



执行切换，修复和切回 ONTAP MetroCluster

NetApp
January 09, 2026

目录

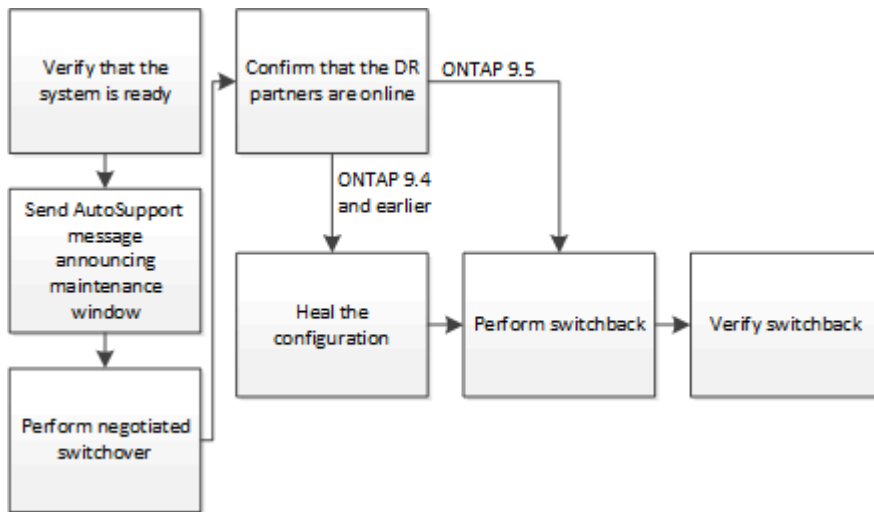
执行切换，修复和切回	1
执行切换以进行测试或维护	1
为测试或维护执行切换	1
MetroCluster配置处于切换状态时的限制	1
验证您的系统是否已做好切换准备	2
在协商切换之前发送自定义 AutoSupport 消息	3
执行协商切换	3
验证SVM是否正在运行且聚合是否联机	5
修复配置	6
执行切回	10
验证切回是否成功	13
用于切换，修复和切回的命令	14
使用系统管理器执行切换和切回(仅限MetroCluster IP配置)	15
切换和切回概述	15
在ONTAP 9.6或9.7中使用System Manager进行切换和切回	15
在ONTAP 9.8或更高版本中使用System Manager进行切换和切回	16
监控 MetroCluster 配置	17
正在检查 MetroCluster 配置	17
用于检查和监控 MetroCluster 配置的命令	20
使用 MetroCluster Tiebreaker 或 ONTAP 调解器监控配置	21
NetApp MetroCluster Tiebreaker 软件如何检测故障	21
使用 NVFAIL 监控和保护文件系统一致性	23
NVFAIL 如何影响对 NFS 卷或 LUN 的访问	23
用于监控数据丢失事件的命令	24
在切换后访问处于 NVFAIL 状态的卷	24
在切换后恢复处于 NVFAIL 状态的 LUN	25
从何处查找追加信息	25
MetroCluster 和其他信息	25

执行切换，修复和切回

执行切换以进行测试或维护

为测试或维护执行切换

如果要测试 MetroCluster 功能或执行计划内维护，您可以执行协商切换，其中一个集群将完全切换到配对集群。然后，您可以修复并切回配置。



从 ONTAP 9.6 开始，可以使用 ONTAP System Manager 对 MetroCluster IP 配置执行切换和切回操作。

MetroCluster配置处于切换状态时的限制

当系统处于切换状态时、不应执行某些操作。了解系统处于切换状态时的受限操作。

切换中的操作受限

当系统处于切换状态时、不支持以下操作：

- 创建或删除聚合和卷
- 创建或删除SVM
- 正在创建或删除此文件
- 添加或删除磁盘(仅当在恢复过程中更换磁盘时)
- 对SnapMirror SVM DR执行配置更改
- 修改现有广播域或创建新广播域
- 修改网络子网

切换时更换硬件

在系统处于切换状态时、请使用以下过程更换控制器硬件：

- 如果您需要在不处于切换状态的站点上更换相同类型的控制器，请按照中的过程进行操作"[从多控制器或存储故障中恢复](#)"。
 - 如果在正常运行的站点上切换节点时需要更换控制器模块和机箱，请关闭这两个控制器，然后执行至的过程"[从多控制器或存储故障中恢复](#)"。
- 如果需要将控制器更换为其他类型的控制器，请按照中适用于您的配置的过程进行操作"[选择控制器升级操作步骤](#)"。
 - 如果您的系统因控制器故障而处于切换状态、或者您在切换期间遇到控制器故障、则必须先更换控制器硬件、执行切回、然后再执行控制器升级：
 - i. 要更换控制器硬件并执行切回，请执行"[从多控制器或存储故障中恢复](#)"。
 - ii. 更换硬件后，请按照中的过程执行控制器升级"[选择控制器升级操作步骤](#)"。

验证您的系统是否已做好切换准备

您可以使用 ` -simulate ` 选项预览切换操作的结果。通过验证检查，您可以在开始操作之前验证是否满足成功运行的大部分前提条件。通过问题描述从站点运行以下命令，这些命令将保持正常运行：

1. 将权限级别设置为 advanced： `set -privilege advanced`
2. 在要保持正常运行的站点上，模拟切换操作： `MetroCluster switchover -simulate`
3. 查看返回的输出。

输出将显示是否存在任何否决来阻止切换操作。每次执行 MetroCluster 操作时，您都必须验证一组标准，以确保操作成功。" veto " 是一种机制，用于在未满足一个或多个条件时禁止此操作。有两种类型的否决： " 软 s " 否决和 " 硬 " 否决。您可以覆盖软否决，但不能覆盖硬否决。例如，要在四节点 MetroCluster 配置中执行协商切换，一个标准是所有节点均已启动且运行状况良好。假设一个节点已关闭并由其 HA 配对节点接管。切换操作将被硬否决，因为所有节点都必须正常运行是一个硬标准。由于这是一种硬否决，因此您无法覆盖此否决。



最好不要覆盖任何否决。

示例：验证结果

以下示例显示了在模拟切换操作时遇到的错误：

```
cluster4::*> metrocluster switchover -simulate

[Job 126] Preparing the cluster for the switchover operation...
[Job 126] Job failed: Failed to prepare the cluster for the switchover
operation. Use the "metrocluster operation show" command to view detailed
error
information. Resolve the errors, then try the command again.
```



在更换所有故障磁盘之前，协商切换和切回将失败。您可以在更换故障磁盘后执行灾难恢复。如果要忽略有关故障磁盘的警告，可以为协商切换和切回添加软否决。

在协商切换之前发送自定义 **AutoSupport** 消息

在执行协商切换之前，您应问题描述发送 AutoSupport 消息以通知 NetApp 技术支持正在进行维护。协商切换可能会导致丛或 MetroCluster 操作失败，从而触发 AutoSupport 消息。告知技术支持正在进行维护，可防止他们在假定已发生中断的情况下创建案例。

必须在每个 MetroCluster 站点上执行此任务。

步骤

1. 登录到 Site_A 上的集群
2. 调用指示维护开始的 AutoSupport 消息：`ssystem node AutoSupport invoke -node * -type all -message MAINT=maintenance-window-in-hours`

`maintenance-window-in-hours` 指定维护时段的长度，最长可为 72 小时。如果在此时间之前完成维护，您可以通过问题描述命令指示维护期已结束：`ssystem node AutoSupport invoke -node * -type all -message MAIN=end`

3. 在配对站点上重复此步骤。

执行协商切换

协商切换会完全关闭配对站点上的进程，然后从配对站点切换操作。您可以使用协商切换在 MetroCluster 站点上执行维护或测试切换功能。

- 在执行切回操作之前，必须完成先前的所有配置更改。

这是为了避免与协商切换或切回操作产生竞争。

- 先前已关闭的所有节点都必须启动并处于集群仲裁状态。

系统管理参考 _ 在 "了解仲裁和 epsilon`" 一节中提供了有关集群仲裁的详细信息。

"系统管理"

- 集群对等网络必须可从两个站点访问。

- MetroCluster 配置中的所有节点都必须运行相同版本的 ONTAP 软件。
- 在创建新的 SnapMirror 关系之前，必须在 MetroCluster 配置中的两个站点上将选项 `replication.create_data_protection_rels.enable` 设置为 on。
- 对于双节点 MetroCluster 配置，如果站点之间的 ONTAP 版本不匹配，则不应在升级期间创建新的 SnapMirror 关系。
- 对于四节点 MetroCluster 配置，不支持站点之间的 ONTAP 版本不匹配。

恢复站点可能需要几个小时才能执行切回操作。

MetroCluster switchover 命令可切换 MetroCluster 配置中所有 DR 组中的节点。例如，在八节点 MetroCluster 配置中，它会切换两个 DR 组中的节点。

在准备和执行协商切换时，您不能更改任一集群的配置，也不能执行任何接管或交还操作。

对于 MetroCluster FC 配置：

- 如果可以访问远程存储，则镜像聚合将保持正常状态。
- 如果无法访问远程存储，则在协商切换后，镜像聚合将降级。
- 如果无法访问远程存储，则位于灾难站点的未镜像聚合将不可用。这可能会导致控制器中断。

对于 MetroCluster IP 配置：



在执行维护任务之前，如果使用 Tiebreaker 或调解器实用程序监控 MetroCluster 配置，则必须删除监控。["在执行维护任务之前，请删除 ONTAP 调解器或 Tiebreaker 监控"](#)

- 对于 ONTAP 9.4 及更早版本：
 - 在协商切换后，镜像聚合将降级。
- 对于 ONTAP 9.5 及更高版本：
 - 如果可以访问远程存储，则镜像聚合将保持正常状态。
 - 如果无法访问远程存储，则在协商切换后，镜像聚合将降级。
- 对于 ONTAP 9.8 及更高版本：
 - 如果无法访问远程存储，则位于灾难站点的未镜像聚合将不可用。这可能会导致控制器中断。
 - i. 使用 `MetroCluster check run`，`MetroCluster check show` 和 `MetroCluster check config-replication show` 命令可确保没有正在进行的或待定的配置更新。通过问题描述访问站点上的这些命令，这些命令将保持正常运行。
 - ii. 在要保持正常运行的站点上，实施切换：`MetroCluster switchover`

此操作可能需要几分钟才能完成。
 - iii. 监控切换完成情况：`MetroCluster operation show`

```
cluster_A::*> metrocluster operation show
Operation: Switchover
Start time: 10/4/2012 19:04:13
State: in-progress
End time: -
Errors:

cluster_A::*> metrocluster operation show
Operation: Switchover
Start time: 10/4/2012 19:04:13
State: successful
End time: 10/4/2012 19:04:22
Errors: -
```

iv. 重新建立任何 SnapMirror 或 SnapVault 配置。

验证SVM是否正在运行且聚合是否联机

切换完成后，您应验证 DR 配对节点是否已取得磁盘的所有权，以及配对 SVM 是否已联机。

在 MetroCluster 切换后运行 storage aggregate plex show 命令时，切换后的根聚合的 plex0 状态不确定，并显示为 failed。在此期间，切换后的根不会更新。只有在 MetroCluster 修复阶段之后才能确定此丛的实际状态。

步骤

1. 使用 storage aggregate show 命令检查聚合是否已切换。

在此示例中，聚合已切换。根聚合（aggr0_B2）处于已降级状态。数据聚合（b2_aggr2）处于镜像正常状态：

```
cluster_A::*> storage aggregate show
```

.

.

.

mccl-b Switched Over Aggregates:

Aggregate	Size	Available	Used%	State	#Vols	Nodes	RAID
aggr0_b2	227.1GB	45.1GB	80%	online	0	node_A_1	
raid_dp,							
mirror							
degraded							
b2_aggr1	227.1GB	200.3GB	20%	online	0	node_A_1	
raid_dp,							
mirrored							
normal							

2. 使用 vservers show 命令确认二级 SVM 已联机。

在此示例中，二级站点上先前处于休眠状态的 sync-destination SVM 已激活，并且 Admin State 为 running：

```
cluster_A::*> vservers show
```

Name	Name	Type	Subtype	Admin State	Operational State	Root Volume
Vserver	Aggregate	Service Mapping				
...						
cluster_B-vs1b-mc	data	sync-destination		running	running	
vs1b_vol	aggr_b1	file	file			

修复配置

修复 MetroCluster FC 配置中的配置

切换后，您必须按特定顺序执行修复操作，才能还原 MetroCluster 功能。

- 必须已执行切换，并且正常运行的站点必须正在提供数据。
- 灾难站点上的节点必须暂停或保持关闭状态。

在修复过程中，不能完全启动它们。

- 灾难站点上的存储必须可访问（磁盘架已启动，正常运行且可访问）。
- 在光纤连接的 MetroCluster 配置中，交换机间链路（ISL）必须已启动且正在运行。
- 在四节点 MetroCluster 配置中，运行正常的站点中的节点不能处于 HA 故障转移状态（对于每个 HA 对，所有节点都必须已启动且正在运行）。

必须先对数据聚合执行修复操作，然后再对根聚合执行此操作。

在协商切换后修复数据聚合

完成任何维护或测试后，您必须修复数据聚合。此过程会重新同步数据聚合，并使灾难站点做好正常运行的准备。在修复根聚合之前，您必须修复数据聚合。

远程集群中的所有配置更新均已成功复制到本地集群。您在此操作步骤中启动灾难站点上的存储，但不会也不能启动灾难站点上的控制器模块。

步骤

1. 运行 MetroCluster operation show 命令，确保切换已完成。

```
controller_A_1::> metrocluster operation show
Operation: switchover
State: successful
Start Time: 7/25/2014 20:01:48
End Time: 7/25/2014 20:02:14
Errors: -
```

2. 在运行正常的集群中运行 MetroCluster heal -phase aggregates 命令，以重新同步数据聚合。

```
controller_A_1::> metrocluster heal -phase aggregates
[Job 130] Job succeeded: Heal Aggregates is successful.
```

如果修复被否决，您可以使用 -override-vetoes 参数重新发出 MetroCluster heal 命令。如果使用此可选参数，则系统将覆盖任何阻止修复操作的软否决。

3. 运行 MetroCluster operation show 命令以验证操作是否已完成。

```

controller_A_1::> metrocluster operation show
  Operation: heal-aggregates
    State: successful
Start Time: 7/25/2014 18:45:55
End Time: 7/25/2014 18:45:56
Errors: -

```

4. 运行 storage aggregate show 命令以检查聚合的状态。

```

controller_A_1::> storage aggregate show
Aggregate      Size Available Used% State   #Vols  Nodes      RAID
Status
-----
...
aggr_b2      227.1GB   227.1GB    0% online    0 mcc1-a2
raid_dp, mirrored, normal...

```

5. 如果已在灾难站点上更换存储，您可能需要重新镜像聚合。

在协商切换后修复根聚合

修复数据聚合之后，您必须修复根聚合，以便为切回操作做好准备。

MetroCluster 修复过程的数据聚合阶段必须已成功完成。

步骤

1. 运行 MetroCluster heal -phase root-aggregates 命令以切回镜像聚合。

```

cluster_A::> metrocluster heal -phase root-aggregates
[Job 137] Job succeeded: Heal Root Aggregates is successful

```

如果修复被否决，您可以使用 -override-vetoes 参数重新发出 MetroCluster heal 命令。如果使用此可选参数，则系统将覆盖任何阻止修复操作的软否决。

2. 在运行正常的集群上运行 MetroCluster operation show 命令，以确认修复操作已完成：

```

cluster_A::> metrocluster operation show
  Operation: heal-root-aggregates
    State: successful
Start Time: 7/29/2014 20:54:41
End Time: 7/29/2014 20:54:42
Errors: -

```

3. 在运行正常的站点上发出以下命令，检查并删除属于灾难站点的任何故障磁盘：`disk show -broken`
4. 启动或启动灾难站点上的每个控制器模块。

如果系统显示 **LOADER** 提示符，请运行 `boot_ontap` 命令。

5. 启动节点后，验证根聚合是否已镜像。

如果两个丛都存在，则在这些丛未同步时，系统会自动进行重新同步。如果一个丛出现故障，则必须销毁该丛，并且必须使用 `storage aggregate mirror -aggregateaggregate-name` 命令重新创建镜像，以重新建立镜像关系。

修复 **MetroCluster IP** 配置中的配置（**ONTAP 9.4** 及更早版本）

您必须修复聚合，以便为切回操作做好准备。



在运行 ONTAP 9.5 的 MetroCluster IP 系统上，会自动执行修复，您可以跳过这些任务。

在执行修复操作步骤之前，必须满足以下条件：

- 必须已执行切换，并且正常运行的站点必须正在提供数据。
- 灾难站点上的存储架必须已启动，正常运行且可访问。
- ISL 必须已启动并正常运行。
- 正常运行的站点中的节点不能处于 HA 故障转移状态（两个节点必须均已启动且正在运行）。

此任务仅限运行 ONTAP 9.5 之前版本的适用场景 MetroCluster IP 配置。

此操作步骤与适用于 MetroCluster FC 配置的修复操作步骤不同。

步骤

1. 打开已切换站点上的每个控制器模块的电源，并使其完全启动。

如果系统显示 **LOADER** 提示符，请运行 `boot_ontap` 命令。

2. 执行根聚合修复阶段：`MetroCluster heal root-aggregates`

```
cluster_A::> metrocluster heal root-aggregates
[Job 137] Job succeeded: Heal Root-Aggregates is successful
```

如果修复被否决，您可以使用 `-override-vetoes` 参数重新发出 `MetroCluster heal root-aggregates` 命令。如果使用此可选参数，则系统将覆盖任何阻止修复操作的软否决。

3. 重新同步聚合：`MetroCluster heal aggregates`

```
cluster_A::> metrocluster heal aggregates
[Job 137] Job succeeded: Heal Aggregates is successful
```

如果修复被否决，您可以使用 `-override-vetoes` 参数重新发出 `MetroCluster heal` 命令。如果使用此可选参数，则系统将覆盖任何阻止修复操作的软否决。

4. 在运行正常的集群上运行 `MetroCluster operation show` 命令，以确认修复操作已完成：

```
cluster_A::> metrocluster operation show
Operation: heal-aggregates
State: successful
Start Time: 7/29/2017 20:54:41
End Time: 7/29/2017 20:54:42
Errors: -
```

执行切回

修复 `MetroCluster` 配置后，您可以执行 `MetroCluster` 切回操作。`MetroCluster` 切回操作会将配置恢复到其正常运行状态，灾难站点上的 `sync-source Storage Virtual Machine`（`SVM`）处于活动状态，并从本地磁盘池提供数据。

- 灾难集群必须已成功切换到正常运行的集群。
- 必须已对数据和根聚合执行修复。
- 正常运行的集群节点不能处于 `HA 故障转移状态`（对于每个 `HA 对`，所有节点都必须已启动且正在运行）。
- 灾难站点控制器模块必须完全启动，而不是处于 `HA 接管模式`。
- 必须镜像根聚合。
- 交换机间链路（`ISL`）必须处于联机状态。
- 必须在系统上安装所有必需的许可证。

- a. 确认所有节点均处于已启用状态：`MetroCluster node show`

以下示例显示了处于已启用状态的节点：

```
cluster_B::> metrocluster node show
```

DR	Configuration	DR
Group Cluster Node	State	Mirroring Mode
-----	-----	-----
1	cluster_A	
	node_A_1	configured enabled heal roots
completed		
	node_A_2	configured enabled heal roots
completed		
	cluster_B	
	node_B_1	configured enabled waiting for
switchback recovery		
	node_B_2	configured enabled waiting for
switchback recovery		
4 entries were displayed.		

- b. 确认所有 SVM 上的重新同步均已完成: MetroCluster SVM show
- c. 验证修复操作执行的任何自动 LIF 迁移是否已成功完成: MetroCluster check lif show
- d. 执行模拟切回以验证系统是否已准备就绪: MetroCluster switchback -simulate
- e. 检查配置:

```
MetroCluster check run
```

此命令作为后台作业运行，可能无法立即完成。

```
cluster_A::> metrocluster check run
The operation has been started and is running in the background. Wait
for
it to complete and run "metrocluster check show" to view the results.
To
check the status of the running metrocluster check operation, use the
command,
"metrocluster operation history show -job-id 2245"
```

```
cluster_A::> metrocluster check show
Last Checked On: 9/13/2018 20:41:37
```

Component	Result
nodes	ok
lifs	ok
config-replication	ok
aggregates	ok
clusters	ok
connections	ok
6 entries were displayed.	

- f. 从运行正常的集群中的任何节点运行 MetroCluster switchback 命令来执行切回： MetroCluster switchback`
- g. 检查切回操作的进度： MetroCluster show

当输出显示 Waiting for-switchback 时，切回操作仍在进行中：

```
cluster_B::> metrocluster show
```

Cluster	Entry Name	State
Local: cluster_B	Configuration state	configured
	Mode	switchover
	AUSO Failure Domain	-
Remote: cluster_A	Configuration state	configured
	Mode	waiting-for-switchback
	AUSO Failure Domain	-

当输出显示正常时，切回操作完成：

```
cluster_B::> metrocluster show
```

Cluster	Entry Name	State
Local: cluster_B	Configuration state	configured
	Mode	normal
	AUSO Failure Domain	-
Remote: cluster_A	Configuration state	configured
	Mode	normal
	AUSO Failure Domain	-

+ 如果切回需要很长时间才能完成，您可以使用 MetroCluster config-replication resync-status show 命令检查正在进行的基线的状态。此命令处于高级权限级别。

- a. 重新建立任何 SnapMirror 或 SnapVault 配置。

在 ONTAP 8.3 中，您需要在执行 MetroCluster 切回操作后手动重新建立丢失的 SnapMirror 配置。在 ONTAP 9.0 及更高版本中，系统会自动重新建立此关系。

验证切回是否成功

执行切回后，您需要确认所有聚合和 Storage Virtual Machine（SVM）均已切回并联机。

1. 验证切换后的数据聚合是否已切回：

s存储聚合显示

在以下示例中，节点 B2 上的 aggr_b2 已切回：

```
node_B_1::> storage aggregate show
Aggregate      Size Available Used% State   #Vols  Nodes           RAID
Status
-----
...
aggr_b2        227.1GB    227.1GB    0% online      0 node_B_2  raid_dp,
mirrored,
normal
```

2. 验证运行正常的集群上的所有同步目标SVM是否均处于休眠状态(显示运行状态为's已')：

vserver show -subtype sync-destination

```
node_B_1::> vserver show -subtype sync-destination
Vserver      Type      Subtype      Admin      Operational  Root
Aggregate
-----
...
cluster_A-vs1a-mc data sync-destination
running      stopped      vs1a_vol    aggr_b2
```

MetroCluster 配置中的 sync-destination 聚合会在其名称中自动附加后缀 "-mc"，以帮助标识它们。

3. 验证灾难集群上的同步源SVM是否已启动且正在运行：

`vserver show -subtype sync-source`

```
node_A_1::> vserver show -subtype sync-source

Vserver      Type      Subtype      Admin      Operational      Root
Aggregate
-----
...
vs1a          data      sync-source
                                running   running         vs1a_vol  aggr_b2
```

4. 使用 `MetroCluster operation show` 命令确认切回操作成功。

如果命令输出显示 ...	那么 ...
切回操作状态为成功。	切回过程已完成，您可以继续操作系统。
切回操作或切回 - 继续 - 代理操作已部分成功。	执行 <code>MetroCluster operation show</code> 命令输出中建议的修复操作。

您必须重复前面的部分，以反向执行切回。如果 `site_A` 已切换 `site_B`，请让 `site_B` 切换 `site_A`

用于切换，修复和切回的命令

您可以使用特定的 ONTAP 命令来执行 MetroCluster 灾难恢复过程。

如果您要 ...	使用此命令 ...
验证是否可以执行切换，而不会出现错误或否决。	<code>MetroCluster switchover -simulate + at the advanced privilege level</code>
验证是否可以执行切回而不出现错误或否决。	<code>MetroCluster switchback -simulate + at the advanced privilege level</code>
切换到配对节点（协商切换）。	<code>MetroCluster switchover</code>
切换到配对节点（强制切换）。	<code>MetroCluster switchover -forced-on-disaster true</code>
执行数据聚合修复。	<code>MetroCluster heal -phase aggregates</code>
执行根聚合修复。	<code>MetroCluster heal -phase root-aggregates</code>

切换回主节点。	MetroCluster 切回
---------	-----------------

使用系统管理器执行切换和切回(仅限MetroCluster IP配置)

您可以将控制权从一个MetroCluster IP站点切换到另一个站点、以执行维护或从问题中恢复。



只有MetroCluster IP配置才支持切换和切回过程。

切换和切回概述

在以下两种情况下，可能会发生切换：

- * 计划内切换 *

此切换由系统管理员使用 System Manager 启动。通过计划内切换，本地集群的系统管理员可以切换控制权，以便远程集群的数据服务由本地集群处理。然后，远程集群位置的系统管理员可以对远程集群执行维护。

- * 计划外切换 *

在某些情况下、当MetroCluster集群关闭或集群之间的连接关闭时、ONTAP会自动启动切换、以便仍在运行的集群处理已关闭的集群的数据处理职责。

在其他情况下、如果ONTAP无法确定其中一个集群的状态、则正在工作的站点的系统管理员将启动切换、以控制另一站点的数据处理职责。

对于任何类型的切换操作步骤，都会使用 *switchback* 进程将数据服务功能返回到集群。

您执行的切换和切回过程取决于您的ONTAP版本：

- [在ONTAP 9.6或9.7中使用System Manager进行切换和切回](#)
- [在ONTAP 9.8或更高版本中使用System Manager进行切换和切回](#)

在ONTAP 9.6或9.7中使用System Manager进行切换和切回

步骤

1. 登录到ONTAP 9.6或9.7中的System Manager。
2. 单击 * (返回经典版本) *。
3. 单击 * 配置 > MetroCluster *。

System Manager 将验证是否可以协商切换。

4. 验证过程完成后，执行以下子步骤之一：
 - a. 如果验证失败，但站点 B 已启动，则表示发生错误。例如，子系统可能存在问题，或者 NVRAM 镜像可能未同步。

- i. 修复导致此错误的问题描述，单击 * 关闭 *，然后在步骤 2 中重新开始。
 - ii. 暂停站点B节点，单击*Close*，然后执行中的步骤 ["执行计划外切换"](#)。
 - b. 如果验证失败、并且站点B已关闭、则很可能出现连接问题。验证站点B是否已关闭，然后执行中的步骤["执行计划外切换"](#)。
5. 单击 * 从站点 B 切换到站点 A * 以启动切换过程。
 6. 单击 * 切换到新体验 *。

在ONTAP 9.8或更高版本中使用System Manager进行切换和切回

执行计划内切换(ONTAP 9.8或更高版本)

步骤

1. 在ONTAP 9.8或更高版本中登录到System Manager。
2. 选择 * 信息板 *。在 * MetroCluster 集群 * 部分中，两个集群显示有一个连接。
3. 在本地集群(如左侧所示)中，单击 ，然后选择*将远程数据服务切换到本地站点*。

验证切换请求后、控制权将从远程站点传输到本地站点。本地站点会为两个集群执行数据服务请求。

远程集群重新启动，但存储组件未处于活动状态，并且集群不会处理数据请求。现在，它可用于计划内维护。



在执行切回之前，不应使用远程集群进行数据服务。

执行计划外切换(ONTAP 9.8或更高版本)

ONTAP 可能会自动启动计划外切换。如果 ONTAP 无法确定是否需要切回，则仍在运行的 MetroCluster 站点的系统管理员将通过以下步骤启动切换：

步骤

1. 在ONTAP 9.8或更高版本中登录到System Manager。
2. 选择 * 信息板 *。

在* MetroCluster *部分中，两个集群之间的连接显示有一个"X"。这意味着无法检测到连接、并且连接或集群已关闭。

3. 在本地集群(如左侧所示)中，单击 ，然后选择*将远程数据服务切换到本地站点*。

如果切换失败并出现错误、请单击错误消息中的"查看详细信息"链接并确认计划外切换。

验证切换请求后、控制权将从远程站点传输到本地站点。本地站点会为两个集群执行数据服务请求。

必须先修复集群，然后才能使其重新联机。



远程集群联机后、在执行切回之前、不应使用它来提供数据服务。

执行切回(ONTAP 9.8或更高版本)

开始之前

如果远程集群因计划内维护或灾难而关闭、则它现在应已启动且正在运行、并等待切回。

步骤

1. 在本地集群上、登录到ONTAP 9.8或更高版本中的System Manager。
2. 选择 * 信息板 *。

在 * MetroCluster 集群 * 部分中，显示了两个集群。

3. 在本地集群(如左侧所示)中，单击 ，然后选择*收回控制权*。

首先、会_heed-heALed.以验证数据是否已在两个集群之间同步和镜像。

4. 数据修复完成后，单击 ，然后选择*Initiate switchback*。

切回完成后，两个集群均处于活动状态并为数据请求提供服务。此外、还会在集群之间镜像和同步数据。

监控 MetroCluster 配置

您可以使用 ONTAP MetroCluster 命令和 Active IQ Unified Manager（以前称为 OnCommand Unified Manager）监控各种软件组件的运行状况以及 MetroCluster 操作的状态。

正在检查 MetroCluster 配置

您可以检查 MetroCluster 配置中的组件和关系是否工作正常。您应在初始配置后以及对 MetroCluster 配置进行任何更改后执行检查。您还应在协商（计划内）切换或切回操作之前执行检查。

关于此任务

如果在任一集群或同时在这两个集群上短时间内发出 MetroCluster check run 命令两次，则可能发生冲突，并且此命令可能无法收集所有数据。后续的 MetroCluster check show 命令不会显示预期输出。

步骤

1. 检查配置：

```
MetroCluster check run
```

此命令作为后台作业运行，可能无法立即完成。

```
cluster_A:> metrocluster check run
```

The operation has been started and is running in the background. Wait for it to complete and run "metrocluster check show" to view the results. To check the status of the running metrocluster check operation, use the command,

```
"metrocluster operation history show -job-id 2245"
```

2. 显示最近的 MetroCluster check run 命令的更详细结果:

```
MetroCluster check aggregate show
```

```
MetroCluster check cluster show
```

```
MetroCluster check config-replication show
```

```
MetroCluster check lif show
```

```
MetroCluster check node show
```

MetroCluster check show 命令可显示最新的 MetroCluster check run 命令的结果。在使用 MetroCluster check show 命令之前，应始终运行 MetroCluster check run 命令，以使显示的信息为最新信息。

以下示例显示了运行正常的四节点 MetroCluster 配置的 MetroCluster check aggregate show 命令输出：

```

cluster_A::> metrocluster check aggregate show

Last Checked On: 8/5/2014 00:42:58

Node                Aggregate                Check
Result
-----
controller_A_1      controller_A_1_aggr0
ok                  mirroring-status
ok                  disk-pool-allocation
ok                  ownership-state
controller_A_1      controller_A_1_aggr1
ok                  mirroring-status
ok                  disk-pool-allocation

```

```

ok
ownership-state
ok
controller_A_1_aggr2
mirroring-status
ok
disk-pool-allocation
ok
ownership-state
ok

controller_A_2      controller_A_2_aggr0
mirroring-status
ok
disk-pool-allocation
ok
ownership-state
ok
controller_A_2_aggr1
mirroring-status
ok
disk-pool-allocation
ok
ownership-state
ok
controller_A_2_aggr2
mirroring-status
ok
disk-pool-allocation
ok
ownership-state

18 entries were displayed.

```

以下示例显示了运行正常的四节点 MetroCluster 配置的 MetroCluster check cluster show 命令输出。它表示集群已准备好在必要时执行协商切换。

Last Checked On: 9/13/2017 20:47:04

Cluster	Check	Result
-----	-----	-----
mccint-fas9000-0102	negotiated-switchover-ready	not-applicable
	switchback-ready	not-applicable
	job-schedules	ok
	licenses	ok
	periodic-check-enabled	ok
mccint-fas9000-0304	negotiated-switchover-ready	not-applicable
	switchback-ready	not-applicable
	job-schedules	ok
	licenses	ok
	periodic-check-enabled	ok

10 entries were displayed.

用于检查和监控 **MetroCluster** 配置的命令

您可以使用特定的 ONTAP 命令来监控 MetroCluster 配置和检查 MetroCluster 操作。

用于检查 **MetroCluster** 操作的命令

如果您要 ...	使用此命令 ...
对 MetroCluster 操作执行检查。 注： * 不应将此命令用作执行灾难恢复前操作系统验证的唯一命令。	<code>MetroCluster check run</code>
查看上次检查 MetroCluster 操作的结果。	<code>MetroCluster show</code>
查看站点间配置复制检查的结果。	<code>MetroCluster check config-replication</code> <code>show MetroCluster check config-replication show-aggregate-eligibility</code>
查看节点配置检查的结果。	<code>MetroCluster check node show</code>
查看对聚合配置的检查结果。	<code>MetroCluster check aggregate show</code>
查看 MetroCluster 配置中的 LIF 故障。	<code>MetroCluster check lif show</code>

用于监控 **MetroCluster** 互连的命令

如果您要 ...	使用此命令 ...
显示集群中 MetroCluster 节点的 HA 和 DR 镜像状态以及信息。	<code>MetroCluster 互连镜像显示</code>

用于监控 **MetroCluster SVM** 的命令

如果您要 ...	使用此命令 ...
查看 MetroCluster 配置中两个站点中的所有 SVM。	<code>MetroCluster SVM show</code>

使用 **MetroCluster Tiebreaker** 或 **ONTAP** 调解器监控配置

请参见 ["ONTAP 调解器与 MetroCluster Tiebreaker 之间的区别"](#) 了解这两种监控 MetroCluster 配置和启动自动切换的方法之间的区别。

使用以下链接安装和配置 Tiebreaker 或调解器：

- ["安装和配置 MetroCluster Tiebreaker 软件："](#)
- ["准备安装 ONTAP 调解器"](#)

NetApp MetroCluster Tiebreaker 软件如何检测故障

Tiebreaker 软件位于 Linux 主机上。只有当您要从第三个站点监控两个集群及其连接状态时，才需要 Tiebreaker 软件。这样可以使集群中的每个配对节点区分站点间链路关闭时的 ISL 故障与站点故障。

在 Linux 主机上安装 Tiebreaker 软件后，您可以在 MetroCluster 配置中配置集群以监控灾难情况。

Tiebreaker 软件如何检测站点间连接故障

如果站点之间的所有连接都丢失， MetroCluster Tiebreaker 软件将向您发出警报。

网络路径的类型

根据配置的不同， MetroCluster 配置中的两个集群之间有三种类型的网络路径：

- * FC 网络（位于光纤连接的 MetroCluster 配置中） *

此类网络由两个冗余 FC 交换机网络结构组成。每个交换机网络结构都有两个 FC 交换机，每个交换机网络结构有一个交换机与一个集群共存。每个集群都有两个 FC 交换机，每个交换机网络结构一个。所有节点都与每个主机代管 IP 交换机建立了 FC （ NV 互连和 FCP 启动程序）连接。数据通过 ISL 从集群复制到集群。

- * 集群间对等网络 *

此类网络由两个集群之间的冗余 IP 网络路径组成。集群对等网络可提供镜像 Storage Virtual Machine （ SVM ） 配置所需的连接。一个集群上所有 SVM 的配置都会由配对集群进行镜像。

- * IP 网络（存在于 MetroCluster IP 配置中） *

此类网络由两个冗余 IP 交换机网络组成。每个网络都有两个 IP 交换机，每个交换机网络结构有一个交换机与一个集群共存。每个集群都有两个 IP 交换机，每个交换机网络结构一个。所有节点均可连接到每个主机代管 FC 交换机。数据通过 ISL 从集群复制到集群。

监控站点间连接

Tiebreaker 软件会定期从节点检索站点间连接的状态。如果 NV 互连连接丢失，并且集群间对等不响应 ping，则集群会假定站点已隔离，Tiebreaker 软件会触发警报 "AllLinksSevered"。如果某个集群发现 "AllLinksSevered" 状态，而另一个集群无法通过网络访问，则 Tiebreaker 软件将触发警报 "disaster"。

Tiebreaker 软件如何检测站点故障

NetApp MetroCluster Tiebreaker 软件会检查 MetroCluster 配置和集群中节点的可访问性，以确定是否发生了站点故障。在某些情况下，Tiebreaker 软件还会触发警报。

Tiebreaker 软件监控的组件

Tiebreaker 软件可通过 IP 网络上托管的节点管理 LIF 和集群管理 LIF 的多条路径建立冗余连接，从而监控 MetroCluster 配置中的每个控制器。

Tiebreaker 软件可监控 MetroCluster 配置中的以下组件：

- 通过本地节点接口连接的节点
- 通过集群指定的接口进行集群
- 正常运行的集群，用于评估它是否与灾难站点（NV 互连，存储和集群间对等）建立连接

如果 Tiebreaker 软件与集群中的所有节点之间以及与集群本身之间的连接断开，Tiebreaker 软件将将此集群声明为 "无法访问"。检测到连接故障大约需要三到五秒。如果无法从 Tiebreaker 软件访问某个集群，则正常运行的集群（仍可访问的集群）必须指示与配对集群的所有链路都已切断，然后 Tiebreaker 软件才会触发警报。



如果正常运行的集群无法再通过 FC（NV 互连和存储）和集群间对等与灾难站点上的集群通信，则所有链路都将切断。

Tiebreaker 软件触发警报的故障情形

如果灾难站点上的集群（所有节点）已关闭或无法访问，并且正常运行的站点上的集群指示 "AllLinksSeved" 状态，则 Tiebreaker 软件将触发警报。

在以下情况下，Tiebreaker 软件不会触发警报（或警报被否决）：

- 在八节点 MetroCluster 配置中，如果灾难站点上的一个 HA 对已关闭
- 在灾难站点上的所有节点均已关闭的集群中，运行正常的站点上的一个 HA 对已关闭，而运行正常的站点上的集群指示 "AllLinksSeved" 状态

Tiebreaker 软件会触发警报，但 ONTAP 会否决此警报。在这种情况下，手动切换也会被否决

- 如果 Tiebreaker 软件至少可以访问灾难站点上的一个节点或集群接口，或者正常运行的站点仍然可以通过 FC（NV 互连和存储）或集群间对等访问灾难站点上的任一节点，则可以执行此操作

["了解ONTAP调解器如何在MetroCluster IP配置中支持自动计划外切换"](#)。

使用 NVFAIL 监控和保护文件系统一致性

使用 `volume modify` 命令的 `-nvfail` 参数，ONTAP 可以在系统启动时或在执行切换操作后检测非易失性 RAM（NVRAM）的不一致性。此外，它还会向您发出警告，并保护系统免受数据访问和修改的影响，直到可以手动恢复卷为止。

如果 ONTAP 检测到任何问题，数据库或文件系统实例将停止响应或关闭。然后，ONTAP 会向控制台发送错误消息，提醒您检查数据库或文件系统的状态。您可以通过 NVFAIL 向数据库管理员发出警告，指出集群节点之间的 NVRAM 不一致可能会影响数据库有效性。

在故障转移或启动恢复期间 NVRAM 数据丢失后，NFS 客户端无法访问任何节点中的数据，直到清除 NVFAIL 状态为止。CIFS 客户端不受影响。

NVFAIL 如何影响对 NFS 卷或 LUN 的访问

如果 ONTAP 在启动时检测到 NVRAM 错误，发生 MetroCluster 切换操作或在卷上设置了 NVFAIL 选项的 HA 接管操作期间检测到 NVRAM 错误，则会设置 NVFAIL 状态。如果在启动时未检测到错误，则文件服务将正常启动。但是，如果检测到 NVRAM 错误或在发生灾难切换时强制执行 NVFAIL 处理，则 ONTAP 会停止数据库实例响应。

启用 NVFAIL 选项后，下表中所述的过程之一将在启动期间发生：

条件	那么 ...
ONTAP 未检测到 NVRAM 错误	文件服务正常启动。
ONTAP 检测到 NVRAM 错误	• ONTAP 会向尝试访问数据库的 NFS 客户端返回陈旧文件句柄（ESTALE）错误，从而导致应用程序停止响应，崩溃或关闭。 然后，ONTAP 会向系统控制台和日志文件发送一条错误消息。 • 应用程序重新启动后，即使尚未验证文件是否有效，CIFS 客户端也可以使用这些文件。 对于 NFS 客户端，在受影响的卷上重置 <code>in-nvfailed-state</code> 选项之前，文件仍不可访问。

如果使用以下参数之一： • 已设置 <code>dr-force-nvfail volume</code> 选项 • 已设置 <code>force-nvfail-all switchover</code> 命令选项。	如果管理员不希望为将来可能的灾难切换操作强制执行 NVFAIL 处理，则可以在切换后取消设置 <code>dr-force-nvfail</code> 选项。对于 NFS 客户端，在受影响的卷上重置 <code>in-nvfailed-state</code> 选项之前，文件仍不可访问。  使用 <code>force-nvfail-all</code> 选项会导致在灾难切换期间处理的所有 DR 卷上设置 <code>dr-force-nvfail</code> 选项。
ONTAP 在包含 LUN 的卷上检测到 NVRAM 错误	该卷中的 LUN 将脱机。必须清除卷上的 <code>in-nvfailed-state</code> 选项，并且必须通过使受影响卷中的每个 LUN 联机来清除 LUN 上的 NVFAIL 属性。您可以执行以下步骤来检查 LUN 的完整性，并根据需要从 Snapshot 副本或备份中恢复 LUN。恢复卷中的所有 LUN 后，受影响卷上的 <code>in-nvfailed-state</code> 选项将被清除。

用于监控数据丢失事件的命令

如果启用了 NVFAIL 选项，则在因 NVRAM 不一致或发生 MetroCluster 切换而导致系统崩溃时，您会收到通知。

默认情况下，NVFAIL 参数未启用。

如果您要 ...	使用此命令 ...
创建启用了 NVFAIL 的新卷	<code>volume create -nvfail on</code>
在现有卷上启用 NVFAIL	<code>volume modify</code> • 注：* 要在创建的卷上启用 NVFAIL，请将 <code>-nvfail</code> 选项设置为 "on"。
显示当前是否已为指定卷启用 NVFAIL	<code>volume show</code> • 注：* 您将 <code>`字段`</code> 参数设置为 "nvfail" 以显示指定卷的 NVFAIL 属性。

相关信息

有关详细信息，请参见每个命令的手册页。

在切换后访问处于 NVFAIL 状态的卷

切换后，您必须通过重置 `volume modify` 命令的 ``in-nvfailed-state`` 参数来清除 NVFAIL 状态，以取消客户端访问数据的限制。

开始之前

数据库或文件系统不得正在运行或尝试访问受影响的卷。

关于此任务

设置 `in-nvfailed-state` 参数需要高级权限。

步骤

1. 使用 `volume modify` 命令并将 `-in-nvfailed-state` 参数设置为 `false`，以恢复卷。

完成后

有关检查数据库文件有效性的说明，请参见特定数据库软件的文档。

如果数据库使用 LUN，请查看相关步骤，以便在 NVRAM 出现故障后使主机可以访问这些 LUN。

相关信息

["使用 NVFAIL 监控和保护文件系统一致性"](#)

在切换后恢复处于 NVFAIL 状态的 LUN

切换后，主机将无法再访问处于 NVFAIL 状态的 LUN 上的数据。在数据库能够访问 LUN 之前，您必须执行许多操作。

开始之前

数据库不得正在运行。

步骤

1. 通过重置 `volume modify` 命令的 `in-nvfailed-state` 参数，清除托管 LUN 的受影响卷上的 NVFAIL 状态。
2. 使受影响的 LUN 联机。
3. 检查 LUN 是否存在任何数据不一致问题并予以解决。

这可能涉及使用 SnapRestore 在存储控制器上执行基于主机的恢复或恢复。

4. 恢复 LUN 后使数据库应用程序联机。

从何处查找追加信息

您可以了解有关 MetroCluster 配置和操作的更多信息。

MetroCluster 和其他信息

信息	主题
"MetroCluster 文档"	• 所有 MetroCluster 信息
"NetApp 技术报告 4375：《适用于 ONTAP 9.3 的 NetApp MetroCluster》"	• MetroCluster 配置和操作的技术概述。 • MetroCluster 配置最佳实践。

" 光纤连接的 MetroCluster 安装和配置 "	• 光纤连接的 MetroCluster 架构 • 为配置布线 • 配置 FC-SAS 网桥 • 配置 FC 交换机 • 在 ONTAP 中配置 MetroCluster
" 延伸型 MetroCluster 安装和配置 "	• 延伸型 MetroCluster 架构 • 为配置布线 • 配置 FC-SAS 网桥 • 在 ONTAP 中配置 MetroCluster
" MetroCluster IP 安装和配置 "	• MetroCluster IP 架构 • 为配置布线 • 在 ONTAP 中配置 MetroCluster
" MetroCluster Tiebreaker 1.21 软件安装和配置 "	• 使用 MetroCluster Tiebreaker 软件监控 MetroCluster 配置
Active IQ Unified Manager 文档 " NetApp 文档：产品指南和资源 "	• 监控 MetroCluster 配置和性能
" 基于副本的过渡 "	• 将数据从 7- 模式存储系统过渡到集群模式存储系统

版权信息

版权所有 © 2026 NetApp, Inc.。保留所有权利。中国印刷。未经版权所有者事先书面许可，本文档中受版权保护的任何部分不得以任何形式或通过任何手段（图片、电子或机械方式，包括影印、录音、录像或存储在电子检索系统中）进行复制。

从受版权保护的 NetApp 资料派生的软件受以下许可和免责声明的约束：

本软件由 NetApp 按“原样”提供，不含任何明示或暗示担保，包括但不限于适销性以及针对特定用途的适用性的隐含担保，特此声明不承担任何责任。在任何情况下，对于因使用本软件而以任何方式造成的任何直接性、间接性、偶然性、特殊性、惩罚性或后果性损失（包括但不限于购买替代商品或服务；使用、数据或利润方面的损失；或者业务中断），无论原因如何以及基于何种责任理论，无论出于合同、严格责任或侵权行为（包括疏忽或其他行为），NetApp 均不承担责任，即使已被告知存在上述损失的可能性。

NetApp 保留在不另行通知的情况下随时对本文档所述的任何产品进行更改的权利。除非 NetApp 以书面形式明确同意，否则 NetApp 不承担因使用本文档所述产品而产生的任何责任或义务。使用或购买本产品不表示获得 NetApp 的任何专利权、商标权或任何其他知识产权许可。

本手册中描述的产品可能受一项或多项美国专利、外国专利或正在申请的专利的保护。

有限权利说明：政府使用、复制或公开本文档受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中“技术数据权利 — 非商用”条款第 (b)(3) 条规定的限制条件的约束。

本文档中所含数据与商业产品和/或商业服务（定义见 FAR 2.101）相关，属于 NetApp, Inc. 的专有信息。根据本协议提供的所有 NetApp 技术数据和计算机软件具有商业性质，并完全由私人出资开发。美国政府对这些数据的使用权具有非排他性、全球性、受限且不可撤销的许可，该许可既不可转让，也不可再许可，但仅限在与交付数据所依据的美国政府合同有关且受合同支持的情况下使用。除本文档规定的情形外，未经 NetApp, Inc. 事先书面批准，不得使用、披露、复制、修改、操作或显示这些数据。美国政府对国防部的授权仅限于 DFARS 的第 252.227-7015(b)（2014 年 2 月）条款中明确的权利。

商标信息

NetApp、NetApp 标识和 <http://www.netapp.com/TM> 上所列的商标是 NetApp, Inc. 的商标。其他公司和产品名称可能是其各自所有者的商标。