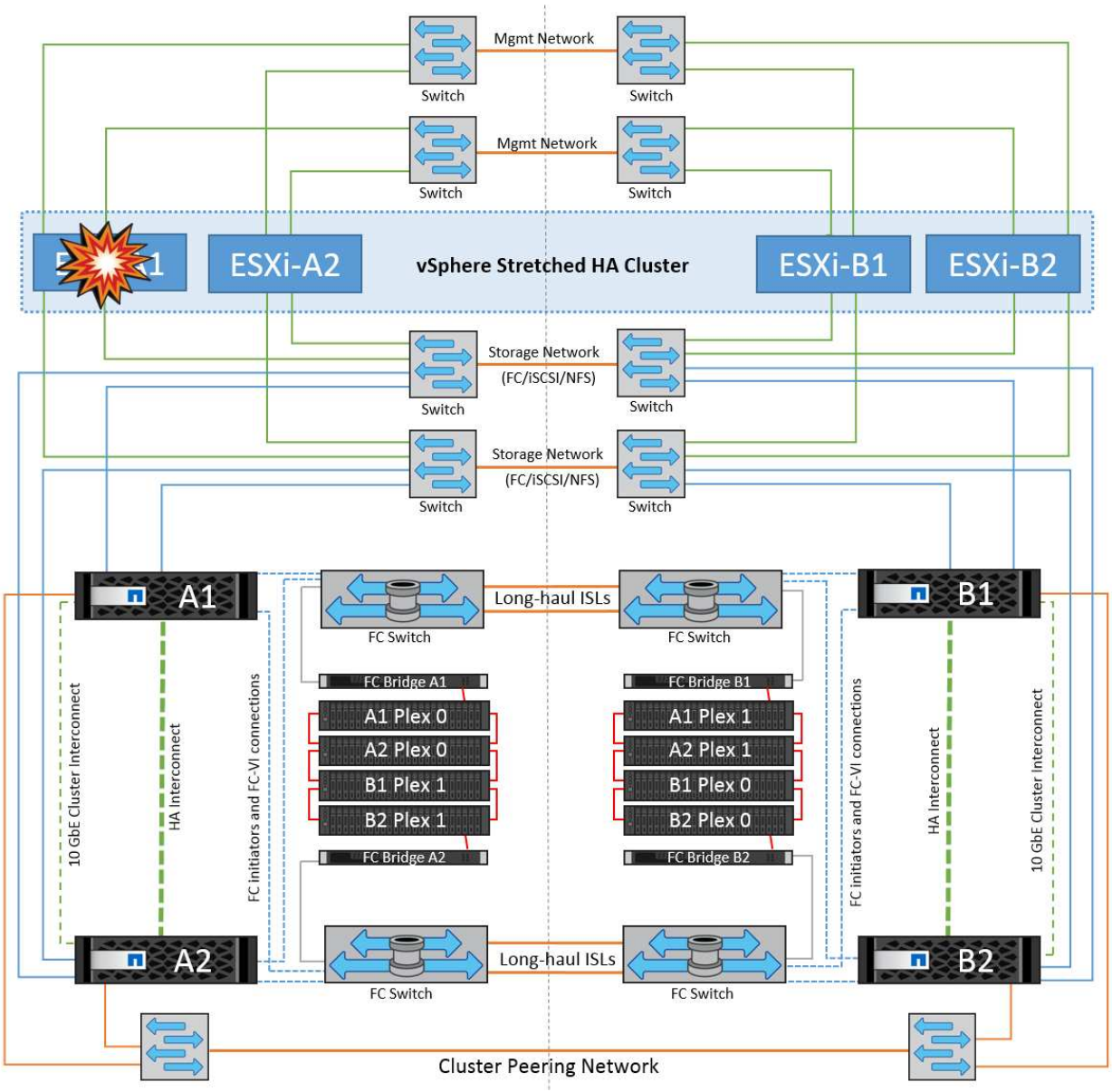
Außerdem muss bei iSCSI LUNs Multipathing konfiguriert werden, indem die VMkernel-Schnittstellen an die iSCSI-Netzwerkadapter gebunden werden. Weitere Informationen finden Sie in der vSphere-Speicherdokumentation.

Best Practice

In Umgebungen mit Fibre-Channel-LUNs empfiehlt NetApp die Verwendung von mindestens zwei HBAs, wodurch Ausfallsicherheit auf HBA-/Port-Ebene garantiert wird. NetApp empfiehlt für das Zoning von einem einzelnen Initiator außerdem als Best Practice zum Konfigurieren des Zoning.

Sie sollten mithilfe der Virtual Storage Console (VSC) Multipathing-Richtlinien festlegen, da sie Richtlinien für alle neuen und vorhandenen NetApp Storage-Geräte definiert.

Ausfall eines einzelnen ESXi-Hosts



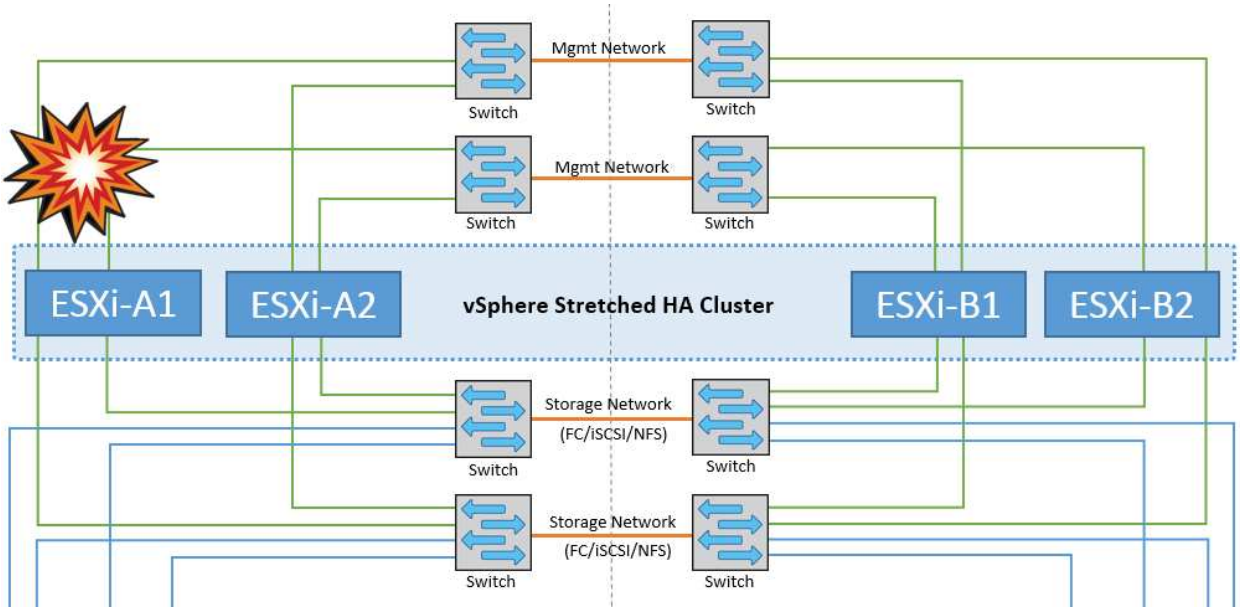
Wenn in diesem Szenario ein ESXi-Host ausfällt, erkennt der Master-Node im VMware HA-Cluster den Host-Ausfall, da er keine Netzwerk-Heartbeats mehr empfängt. Um festzustellen, ob der Host wirklich ausgefallen ist oder nur eine Netzwerkpartition ist, überwacht der Master-Knoten die Datastore-Heartbeats und führt, falls sie nicht vorhanden sind, eine abschließende Prüfung durch, indem er die Management-IP-Adressen des ausgefallenen Hosts anpingt. Wenn alle Prüfungen negativ sind, erklärt der Master-Node diesen Host als ausgefallenen Host, und alle virtuellen Maschinen, die auf diesem ausgefallenen Host ausgeführt wurden, werden auf dem noch verbleibenden Host im Cluster neu gestartet.

Wenn DRS VM und Host Affinity Regeln konfiguriert wurden (VMs in VM Gruppe sitea_vms sollten Hosts in Host Gruppe sitea_Hosts laufen lassen), dann prüft der HA Master zunächst auf verfügbare Ressourcen an Standort A. Wenn an Standort A keine verfügbaren Hosts vorhanden sind, versucht der Master, die VMs auf den Hosts an Standort B neu zu starten

Es ist möglich, dass die virtuellen Maschinen auf den ESXi-Hosts am anderen Standort gestartet werden, wenn am lokalen Standort eine Ressourcenbeschränkung vorhanden ist. Die definierten Regeln für die DRS-VM und Host-Affinität werden jedoch korrigiert, wenn Regeln verletzt werden, indem die virtuellen Maschinen zurück zu den noch verbleibenden ESXi-Hosts am lokalen Standort migriert werden. In Fällen, in denen DRS auf manuell festgelegt ist, empfiehlt NetApp, DRS zu aktivieren und die Empfehlungen anzuwenden, um die Platzierung der Virtual Machines zu korrigieren.

Es gibt keine Änderung im MetroCluster Verhalten in diesem Szenario und alle Datenspeicher bleiben von ihren jeweiligen Seiten intakt.

ESXi-Host-Isolierung

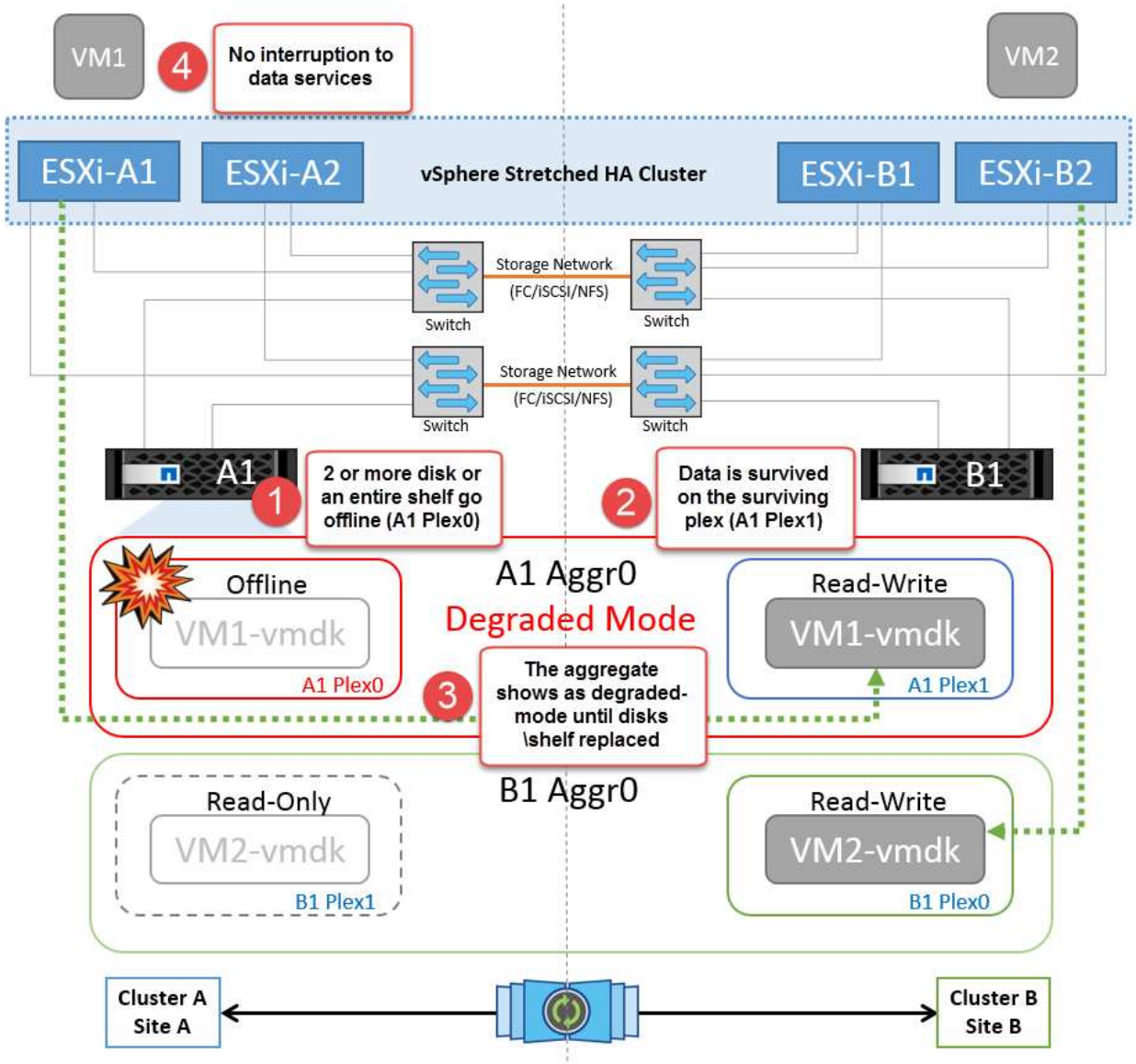


Wenn in diesem Szenario das Managementnetzwerk des ESXi-Hosts ausgefallen ist, erhält der Master-Node im HA-Cluster keine Heartbeats, wodurch dieser Host im Netzwerk isoliert wird. Um festzustellen, ob es ausgefallen ist oder nur isoliert ist, beginnt der Master-Node mit der Überwachung des Datastore-Herzschlags. Wenn er vorhanden ist, wird der Host vom Master-Knoten isoliert deklariert. Je nach konfigurierter Isolationsantwort kann der Host sich entscheiden, die virtuellen Maschinen auszuschalten, herunterzufahren oder die virtuellen Maschinen sogar eingeschaltet zu lassen. Das Standardintervall für die Isolationsantwort beträgt 30 Sekunden.

Es gibt keine Änderung im MetroCluster Verhalten in diesem Szenario und alle Datenspeicher bleiben von ihren jeweiligen Seiten intakt.

Platten-Shelf-Fehler

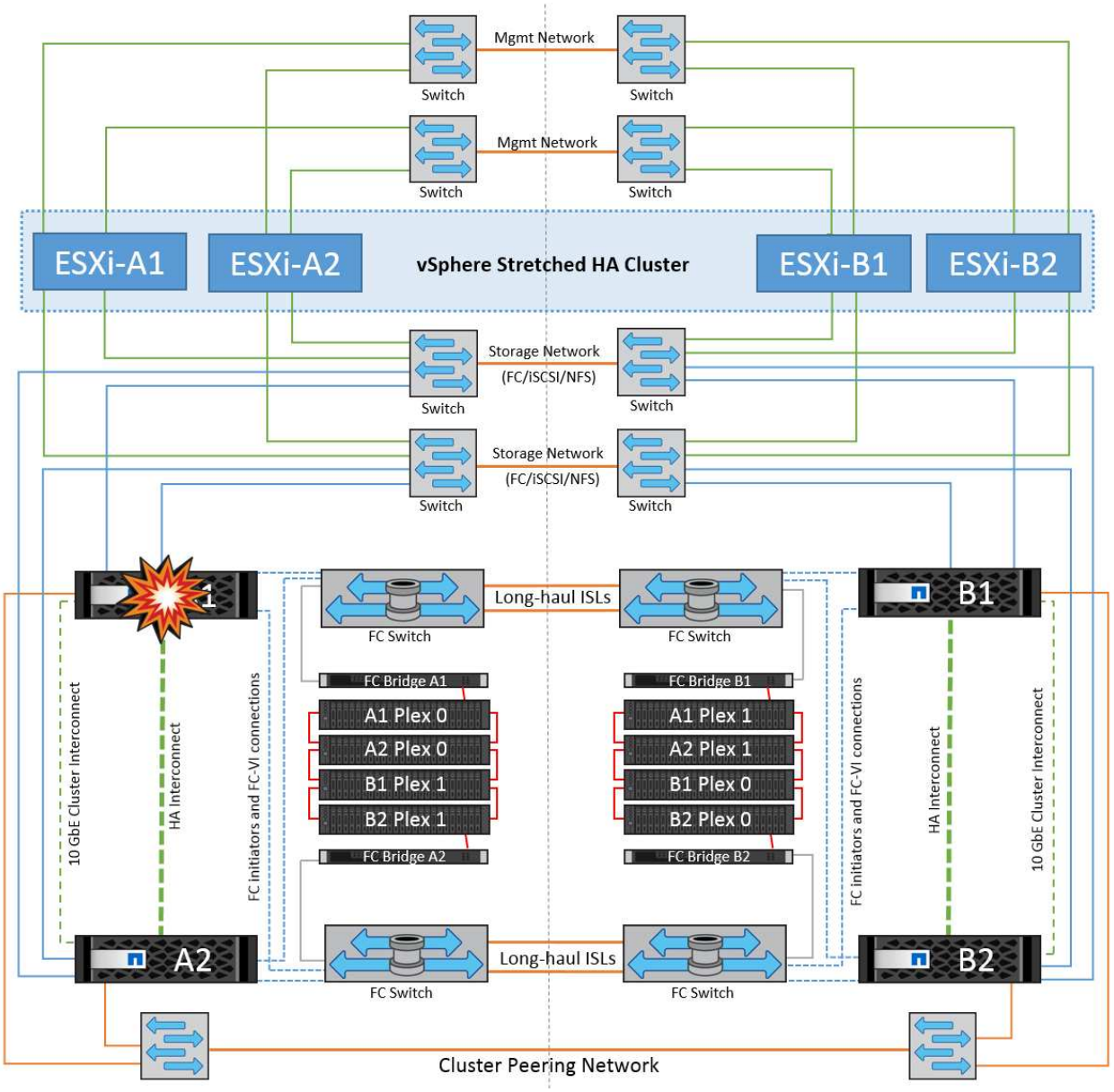
In diesem Szenario kommt es zu einem Ausfall von mehr als zwei Festplatten oder eines gesamten Shelf. Daten werden vom verbleibenden Plex ohne Unterbrechung der Datenservices bereitgestellt. Der Festplattenausfall kann sich auf einen lokalen oder einen Remote-Plex auswirken. Die Aggregate werden als degradierter Modus angezeigt, da nur ein Plex aktiv ist. Sobald die ausgefallenen Festplatten ersetzt wurden, werden die betroffenen Aggregate automatisch neu synchronisiert, um die Daten neu aufzubauen. Nach der Neusynchronisierung kehren die Aggregate automatisch in den normalen gespiegelten Modus zurück. Wenn mehr als zwei Laufwerke innerhalb einer einzelnen RAID-Gruppe ausgefallen sind, muss der Plex neu erstellt werden.



Hinweis: während dieses Zeitraums gibt es keine Auswirkungen auf die I/O-Vorgänge der virtuellen Maschine, aber die Performance ist beeinträchtigt, da die Daten vom Remote-Festplatten-Shelf über ISL-Links abgerufen werden.

Ausfall Eines Einzelnen Storage Controllers

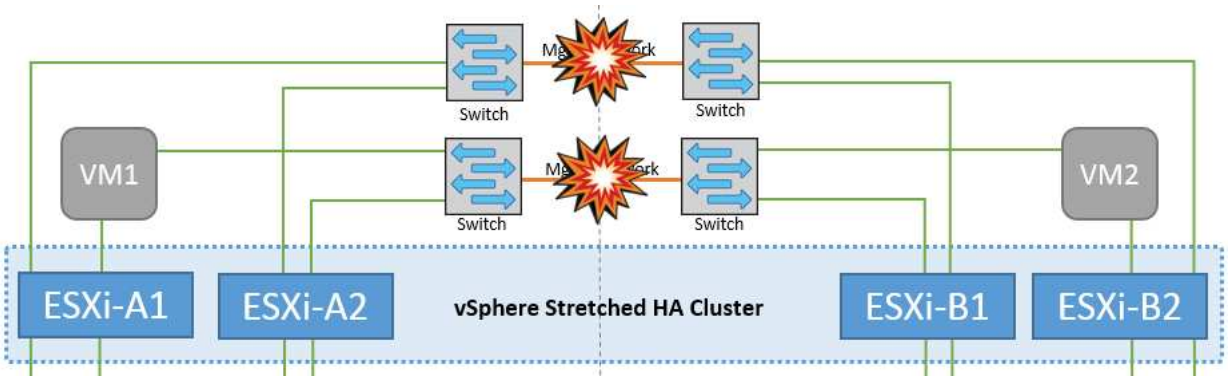
In diesem Szenario fällt einer der beiden Storage Controller an einem Standort aus. Da an jedem Standort ein HA-Paar vorhanden ist, wird bei einem Ausfall eines Node automatisch ein Failover auf den anderen Node ausgelöst. Wenn beispielsweise Node A1 ausfällt, werden dessen Storage und Workloads automatisch auf Node A2 übertragen. Virtuelle Maschinen sind nicht betroffen, da alle Plexe verfügbar bleiben. Die Knoten des zweiten Standorts (B1 und B2) sind davon nicht betroffen. Außerdem führt vSphere HA keine Aktion durch, da der Master-Node im Cluster weiterhin Netzwerk-Heartbeats empfängt.



Wenn der Failover Teil eines rollierenden Disaster ist (Node A1 führt ein Failover auf A2 durch) und ein nachfolgender Ausfall von A2 oder ein vollständiger Ausfall von Standort A auftritt, kann an Standort B das Umschalten nach einem Ausfall stattfinden

Verbindungsfehler Zwischen Switches

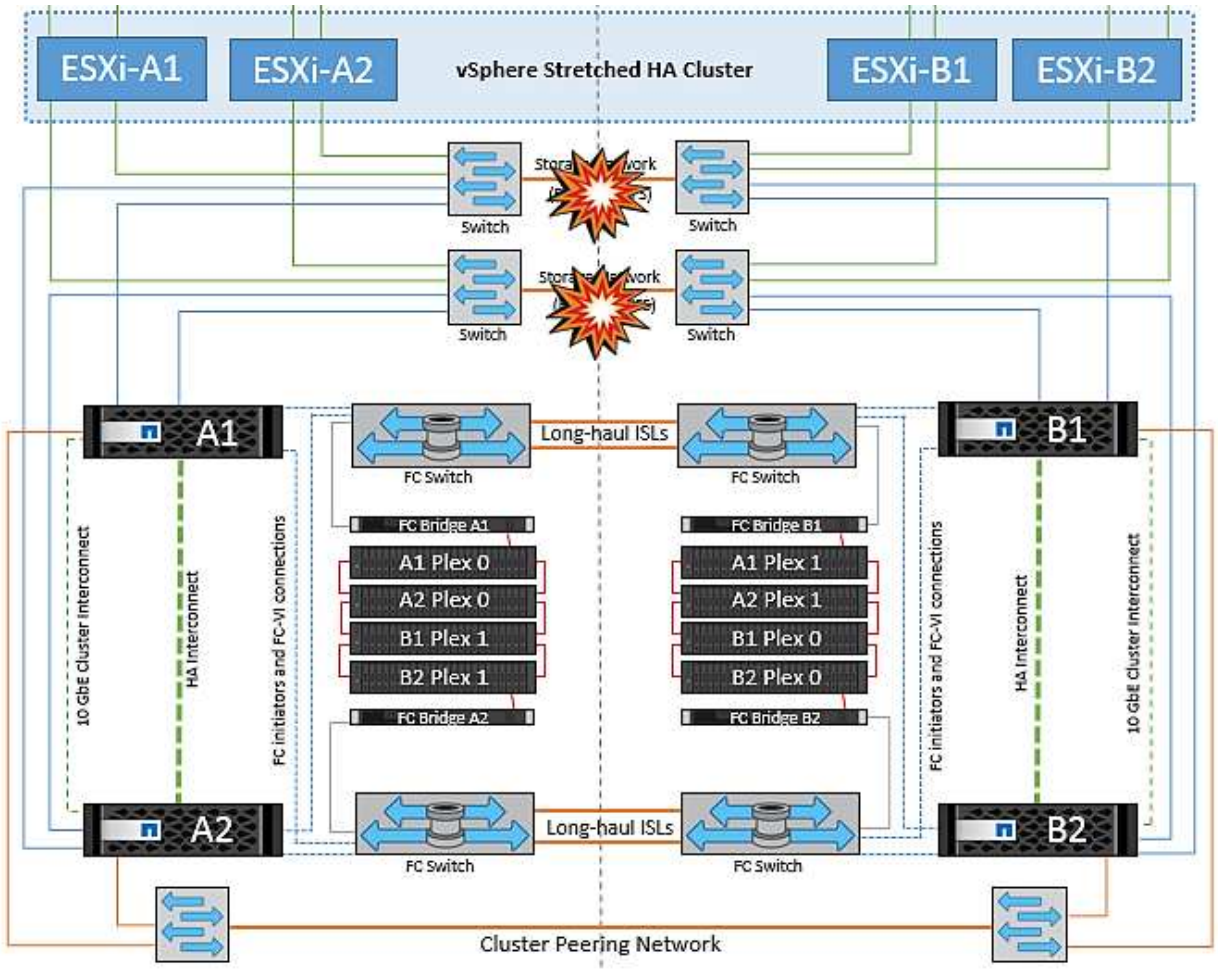
Verbindungsfehler zwischen Switches im Managementnetzwerk



In diesem Szenario können die ESXi-Hosts an Standort A nicht mit ESXi-Hosts an Standort B kommunizieren, wenn die ISL-Links am Front-End-Hostverwaltungsnetzwerk fehlschlagen Dies führt zu einer Netzwerkpartition, da ESXi-Hosts an einem bestimmten Standort die Netzwerk-Heartbeats nicht an den Master-Node im HA-Cluster senden können. Daher gibt es aufgrund der Partition zwei Netzwerksegmente, und in jedem Segment gibt es einen Master-Knoten, der die VMs vor Host-Ausfällen innerhalb des jeweiligen Standorts schützt.

Hinweis: während dieser Zeit bleiben die virtuellen Maschinen aktiv und es gibt keine Änderung im MetroCluster-Verhalten in diesem Szenario. Alle Datenspeicher bleiben von ihren jeweiligen Seiten intakt.

Verbindungsfehler zwischen Switches im Speichernetzwerk



Wenn in diesem Szenario die ISL-Verbindungen im Back-End-Speichernetzwerk ausfallen, verlieren die Hosts an Standort A den Zugriff auf die Speicher-Volumes oder LUNs von Cluster B an Standort B und umgekehrt.

Die VMware DRS Regeln sind so definiert, dass die Host-Storage-Standortaffinität die Ausführung der Virtual Machines ohne Auswirkungen auf den Standort erleichtert.

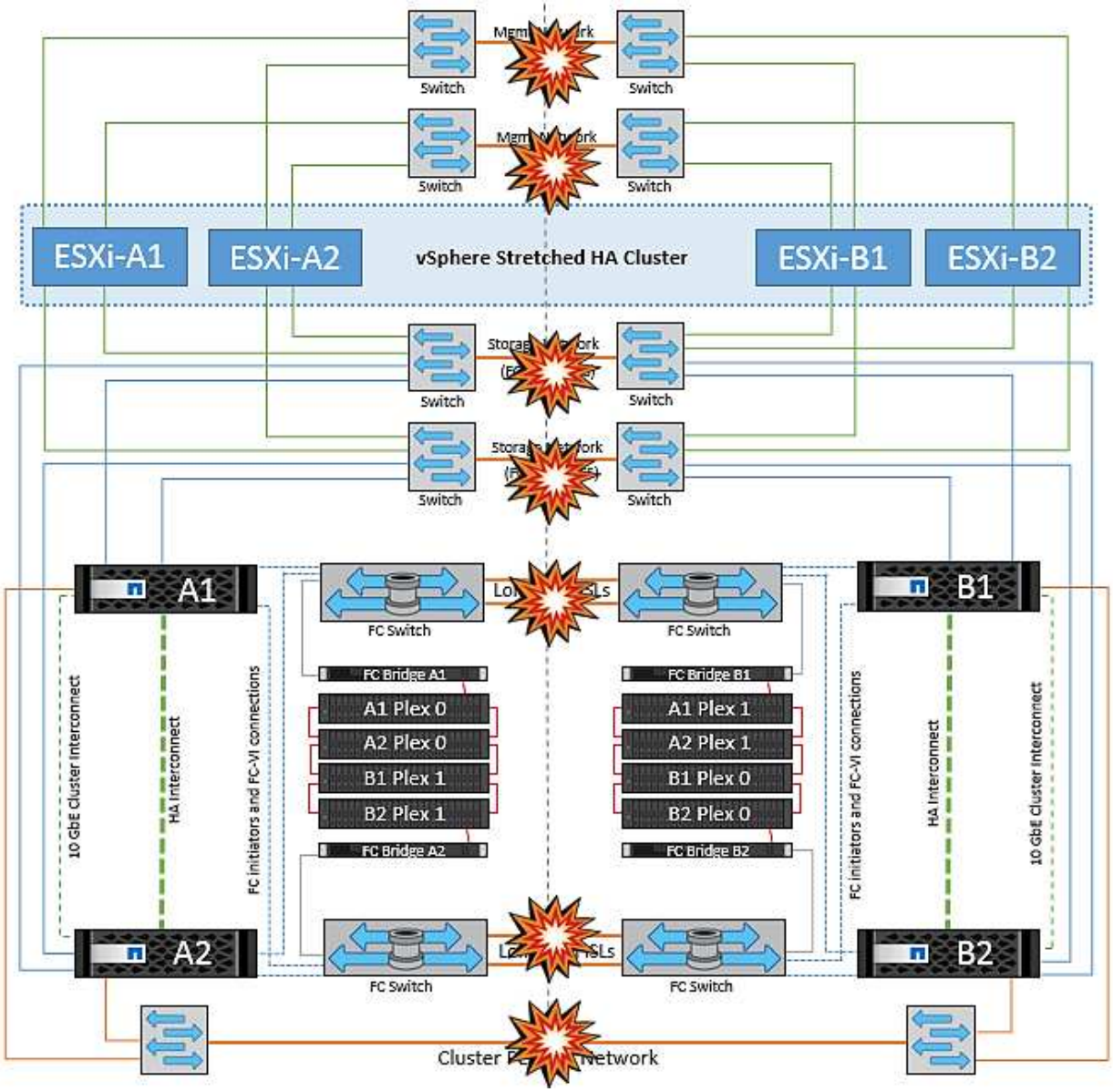
Während dieses Zeitraums bleiben die virtuellen Maschinen an ihren jeweiligen Standorten in Betrieb und es gibt keine Änderung im MetroCluster-Verhalten in diesem Szenario. Alle Datenspeicher bleiben von ihren jeweiligen Seiten intakt.

Wenn aus irgendeinem Grund die Affinitätsregel verletzt wurde (z. B. VM1, das von Standort A ausgeführt werden sollte, wo sich seine Festplatten auf lokalen Cluster A-Knoten befinden, auf einem Host an Standort B ausgeführt wird), wird der Remote-Zugriff auf das Laufwerk der virtuellen Maschine über ISL-Links erfolgen. Aufgrund eines ISL-Verbindungsfehlers kann VM1, der an Standort B ausgeführt wird, nicht auf seine Festplatten schreiben, da die Pfade zum Storage-Volume ausgefallen sind und die jeweilige Virtual Machine nicht verfügbar ist. In diesen Situationen nimmt VMware HA keine Aktion vor, da die Hosts aktiv Heartbeats senden. Diese Virtual Machines müssen an den jeweiligen Standorten manuell ausgeschaltet und eingeschaltet werden. Die folgende Abbildung zeigt eine VM, die gegen eine DRS Affinitätsregel verstößt.

Alle Interswitch-Fehler oder komplette Rechenzentrumspartition

In diesem Szenario sind alle ISL-Verbindungen zwischen den Standorten ausgefallen und beide Standorte voneinander isoliert. Wie bereits in früheren Szenarien erläutert, wie z. B. ISL-Fehler im Managementnetzwerk und im Speichernetzwerk, werden die virtuellen Maschinen bei einem vollständigen ISL-Ausfall nicht beeinträchtigt.

Nachdem ESXi-Hosts zwischen Standorten partitioniert wurden, prüft der vSphere HA-Agent auf Datastore-Heartbeats. An jedem Standort sind die lokalen ESXi-Hosts in der Lage, die Datastore-Heartbeats auf ihr jeweiliges Lese-/Schreibvolumen/LUN zu aktualisieren. Hosts an Standort A gehen davon aus, dass die anderen ESXi-Hosts an Standort B ausgefallen sind, da keine Netzwerk-/Datastore-Heartbeats vorhanden sind. VSphere HA an Standort A versucht, die virtuellen Maschinen von Standort B neu zu starten. Dies schlägt schließlich fehl, da der Zugriff auf die Datenspeicher von Standort B aufgrund eines Fehlers in der Storage-ISL nicht möglich ist. Eine ähnliche Situation wiederholt sich in Standort B.



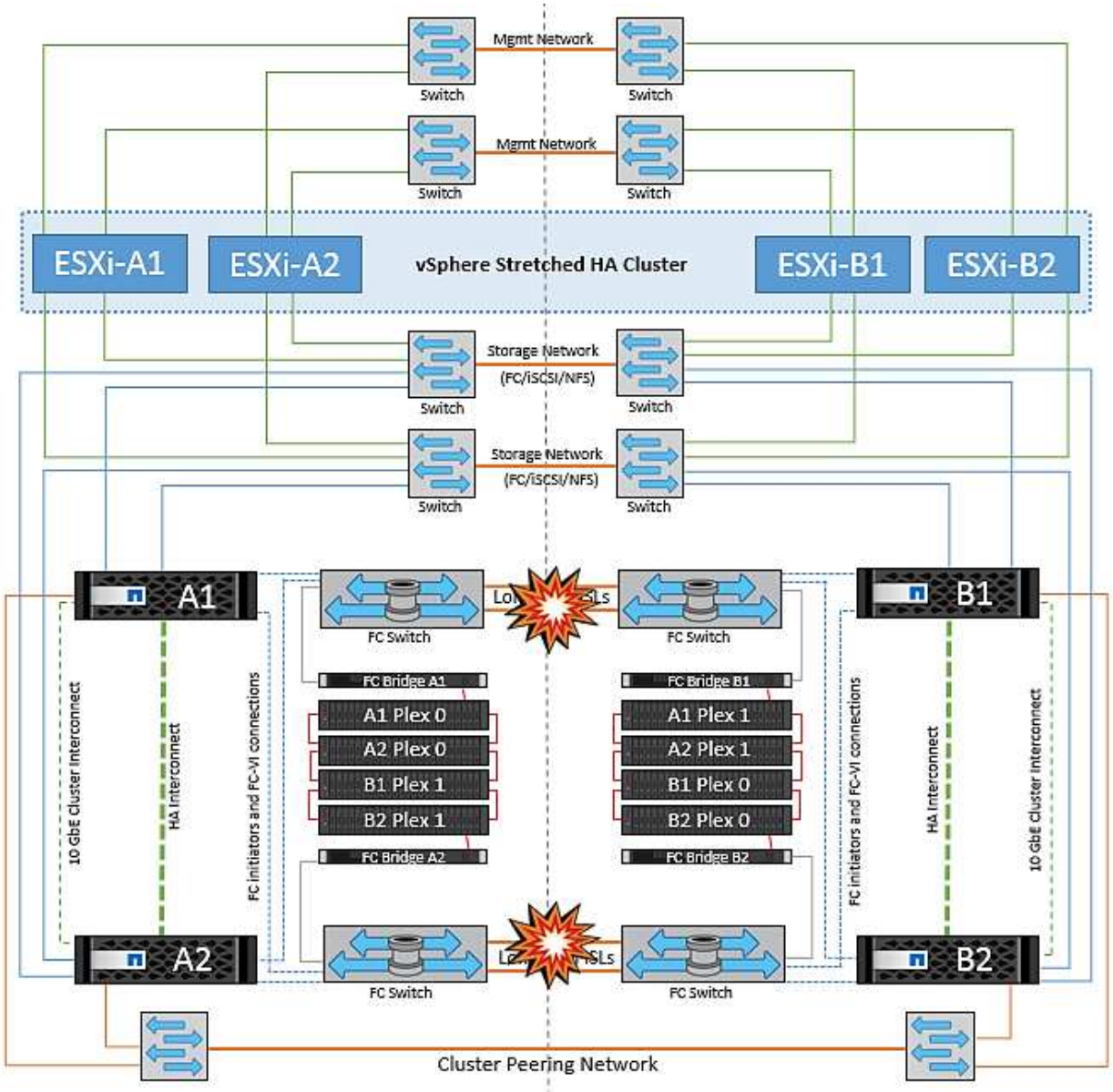
NetApp empfiehlt, festzustellen, ob eine Virtual Machine gegen die DRS Regeln verstoßen hat. Alle virtuellen Maschinen, die von einem Remote-Standort aus ausgeführt werden, sind ausgefallen, da sie nicht auf den Datastore zugreifen können, und vSphere HA startet diese virtuelle Maschine am lokalen Standort neu. Nachdem die ISL-Links wieder online sind, wird die virtuelle Maschine, die am Remote-Standort ausgeführt wurde, abgebrochen, da es nicht zwei Instanzen virtueller Maschinen geben kann, die mit denselben MAC-Adressen ausgeführt werden.

Verbindungsfehler zwischen Switches auf beiden Fabrics in NetApp MetroCluster

In einem Szenario, in dem ein oder mehrere ISLs ausfallen, wird der Datenverkehr über die verbleibenden Links fortgesetzt. Wenn alle ISLs auf beiden Fabrics ausfallen, sodass kein Link zwischen den Standorten für die Storage- und NVRAM-Replizierung vorhanden ist, stellt jeder Controller weiterhin seine lokalen Daten bereit. Bei der Wiederherstellung eines Minimums von einer ISL wird die Resynchronisierung aller Plexe automatisch durchgeführt.

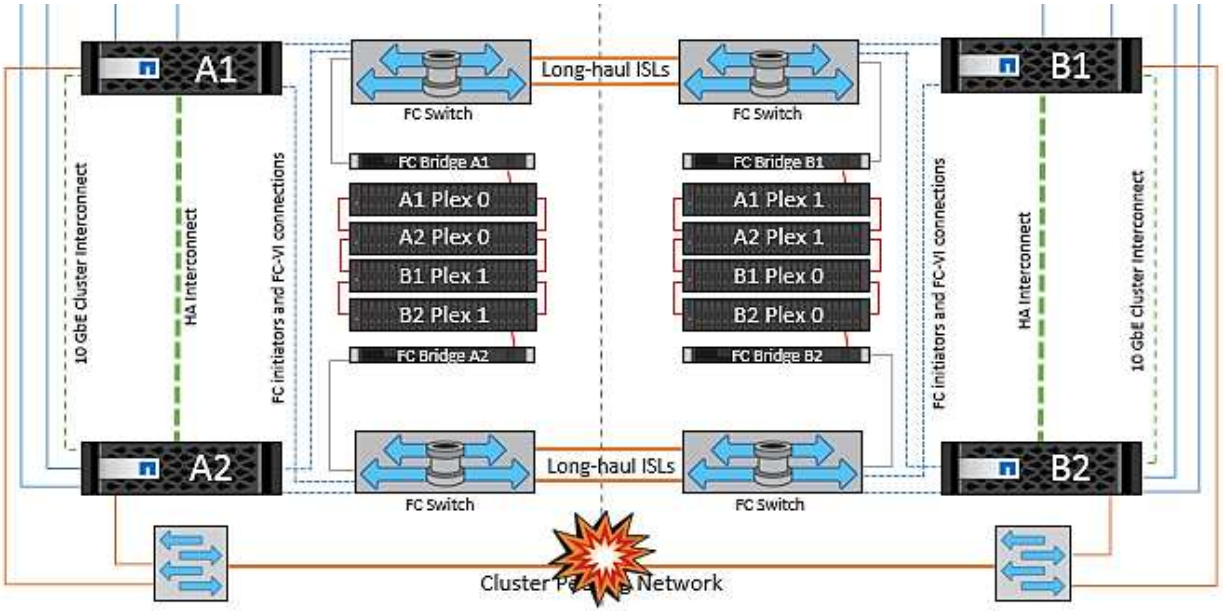
Alle Schreibvorgänge, die nach einem Ausfall aller ISLs stattfinden, werden nicht auf den anderen Standort gespiegelt. Bei einem Disaster-Switchover käme es, während sich die Konfiguration in diesem Zustand befindet, zu einem Verlust der nicht synchronisierten Daten. In diesem Fall ist ein manueller Eingriff für die

Wiederherstellung nach der Umschaltung erforderlich. Wenn es wahrscheinlich ist, dass über einen längeren Zeitraum keine ISLs verfügbar sind, kann ein Administrator alle Datenservices herunterfahren, um bei Bedarf ein Switchover im Notfall zu verhindern, dass Daten verloren gehen. Die Durchführung dieser Maßnahme sollte mit der Wahrscheinlichkeit einer Katastrophe abgewogen werden, die eine Umschaltung erfordert, bevor mindestens eine ISL verfügbar wird. Wenn ISLs in einem kaskadierenden Szenario ausfallen, könnte ein Administrator alternativ eine geplante Umschaltung zu einem der Standorte auslösen, bevor alle Links fehlgeschlagen sind.



Verbindungsfehler Bei Peered Cluster

In einem Peering-Cluster-Link-Ausfallszenario, da die Fabric-ISLs noch aktiv sind, werden die Datenservices (Lese- und Schreibvorgänge) an beiden Standorten auf beiden Plexen fortgesetzt. Jegliche Änderungen an der Cluster-Konfiguration (beispielsweise das Hinzufügen einer neuen SVM, die Bereitstellung eines Volumes oder einer LUN in einer vorhandenen SVM) können nicht an den anderen Standort weitergegeben werden. Diese werden in den lokalen CRS-Metadaten-Volumes aufbewahrt und bei Wiederherstellung der Peering-Cluster-Verbindung automatisch auf das andere Cluster übertragen. Wenn eine erzwungene Umschaltung erforderlich ist, bevor der Peered Cluster-Link wiederhergestellt werden kann, werden ausstehende Cluster-Konfigurationsänderungen automatisch von der replizierten Remote-Kopie der Metadaten-Volumes am noch verbleibenden Standort im Rahmen der Umschaltung eingespielt.



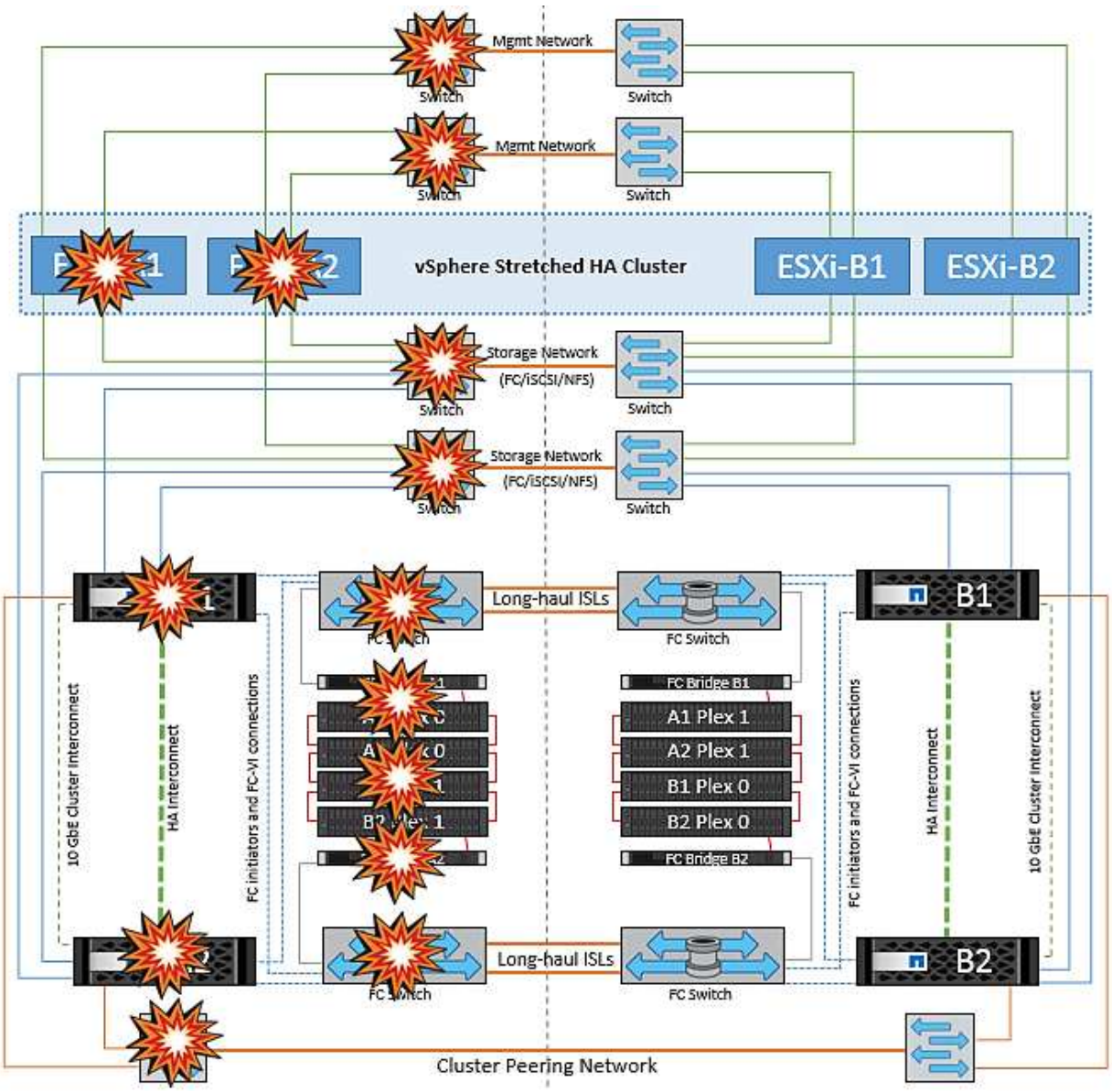
Kompletter Standortausfall

In einem kompletten Standort-A-Fehlerszenario erhalten die ESXi-Hosts an Standort B keinen Netzwerk-Heartbeat von den ESXi-Hosts an Standort A, weil sie ausgefallen sind. Der HA-Master an Standort B überprüft, ob die Datastore-Heartbeats nicht vorhanden sind, deklariert die Hosts an Standort A als fehlgeschlagen und versucht, die virtuellen Maschinen an Standort A an Standort B neu zu starten In diesem Zeitraum führt der Speicheradministrator eine Umschaltung durch, um die Dienste der ausgefallenen Nodes am noch intakten Standort wiederaufzunehmen. Dadurch werden alle Speicherservices von Standort A an Standort B wiederhergestellt Nachdem die Volumes oder LUNs an Standort A an Standort B verfügbar sind, versucht der HA-Master-Agent, die virtuellen Maschinen am Standort A an Standort B neu zu starten

Wenn der Versuch des vSphere HA Master-Agenten, eine VM neu zu starten, fehlschlägt (d. h. sie wird registriert und eingeschaltet), wird der Neustart nach einer Verzögerung erneut durchgeführt. Die Verzögerung zwischen den Neustarts kann auf maximal 30 Minuten konfiguriert werden. VSphere HA versucht diese Neustarts für eine maximale Anzahl von Versuchen (standardmäßig sechs Versuche).

Hinweis: der HA-Master startet die Neustartversuche erst, wenn der Platzierungsmanager geeigneten Speicher findet. Im Falle eines vollständigen Standortausfalls würde dies nach der Umschaltung der Fall sein.

Wenn Standort A umgeschaltet wurde, kann ein nachträglicher Ausfall eines der noch intakten Knoten Standort B nahtlos durch einen Failover auf den noch intakten Knoten bewältigt werden. In diesem Fall wird die Arbeit von vier Nodes jetzt nur von einem Node ausgeführt. Die Wiederherstellung würde in diesem Fall eine Rückgabe an den lokalen Knoten bedeuten. Wenn Standort A wiederhergestellt wird, wird ein Switchback-Vorgang durchgeführt, um den stabilen Konfigurationsbetrieb wiederherzustellen.


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
Long-haul ISLs
FC Switch
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
Cluster Peering Network