



# Ausfallszenarien

## Enterprise applications

NetApp  
December 17, 2024

# Inhalt

- Ausfallszenarien ..... 1
  - Überblick ..... 1
  - Beispielarchitektur ..... 2
  - RAC-Verbindungsfehler ..... 4
  - SnapMirror-Kommunikationsfehler ..... 4
  - Totaler Fehler bei der Netzwerkverbindung ..... 5
  - Standortausfall ..... 7
  - Fehler beim Mediator ..... 9
  - Servicewiederherstellung ..... 10
  - Manueller Failover ..... 10

# Ausfallszenarien

## Überblick

Die Planung einer vollständigen Applikationsarchitektur für die aktive Synchronisierung von SnapMirror erfordert ein Verständnis dafür, wie SM-AS in verschiedenen geplanten und ungeplanten Failover-Szenarien reagiert.

In den folgenden Beispielen wird davon ausgegangen, dass Standort A als bevorzugter Standort konfiguriert ist.

### Verlust der Replikationskonnektivität

Wenn die SM-AS-Replikation unterbrochen wird, kann die Schreib-I/O nicht abgeschlossen werden, da ein Cluster Änderungen nicht auf den anderen Standort replizieren kann.

#### Standort A (bevorzugte Website)

Das Ergebnis eines Ausfalls der Replikationsverbindung auf dem bevorzugten Standort ist eine ca. 15-Sekunden-Pause bei der Schreib-I/O-Verarbeitung, da ONTAP erneut replizierte Schreibvorgänge versucht, bevor festgestellt wird, dass die Replikationsverbindung wirklich nicht erreichbar ist. Nach 15 Sekunden wird die I/O-Verarbeitung von Lese- und Schreibzugriffen von Standort A fortgesetzt. Die SAN-Pfade ändern sich nicht, und die LUNs bleiben online.

#### Standort B

Da Standort B nicht der bevorzugte Standort für SnapMirror Active Sync ist, sind die LUN-Pfade nach ca. 15 Sekunden nicht mehr verfügbar.

### Ausfall des Storage-Systems

Das Ergebnis eines Storage-Systemausfalls ist nahezu identisch mit dem Ergebnis des Verlusts der Replizierungsverbindung. Am überlebenden Standort sollte eine I/O-Pause von etwa 15 Sekunden stattfinden. Nach Ablauf dieses Zeitraums von 15 Sekunden wird die E/A-Vorgänge wie gewohnt an diesem Standort fortgesetzt.

### Verlust des Mediators

Der Mediator hat keine direkte Kontrolle über Storage-Vorgänge. Er fungiert als alternativer Kontrollpfad zwischen Clustern. Die Lösung bietet insbesondere automatisierte Failover-Prozesse ohne Split-Brain-Szenario. Im normalen Betrieb repliziert jedes Cluster Änderungen an seinem Partner. Daher kann jedes Cluster überprüfen, ob das Partner-Cluster online ist und Daten bereitstellt. Wenn die Replikationsverbindung fehlschlägt, wird die Replikation beendet.

Der Grund für einen sicheren automatisierten Failover ist der Mediator, der darauf zurückzuführen ist, dass ein Storage-Cluster andernfalls nicht feststellen kann, ob der Ausfall einer bidirektionalen Kommunikation auf einen Netzwerkausfall oder einen tatsächlichen Storage-Ausfall zurückzuführen ist.

Der Mediator bietet jedem Cluster einen alternativen Pfad zur Überprüfung der Integrität seines Partners. Die Szenarien sind wie folgt:

- Wenn ein Cluster seinen Partner direkt kontaktieren kann, sind die Replizierungsservices betriebsbereit. Keine Aktion erforderlich.
- Wenn ein bevorzugter Standort nicht direkt mit dem Partner oder über den Mediator in Kontakt treten kann, wird davon ausgegangen, dass der Partner entweder tatsächlich nicht verfügbar ist oder isoliert wurde und seine LUN-Pfade offline geschaltet hat. Der bevorzugte Standort setzt dann den Status RPO=0 frei und setzt die Verarbeitung von Lese- und Schreib-I/O fort.
- Wenn ein nicht bevorzugter Standort seinen Partner nicht direkt kontaktieren kann, ihn aber über den Mediator kontaktieren kann, nimmt er seine Pfade offline und wartet auf die Rückkehr der Replikationsverbindung.
- Wenn ein nicht bevorzugter Standort keine direkte Kontaktaufnahme mit dem Partner oder über einen betrieblichen Mediator bietet, nimmt er an, dass der Partner entweder tatsächlich nicht verfügbar ist oder isoliert war und seine LUN-Pfade offline geschaltet hat. Der nicht bevorzugte Standort setzt dann den Status RPO=0 frei und verarbeitet sowohl Lese- als auch Schreib-I/O weiter. Er übernimmt die Rolle der Replikationsquelle und wird der neue bevorzugte Standort.

Wenn der Mediator vollständig nicht verfügbar ist:

- Wenn keine Replizierungsservices aus irgendeinem Grund verfügbar sind, beispielsweise der Ausfall des nicht bevorzugten Standorts oder des Storage-Systems, wird der bevorzugte Standort den Zustand RPO=0 freigeben und die I/O-Verarbeitung für Lese- und Schreibvorgänge wird wieder aufgenommen. Der nicht bevorzugte Standort nimmt seine Pfade offline.
- Ein Ausfall des bevorzugten Standorts führt zu einem Ausfall, da der nicht bevorzugte Standort nicht verifizieren kann, dass der gegenteilige Standort wirklich offline ist. Daher ist es für den nicht bevorzugten Standort nicht sicher, die Services wieder aufzunehmen.

## Dienste werden wiederhergestellt

Wenn ein Fehler behoben wurde, wie z. B. die Wiederherstellung der Site-to-Site-Verbindung oder das Einschalten eines ausgefallenen Systems, erkennen die SnapMirror Active Sync-Endpunkte automatisch, dass eine fehlerhafte Replikationsbeziehung vorhanden ist, und versetzen sie wieder in den Zustand RPO=0. Sobald die synchrone Replizierung wiederhergestellt ist, werden die fehlerhaften Pfade wieder online geschaltet.

In vielen Fällen erkennen Cluster-Applikationen automatisch die Rückgabe ausgefallener Pfade, und diese Applikationen sind ebenfalls wieder online. In anderen Fällen ist möglicherweise ein SAN-Scan auf Host-Ebene erforderlich oder Applikationen müssen manuell wieder online geschaltet werden. Es hängt von der Anwendung und ihrer Konfiguration ab, und im Allgemeinen lassen sich solche Aufgaben leicht automatisieren. ONTAP selbst behebt selbstständig und sollte keinen Benutzereingriff erfordern, um den RPO=0-Storage-Betrieb wiederaufzunehmen.

## Manueller Failover

Das Ändern des bevorzugten Standorts erfordert eine einfache Bedienung. I/O-Vorgänge werden für eine oder zwei Sekunden angehalten, da zwischen den Clustern die Berechtigung für das Replikationsverhalten wechselt, die E/A-Vorgänge sind jedoch ansonsten nicht betroffen.

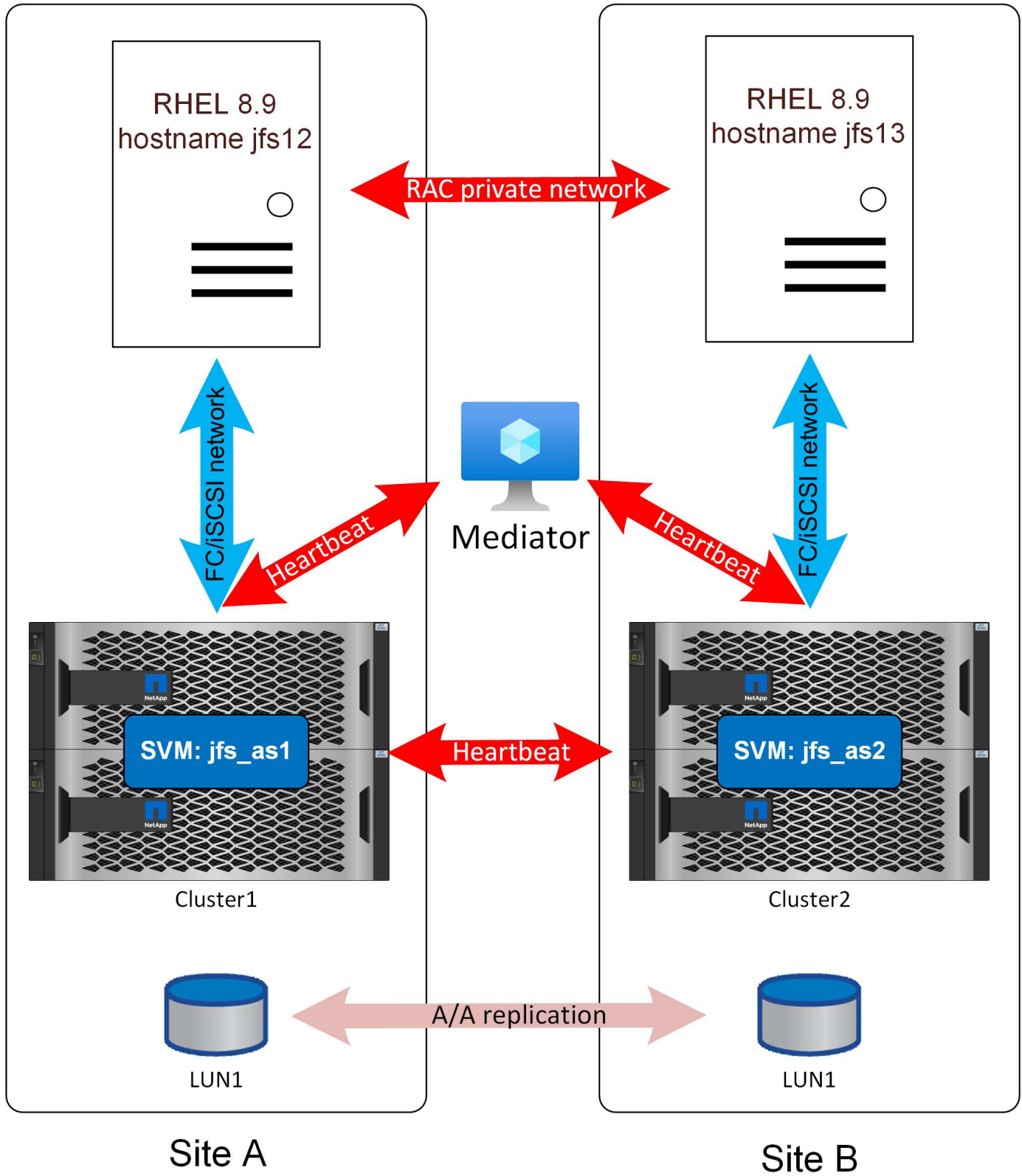
## Beispielarchitektur

Die in diesen Abschnitten gezeigten detaillierten Fehlerbeispiele basieren auf der unten dargestellten Architektur.



Dies ist nur eine von vielen Optionen für Oracle Datenbanken auf SnapMirror Active Sync. Dieses Design wurde gewählt, weil es einige der komplizierteren Szenarien illustriert.

Bei diesem Design wird davon ausgegangen, dass Standort A auf der eingestellt ist "Bevorzugter Standort".



# RAC-Verbindungsfehler

Der Verlust der Oracle RAC-Replikationsverbindung führt zu einem ähnlichen Ergebnis wie der Verlust der SnapMirror-Konnektivität, mit Ausnahme der standardmäßig kürzeren Timeouts. In den Standardeinstellungen wartet ein Oracle RAC-Knoten 200 Sekunden nach Verlust der Speicherverbindung, bevor er entfernt wird, aber er wartet nur 30 Sekunden nach Verlust des RAC-Netzwerk-Heartbeat.

Die CRS-Meldungen ähneln denen unten. Sie können die Zeitlimitüberschreitung von 30 Sekunden sehen. Da `css_Critical` auf `jfs12` gesetzt wurde, befindet sich an Ort A, das wird die Website zu überleben und `jfs13` auf Standort B wird entfernt werden.

```
2024-09-12 10:56:44.047 [ONMD(3528)]CRS-1611: Network communication with
node jfs13 (2) has been missing for 75% of the timeout interval. If this
persists, removal of this node from cluster will occur in 6.980 seconds
2024-09-12 10:56:48.048 [ONMD(3528)]CRS-1610: Network communication with
node jfs13 (2) has been missing for 90% of the timeout interval. If this
persists, removal of this node from cluster will occur in 2.980 seconds
2024-09-12 10:56:51.031 [ONMD(3528)]CRS-1607: Node jfs13 is being evicted
in cluster incarnation 621599354; details at (:CSSNM00007:) in
/gridbase/diag/crs/jfs12/crs/trace/onmd.trc.
2024-09-12 10:56:52.390 [CRSD(6668)]CRS-7503: The Oracle Grid
Infrastructure process 'crsd' observed communication issues between node
'jfs12' and node 'jfs13', interface list of local node 'jfs12' is
'192.168.30.1:33194;', interface list of remote node 'jfs13' is
'192.168.30.2:33621;'.
2024-09-12 10:56:55.683 [ONMD(3528)]CRS-1601: CSSD Reconfiguration
complete. Active nodes are jfs12 .
2024-09-12 10:56:55.722 [CRSD(6668)]CRS-5504: Node down event reported for
node 'jfs13'.
2024-09-12 10:56:57.222 [CRSD(6668)]CRS-2773: Server 'jfs13' has been
removed from pool 'Generic'.
2024-09-12 10:56:57.224 [CRSD(6668)]CRS-2773: Server 'jfs13' has been
removed from pool 'ora.NTAP'.
```

# SnapMirror-Kommunikationsfehler

Wenn der SnapMirror Active Sync Replication Link verwendet wird, kann die Schreib-I/O nicht abgeschlossen werden, da ein Cluster Änderungen nicht am anderen Standort replizieren könnte.

## Standort A

Das Ergebnis eines Ausfalls einer Replikationsverbindung an Standort A ist eine ca. 15-Sekunden-Pause bei der Schreib-I/O-Verarbeitung, da ONTAP versucht, Schreibvorgänge zu replizieren, bevor es feststellt, dass die Replikationsverbindung wirklich nicht funktionsfähig ist. Nach 15 Sekunden wird das ONTAP Cluster vor Ort A

mit der Lese- und Schreib-I/O-Verarbeitung fortgesetzt. Die SAN-Pfade ändern sich nicht, und die LUNs bleiben online.

## Standort B

Da Standort B nicht der bevorzugte Standort für SnapMirror Active Sync ist, sind die LUN-Pfade nach ca. 15 Sekunden nicht mehr verfügbar.

Die Replikationsverbindung wurde mit dem Zeitstempel 15:19:44 geschnitten. Die erste Warnung von Oracle RAC kommt 100 Sekunden später, da sich das 200-Sekunden-Timeout (gesteuert durch den Oracle RAC Parameter disktimeout) nähert.

```
2024-09-10 15:21:24.702 [ONMD(2792)]CRS-1615: No I/O has completed after
50% of the maximum interval. If this persists, voting file
/dev/mapper/grid2 will be considered not functional in 99340 milliseconds.
2024-09-10 15:22:14.706 [ONMD(2792)]CRS-1614: No I/O has completed after
75% of the maximum interval. If this persists, voting file
/dev/mapper/grid2 will be considered not functional in 49330 milliseconds.
2024-09-10 15:22:44.708 [ONMD(2792)]CRS-1613: No I/O has completed after
90% of the maximum interval. If this persists, voting file
/dev/mapper/grid2 will be considered not functional in 19330 milliseconds.
2024-09-10 15:23:04.710 [ONMD(2792)]CRS-1604: CSSD voting file is offline:
/dev/mapper/grid2; details at (:CSSNM00058:) in
/gridbase/diag/crs/jfs13/crs/trace/onmd.trc.
2024-09-10 15:23:04.710 [ONMD(2792)]CRS-1606: The number of voting files
available, 0, is less than the minimum number of voting files required, 1,
resulting in CSSD termination to ensure data integrity; details at
(:CSSNM00018:) in /gridbase/diag/crs/jfs13/crs/trace/onmd.trc
2024-09-10 15:23:04.716 [ONMD(2792)]CRS-1699: The CSS daemon is
terminating due to a fatal error from thread:
clssnmvDiskPingMonitorThread; Details at (:CSSSC00012:) in
/gridbase/diag/crs/jfs13/crs/trace/onmd.trc
2024-09-10 15:23:04.731 [OCSSD(2794)]CRS-1652: Starting clean up of CRSD
resources.
```

Sobald das 200-Sekunden-Zeitlimit für Abstimmdateiträger erreicht wurde, wird dieser Oracle RAC-Knoten selbst aus dem Cluster entfernt und neu gestartet.

## Totaler Fehler bei der Netzwerkverbindung

Wenn die Replikationsverbindung zwischen den Standorten vollständig unterbrochen wird, werden sowohl die aktive SnapMirror-Synchronisierung als auch die Oracle RAC-Verbindung unterbrochen.

Die Split-Brain-Erkennung von Oracle RAC ist vom Heartbeat des Oracle RAC Storage abhängig. Wenn der Verlust der Site-to-Site-Konnektivität zu einem gleichzeitigen Verlust sowohl des RAC-Netzwerk-Heartbeat als auch der Speicherreplikationsdienste führt, können die RAC-Standorte weder über das RAC-Interconnect noch über die RAC-Abstimmungs-Laufwerke standortübergreifend kommunizieren. Das Ergebnis einer geraden

Anzahl von Knoten kann die Entfernung beider Standorte unter den Standardeinstellungen sein. Das genaue Verhalten hängt von der Reihenfolge der Ereignisse und dem Timing des RAC-Netzwerks und der Disk-Heartbeat-Abfragen ab.

Das Risiko eines Ausfalls von 2 Standorten kann auf zwei Arten behoben werden. Zunächst kann eine "Tiebreaker" Konfiguration verwendet werden.

Wenn kein dritter Standort verfügbar ist, kann dieses Risiko durch Anpassung des Parameters für die Fehlzählung im RAC-Cluster behoben werden. Unter den Standardeinstellungen beträgt das Heartbeat-Timeout des RAC-Netzwerks 30 Sekunden. Dies wird normalerweise von RAC verwendet, um fehlerhafte RAC-Knoten zu identifizieren und aus dem Cluster zu entfernen. Es hat auch eine Verbindung zum Abstimmmedium Heartbeat.

Wenn beispielsweise das Verbindungsrohr, das den Datenverkehr zwischen den Standorten für Oracle RAC und Speicherreplikationsdienste transportiert, durch einen Bagger gekürzt wird, beginnt der 30-Sekunden-Countdown für die Fehlzählung. Wenn der bevorzugte RAC-Standortknoten den Kontakt zum anderen Standort nicht innerhalb von 30 Sekunden wiederherstellen kann und er auch nicht die Abstimmdisks verwenden kann, um zu bestätigen, dass sich der entgegengesetzte Standort innerhalb desselben 30-Sekunden-Fensters befindet, werden die bevorzugten Standortknoten ebenfalls entfernt. Das Ergebnis ist ein vollständiger Ausfall der Datenbank.

Je nachdem, wann die Abfrage der Fehlzählung erfolgt, sind 30 Sekunden möglicherweise nicht genügend Zeit für die SnapMirror Active Sync, um die Zeit zu verkürzen und die Speicherung auf dem bevorzugten Standort zu ermöglichen, um die Dienste wieder aufzunehmen, bevor das 30-Sekunden-Fenster abläuft. Dieses 30-Sekunden-Fenster kann vergrößert werden.

```
[root@jfs12 ~]# /grid/bin/crsctl set css misscount 100
CRS-4684: Successful set of parameter misscount to 100 for Cluster
Synchronization Services.
```

Mit diesem Wert kann das Speichersystem am bevorzugten Standort den Betrieb wieder aufnehmen, bevor das Timeout für die Fehlzählung abläuft. Das Ergebnis ist eine Entfernung nur der Knoten am Standort, an dem die LUN-Pfade entfernt wurden. Beispiel unten:



```
2024-09-12 09:50:59.352 [ONMD(681360)]CRS-1612: Network communication with
node jfs13 (2) has been missing for 50% of the timeout interval. If this
persists, removal of this node from cluster will occur in 49.570 seconds
2024-09-12 09:51:10.082 [CRSD(682669)]CRS-7503: The Oracle Grid
Infrastructure process 'crsd' observed communication issues between node
'jfs12' and node 'jfs13', interface list of local node 'jfs12' is
'192.168.30.1:46039;', interface list of remote node 'jfs13' is
'192.168.30.2:42037;'.
2024-09-12 09:51:24.356 [ONMD(681360)]CRS-1611: Network communication with
node jfs13 (2) has been missing for 75% of the timeout interval. If this
persists, removal of this node from cluster will occur in 24.560 seconds
2024-09-12 09:51:39.359 [ONMD(681360)]CRS-1610: Network communication with
node jfs13 (2) has been missing for 90% of the timeout interval. If this
persists, removal of this node from cluster will occur in 9.560 seconds
2024-09-12 09:51:47.527 [OHASD(680884)]CRS-8011: reboot advisory message
from host: jfs13, component: cssagent, with time stamp: L-2024-09-12-
09:51:47.451
2024-09-12 09:51:47.527 [OHASD(680884)]CRS-8013: reboot advisory message
text: oracssdagent is about to reboot this node due to unknown reason as
it did not receive local heartbeats for 10470 ms amount of time
2024-09-12 09:51:48.925 [ONMD(681360)]CRS-1632: Node jfs13 is being
removed from the cluster in cluster incarnation 621596607
```

Der Oracle Support rät dringend davon ab, die Parameter „Fehlstellen“ oder „Disktimeout“ zu ändern, um Konfigurationsprobleme zu lösen. Eine Änderung dieser Parameter kann jedoch in vielen Fällen gerechtfertigt und unvermeidbar sein, einschließlich Konfigurationen für SAN-Bootvorgänge, virtualisierte Konfigurationen und Speicherreplikation. Wenn Sie beispielsweise Stabilitätsprobleme mit einem SAN- oder IP-Netzwerk hatten, das zu RAC-Räumungen führte, sollten Sie das zugrunde liegende Problem beheben und die Werte des Misscount- oder Disktimeout nicht aufladen. Durch das Ändern von Timeouts zur Behebung von Konfigurationsfehlern wird ein Problem maskiert und kein Problem gelöst. Die Änderung dieser Parameter zur ordnungsgemäßen Konfiguration einer RAC-Umgebung basierend auf Designaspekten der zugrunde liegenden Infrastruktur unterscheidet sich und entspricht den Oracle-Support-Anweisungen. Bei SAN-Bootvorgängen ist es üblich, Fehlstellen bis zu 200 anzupassen, um Disktimeout zu entsprechen. Weitere Informationen finden Sie unter ["Dieser Link"](#).

## Standortausfall

Das Ergebnis eines Storage-System- oder Standortausfalls ist nahezu identisch mit dem Ergebnis des Verlusts der Replizierungsverbindung. Am verbleibenden Standort sollte eine I/O-Pause von etwa 15 Sekunden bei Schreibvorgängen stattfinden. Nach Ablauf dieses Zeitraums von 15 Sekunden wird die E/A-Vorgänge wie gewohnt an diesem Standort fortgesetzt.

Wenn nur das Speichersystem betroffen war, gehen die Speicherdienste des Oracle RAC-Knotens am ausgefallenen Standort verloren und führen vor der Entfernung und dem anschließenden Neustart denselben Countdown für die 200-Sekunden-Zeitüberschreitung für die Festplatte ein.

```

2024-09-11 13:44:38.613 [ONMD(3629)]CRS-1615: No I/O has completed after
50% of the maximum interval. If this persists, voting file
/dev/mapper/grid2 will be considered not functional in 99750 milliseconds.
2024-09-11 13:44:51.202 [ORAAGENT(5437)]CRS-5011: Check of resource "NTAP"
failed: details at "(:CLSN00007:)" in
"/gridbase/diag/crs/jfs13/crs/trace/crsd_oraagent_oracle.trc"
2024-09-11 13:44:51.798 [ORAAGENT(75914)]CRS-8500: Oracle Clusterware
ORAAGENT process is starting with operating system process ID 75914
2024-09-11 13:45:28.626 [ONMD(3629)]CRS-1614: No I/O has completed after
75% of the maximum interval. If this persists, voting file
/dev/mapper/grid2 will be considered not functional in 49730 milliseconds.
2024-09-11 13:45:33.339 [ORAAGENT(76328)]CRS-8500: Oracle Clusterware
ORAAGENT process is starting with operating system process ID 76328
2024-09-11 13:45:58.629 [ONMD(3629)]CRS-1613: No I/O has completed after
90% of the maximum interval. If this persists, voting file
/dev/mapper/grid2 will be considered not functional in 19730 milliseconds.
2024-09-11 13:46:18.630 [ONMD(3629)]CRS-1604: CSSD voting file is offline:
/dev/mapper/grid2; details at (:CSSNM00058:) in
/gridbase/diag/crs/jfs13/crs/trace/onmd.trc.
2024-09-11 13:46:18.631 [ONMD(3629)]CRS-1606: The number of voting files
available, 0, is less than the minimum number of voting files required, 1,
resulting in CSSD termination to ensure data integrity; details at
(:CSSNM00018:) in /gridbase/diag/crs/jfs13/crs/trace/onmd.trc
2024-09-11 13:46:18.638 [ONMD(3629)]CRS-1699: The CSS daemon is
terminating due to a fatal error from thread:
clssnmvDiskPingMonitorThread; Details at (:CSSSC00012:) in
/gridbase/diag/crs/jfs13/crs/trace/onmd.trc
2024-09-11 13:46:18.651 [OCSSD(3631)]CRS-1652: Starting clean up of CRS
resources.

```

Der SAN-Pfadstatus auf dem RAC-Knoten, der die Speicherdienste verloren hat, sieht wie folgt aus:

```

oradata7 (3600a0980383041334a3f55676c697347) dm-20 NETAPP,LUN C-Mode
size=128G features='3 queue_if_no_path pg_init_retries 50' hwhandler='1
alua' wp=rw
|+- policy='service-time 0' prio=0 status=enabled
|`- 34:0:0:18 sdam 66:96 failed faulty running
`+- policy='service-time 0' prio=0 status=enabled
  `- 33:0:0:18 sdaj 66:48 failed faulty running

```

Der linux-Host hat den Verlust der Pfade viel schneller als 200 Sekunden erkannt, aber aus Sicht der Datenbank werden die Clientverbindungen zum Host auf dem ausgefallenen Standort unter den standardmäßigen Oracle RAC-Einstellungen weiterhin 200 Sekunden lang eingefroren. Die vollständigen Datenbankvorgänge werden erst nach Abschluss der Entfernung fortgesetzt.

In der Zwischenzeit zeichnet der Oracle RAC-Knoten am gegenüberliegenden Standort den Verlust des anderen RAC-Knotens auf. Ansonsten funktioniert es wie gewohnt.

```
2024-09-11 13:46:34.152 [ONMD(3547)]CRS-1612: Network communication with
node jfs13 (2) has been missing for 50% of the timeout interval. If this
persists, removal of this node from cluster will occur in 14.020 seconds
2024-09-11 13:46:41.154 [ONMD(3547)]CRS-1611: Network communication with
node jfs13 (2) has been missing for 75% of the timeout interval. If this
persists, removal of this node from cluster will occur in 7.010 seconds
2024-09-11 13:46:46.155 [ONMD(3547)]CRS-1610: Network communication with
node jfs13 (2) has been missing for 90% of the timeout interval. If this
persists, removal of this node from cluster will occur in 2.010 seconds
2024-09-11 13:46:46.470 [OHASD(1705)]CRS-8011: reboot advisory message
from host: jfs13, component: cssmonit, with time stamp: L-2024-09-11-
13:46:46.404
2024-09-11 13:46:46.471 [OHASD(1705)]CRS-8013: reboot advisory message
text: At this point node has lost voting file majority access and
oracssdmonitor is rebooting the node due to unknown reason as it did not
receive local hearbeats for 28180 ms amount of time
2024-09-11 13:46:48.173 [ONMD(3547)]CRS-1632: Node jfs13 is being removed
from the cluster in cluster incarnation 621516934
```

## Fehler beim Mediator

Der Mediator hat keine direkte Kontrolle über Storage-Vorgänge. Er fungiert als alternativer Kontrollpfad zwischen Clustern. Die Lösung bietet insbesondere automatisierte Failover-Prozesse ohne Split-Brain-Szenario.

Im normalen Betrieb repliziert jedes Cluster Änderungen an seinem Partner. Daher kann jedes Cluster überprüfen, ob das Partner-Cluster online ist und Daten bereitstellt. Wenn die Replikationsverbindung fehlschlägt, wird die Replikation beendet.

Der Grund für den sicheren automatisierten Betrieb ist ein Mediator, der andernfalls nicht feststellen kann, ob der Ausfall einer bidirektionalen Kommunikation auf einen Netzwerkausfall oder einen tatsächlichen Storage-Ausfall zurückzuführen ist.

Der Mediator bietet jedem Cluster einen alternativen Pfad zur Überprüfung der Integrität seines Partners. Die Szenarien sind wie folgt:

- Wenn ein Cluster seinen Partner direkt kontaktieren kann, sind die Replizierungsservices betriebsbereit. Keine Aktion erforderlich.
- Wenn ein bevorzugter Standort nicht direkt mit dem Partner oder über den Mediator in Kontakt treten kann, wird davon ausgegangen, dass der Partner entweder tatsächlich nicht verfügbar ist oder isoliert wurde und seine LUN-Pfade offline geschaltet hat. Der bevorzugte Standort setzt dann den Status RPO=0 frei und setzt die Verarbeitung von Lese- und Schreib-I/O fort.
- Wenn ein nicht bevorzugter Standort seinen Partner nicht direkt kontaktieren kann, ihn aber über den Mediator kontaktieren kann, nimmt er seine Pfade offline und wartet auf die Rückkehr der Replikationsverbindung.

- Wenn ein nicht bevorzugter Standort keine direkte Kontaktaufnahme mit dem Partner oder über einen betrieblichen Mediator bietet, nimmt er an, dass der Partner entweder tatsächlich nicht verfügbar ist oder isoliert war und seine LUN-Pfade offline geschaltet hat. Der nicht bevorzugte Standort setzt dann den Status RPO=0 frei und verarbeitet sowohl Lese- als auch Schreib-I/O weiter. Er übernimmt die Rolle der Replikationsquelle und wird der neue bevorzugte Standort.

Wenn der Mediator vollständig nicht verfügbar ist:

- Ein Ausfall der Replikationsdienste führt aus irgendeinem Grund dazu, dass der bevorzugte Standort den Zustand RPO=0 freigibt und die Lese- und Schreib-I/O-Verarbeitung wieder aufgenommen wird. Der nicht bevorzugte Standort nimmt seine Pfade offline.
- Ein Ausfall des bevorzugten Standorts führt zu einem Ausfall, da der nicht bevorzugte Standort nicht verifizieren kann, dass der gegenteilige Standort wirklich offline ist. Daher ist es für den nicht bevorzugten Standort nicht sicher, die Services wieder aufzunehmen.

## Servicewiederherstellung

SnapMirror bietet Selbstreparatur. SnapMirror Active Sync erkennt automatisch eine fehlerhafte Replikationsbeziehung und versetzt sie zurück in den Zustand RPO=0. Sobald die synchrone Replikation wiederhergestellt ist, werden die Pfade wieder online geschaltet.

In vielen Fällen erkennen Cluster-Applikationen automatisch die Rückgabe ausgefallener Pfade, und diese Applikationen sind ebenfalls wieder online. In anderen Fällen ist möglicherweise ein SAN-Scan auf Host-Ebene erforderlich oder Applikationen müssen manuell wieder online geschaltet werden.

Es hängt von der Anwendung und ihrer Konfiguration ab, und im Allgemeinen können solche Aufgaben leicht automatisiert werden. Die SnapMirror Active Sync Software selbst wird automatisch behoben und sollte nach der Wiederherstellung der Stromversorgung und Konnektivität keinen Benutzereingriff erfordern, um die RPO=0-Speichervorgänge wiederaufzunehmen.

## Manueller Failover

Der Begriff „Failover“ bezieht sich nicht auf die Richtung der Replizierung mit SnapMirror Active Sync, da es sich um eine bidirektionale Replizierungstechnologie handelt. Stattdessen bezieht sich „Failover“ darauf, welches Speichersystem bei einem Ausfall der bevorzugte Standort ist.

Möglicherweise möchten Sie beispielsweise ein Failover ausführen, um den bevorzugten Standort zu ändern, bevor Sie einen Standort zu Wartungszwecken herunterfahren oder bevor Sie einen DR-Test durchführen.

Das Ändern des bevorzugten Standorts erfordert eine einfache Bedienung. I/O-Vorgänge werden für eine oder zwei Sekunden angehalten, da zwischen den Clustern die Berechtigung für das Replikationsverhalten wechselt, die E/A-Vorgänge sind jedoch ansonsten nicht betroffen.

GUI-Beispiel:

# Relationships

Local destinations

Local sources

Search Download Show/hide Filter

Source	Destination	Policy type
▼ jfs_as1:/cg/jfsAA	⋮ jfs_as2:/cg/jfsAA	Synchronous

- Edit
- Update
- Delete
- Failover

Beispiel für eine Rückänderung über die CLI:

```
Cluster2::> snapmirror failover start -destination-path jfs_as2:/cg/jfsAA
[Job 9575] Job is queued: SnapMirror failover for destination
"jfs_as2:/cg/jfsAA".
```

```
Cluster2::> snapmirror failover show
```

Source Path	Destination Path	Type	Status	start-time	end-time	Error Reason
jfs_as1:/cg/jfsAA	jfs_as2:/cg/jfsAA	planned	completed	9/11/2024 09:29:22	9/11/2024 09:29:32	

The new destination path can be verified as follows:

```
Cluster1::> snapmirror show -destination-path jfs_as1:/cg/jfsAA
```

```
Source Path: jfs_as2:/cg/jfsAA
Destination Path: jfs_as1:/cg/jfsAA
Relationship Type: XDP
Relationship Group Type: consistencygroup
SnapMirror Policy Type: automated-failover-duplex
SnapMirror Policy: AutomatedFailOverDuplex
Tries Limit: -
Mirror State: Snapmirrored
Relationship Status: InSync
```

## Copyright-Informationen

Copyright © 2024 NetApp. Alle Rechte vorbehalten. Gedruckt in den USA. Dieses urheberrechtlich geschützte Dokument darf ohne die vorherige schriftliche Genehmigung des Urheberrechtinhabers in keiner Form und durch keine Mittel – weder grafische noch elektronische oder mechanische, einschließlich Fotokopieren, Aufnehmen oder Speichern in einem elektronischen Abrufsystem – auch nicht in Teilen, vervielfältigt werden.

Software, die von urheberrechtlich geschütztem NetApp Material abgeleitet wird, unterliegt der folgenden Lizenz und dem folgenden Haftungsausschluss:

DIE VORLIEGENDE SOFTWARE WIRD IN DER VORLIEGENDEN FORM VON NETAPP ZUR VERFÜGUNG GESTELLT, D. H. OHNE JEGLICHE EXPLIZITE ODER IMPLIZITE GEWÄHRLEISTUNG, EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE STILLSCHWEIGENDE GEWÄHRLEISTUNG DER MARKTGÄNGIGKEIT UND EIGNUNG FÜR EINEN BESTIMMTEN ZWECK, DIE HIERMIT AUSGESCHLOSSEN WERDEN. NETAPP ÜBERNIMMT KEINERLEI HAFTUNG FÜR DIREKTE, INDIREKTE, ZUFÄLLIGE, BESONDERE, BEISPIELHAFT SCHÄDEN ODER FOLGESCHÄDEN (EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE BESCHAFFUNG VON ERSATZWAREN ODER -DIENSTLEISTUNGEN, NUTZUNGS-, DATEN- ODER GEWINNVERLUSTE ODER UNTERBRECHUNG DES GESCHÄFTSBETRIEBS), UNABHÄNGIG DAVON, WIE SIE VERURSACHT WURDEN UND AUF WELCHER HAFTUNGSTHEORIE SIE BERUHEN, OB AUS VERTRAGLICH FESTGELEGTER HAFTUNG, VERSCHULDENSUNABHÄNGIGER HAFTUNG ODER DELIKTSHAFTUNG (EINSCHLIESSLICH FAHRLÄSSIGKEIT ODER AUF ANDEREM WEGE), DIE IN IRGEND EINER WEISE AUS DER NUTZUNG DIESER SOFTWARE RESULTIEREN, SELBST WENN AUF DIE MÖGLICHKEIT DERARTIGER SCHÄDEN HINGEWIESEN WURDE.

NetApp behält sich das Recht vor, die hierin beschriebenen Produkte jederzeit und ohne Vorankündigung zu ändern. NetApp übernimmt keine Verantwortung oder Haftung, die sich aus der Verwendung der hier beschriebenen Produkte ergibt, es sei denn, NetApp hat dem ausdrücklich in schriftlicher Form zugestimmt. Die Verwendung oder der Erwerb dieses Produkts stellt keine Lizenzierung im Rahmen eines Patentrechts, Markenrechts oder eines anderen Rechts an geistigem Eigentum von NetApp dar.

Das in diesem Dokument beschriebene Produkt kann durch ein oder mehrere US-amerikanische Patente, ausländische Patente oder anhängige Patentanmeldungen geschützt sein.

ERLÄUTERUNG ZU „RESTRICTED RIGHTS“: Nutzung, Vervielfältigung oder Offenlegung durch die US-Regierung unterliegt den Einschränkungen gemäß Unterabschnitt (b)(3) der Klausel „Rights in Technical Data – Noncommercial Items“ in DFARS 252.227-7013 (Februar 2014) und FAR 52.227-19 (Dezember 2007).

Die hierin enthaltenen Daten beziehen sich auf ein kommerzielles Produkt und/oder einen kommerziellen Service (wie in FAR 2.101 definiert) und sind Eigentum von NetApp, Inc. Alle technischen Daten und die Computersoftware von NetApp, die unter diesem Vertrag bereitgestellt werden, sind gewerblicher Natur und wurden ausschließlich unter Verwendung privater Mittel entwickelt. Die US-Regierung besitzt eine nicht ausschließliche, nicht übertragbare, nicht unterlizenzierbare, weltweite, limitierte unwiderrufliche Lizenz zur Nutzung der Daten nur in Verbindung mit und zur Unterstützung des Vertrags der US-Regierung, unter dem die Daten bereitgestellt wurden. Sofern in den vorliegenden Bedingungen nicht anders angegeben, dürfen die Daten ohne vorherige schriftliche Genehmigung von NetApp, Inc. nicht verwendet, offengelegt, vervielfältigt, geändert, aufgeführt oder angezeigt werden. Die Lizenzrechte der US-Regierung für das US-Verteidigungsministerium sind auf die in DFARS-Klausel 252.227-7015(b) (Februar 2014) genannten Rechte beschränkt.

## Markeninformationen

NETAPP, das NETAPP Logo und die unter <http://www.netapp.com/TM> aufgeführten Marken sind Marken von NetApp, Inc. Andere Firmen und Produktnamen können Marken der jeweiligen Eigentümer sein.