



# MetroCluster 構成をテストする

## ONTAP MetroCluster

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

# 目次

MetroCluster 構成をテストする .....	1
ネゴシエートスイッチオーバーを検証中 .....	1
修復と手動スイッチバックの検証 .....	3
電源回線切断後の動作確認 .....	6
単一のストレージシェルフが停止したあとの動作確認 .....	8

# MetroCluster 構成をテストする

障害シナリオをテストして、MetroCluster 構成が正しく動作していることを確認できます。

## ネゴシエートスイッチオーバーを検証中

ネゴシエート（計画的）スイッチオーバー処理をテストして、データが中断なく提供されることを確認できます。

このテストでは、クラスタを 2 つ目のデータセンターに切り替えてデータの可用性に影響がないことを検証します。ただし、Microsoft Server Message Block（SMB；サーバメッセージブロック）プロトコルおよび Solaris Fibre Channel プロトコルは除きます。

このテストには約 30 分かかります。

この手順の想定される結果は次のとおりです。

- MetroCluster switchover コマンドは ' 警告プロンプトを表示します  
プロンプトに「yes」と応答すると、コマンドが発行されたサイトがパートナーサイトを切り替えます。

MetroCluster IP 構成の場合：

- ONTAP 9.4 以前：
  - ミラーアグリゲートはネゴシエートスイッチオーバー後にデグレード状態になります。
- ONTAP 9.5 以降：
  - リモートストレージがアクセス可能である場合、ミラーされたアグリゲートは通常の状態のままになります。
  - リモートストレージへのアクセスが失われると、ネゴシエートスイッチオーバー後にミラーアグリゲートがデグレード状態になります。
- ONTAP 9.8 以降の場合：
  - リモートストレージへのアクセスが失われると、ディザスタサイトにあるミラーされていないアグリゲートは使用できなくなります。これにより、コントローラが停止する可能性があります。

手順

1. すべてのノードが設定済みの状態で通常モードになっていることを確認します。

```
MetroCluster node show
```

```
cluster_A::> metrocluster node show
```

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

## 2. スイッチオーバー処理を開始します。

### MetroCluster スイッチオーバー

```
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. ローカルクラスタが設定済みの状態でスイッチオーバーモードになっていることを確認します。

### 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. スイッチオーバー処理が成功したことを確認します。

「MetroCluster operation show」を参照してください

```
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. 「 vserver show 」コマンドと「 network interface show 」コマンドを使用して、DR SVM と LIF がオンラインになったことを確認します。

## 修復と手動スイッチバックの検証

修復処理と手動スイッチバック処理をテストするには、ネゴシエートスイッチオーバー後にクラスタを元のデータセンターにスイッチバックして、データの可用性に影響がないことを検証します（SMB および Solaris FC 構成を除く）。

このテストには約 30 分かかります。

この手順では、想定される結果として、サービスがホームノードにスイッチバックされることを確認します。

ONTAP 9.5 以降を実行しているシステムでは、ネゴシエートスイッチオーバー後に自動的に修復が実行されるため、修復の手順は必要ありません。ONTAP 9.6 以降を実行するシステムでは、計画外のスイッチオーバー後にも自動的に修復が実行されます。

### 手順

1. ONTAP 9.4 以前を実行しているシステムの場合は、データアグリゲートを修復します。

MetroCluster はアグリゲートを修復します

次の例は、コマンドが正常に完了したことを示しています。

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

2. ONTAP 9.4 以前を実行しているシステムの場合は、ルートアグリゲートを修復します。

MetroCluster はルートアグリゲートを修復します

この手順は、次の構成が必要です。

- MetroCluster FC 構成
- ONTAP 9.4 以前を実行している MetroCluster IP 構成次の例は、コマンドが正常に完了したことを示しています。

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

3. 修復が完了したことを確認します。

MetroCluster node show

次の例は、コマンドが正常に完了したことを示しています。

```

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

```

何らかの理由で自動修復操作が失敗した場合は、ONTAP 9.5 より前のバージョンの ONTAP の場合と同様に、MetroCluster heal コマンドを手動で問題する必要があります。修復のステータスを監視し、障害の原因を判別するには、MetroCluster operation show および MetroCluster operation history show -instance コマンドを使用します。

4. すべてのアグリゲートがミラーされたことを確認します。

「 storage aggregate show

次の例では、すべてのアグリゲートの RAID ステータスが mirrored になっています。

```

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. ディザスタサイトからノードをブートします。
6. スイッチバックリカバリのステータスを確認します。

#### 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. スイッチバックを実行します。

### MetroCluster スイッチバック

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

## 8. ノードのステータスを確認します。

### MetroCluster node show

```
cluster_A::> metrocluster node show
DR
Group Cluster Node Configuration State DR 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 処理のステータスを確認します。

「 MetroCluster operation show 」を参照してください

正常に完了したことを示す出力が表示されます。

```
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: -
```

## 電源回線切断後の動作確認

MetroCluster 構成で PDU の障害が発生した場合の対応をテストできます。

ベストプラクティスとして、コンポーネントの各電源装置（ PSU ）を別々の電源装置に接続することを推奨します。両方の PSU を同じ配電ユニット（ PDU ）に接続している場合、停電が発生したときに、サイトが停止したりシェルフ全体が使用できなくなったりする可能性があります。1 本原因の電源回線に障害が発生した場合に、サービスが中断する可能性があるケーブル接続の不一致がないかどうかをテストします。

このテストには約 15 分かかります。

このテストでは、MetroCluster コンポーネントが格納されたすべてのラックで、左側のすべての PDU の電源をオフにしてから、右側のすべての PDU の電源をオフにする必要があります。

この手順の想定される結果は次のとおりです。

- PDU の接続が切断されるため、エラーが生成されます。
- フェイルオーバーやサービスの中断は発生しません。

#### 手順

1. MetroCluster コンポーネントが格納されたラックの左側の PDU の電源をオフにします。
2. コンソールで結果を監視します。

「 system environment sensors show -state fault 」というエラーが表示されます

「 storage shelf show -errors 」を参照してください

```
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. 左側の PDU の電源を再度オンにします。
4. ONTAP がエラー状態をクリアすることを確認します。
5. 右側の PDU で上記の手順を繰り返します。

## 単一のストレージシェルフが停止したあとの動作確認

単一のストレージシェルフの障害をテストして、単一点障害がないことを確認できます。

この手順の想定される結果は次のとおりです。

- 監視ソフトウェアからエラーメッセージが報告されます。
- フェイルオーバーやサービスの中断は発生しません。
- ハードウェア障害がリストアされたあと、ミラーの再同期が自動的に開始されます。

### 手順

1. ストレージフェイルオーバーのステータスを確認します。

「 storage failover show 」をクリックします

```
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. アグリゲートのステータスを確認します。

「 storage aggregate show

```
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. すべてのデータ SVM とデータボリュームがオンラインで、データを提供していることを確認します。

```
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. ハードウェア障害をシミュレートするために電源をオフにするノード node\_A\_1 のプール 1 内のシェルフを特定します。

「 storage aggregate show -r -node node\_name !!\* root

ミラーされたデータアグリゲートを構成するドライブが格納されたシェルフを選択する必要があります。

次の例では、停止する対象としてシェルフ ID 31 を選択します。

```
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   Size
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   Size
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)

Physical
Position Disk Pool Type RPM Usable
Size Status Size
-----
-----
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. 選択したシェルフの電源を物理的にオフにします。
6. アグリゲートのステータスを再度確認します。

「 storage aggregate show

「 storage aggregate show -r -node node\_name ! \* root 」

次の例に示すように、電源をオフにしたシェルフにドライブがあるアグリゲートの RAID ステータスが「 degraded 」になり、影響を受けたプレックスのドライブのステータスが「 failed 」になります。

```

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				
-----					
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 (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. データが提供されていること、およびすべてのボリュームがオンラインのままであることを確認します。

```
vserver show -type data
```

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

```
volume show ! vol0、 ! MDV *
```

```

cluster_A::> vservers show -type data

cluster_A::> vservers 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. シェルフの電源を物理的にオンにします。

再同期が自動的に開始されます

9. 再同期が開始されたことを確認します。

「 storage aggregate show

次の例に示すように、影響を受けたアグリゲートの RAID ステータスが「resyncing」になります。

```
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. アグリゲートを監視して、再同期が完了したことを確認します。

「 storage aggregate show

次の例に示すように、影響を受けるアグリゲートの RAID ステータスが「 normal 」になっている必要があります。

```
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.