



## 配置 MetroCluster

### ONTAP MetroCluster

Thom Illingworth, Zachary Wambold, Megan Bock  
August 31, 2021

# 目 录

配置 MetroCluster	1
商的切换	1
修和手切回	2
在源中断后操作	6
在失一个存架后操作	7

## □□ MetroCluster 配置

□可以□□故障情形，以□□ MetroCluster 配置是否正常□行。

### □□□商的切□

□可以□□□商（□□内）切□操作，以□□无中断数据可用性。

此□□通□将集群切□到第二个数据中心来□□数据可用性不受影□（ Microsoft 服□器消息□（ SMB ）和 Solaris 光□通道□□除外）。

此□□需要大□ 30 分□。

此操作□□具有以下□期□果：

- `MetroCluster switchover` 命令将□示警告提示符。

如果□提示符回答 `yes`， □□出命令的站点将切□配□站点。

□于 MetroCluster IP 配置：

- □于 ONTAP 9.4 及更早版本：
  - 在□商切□后， □像聚合将降□。
- □于 ONTAP 9.5 及更高版本：
  - 如果可以□□□程存□， □□像聚合将保持正常状□。
  - 如果无法□□□程存□， □在□商切□后， □像聚合将降□。
- □于 ONTAP 9.8 及更高版本：
  - 如果无法□□□程存□， □位于□□站点的未□像聚合将不可用。 □可能会□致控制器中断。

□□

1. □□所有□点均□于已配置状□和正常模式：

```
MetroCluster node show
```

```
cluster_A::> metrocluster node show

Cluster                               Configuration State      Mode
-----
Local: cluster_A                       configured                normal
Remote: cluster_B                       configured                normal
```

2. □始切□操作：

### MetroCluster switchover

```
cluster_A::> metrocluster switchover
Warning: negotiated switchover is about to start. It will stop all the
data Vservers on cluster "cluster_B" and
automatically re-start them on cluster "cluster_A". It will finally
gracefully shutdown cluster "cluster_B".
```

### 3. 本地集群于已配置状态和切换模式：

#### MetroCluster node show

```
cluster_A::> metrocluster node show

Cluster                               Configuration State      Mode
-----                               -
-----
Local: cluster_A                      configured                switchover
Remote: cluster_B                     not-reachable            -
                                     configured                normal
```

### 4. 切换操作已成功：

#### MetroCluster 操作显示

```
cluster_A::> metrocluster operation show

cluster_A::> metrocluster operation show
  Operation: switchover
    State: successful
  Start Time: 2/6/2016 13:28:50
  End Time: 2/6/2016 13:29:41
  Errors: -
```

### 5. 使用 `vserver show` 和 `network interface show` 命令恢复 SVM 和 LIF 是否已联机。

## 修复和手动切换回

可以通过在切换后将集群切换回原始数据中心来修复和手动切换回操作，以确保数据可用性不受影响（SMB 和 Solaris FC 配置除外）。

此操作需要大约 30 分钟。

此操作的结果是，将服务器切换回其主点。

在运行 ONTAP 9.5 或更高版本的系统上不需要执行修复操作，而是在商切后自行完成其修复。在运行 ONTAP 9.6 及更高版本的系统上，也会在商切后自行完成修复。

□□

1. 如果系统是运行的是 ONTAP 9.4 或更早版本，修复数据聚合：

MetroCluster 修复聚合`

以下示例显示了命令的成功完成：

```
cluster_A::> metrocluster heal aggregates
[Job 936] Job succeeded: Heal Aggregates is successful.
```

2. 如果系统是运行的是 ONTAP 9.4 或更早版本，修复根聚合：

MetroCluster 修复根聚合`

以下配置需要执行此操作：

- MetroCluster FC 配置。
- 运行 ONTAP 9.4 或更早版本的 MetroCluster IP 配置。以下示例显示了命令的成功完成：

```
cluster_A::> metrocluster heal root-aggregates
[Job 937] Job succeeded: Heal Root Aggregates is successful.
```

3. 修复是否已完成：

MetroCluster node show

以下示例显示了命令的成功完成：

```
cluster_A::> metrocluster node show
DR                               Configuration  DR
Group Cluster Node                State           Mirroring Mode
-----
1      cluster_A
      node_A_1          configured     enabled      heal roots
completed
      cluster_B
      node_B_2          unreachable   -            switched over
42 entries were displayed.metrocluster operation show
```

如果自行修复操作因任何原因失败，必须 MetroCluster 按照 ONTAP 9.5 之前的 ONTAP 版本中的手册进行操作描述 heal` 命令。可以使用 MetroCluster operation show` 和 MetroCluster operation history show -instance` 命令监控修复状态并确定故障的生成原因。

4. 所有聚合是否均已镜像：

查看聚合显示

以下示例显示所有聚合的 RAID 状态均已镜像：

```
cluster_A::> storage aggregate show
cluster Aggregates:
Aggregate Size      Available Used% State  #Vols  Nodes      RAID
Status
-----
data_cluster
      4.19TB      4.13TB    2% online    8 node_A_1  raid_dp,
mirrored,
normal

root_cluster
      715.5GB    212.7GB   70% online    1 node_A_1  raid4,
mirrored,
normal

cluster_B Switched Over Aggregates:
Aggregate Size      Available Used% State  #Vols  Nodes      RAID
Status
-----
data_cluster_B
      4.19TB      4.11TB    2% online    5 node_A_1  raid_dp,
mirrored,
normal

root_cluster_B    -          -        - unknown    - node_A_1  -
```

5. 从站点点。

6. 切换回恢复的状态：

MetroCluster node show

```

cluster_A::> metrocluster node show
DR
Group Cluster Node          Configuration  DR
State          Mirroring Mode
-----
1      cluster_A
      node_A_1      configured    enabled    heal roots
completed
      cluster_B
      node_B_2      configured    enabled    waiting for
switchback                                           recovery
2 entries were displayed.

```

7. 行切回：

MetroCluster 切回

```

cluster_A::> metrocluster switchback
[Job 938] Job succeeded: Switchback is successful.Verify switchback

```

8. 点的状：

MetroCluster node show

```

cluster_A::> metrocluster node show
DR
Group Cluster Node          Configuration  DR
State          Mirroring Mode
-----
1      cluster_A
      node_A_1      configured    enabled    normal
      cluster_B
      node_B_2      configured    enabled    normal
2 entries were displayed.

```

9. MetroCluster 操作的状：

MetroCluster 操作示

出示成功状。

```
cluster_A::> metrocluster operation show
Operation: switchback
State: successful
Start Time: 2/6/2016 13:54:25
End Time: 2/6/2016 13:56:15
Errors: -
```

## 在 P 源 P 中断后 P P 操作

P 可以 P P MetroCluster 配置 P PDU 故障的 P P。

最佳做法是，将 P 件中的 P 个 P 源 P P ( PSU ) P 接到 P 独的 P 源。如果 P 个 PSU 都 P 接到同一个配 P P 元 ( PDU )，并且 P 生 P 气中断，P 站点可能会 P P，或者整个磁 P 架可能不可用。P P 一条 P 源 P 故障，以 P P 没有布 P 不匹配，从而 P 生原因可能 P 致服 P 中断。

此 P P 需要大 P 15 分 P。

此 P P 需要 P P 所有左 P PDU 的 P 源，然后 P P 包含 MetroCluster P 件的所有机架上的所有右 P PDU 的 P 源。

此操作 P P 具有以下 P 期 P 果：

- 当 PDU 断 P P 接 P，P 生成 P P。
- 不 P P 生故障 P 移或服 P P 失。

P P

1. P P 包含 MetroCluster P 件的机架左 P PDU 的 P 源。
2. 在控制台上 P 控 P 果：

```
s系 P P 境 P 感器 P 示 -state fault
```

```
s存 P 架 show -errors
```



```

cluster_A::> system environment sensors show -state fault

Node Sensor                State Value/Units Crit-Low Warn-Low Warn-Hi
Crit-Hi
-----
node_A_1
    PSU1                    fault
                               PSU_OFF
    PSU1 Pwr In OK          fault
                               FAULT
node_A_2
    PSU1                    fault
                               PSU_OFF
    PSU1 Pwr In OK          fault
                               FAULT

4 entries were displayed.

cluster_A::> storage shelf show -errors
    Shelf Name: 1.1
    Shelf UID: 50:0a:09:80:03:6c:44:d5
    Serial Number: SHFHU1443000059

Error Type                Description
-----
Power                      Critical condition is detected in storage shelf
power supply unit "1". The unit might fail.Reconnect PSU1

```

3. 重新打开左 PDU 的电源。
4. 保持 ONTAP 清除故障情况。
5. 对右 PDU 重复上述操作。

## 在丢失一个存储架后操作

可以清除存储架的故障，以确认是否没有故障。

此操作具有以下预期结果：

- 控制件报告消息。
- 不产生故障移除或服务丢失。
- 硬件故障恢复后，像重新同步将自。

□□

1. 清除故障移除状态：

## s存故障移示

```
cluster_A::> storage failover show
```

Node	Partner	Possible	State Description
node_A_1	node_A_2	true	Connected to node_A_2
node_A_2	node_A_1	true	Connected to node_A_1

2 entries were displayed.

## 2. 聚合状 :

## s存聚合示

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

```
cluster Aggregates:
```

Aggregate	Size	Available	Used%	State	#Vols	Nodes	RAID
node_A_1data01_mirrored	4.15TB	3.40TB	18%	online	3	node_A_1	
raid_dp,							
mirrored,							
normal							
node_A_1root	707.7GB	34.29GB	95%	online	1	node_A_1	
raid_dp,							
mirrored,							
normal							
node_A_2_data01_mirrored	4.15TB	4.12TB	1%	online	2	node_A_2	
raid_dp,							
mirrored,							
normal							
node_A_2_data02_unmirrored	2.18TB	2.18TB	0%	online	1	node_A_2	
raid_dp,							
normal							
node_A_2_root	707.7GB	34.27GB	95%	online	1	node_A_2	
raid_dp,							
mirrored,							
normal							

3. 所有数据 SVM 和数据卷是否均已机并提供数据：

```
vserver show -type data
```

```
network interface show -fields is-home false
```

```
volume show ! vol0 , ! mdv*
```

```
cluster_A::> vserver show -type data
Vserver      Type      Subtype      Admin      Operational  Root
Aggregate
-----
SVM1         data      sync-source      running      SVM1_root
node_A_1_data01_mirrored
SVM2         data      sync-source      running      SVM2_root
node_A_2_data01_mirrored
```

```
cluster_A::> network interface show -fields is-home false
There are no entries matching your query.
```

```
cluster_A::> volume show !vol0,!MDV*
```

```
Vserver      Volume      Aggregate      State      Type      Size
Available Used%
-----
SVM1
          SVM1_root
                    node_A_1data01_mirrored
                    online      RW      10GB
9.50GB      5%
SVM1
          SVM1_data_vol
                    node_A_1data01_mirrored
                    online      RW      10GB
9.49GB      5%
SVM2
          SVM2_root
                    node_A_2_data01_mirrored
                    online      RW      10GB
9.49GB      5%
SVM2
          SVM2_data_vol
                    node_A_2_data02_unmirrored
                    online      RW      1GB
972.6MB      5%
```

4. 定池 1 中用于点 node\_A\_2 的磁架以源以模突然生的硬件故障：

```
storage aggregate show -r -node node-name ! * root
```

磁架必包含像数据聚合中的器。

在以下示例中，磁架 ID 31 失。

```
cluster_A::> storage aggregate show -r -node node_A_2 !*root
Owner Node: node_A_2
Aggregate: node_A_2_data01_mirrored (online, raid_dp, mirrored) (block
checksums)
Plex: /node_A_2_data01_mirrored/plex0 (online, normal, active, pool0)
RAID Group /node_A_2_data01_mirrored/plex0/rg0 (normal, block
checksums)

Physical
Position Disk Pool Type RPM Usable
Size Status
-----
-----
dparity 2.30.3 0 BSAS 7200 827.7GB
828.0GB (normal)
parity 2.30.4 0 BSAS 7200 827.7GB
828.0GB (normal)
data 2.30.6 0 BSAS 7200 827.7GB
828.0GB (normal)
data 2.30.8 0 BSAS 7200 827.7GB
828.0GB (normal)
data 2.30.5 0 BSAS 7200 827.7GB
828.0GB (normal)

Plex: /node_A_2_data01_mirrored/plex4 (online, normal, active, pool1)
RAID Group /node_A_2_data01_mirrored/plex4/rg0 (normal, block
checksums)

Physical
Position Disk Pool Type RPM Usable
Size Status
-----
-----
dparity 1.31.7 1 BSAS 7200 827.7GB
828.0GB (normal)
parity 1.31.6 1 BSAS 7200 827.7GB
828.0GB (normal)
data 1.31.3 1 BSAS 7200 827.7GB
828.0GB (normal)
data 1.31.4 1 BSAS 7200 827.7GB
```

```

828.0GB (normal)
  data      1.31.5          1   BSAS    7200   827.7GB
828.0GB (normal)

Aggregate: node_A_2_data02_unmirrored (online, raid_dp) (block
checksums)
  Plex: /node_A_2_data02_unmirrored/plex0 (online, normal, active,
pool0)
  RAID Group /node_A_2_data02_unmirrored/plex0/rg0 (normal, block
checksums)

Usable
Physical
  Position Disk          Pool Type    RPM    Size
Size Status
-----
-----
  dparity  2.30.12          0   BSAS    7200   827.7GB
828.0GB (normal)
  parity   2.30.22          0   BSAS    7200   827.7GB
828.0GB (normal)
  data     2.30.21          0   BSAS    7200   827.7GB
828.0GB (normal)
  data     2.30.20          0   BSAS    7200   827.7GB
828.0GB (normal)
  data     2.30.14          0   BSAS    7200   827.7GB
828.0GB (normal)
15 entries were displayed.

```

5. 物理磁盘定磁架的源。

6. 再次聚合状态：

s存聚合示

```
storage aggregate show -r -node node_A_2 ! * root
```

器位于已源架上的聚合具有 " 已降 " RAID 状，而受影上的器具有 " 故障 " 状，如以下示例所示：

```

cluster_A::> storage aggregate show
Aggregate      Size Available Used% State    #Vols  Nodes      RAID
Status
-----
-----
node_A_1data01_mirrored
      4.15TB    3.40TB    18% online      3 node_A_1
raid_dp,

```

```

mirrored,

normal
node_A_1root
      707.7GB   34.29GB   95% online      1 node_A_1
raid_dp,

```

```

mirrored,

normal
node_A_2_data01_mirrored
      4.15TB    4.12TB    1% online      2 node_A_2
raid_dp,

```

```

mirror

degraded
node_A_2_data02_unmirrored
      2.18TB    2.18TB    0% online      1 node_A_2
raid_dp,

```

```

normal
node_A_2_root
      707.7GB   34.27GB   95% online      1 node_A_2
raid_dp,

```

```

mirror

degraded
cluster_A::> storage aggregate show -r -node node_A_2 !*root
Owner Node: node_A_2
Aggregate: node_A_2_data01_mirrored (online, raid_dp, mirror degraded)
(block checksums)
Plex: /node_A_2_data01_mirrored/plex0 (online, normal, active, pool0)
RAID Group /node_A_2_data01_mirrored/plex0/rg0 (normal, block
checksums)

```

					Usable	
Physical	Position	Disk	Pool	Type	RPM	Size
Size	Status					
-----						
828.0GB	dparity	2.30.3	0	BSAS	7200	827.7GB
						(normal)
	parity	2.30.4	0	BSAS	7200	827.7GB

```

828.0GB (normal)
  data      2.30.6          0   BSAS    7200   827.7GB
828.0GB (normal)
  data      2.30.8          0   BSAS    7200   827.7GB
828.0GB (normal)
  data      2.30.5          0   BSAS    7200   827.7GB
828.0GB (normal)

```

Plex: /node\_A\_2\_data01\_mirrored/plex4 (offline, failed, inactive, pool1)

RAID Group /node\_A\_2\_data01\_mirrored/plex4/rg0 (partial, none checksums)

					Usable
Physical					
Position	Disk	Pool	Type	RPM	Size
Size	Status				
-----					
dparity	FAILED	-	-	-	827.7GB
- (failed)					
parity	FAILED	-	-	-	827.7GB
- (failed)					
data	FAILED	-	-	-	827.7GB
- (failed)					
data	FAILED	-	-	-	827.7GB
- (failed)					
data	FAILED	-	-	-	827.7GB
- (failed)					

Aggregate: node\_A\_2\_data02\_unmirrored (online, raid\_dp) (block checksums)

Plex: /node\_A\_2\_data02\_unmirrored/plex0 (online, normal, active, pool0)

RAID Group /node\_A\_2\_data02\_unmirrored/plex0/rg0 (normal, block checksums)

					Usable
Physical					
Position	Disk	Pool	Type	RPM	Size
Size	Status				
-----					
dparity	2.30.12	0	BSAS	7200	827.7GB
828.0GB (normal)					
parity	2.30.22	0	BSAS	7200	827.7GB
828.0GB (normal)					
data	2.30.21	0	BSAS	7200	827.7GB



```
828.0GB (normal)
  data      2.30.20          0   BSAS   7200  827.7GB
828.0GB (normal)
  data      2.30.14          0   BSAS   7200  827.7GB
828.0GB (normal)
15 entries were displayed.
```

7. 是否正在提供数据，以及所有卷是否仍处于机状态：

```
vserver show -type data
```

```
network interface show -fields is-home false
```

```
volume show ! vol0 , ! mdv*
```

```

cluster_A::> vserver show -type data

cluster_A::> vserver show -type data
Admin      Operational Root
Vserver    Type      Subtype    State      State      Volume
Aggregate
-----
-----
SVM1       data      sync-source      running     SVM1_root
node_A_1_data01_mirrored
SVM2       data      sync-source      running     SVM2_root
node_A_1_data01_mirrored

cluster_A::> network interface show -fields is-home false
There are no entries matching your query.

cluster_A::> volume show !vol0,!MDV*
Vserver    Volume      Aggregate    State      Type      Size
Available Used%
-----
-----
SVM1
      SVM1_root
      node_A_1data01_mirrored
      online      RW      10GB
9.50GB    5%
SVM1
      SVM1_data_vol
      node_A_1data01_mirrored
      online      RW      10GB
9.49GB    5%
SVM2
      SVM2_root
      node_A_1data01_mirrored
      online      RW      10GB
9.49GB    5%
SVM2
      SVM2_data_vol
      node_A_2_data02_unmirrored
      online      RW      1GB
972.6MB   5%

```

## 8. 物理磁架。

重新同将自。

9. 重新同步是否已完成：

存储聚合指示

受影响聚合的 RAID 状态为 "resyncing"，如以下示例所示：

```
cluster_A::> storage aggregate show
cluster Aggregates:
Aggregate      Size Available Used% State  #Vols  Nodes      RAID
Status
-----
node_A_1_data01_mirrored
      4.15TB      3.40TB      18% online    3 node_A_1
raid_dp,
mirrored,
normal
node_A_1_root
      707.7GB      34.29GB      95% online    1 node_A_1
raid_dp,
mirrored,
normal
node_A_2_data01_mirrored
      4.15TB      4.12TB       1% online    2 node_A_2
raid_dp,
resyncing
node_A_2_data02_unmirrored
      2.18TB      2.18TB       0% online    1 node_A_2
raid_dp,
normal
node_A_2_root
      707.7GB      34.27GB      95% online    1 node_A_2
raid_dp,
resyncing
```

10. 控制聚合以重新同步已完成：

存储聚合指示

受影响的聚合的 RAID 状态为 "Normal"，如以下示例所示：

```
cluster_A::> storage aggregate show
cluster Aggregates:
Aggregate      Size Available Used% State  #Vols  Nodes      RAID
Status
-----
node_A_1data01_mirrored
      4.15TB      3.40TB      18% online      3 node_A_1
raid_dp,
mirrored,
normal
node_A_1root
      707.7GB      34.29GB      95% online      1 node_A_1
raid_dp,
mirrored,
normal
node_A_2_data01_mirrored
      4.15TB      4.12TB      1% online      2 node_A_2
raid_dp,
normal
node_A_2_data02_unmirrored
      2.18TB      2.18TB      0% online      1 node_A_2
raid_dp,
normal
node_A_2_root
      707.7GB      34.27GB      95% online      1 node_A_2
raid_dp,
resyncing
```

## Copyright Information

Copyright © 2021 NetApp, Inc. All rights reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means-graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system- without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

RESTRICTED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.277-7103 (October 1988) and FAR 52-227-19 (June 1987).

## Trademark Information

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.