スイッチオーバー、修復、スイッチバックを実行

ONTAP MetroCluster

NetApp
April 19, 2024

This PDF was generated from https://docs.netapp.com/ja-jp/ontap-metrocluster/manage/task_perform_switchover_for_tests_or_maintenance.html on April 19, 2024. Always check docs.netapp.com for the latest.

目次

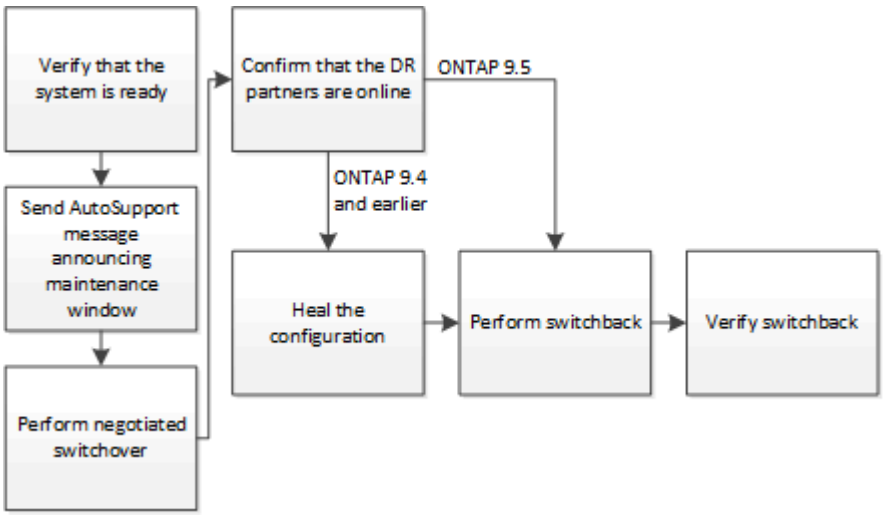
スイッチオーバー、修復、スイッチバックを実行 1
テストやメンテナンスのためにスイッチオーバーを実行 1
スイッチオーバー、修復、スイッチバック用のコマンドです 14
MetroCluster 設定を監視する 14
NVFAIL を使用したファイルシステム整合性の監視および保護 20
追加情報の参照先 23

スイッチオーバー、修復、スイッチバックを実行

テストやメンテナンスのためにスイッチオーバーを実行

テストまたはメンテナンスのためのスイッチオーバーの実行

MetroCluster 機能のテストや計画的メンテナンスを実行する場合は、一方のクラスタをパートナークラスタに完全にスイッチオーバーする、ネゴシエートスイッチオーバーを実行します。その後、構成を修復してスイッチバックします。



ONTAP 9.6 以降では、 ONTAP System Manager を使用して MetroCluster IP 構成でスイッチオーバーとスイッチバックの処理を実行できます。

スイッチオーバーに向けたシステムの事前チェック

--simulate オプションを使用して ' スイッチオーバー操作の結果をプレビューできますこれにより、処理を開始する前に、処理が成功するために必要なほとんどの前提条件が満たされていることを確認できます。問題：稼働したままになる、サイトからの次のコマンド。

1. 権限レベルを advanced に設定します。「 set -privilege advanced 」
2. 稼働したままになるサイトから ' スイッチオーバー操作をシミュレートします MetroCluster switchover-simulate
3. 返された出力を確認します。

   出力には、スイッチオーバー処理を阻止する拒否の有無が示されます。MetroCluster 処理を実行するたびに、処理が成功するための一連の条件を確認する必要があります。「 "" まで」は、 1 つ以上の条件が満たされない場合に操作を禁止するメカニズムです。拒否には「ソフト」と「ハード」の 2 種類があります。ソフトな拒否は無視できますが、ハードな拒否は無視できません。たとえば、 4 ノード MetroCluster 構成でネゴシエートスイッチオーバーを実行する場合、条件の 1 つとしてすべてのノードが正常に稼働していることがあります。たとえば、 1 つのノードが停止し、 HA パートナーにテイクオーバーされたとします。すべてのノードが正常に稼働している必要があるという条件はハードな条件であるため、スイッチオ

ーバー処理にはハードな拒否が適用されます。ハードな拒否は無視できません。

> 拒否は無視しないことを推奨します。

例：検証結果

次の例は、スイッチオーバー処理のシミュレーションで発生したエラーを示しています。

```
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.
```

> 障害が発生したすべてのディスクを交換するまで、ネゴシエートスイッチオーバーとスイッチバックは失敗します。障害が発生したディスクを交換したあとにディザスタリカバリを実行できます。障害が発生したディスクに関する警告を無視するには、ネゴシエートスイッチオーバーとスイッチバックにソフトな拒否を追加します。

カスタム **AutoSupport** メッセージの送信 **-** ネゴシエートスイッチオーバー前

ネゴシエートスイッチオーバーを実行する前に、 AutoSupport メッセージを問題で送信して、メンテナンスが進行中であることをネットアップテクニカルサポートに通知する必要があります。ネゴシエートスイッチオーバーでは、プレックスや MetroCluster の処理が失敗して AutoSupport メッセージがトリガーされることがあります。システム停止が発生したとみなしてテクニカルサポートがケースをオープンしないように、メンテナンスが進行中であることを通知する必要があります。

このタスクは MetroCluster サイトごとに実行する必要があります。

手順

1. Site_A のクラスタにログインします
2. メンテナンスの開始を示す AutoSupport メッセージを起動します。「 system node AutoSupport invoke -node * -type all -message MAINT= maintenance-window-in-hours

   maintenance-window-in-hours にはメンテナンス時間の長さを指定します。最大値は 72 時間です。この時間が経過する前にメンテナンスが完了した場合は、メンテナンス期間が終了したことを示すコマンドを問題に送信できます。「 system node AutoSupport invoke -node * -type all -message MAINT=end 」

3. 同じ手順をパートナーサイトでも実行します。

ネゴシエートスイッチオーバーの実行

ネゴシエートスイッチオーバーは、パートナーサイトのプロセスをクリーンシャットダ

ウンしてから、パートナーサイトの処理をスイッチオーバーします。ネゴシエートスイッチオーバーは、 MetroCluster サイトでのメンテナンスの実施、またはスイッチオーバー機能のテストに使用できます。

- スイッチバック処理を実行する前に、設定の変更がすべて完了している必要があります。

  これにより、ネゴシエートスイッチオーバーまたはスイッチバックの処理との競合が回避されます。

- 停止していたノードがブートし、クラスタクォーラムに参加している必要があります。

  クラスタ・クォーラムの詳細については ' 定足数と Epsilon についての項を参照してください

  "システム管理"

- 両方のサイトからクラスタピアリングネットワークにアクセスできる必要があります。
- MetroCluster 構成のすべてのノードで、同じバージョンの ONTAP ソフトウェアが実行されている必要があります。
- 新しい SnapMirror 関係を作成する前に、 MetroCluster 構成の両方のサイトで replication.create_data_protection_rels.enable オプションを ON に設定しておく必要があります。
- 2 ノード MetroCluster 構成の場合、サイト間で ONTAP のバージョンが一致していないときは、アップグレード時に新しい SnapMirror 関係を作成しないようにする必要があります。
- 4 ノード MetroCluster 構成の場合、サイト間で ONTAP のバージョンが一致している必要があります。

リカバリサイトがスイッチバック処理を実行できるようになるまでに数時間かかることがあります。

MetroCluster switchover コマンドは、 MetroCluster 構成のすべての DR グループで、ノードをスイッチオーバーします。たとえば、 8 ノードの MetroCluster 構成では、両方の DR グループでノードがスイッチオーバーされます。

ネゴシエートスイッチオーバーの準備中および実行中は、クラスタの設定を変更したり、テイクオーバーまたはギブバック処理を実行したりしないでください。

MetroCluster FC 構成の場合：

- 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。
- 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。

MetroCluster IP 構成の場合：

> メンテナンスタスクを実行する前に、 Tiebreaker またはメディエーターユーティリティで MetroCluster 設定を監視している場合は、監視を解除する必要があります。"メンテナンス作業を実行する前に、 ONTAP メディエーターまたは Tiebreaker の監視を削除します"

- 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。

    i. MetroCluster check run 、 MetroCluster check show 、および MetroCluster check config-replication show の各コマンドを使用して、進行中または保留中の設定更新がないことを確認します。問題：稼働したままにするサイトからのコマンドです。

    ii. 稼働したままにするサイトから ' スイッチオーバーを実行します MetroCluster switchover

    この処理が完了するまでに数分かかることがあります。

    iii. スイッチオーバーの完了を監視します 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. SnapMirror 構成または SnapVault 構成があれば、再確立します。

SVMが実行されていて、アグリゲートがオンラインになっていることを確認してください

スイッチオーバーが完了したら、 DR パートナーがディスクの所有権を取得し、パートナー SVM がオンラインになったことを確認する必要があります。

MetroCluster のスイッチオーバー後に storage aggregate plex show コマンドを実行すると、スイッチオーバーされたルートアグリゲートの plex0 のステータスが確定していないため、失敗と表示されます。この間、スイッチオーバーされたルートは更新されません。このプレックスの実際のステータスは、 MetroCluster 修復フェーズ後に確定します。

手順

1. storage aggregate show コマンドを使用して、アグリゲートがスイッチオーバーされたことを確認します。

   この例では、アグリゲートがスイッチオーバーされています。ルートアグリゲート（ aggr0_b2 ）はデグレード状態、データアグリゲート（ b2_aggr2 ）はミラーされた正常な状態になっています。

```
cluster_A::*> storage aggregate show

.
.
.
mcc1-b Switched Over Aggregates:
Aggregate     Size Available Used% State   #Vols  Nodes            RAID
Status
--------- -------- --------- ----- ------- ------ ----------------
------------
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. vserver show コマンドを使用して、セカンダリ SVM がオンラインになったことを確認します。

   この例では、これまで休止状態だったセカンダリサイトの同期先 SVM がアクティブで、 Admin State は running になっています。

```
cluster_A::*> vserver show
                                          Admin      Operational  Root
Name    Name
Vserver           Type  Subtype           State      State        Volume
Aggregate  Service Mapping
-----------       ----- ----------        ---------- -----------
--------- ---------- ------- -------
...
cluster_B-vs1b-mc data   sync-destination  running   running
vs1b_vol   aggr_b1   file    file
```

構成を修復

MetroCluster FC 構成の構成を修復します

MetroCluster FC 構成での構成の修復

スイッチオーバーに続いて、特定の順序で修復処理を実行して MetroCluster 機能をリストアする必要があります。

- スイッチオーバーを実行済みで、サバイバーサイトがデータを提供している必要があります。
- ディザスタサイトのノードが停止しているか、電源がオフのままになっている必要があります。

  修復プロセスでは、これらのノードを完全にブートしないでください。

- ディザスタサイトのストレージにアクセスできる（シェルフは電源がオンで稼働しており、アクセス可能である）必要があります。
- ファブリック接続 MetroCluster 構成では、スイッチ間リンク（ ISL ）が稼働している必要があります。
- 4 ノード MetroCluster 構成では、サバイバーサイトのノードが HA フェイルオーバー状態でない（各 HA ペアのすべてのノードが稼働中である）必要があります。

修復処理は、最初にデータアグリゲートで実行し、次にルートアグリゲートで実行する必要があります。

ネゴシエートスイッチオーバー後のデータアグリゲートの修復

メンテナンスまたはテストが完了したら、データアグリゲートを修復する必要があります。この処理では、データアグリゲートを再同期し、ディザスタサイトを通常の動作用に準備します。データアグリゲートの修復は、ルートアグリゲートの修復の前に行う必要があります。

リモートクラスタでの構成の更新はすべてローカルクラスタにレプリケートされます。この手順の一部としてディザスタサイトのストレージに電源を投入しますが、ディザスタサイトにあるコントローラモジュールの電源はオンにしないでください。

手順

1. MetroCluster operation show コマンドを実行して、スイッチオーバーが完了していることを確認します。

```
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. サバイバークラスタから MetroCluster heal-phase aggregates コマンドを実行して、データアグリゲートを再同期します。

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

修復が拒否された場合は、 -override-vetoes パラメータを指定して MetroCluster heal コマンドを再実行できます。このオプションパラメータを使用すると、修復処理を妨げるソフトな拒否はすべて無視されます。

3. MetroCluster operation show コマンドを実行して、処理が完了したことを確認します。

```
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. storage aggregate show コマンドを実行して、アグリゲートの状態を確認します。

```
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. ディザスタサイトのストレージを交換した場合は、アグリゲートの再ミラーリングが必要になることがあります。

ネゴシエートスイッチオーバー後のルートアグリゲートの修復

データアグリゲートの修復が完了したら、スイッチバック処理に備えてルートアグリゲートを修復する必要があります。

MetroCluster 修復プロセスのデータアグリゲートの修復が完了している必要があります。

手順

1. MetroCluster heal-phase root-aggregates コマンドを実行して、ミラーされたアグリゲートをスイッチバックします。

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

修復が拒否された場合は、 -override-vetoes パラメータを指定して MetroCluster heal コマンドを再実行できます。このオプションパラメータを使用すると、修復処理を妨げるソフトな拒否はすべて無視されます。

2. 正常なクラスタで MetroCluster operation show コマンドを実行して、修復処理が完了したことを確認します。

```
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. 正常なサイトで次のコマンドを実行して、ディザスタサイトに障害が発生したディスクがないかどうかを確認し、ある場合は取り外します。「 Disk show -broken 」
4. ディザスタサイトの各コントローラモジュールの電源をオンにするか、ブートします。

   システムに LOADER プロンプトが表示されたら、「 boot_ontap 」コマンドを実行します。

5. ノードがブートしたら、ルートアグリゲートがミラーされていることを確認します。

   両方のプレックスが存在する場合は、プレックスが同期されていないと自動的に再同期が実行されます。1 つのプレックスで障害が発生している場合は、そのプレックスを削除してから、 storage aggregate mirror -aggregate aggregate-name コマンドを使用してミラー関係を再確立し、ミラーを再作成する必要があります。

MetroCluster IP 構成での構成の修復（ **ONTAP 9.4** 以前）

スイッチバック処理の準備として、アグリゲートを修復する必要があります。

> ONTAP 9.5 を実行している MetroCluster IP システムでは、自動的に修復が実行されるため、これらのタスクを省略できます。

修復手順を実行するには、次の条件を満たしている必要があります。

- スイッチオーバーを実行済みで、サバイバーサイトがデータを提供している必要があります。
- ディザスタサイトのストレージシェルフの電源がオンで稼働しており、アクセス可能である必要があります。
- ISL が稼働中である必要があります。
- サバイバーサイトのノードが HA フェイルオーバー状態でない（両方のノードが稼働中である）必要があります。

このタスクでは、 ONTAP 9.5 より前のバージョンを実行している環境 MetroCluster IP 構成のみを行います。

この手順は、 MetroCluster FC 構成の修復手順とは異なります。

手順

1. スイッチオーバーされたサイトの各コントローラモジュールの電源をオンにして、完全にブートします。

   システムに LOADER プロンプトが表示されたら、「 boot_ontap 」コマンドを実行します。

2. ルートアグリゲートの修復フェーズ「 MetroCluster heal root-aggregates 」を実行します

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

   修復が拒否された場合は、 MetroCluster heal-rootaggregates コマンドを --override-vetoes パラメータで再実行できます。このオプションパラメータを使用すると、修復処理を妨げるソフトな拒否はすべて無視されます。

3. アグリゲートを再同期します。「 MetroCluster heal aggregates 」

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

   修復が拒否された場合は、 -override-vetoes パラメータを指定して MetroCluster heal コマンドを再実行できます。このオプションパラメータを使用すると、修復処理を妨げるソフトな拒否はすべて無視されます。

4. 正常なクラスタで MetroCluster operation show コマンドを実行して、修復処理が完了したことを確認します。

   ```
   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: -
   ```

スイッチバックを実行しています

MetroCluster 構成の修復が完了したら、 MetroCluster のスイッチバック処理を実行できます。MetroCluster のスイッチバック処理を実行すると、構成が通常の動作状態に戻ります。ディザスタサイトにある同期元の Storage Virtual Machine （ SVM ）がアクティブになり、ローカルディスクプールからデータを提供します。

- ディザスタクラスタからサバイバークラスタへのスイッチオーバーが正常に完了している必要があります。
- データアグリゲートとルートアグリゲートに対して修復が実行されている必要があります。
- サバイバークラスタノードが HA フェイルオーバー状態ではない（各 HA ペアのすべてのノードが稼働中

である）必要があります。

- ディザスタサイトのコントローラモジュールが完全にブートしていること、および HA テイクオーバーモードでないことが必要です。
- ルートアグリゲートがミラーされている必要があります。
- スイッチ間リンク（ ISL ）がオンラインになっている必要があります。
- 必要なライセンスがシステムにインストールされている必要があります。

a. すべてのノードが enabled 状態であることを確認します MetroCluster node show

次の例は、ノードが有効な状態であることを表示します。

```
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. すべての SVM で再同期が完了したことを確認します。「 MetroCluster vserver show 」

c. 修復処理によって実行される LIF の自動移行が正常に完了したことを確認します。 MetroCluster check lif show

d. シミュレートされたスイッチバックを実行して、システムが「 MetroCluster switchback -simulate 」状態であることを確認します

e. 構成を確認します。

「 MetroCluster check run 」のようになります

このコマンドはバックグラウンドジョブとして実行され、すぐに完了しない場合があります。

```
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. サバイバークラスタ内の任意のノードから MetroCluster switchback コマンドを実行して、スイッチバックを実行します。 MetroCluster switchback

g. スイッチバック操作の進行状況を確認します MetroCluster show

出力に waiting-for-switchback と表示された場合は、スイッチバック処理がまだ進行中です。

```
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 -
```

出力に normal と表示された場合、スイッチバック処理は完了しています。

```
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 -
```

+ スイッチバックが完了するまでに時間がかかる場合は、「 MetroCluster config-replication resync-status show 」コマンドを使用することで、進行中のベースラインのステータスを確認できます。このコマンドは、 advanced 権限レベルで実行します。

a. SnapMirror 構成または SnapVault 構成があれば、再確立します。

   ONTAP 8.3 では、失われた SnapMirror 構成を MetroCluster スイッチバック処理のあとに手動で再確立する必要があります。ONTAP 9.0 以降では、関係が自動的に再確立されます。

スイッチバックが成功したことを確認する

スイッチバックの実行後に、すべてのアグリゲートと Storage Virtual Machine （ SVM ）がスイッチバックされてオンラインになっていることを確認します。

1. スイッチオーバーされたデータアグリゲートがスイッチバックされたことを確認します。

   「 storage aggregate show

   次の例では、ノード B2 の aggr_b2 がスイッチバックされています。

```
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. サバイバークラスタにある同期先のすべての SVM が休止状態（管理状態が「 stopped 」と表示されている）であり、ディザスタクラスタにある同期元の SVM が稼働していることを確認します。

   「 vserver show -subtype sync-source 」のようになります

```
node_B_1::> vserver show -subtype sync-source
                                Admin      Root
Name    Name
Vserver     Type    Subtype    State      Volume     Aggregate
Service Mapping
----------- ------- ---------- ---------- ---------- ----------
------- -------
...
vs1a        data    sync-source
                                running    vs1a_vol   node_B_2
file    file

aggr_b2

node_A_1::> vserver show -subtype sync-destination
                                Admin      Root
Name    Name
Vserver            Type    Subtype    State      Volume     Aggregate
Service Mapping
-----------        ------- ---------- ---------- ---------- ----------
------- -------
...
cluster_A-vs1a-mc  data    sync-destination
                                      stopped    vs1a_vol   sosb_
file    file

aggr_b2
```

MetroCluster 構成の同期先アグリゲートの名前には、識別しやすいように「 -mc 」というサフィックスが自動的に付加されます。

3. MetroCluster operation show コマンドを使用して、スイッチバック操作が成功したことを確認します。

出力内容	作業
スイッチバック処理の状態が「 successful 」である	スイッチバックプロセスは完了しており、システムの処理を続行できます。
スイッチバック処理またはスイッチバック継続エージェントの処理が「 partially successful 」である	MetroCluster operation show コマンドの出力に記載されている推奨修正を実行します

上記の手順を繰り返して、逆方向へのスイッチバックを実行する必要があります。site_A が site_B のスイッチオーバーを行った場合は、 site_B で site_A のスイッチオーバーを行います

スイッチオーバー、修復、スイッチバック用のコマンドです

ONTAP には、 MetroCluster のディザスタリカバリプロセスを実行するためのコマンドが用意されています。

状況	使用するコマンド
スイッチオーバーでエラーや拒否が発生しないことを確認します。	MetroCluster switchover -simulate + を advanced 権限レベルで実行します
スイッチバックを実行してもエラーや拒否が発生しないことを確認します。	MetroCluster switchback -simulate + at the advanced privilege level
パートナーノードにスイッチオーバーする（ネゴシエートスイッチオーバー）	MetroCluster スイッチオーバー
パートナーノードにスイッチオーバーする（強制スイッチオーバー）	MetroCluster switchover -forced-on-disaster true
データアグリゲートの修復を実行する	「 MetroCluster heal-phase aggregates 」
ルートアグリゲートの修復を実行する	「 MetroCluster heal-phase root-aggregates 」
ホームノードにスイッチバックする	MetroCluster スイッチバック

MetroCluster 設定を監視する

ONTAP MetroCluster コマンドと Active IQ Unified Manager （旧 OnCommand Unified Manager ）を使用して、各種ソフトウェアコンポーネントの健常性および MetroCluster 処理の状態を監視できます。

MetroCluster の設定を確認しています

MetroCluster 構成内のコンポーネントおよび関係が正しく機能していることを確認できます。チェックは、初期設定後と、 MetroCluster 設定に変更を加えたあとに実施する必要があります。また、ネゴシエート（計画的）スイッチオーバーやスイッチバックの処理の前にも実施します。

このタスクについて

いずれかまたは両方のクラスタに対して短時間に MetroCluster check run コマンドを 2 回発行すると ' 競合が発生し ' コマンドがすべてのデータを収集しない場合がありますそれ以降の「 MetroCluster check show 」コマンドでは、期待される出力が表示されません。

手順

1. 構成を確認します。

   「 MetroCluster check run 」のようになります

   このコマンドはバックグラウンドジョブとして実行され、すぐに完了しない場合があります。

```
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. 最新の MetroCluster check run コマンドから ' より詳細な結果を表示します

   MetroCluster check aggregate show

   MetroCluster check cluster show

   MetroCluster check config-replication show

   MetroCluster check lif show

   MetroCluster check node show

   「 MetroCluster check show 」コマンドは、最新の「 MetroCluster check run 」コマンドの結果を表示します。MetroCluster check show コマンドを使用する前に ' 必ず MetroCluster check run コマンドを実行して ' 表示されている情報が最新であることを確認してください

   次に、正常な 4 ノード MetroCluster 構成の MetroCluster check aggregate show コマンドの出力例を示します。

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

```
ok
                      controller_A_1_aggr1
                                                  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
ok

18 entries were displayed.
```

次に、正常な 4 ノード MetroCluster 構成の MetroCluster check cluster show コマンドの出力例を示します。この出力は、必要に応じてネゴシエートスイッチオーバーを実行できる状態であることを示しています。

```
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.
```

MetroCluster 設定を確認および監視するためのコマンド

ONTAP には、 MetroCluster 構成を監視し、 MetroCluster 処理を確認するためのコマンドが用意されています。

MetroCluster 処理を確認するためのコマンド

状況	使用するコマンド
MetroCluster 処理のチェックを実行します。 • 注：このコマンドは、 DR 処理実行前のシステムの事前検証に使用できるコマンドとしては使用しないでください。	「 MetroCluster check run 」のようになります
MetroCluster 処理に対する前回のチェック結果を表示する	「 MetroCluster show 」
サイト間の設定レプリケーションに対するチェック結果を表示する	「 MetroCluster check config-replication show MetroCluster check config-replication show-aggregate-eligibility -
ノード設定に対するチェック結果を表示する	MetroCluster check node show
アグリゲート設定に対するチェック結果を表示する	MetroCluster check aggregate show
MetroCluster 構成の LIF 配置エラーを表示する	MetroCluster check lif show

MetroCluster インターコネクトの監視用コマンドです

状況	使用するコマンド
HA および DR のミラーリングステータスと、クラスタ内の MetroCluster ノードに関する情報を表示する	MetroCluster interconnect mirror show

MetroCluster SVM の監視用コマンド

状況	使用するコマンド
MetroCluster 構成の両方のサイトにあるすべての SVM を表示する	MetroCluster vserver show

MetroCluster Tiebreaker または ONTAP メディエーターを使用して設定を監視する

を参照してください "ONTAP メディエーターと MetroCluster Tiebreaker の違い" MetroCluster 構成を監視する方法と自動スイッチオーバーを開始する方法の違いを理解するために、

Tiebreaker またはメディエーターをインストールおよび設定するには、次のリンクを使用します。

- "MetroCluster Tiebreaker ソフトウェアをインストールして設定します。"
- "ONTAP メディエーターサービスをインストールする準備をします"

NetApp MetroCluster Tiebreaker ソフトウェアで障害を検出する方法

Tiebreaker ソフトウェアは Linux ホストにインストールされます。Tiebreaker ソフトウェアは、 2 つのクラスタおよびクラスタ間の接続ステータスを第 3 のサイトから監視する場合にのみ使用します。これにより、サイト間リンクが停止した場合に、クラスタ内の各パートナーで ISL 障害とサイト障害を区別することができます。

Linux ホストに Tiebreaker ソフトウェアをインストールしたら、災害状況を監視するように MetroCluster 構成内のクラスタを設定できます。

Tiebreaker ソフトウェアでサイト間接続障害を検出する方法

MetroCluster Tiebreaker ソフトウェアは、サイト間のすべての接続が失われると警告します。

ネットワークパスのタイプ

構成によっては、 MetroCluster 構成の 2 つのクラスタ間のネットワークパスに次の 3 つの種類が存在します。

- * FC ネットワーク（ファブリック接続 MetroCluster 構成に存在） *

  この種類のネットワークは、 2 つの冗長 FC スイッチファブリックで構成されます。各スイッチファブリックには 2 つの FC スイッチがあり、各スイッチファブリックの 1 つのスイッチはクラスタと同じ場所に配置されます。各クラスタには、各スイッチファブリックから 1 つずつ、 2 つの FC スイッチがあります。すべてのノードは、同じ場所に配置されている各 IP スイッチに FC （ NV インターコネクトおよび FCP イニシエータ）接続されています。データは、クラスタからクラスタへ、 ISL 経由でレプリケートさ

れます。

- * クラスタ間ピアリングネットワーク *

  この種類のネットワークは、 2 つのクラスタ間の冗長 IP ネットワークパスで構成されます。クラスタピアリングネットワークは、 Storage Virtual Machine （ SVM ）構成をミラーするために必要な接続を提供します。一方のクラスタのすべての SVM の設定が、パートナークラスタにミラーされます。

- * IP ネットワーク（ MetroCluster IP 構成に存在） *

  この種類のネットワークは、 2 つの冗長 IP スイッチネットワークで構成されます。各ネットワークには 2 つの IP スイッチがあり、各スイッチファブリックの 1 つのスイッチはクラスタと同じ場所に配置されます。各クラスタには、各スイッチファブリックから 1 つずつ、 2 つの IP スイッチがあります。すべてのノードは、同じ場所に配置されている各 FC スイッチに接続されています。データは、クラスタからクラスタへ、 ISL 経由でレプリケートされます。

サイト間接続を監視する

Tiebreaker ソフトウェアは、サイト間接続のステータスをノードから定期的に取得します。NV インターコネクト接続が失われ、クラスタ間ピアリングが ping に応答しない場合、クラスタはサイトが分離されているとみなし、 Tiebreaker ソフトウェアは「 AllLinksSevered 」としてアラートをトリガーします。クラスタが「 AllLinksSevered 」ステータスを識別し、もう一方のクラスタがネットワーク経由で到達できない場合、 Tiebreaker ソフトウェアは「 disaster 」としてアラートをトリガーします。

Tiebreaker ソフトウェアでサイト障害を検出する方法

NetApp MetroCluster Tiebreaker ソフトウェアは、 MetroCluster 構成のノードおよびクラスタに到達できるかどうかをチェックして、サイト障害の有無を判断します。また、 Tiebreaker ソフトウェアは、特定の状況でアラートをトリガーします。

Tiebreaker ソフトウェアで監視されるコンポーネント

Tiebreaker ソフトウェアは、 IP ネットワークでホストされるノード管理 LIF およびクラスタ管理 LIF への複数のパスを経由する冗長接続を確立することによって、 MetroCluster 構成内の各コントローラを監視します。

Tiebreaker ソフトウェアで監視される MetroCluster 構成のコンポーネントは次のとおりです。

- ローカルノードインターフェイスを介してノードを追加します
- クラスタ指定インターフェイスを介したクラスタ
- サバイバークラスタ - ディザスタサイトとの接続の有無を評価（ NV インターコネクト、ストレージ、クラスタ間ピアリング）

Tiebreaker ソフトウェアとクラスタ内のすべてのノードおよびクラスタ自体との接続が失われると、クラスタは Tiebreaker ソフトウェアによって「到達不能」と宣言されます。接続障害は 3~5 秒で検出されます。Tiebreaker ソフトウェアからクラスタに到達できない場合、障害が発生していないクラスタ（到達可能なクラスタ）は、 Tiebreaker ソフトウェアがアラートをトリガーする前に、パートナークラスタへのすべてのリンクが切断されていることを示す必要があります。

サバイバークラスタが FC （ NV インターコネクトとストレージ）とクラスタ間ピアリングを介してディザスタサイトのクラスタと通信できなくなると、すべてのリンクが切断されたとみなされます。

Tiebreaker ソフトウェアがアラートをトリガーする障害シナリオ

Tiebreaker ソフトウェアは、ディザスタサイトのクラスタ（すべてのノード）が停止しているか到達不能であり、サバイバーサイトのクラスタが「 AllLinksSevered 」ステータスを示している場合、アラートをトリガーします。

次のシナリオでは、 Tiebreaker ソフトウェアはアラートをトリガーしません（またはアラートが拒否されます）。

- 8 ノード MetroCluster 構成で、ディザスタサイトの HA ペアの 1 つが停止している場合
- ディザスタサイトのすべてのノードを含むクラスタが停止し、サバイバーサイトの 1 つの HA ペアが停止し、サバイバーサイトのクラスタが「 AllLinksSevered 」ステータスを示している場合

  Tiebreaker ソフトウェアはアラートをトリガーしますが、 ONTAP はアラートを拒否します。この場合、手動によるスイッチオーバーも拒否されます

- Tiebreaker ソフトウェアがディザスタサイトの少なくとも 1 つのノードまたはクラスタインターフェイスに到達できる、またはサバイバーサイトが FC （ NV インターコネクトとストレージ）またはクラスタ間ピアリングを介してディザスタサイトのいずれかのノードに到達できる場合

ONTAP メディエーターでの自動計画外スイッチオーバーのサポート

"MetroCluster IP構成での自動計画外スイッチオーバーのONTAPメディエーターによるサポートの詳細"。

NVFAIL を使用したファイルシステム整合性の監視および保護

volume modify コマンドの -nvfail` パラメータを使用すると、 ONTAP は、システムの起動時またはスイッチオーバー操作後に NVRAM （不揮発性 RAM ）の不整合を検出できます。また、警告を表示し、ボリュームを手動でリカバリできるようになるまでデータのアクセスと変更ができないようにシステムを保護します。

ONTAP が何らかの問題を検出すると、データベースインスタンスまたはファイルシステムインスタンスは応答を停止するか、シャットダウンされます。ONTAP はコンソールにエラーメッセージを送信して、データベースまたはファイルシステムの状態をチェックするようユーザに警告します。NVFAIL を有効にすると、データベースの有効性を侵害する可能性のある、クラスタ化されたノード間での NVRAM の不整合をデータベース管理者に警告できます。

フェイルオーバーまたはブートリカバリのあとは、 NVFAIL 状態が解消されるまで NFS クライアントはどのノードからもデータにアクセスできません。CIFS クライアントには影響はありません。

NFS ボリュームまたは LUN へのアクセスに対する NVFAIL の影響

NVFAIL 状態は、起動時に ONTAP が NVRAM エラーを検出した場合、 MetroCluster スイッチオーバー処理が実行された場合、または NVFAIL オプションが設定されたボリュームの HA テイクオーバー処理中に発生します。起動時にエラーが検出されなければ、ファイルサービスは正常に開始されます。ただし、 NVRAM エラーが検出された場合、またはディザスタスイッチオーバーで NVFAIL の処理が実行された場合、 ONTAP はデ

ータベースインスタンスの応答を停止します。

NVFAIL オプションを有効にすると、起動時に次の表に示すいずれかのプロセスが開始されます。

状況	作業
ONTAP は NVRAM エラーを検出しませんでした	ファイルサービスが正常に開始されます。
ONTAP が NVRAM エラーを検出した場合	• ONTAP がデータベースにアクセスしようとする NFS クライアントに古いファイルハンドル（ ESTALE ）エラーを返します。これにより、アプリケーションは、応答を中断するか、クラッシュまたは停止します。 ONTAP からシステムコンソールおよびログファイルにエラーメッセージが送信されます。 • アプリケーションが再起動されると、データベースの妥当性が確認されていなくても、 CIFS クライアントからのアクセスが可能になります。 NFS クライアントの場合は、影響を受けたボリューム上の in-nvfailed-state オプションをリセットするまでファイルにアクセスできません。
次のいずれかのパラメータを使用する場合： • 「 dr-force-nvfail 」ボリュームオプションが設定されている • force-nvfail-all 'switchover コマンド・オプションが設定されている	管理者が将来のディザスタスイッチオーバーで NVFAIL 処理を強制的に実行する予定がない場合は、スイッチオーバー後に「 dr-force-nvfail 」オプションを解除できます。NFS クライアントの場合は、影響を受けたボリューム上の in-nvfailed-state オプションをリセットするまでファイルにアクセスできません。 「 force-nvfail-all 」オプションを使用すると、ディザスタスイッチオーバー時に処理されるすべての DR ボリュームで「 dr -force-nvfail 」オプションが設定されます。
ONTAP は、 LUN を含むボリューム上で NVRAM エラーを検出します	そのボリュームの LUN がオフラインになります。ボリュームの in-nvfailed-state オプションをクリアし、さらに影響を受けたボリュームの各 LUN をオンラインにして LUN 上の NVFAIL 属性をクリアする必要があります。LUN の整合性をチェックする手順を実行し、必要に応じて Snapshot コピーまたはバックアップから LUN をリカバリすることができます。ボリューム内のすべての LUN がリカバリされると、影響を受けたボリュームの in-nvfailed-state オプションがクリアされます。

データ損失イベントを監視するコマンド

NVFAIL オプションを有効にすると、 NVRAM の不整合によってシステムクラッシュが発生した場合、または MetroCluster スイッチオーバーが発生した場合に、通知を受け取ることができます。

デフォルトでは、 NVFAIL パラメータは無効になっています。

状況	使用するコマンド
NVFAIL を有効にして新しいボリュームを作成します	'volume create-nvfail on `
既存のボリュームで NVFAIL を有効にする	'volume modify' • 注： * 作成されたボリュームで NVFAIL を有効にするには、 -nvfail オプションを on に設定します。
指定したボリュームで NVFAIL が有効になっているかどうかを表示します	volume show • 注： * 指定したボリュームの NVFAIL 属性を表示するには、「 -fields 」パラメータを「 nvfail 」に設定します。

関連情報

詳細については、各コマンドのマニュアルページを参照してください。

スイッチオーバー後に **NVFAIL** 状態のボリュームにアクセスする場合

スイッチオーバー後に 'volume modify コマンドの -in-nvfailed-state パラメータをリセットし ' クライアントがデータにアクセスする制限を解除することにより 'NVFAIL 状態をクリアする必要があります

作業を開始する前に

データベースまたはファイルシステムが実行されていない、または影響を受けたボリュームにアクセスしようとしていないことが必要です。

このタスクについて

in-nvfailed-state パラメータを設定するには、 advanced 権限が必要です。

ステップ

1. in-nvfailed-state パラメータを false に設定して volume modify コマンドを実行し、ボリュームをリカバリします。

完了後

データベースファイルの有効性を調べる方法については、ご使用のデータベースソフトウェアのマニュアルを参照してください。

データベースで LUN が使用されている場合は、 NVRAM 障害後にホストから LUN にアクセスできるようにする手順を確認してください。

関連情報

"NVFAIL を使用したファイルシステム整合性の監視および保護"

スイッチオーバー後の NVFAIL 状態にある LUN のリカバリ

スイッチオーバーが発生すると、 NVFAIL 状態にある LUN 上のデータにホストがアクセスできなくなります。データベースが LUN にアクセスできるようにするには、いくつかの作業が必要です。

作業を開始する前に

データベースが実行されていないことを確認します。

手順

1. volume modify コマンドの -in-nvfailed-state パラメータをリセットし、 LUN をホストする、影響を受けたボリュームの NVFAIL 状態をクリアします。
2. 影響を受けた LUN をオンラインにします。
3. LUN でデータの不整合の有無を確認し、これを解決します。

   これには、ホストベースのリカバリか、 SnapRestore を使用してストレージコントローラ上で実行するリカバリが含まれる場合があります。

4. LUN をリカバリしたあと、データベースアプリケーションをオンラインにします。

追加情報の参照先

MetroCluster の設定と処理の詳細については、を参照してください。

MetroCluster およびその他の情報

情報	件名
"MetroCluster のドキュメント"	• すべての MetroCluster 情報
"ネットアップテクニカルレポート 4375 ：『 NetApp MetroCluster for ONTAP 9.3 』"	• MetroCluster の設定と運用の技術概要 • MetroCluster 構成のベストプラクティス
"ファブリック接続 MetroCluster のインストールと設定"	• ファブリック接続 MetroCluster アーキテクチャ • 構成のケーブル接続 • FC-to-SAS ブリッジの設定 • FC スイッチを設定します • ONTAP で MetroCluster を設定します

"ストレッチ MetroCluster のインストールと設定"	• ストレッチ MetroCluster アーキテクチャ • 構成のケーブル接続 • FC-to-SAS ブリッジの設定 • ONTAP で MetroCluster を設定します
"MetroCluster IP のインストールと設定"	• MetroCluster IP アーキテクチャ • 構成のケーブル接続 • ONTAP で MetroCluster を設定します
"MetroCluster Tiebreaker 1.21 ソフトウェアのインストールと設定"	• MetroCluster Tiebreaker ソフトウェアを使用して MetroCluster 構成を監視する
Active IQ Unified Manager のドキュメント "ネットアップのマニュアル：製品ガイドとリソース"	• MetroCluster の設定とパフォーマンスを監視する
"コピーベースの移行"	• 7-Mode ストレージシステムから clustered Data ONTAP ストレージシステムへデータを移行する

著作権に関する情報

Copyright © 2024 NetApp, Inc. All Rights Reserved. Printed in the U.S.このドキュメントは著作権によって保護されています。著作権所有者の書面による事前承諾がある場合を除き、画像媒体、電子媒体、および写真複写、記録媒体、テープ媒体、電子検索システムへの組み込みを含む機械媒体など、いかなる形式および方法による複製も禁止します。

ネットアップの著作物から派生したソフトウェアは、次に示す使用許諾条項および免責条項の対象となります。

このソフトウェアは、ネットアップによって「現状のまま」提供されています。ネットアップは明示的な保証、または商品性および特定目的に対する適合性の暗示的保証を含み、かつこれに限定されないいかなる暗示的な保証も行いません。ネットアップは、代替品または代替サービスの調達、使用不能、データ損失、利益損失、業務中断を含み、かつこれに限定されない、このソフトウェアの使用により生じたすべての直接的損害、間接的損害、偶発的損害、特別損害、懲罰的損害、必然的損害の発生に対して、損失の発生の可能性が通知されていたとしても、その発生理由、根拠とする責任論、契約の有無、厳格責任、不法行為（過失またはそうでない場合を含む）にかかわらず、一切の責任を負いません。

ネットアップは、ここに記載されているすべての製品に対する変更を随時、予告なく行う権利を保有します。ネットアップによる明示的な書面による合意がある場合を除き、ここに記載されている製品の使用により生じる責任および義務に対して、ネットアップは責任を負いません。この製品の使用または購入は、ネットアップの特許権、商標権、または他の知的所有権に基づくライセンスの供与とはみなされません。

このマニュアルに記載されている製品は、1つ以上の米国特許、その他の国の特許、および出願中の特許によって保護されている場合があります。

権利の制限について：政府による使用、複製、開示は、DFARS 252.227-7013（2014年2月）およびFAR 5252.227-19（2007年12月）のRights in Technical Data -Noncommercial Items（技術データ - 非商用品目に関する諸権利）条項の(b)(3)項、に規定された制限が適用されます。

本書に含まれるデータは商用製品および / または商用サービス（FAR 2.101の定義に基づく）に関係し、データの所有権はNetApp, Inc.にあります。本契約に基づき提供されるすべてのネットアップの技術データおよびコンピュータ ソフトウェアは、商用目的であり、私費のみで開発されたものです。米国政府は本データに対し、非独占的かつ移転およびサブライセンス不可で、全世界を対象とする取り消し不能の制限付き使用権を有し、本データの提供の根拠となった米国政府契約に関連し、当該契約の裏付けとする場合にのみ本データを使用できます。前述の場合を除き、NetApp, Inc.の書面による許可を事前に得ることなく、本データを使用、開示、転載、改変するほか、上演または展示することはできません。国防総省にかかる米国政府のデータ使用権については、DFARS 252.227-7015(b)項（2014年2月）で定められた権利のみが認められます。

商標に関する情報

NetApp、NetAppのロゴ、http://www.netapp.com/TMに記載されているマークは、NetApp, Inc.の商標です。その他の会社名と製品名は、それを所有する各社の商標である場合があります。