管理BeeGFS叢集

BeeGFS on NetApp with E-Series Storage

NetApp
December 01, 2025

This PDF was generated from https://docs.netapp.com/zh-tw/beegfs/administer/clusters-overview.html on December 01, 2025. Always check docs.netapp.com for the latest.

目錄

管理BeeGFS叢集 1
概述、主要概念和術語 1
總覽 1
重要概念 1
通用術語 1
使用Ansible與PCS工具的時機 2
檢查叢集的狀態 2
總覽 2
瞭解輸出來源 `pcs status` 2
重新設定HA叢集和BeeGFS 4
總覽 4
如何停用和啟用屏障 4
更新 HA 叢集元件 5
更新 BeeGFS 版本 5
升級 HA 叢集中的 Pacemaker 和 corosync 套件 7
更新檔案節點介面卡韌體 10
升級 E-Series 儲存陣列 14
服務與維護 16
容錯移轉和容錯回復服務 16
將叢集置於維護模式 19
停止並啟動叢集 19
取代檔案節點 20
擴充或縮小叢集 21
疑難排解 22
總覽 22
疑難排解指南 23
常見問題 26
常見疑難排解工作 27

管理BeeGFS叢集

概述、主要概念和術語

瞭解如何在部署BeeGFS HA叢集之後管理這些叢集。

總覽

本節適用於部署BeeGFS HA叢集之後、需要管理的叢集管理員。即使是熟悉Linux HA叢集的使用者、也應該徹底閱讀本指南、因為管理叢集的方式有許多差異、尤其是使用Ansible進行重新設定。

重要概念

雖然在主"詞彙與概念"頁上介紹了其中一些概念、但在 BeeGFS HA 叢集的背景下重新介紹這些概念是很有幫助的：

**叢集節點：*執行起搏器和電暈器同步服務的伺服器、並參與HA叢集。

**檔案節點：*用於執行一或多個BeeGFS管理、中繼資料或儲存服務的叢集節點。

*區塊節點： NetApp E系列儲存系統、可為檔案節點提供區塊儲存。這些節點不會參與BeeGFS HA叢集、因為它們提供自己的獨立HA功能。每個節點包含兩個儲存控制器、可在區塊層提供高可用度。

- BeeGFS服務：* BeeGFS管理、中繼資料或儲存服務。每個檔案節點都會執行一或多項服務、這些服務會使用區塊節點上的磁碟區來儲存其資料。

**建置區塊：*標準化部署BeeGFS檔案節點、E系列區塊節點、以及其上執行的BeeGFS服務、可簡化BeeGFS HA叢集/檔案系統的擴充、並遵循NetApp驗證架構。也支援自訂HA叢集、但通常採用類似的建置區塊方法來簡化擴充。

- BeeGFS HA叢集：*一組可擴充的檔案節點、用於執行BeeGFS服務、並以區塊節點作為後盾、以高可用度的方式儲存BeeGFS資料。以業界公認的開放原始碼元件為基礎、採用Ansible進行包裝與部署。

**叢集服務：*是指叢集內每個節點上執行的起搏器和電暈器同步服務。請注意、如果只需要兩個檔案節點、節點可能無法執行任何BeeGFS服務、只能以「tiebreaker」節點的形式參與叢集。

**叢集資源：*針對叢集中執行的每個BeeGFS服務、您將會看到BeeGFS監控資源、以及一個資源群組、其中包含BeeGFS目標、IP位址（浮動IP）及BeeGFS服務本身的資源。

- Ansible：*軟體資源配置、組態管理及應用程式部署工具、以程式碼形式提供基礎架構。這是BeeGFS HA叢集的封裝方式、可簡化在NetApp上部署、重新設定及更新BeeGFS的程序。
- PCS：*叢集中任何檔案節點均提供命令列介面、用於查詢及控制叢集中節點和資源的狀態。

通用術語

**容錯移轉：*每個BeeGFS服務都有一個優先要執行的檔案節點、除非該節點故障。當BeeGFS服務在非慣用/次要檔案節點上執行時、就表示該服務正在進行容錯移轉。

*容錯回復：*將BeeGFS服務從非偏好的檔案節點移回偏好的節點。

- HA配對：*可存取相同區塊節點集的兩個檔案節點、有時稱為HA配對。這是整個NetApp常用的詞彙、用來指可「接管」彼此的兩個儲存控制器或節點。
- 維護模式： ** 停用所有資源監控"維護模式"功能、並防止 Pacemaker 移動或以其他方式管理叢集中的資源（請同時參閱上的一節）。
- HA叢集：*執行BeeGFS服務的一或多個檔案節點、可在叢集中的多個節點之間容錯移轉、以建立高可用度的BeeGFS檔案系統。檔案節點通常會設定成HA配對、以便在叢集中執行BeeGFS服務的子集。

使用Ansible與PCS工具的時機

您應該何時使用Ansible與PCS命令列工具來管理HA叢集？

所有叢集部署和重新設定工作都應使用外部可控制節點的Ansible來完成。叢集狀態的暫時性變更（例如、將節點置入和移出待命）通常是透過登入叢集的一個節點（最好是未降級或即將進行維護的節點）、以及使用PCS命令列工具來執行。

修改任何叢集組態、包括資源、限制、內容及BeeGFS服務本身、都應該使用Ansible來完成。維護叢集的一部分、是維護「Ansible Inventory and playbook（Ansible Inventory and playbook）」的最新複本（理想的來源控制方式是追蹤變更）。當您需要變更組態時、請更新詳細目錄、然後重新執行匯入BeeGFS HA角色的Ansible教戰手冊。

HA角色會處理將叢集置於維護模式、然後在重新啟動BeeGFS或叢集服務以套用新組態之前進行任何必要的變更。由於在初始部署之外通常不需要完整的節點重新開機、因此重新執行Ansible通常被視為「安全」程序、但如果任何BeeGFS服務需要重新啟動、則建議在維護期間或下班時間重新執行。這些重新啟動通常不會造成應用程式錯誤、但可能會影響效能（某些應用程式處理效能可能優於其他應用程式）。

當您想要將整個叢集恢復到完全最佳狀態時、重新執行Ansible也是一個選項、而且在某些情況下、可能比使用PC更容易恢復叢集狀態。尤其是在叢集因某種原因而停機的緊急情況下、一旦所有節點都備份重新執行Ansible、就可能比嘗試使用PC更快且更可靠地還原叢集。

檢查叢集的狀態

使用PC檢視叢集的狀態。

總覽

執行中 `pcs status` 從任何叢集節點、都是查看叢集整體狀態及每個資源狀態（例如BeeGFS服務及其相依性）的最簡單方法。本節將說明您在的輸出中所能找到的內容 `pcs status` 命令。

瞭解輸出來源 `pcs status`

執行 `pcs status` 在任何叢集服務（起搏器和電暈器同步）啟動的叢集節點上。輸出頂端會顯示叢集摘要：

```
[root@beegfs_01 ~]# pcs status
Cluster name: hacluster
Cluster Summary:
  * Stack: corosync
  * Current DC: beegfs_01 (version 2.0.5-9.el8_4.3-ba59be7122) - partition
with quorum
  * Last updated: Fri Jul  1 13:37:18 2022
  * Last change:  Fri Jul  1 13:23:34 2022 by root via cibadmin on
beegfs_01
  * 6 nodes configured
  * 235 resource instances configured
```

以下章節列出叢集中的節點：

```
Node List:
  * Node beegfs_06: standby
  * Online: [ beegfs_01 beegfs_02 beegfs_04 beegfs_05 ]
  * OFFLINE: [ beegfs_03 ]
```

這特別代表任何處於待命或離線狀態的節點。待命中的節點仍在參與叢集、但標記為不符合執行資源的資格。離線的節點表示叢集服務未在該節點上執行、可能是因為手動停止、或是因為節點已重新開機/關機。

> 當節點初次啟動時、叢集服務將會停止、需要手動啟動、以避免意外將資源還原至不正常的節點。

如果節點因為非管理原因（例如故障）而處於待命或離線狀態、則會在節點狀態旁以括弧顯示其他文字。例如、如果隔離功能已停用、而資源遇到您將會看到的故障 `Node <HOSTNAME>: standby (on-fail)`。另一個可能的狀態是 `Node <HOSTNAME>: UNCLEAN (offline)`（這會短暫地被視為節點正在被圍起來、但如果隔離失敗、表示叢集無法確認節點狀態、則會持續存在（這可能會封鎖其他節點上的資源啟動）。

下一節顯示叢集中所有資源及其狀態的清單：

```
Full List of Resources:
  * mgmt-monitor    (ocf::eseries:beegfs-monitor):   Started beegfs_01
  * Resource Group: mgmt-group:
    * mgmt-FS1  (ocf::eseries:beegfs-target):    Started beegfs_01
    * mgmt-IP1  (ocf::eseries:beegfs-ipaddr2):   Started beegfs_01
    * mgmt-IP2  (ocf::eseries:beegfs-ipaddr2):   Started beegfs_01
    * mgmt-service  (systemd:beegfs-mgmtd):  Started beegfs_01
[...]
```

與節點類似、如果資源有任何問題、資源狀態旁會顯示其他文字、並以括弧表示。例如、如果心臟起搏器要求資源停止、但無法在分配的時間內完成、則心臟起搏器會嘗試隔離節點。如果禁用隔離功能或隔離操作失敗，則資源狀態將是 `FAILED <HOSTNAME> (blocked)` 而起搏器將無法在不同的節點上啟動。

值得一提的是、BeeGFS HA叢集運用了許多BeeGFS最佳化的自訂OCF資源代理程式。特別是BeeGFS監視器、負責在特定節點上的BeeGFS資源無法使用時觸發容錯移轉。

重新設定HA叢集和BeeGFS

使用Ansible重新設定叢集。

總覽

一般而言、只要更新 Ansible 清查並重新執行 `ansible-playbook`命令、即可重新設定 BeeGFS HA 叢集的任何層面。這包括更新警示、變更永久隔離組態或調整BeeGFS服務組態。您 `group_vars/ha_cluster.yml`可以使用檔案來調整這些選項、並在"指定通用檔案節點組態"一節中找到完整的選項清單。

如需管理員在執行維護或服務叢集時應注意的特定組態選項詳細資料、請參閱下方。

如何停用和啟用屏障

設定叢集時、預設會啟用/需要隔離功能。在某些情況下、可能需要暫時停用隔離功能、以確保在執行某些維護作業（例如升級作業系統）時、不會意外關閉節點。雖然可以手動停用此功能、但系統管理員仍應注意取捨。

選項**1**：使用**Ansible**（建議）停用隔離功能。

使用Ansible停用隔離功能時、BeeGFS監視器的失敗動作會從「Fence」變更為「standby」（待命）。這表示、如果BeeGFS監視器偵測到故障、就會嘗試將節點置於待命狀態、並容錯移轉所有BeeGFS服務。在主動式疑難排解/測試之外、這通常比選項2更為理想。缺點是、如果資源無法在原始節點上停止、就會被封鎖、無法從其他位置啟動（這也是為什麼正式作業叢集通常需要隔離）。

1. 在您的Ansible庫存中 `groups_vars/ha_cluster.yml` 新增下列組態：

```
beegfs_ha_cluster_crm_config_options:
    stonith-enabled: False
```

2. 重新執行「Ansible」方針、將變更套用至叢集。

選項**2**：手動停用隔離功能。

在某些情況下、您可能想要暫時停用隔離功能、而不重新執行Ansible、或許是為了協助疑難排解或測試叢集。

> 在此組態中、如果BeeGFS監視器偵測到故障、叢集將嘗試停止對應的資源群組。它不會觸發完整容錯移轉、也不會嘗試重新啟動受影響的資源群組、或將其移至其他主機。若要恢復、請先解決所有問題、然後再執行 `pcs resource cleanup` 或手動將節點置於待命狀態。

步驟：

1. 若要判斷隔離（stonith）是否已全域啟用或停用、請執行： `pcs property show stonith-enabled`
2. 若要停用隔離執行： `pcs property set stonith-enabled=false`
3. 若要啟用隔離執行： `pcs property set stonith-enabled=true`

下次執行 Ansible 劇本時，此設定將被覆蓋。

更新 HA 叢集元件

更新 BeeGFS 版本

請依照下列步驟、使用 Ansible 更新 HA 叢集的 BeeGFS 版本。

總覽

BeeGFS 遵循 `major.minor.patch`版本管理方案。`BeeGFS HA Ansible` 角色適用於每個支援的 `major.minor`版本（例如、 `beegfs_ha_7_2`和 `beegfs_ha_7_3`）。每個 HA 角色都固定在 Ansible 集合發行時可用的最新 BeeGFS 修補程式版本上。

Ansible應用於所有BeeGFS升級、包括在BeeGFS的主要、次要和修補版本之間移動。若要更新BeeGFS、您必須先更新BeeGFS Ansible集合、這也會為部署/管理自動化和基礎HA叢集提供最新的修正與增強功能。即使更新至最新版本的收藏、BeeGFS也不會升級至 `ansible-playbook` 與一起執行 `-e "beegfs_ha_force_upgrade=true"` 設定。

如需BeeGFS版本的詳細資訊、請參閱 "BeeGFS升級文件"。

已測試的升級途徑

每個版本的BeeGFS集合都會與特定版本的BeeGFS進行測試、以確保所有元件之間的互通性。也會執行測試、確保可從最新版集合所支援的BeeGFS版本升級至最新版所支援的版本。

原始版本	升級版本	多重軌道	詳細資料
7.2.6	7.3.2	是的	將beegfs集合從v3.0.1升級至v3.1.0、新增多重軌道
7.2.6	7.2.8	否	將beegfs集合從v3.0.1升級至v3.1.0
7.2.8	7.3.1	是的	使用beegfs集合v3.1.0升級、新增多重軌道
7.3.1	7.3.2	是的	使用beegfs集合v3.1.0升級
7.3.2	7.4.1..	是的	使用beegfs集合v3.2.0升級
7.4.1..	7.4.2..	是的	使用beegfs集合v3.2.0升級

BeeGFS 升級步驟

下列各節提供更新 BeeGFS Ansible 系列和 BeeGFS 本身的步驟。請特別注意任何額外步驟、以更新BeeGFS主要或次要版本。

步驟 **1** ：升級 **BeeGFS** 系列

可存取的集合升級 "Ansible Galaxy"，執行下列命令：

```
ansible-galaxy collection install netapp_eseries.beegfs --upgrade
```

如需離線收藏升級、請從下載收藏 "Ansible Galaxy" 按一下所需的 `Install Version`` 然後 `Download tarball`。將tar傳輸到Ansible控制節點、然後執行下列命令。

```
ansible-galaxy collection install netapp_eseries-beegfs-<VERSION>.tar.gz
--upgrade
```

請參閱 "安裝集合" 以取得更多資訊。

步驟 2 ：更新 Ansible 庫存

對叢集的Ansible庫存檔案進行必要或想要的更新。如"版本升級附註"需您特定升級需求的詳細資訊、請參閱以下章節。"Ansible Inventory Overview"如需設定 BeeGFS HA 清查的一般資訊、請參閱一節。

步驟 3 ：更新 Ansible 教戰手冊（僅更新主要或次要版本時）

如果您要在主要或次要版本之間移動、請在 `playbook.yml` 用於部署和維護叢集的檔案中、更新角色名稱 `beegfs_ha_<VERSION>` 以反映所需的版本。例如，如果您想部署 BeeGFS 7.4 ，這將是 `beegfs_ha_7_4`：

```
- hosts: all
  gather_facts: false
  any_errors_fatal: true
  collections:
    - netapp_eseries.beegfs
  tasks:
    - name: Ensure BeeGFS HA cluster is setup.
      ansible.builtin.import_role:  # import_role is required for tag
availability.
        name: beegfs_ha_7_4
```

如需此教戰手冊檔案內容的詳細資訊"部署BeeGFS HA叢集"、請參閱一節。

步驟 4 ：執行 BeeGFS 升級

若要套用BeeGFS更新：

```
ansible-playbook -i inventory.yml beegfs_ha_playbook.yml -e
"beegfs_ha_force_upgrade=true" --tags beegfs_ha
```

BeeGFS HA角色將在幕後處理：

- 確保叢集處於最佳狀態、且每個BeeGFS服務都位於其偏好的節點上。
- 將叢集置於維護模式。
- 更新HA叢集元件（如有需要）。
- 每次升級一個檔案節點、如下所示：

    - 將其置於待命狀態、並將其服務容錯移轉至次要節點。
    - 升級BeeGFS套件。
    - 回復服務：
- 將叢集移出維護模式。

版本升級注意事項

從BeeGFS 7.2.6或7.3.0版升級

連線型驗證變更

在7.3.1之後發佈的BeeGFS版本將不再允許服務在未指定任何一項的情況下啟動 `connAuthFile` 或設定 `connDisableAuthentication=true` 在服務的組態檔中。強烈建議啟用連線型驗證安全性。請參閱 "BeeGFS連線型驗證" 以取得更多資訊。

根據預設 `beegfs_ha*` 角色將產生及發佈此檔案、並將其新增至位於的Ansible控制節點 `<playbook_directory>/files/beegfs/<beegfs_mgmt_ip_address>_connAuthFile`。。 `beegfs_client` 角色也會檢查此檔案是否存在、並將其提供給用戶端（若有）。

> 如果是 `beegfs_client` 角色未用於設定用戶端、此檔案必須手動散佈至每個用戶端和 `connAuthFile` 中的組態 `beegfs-client.conf` 檔案集以供使用。從未啟用連線型驗證的舊版BeeGFS升級時、除非設定停用連線型驗證、否則用戶端將會失去存取權 `beegfs_ha_conn_auth_enabled: false` 在中 `group_vars/ha_cluster.yml` （不建議）。

如需其他詳細資料和其他組態選項"指定通用檔案節點組態"、請參閱一節中的設定連線驗證步驟。

升級 HA 叢集中的 Pacemaker 和 corosync 套件

請依照下列步驟升級 HA 叢集中的節律器和電暈器同步套件。

總覽

升級 Pacemaker 和 corosync 可確保叢集從新功能，安全修補程式和效能改善中獲益。

升級方法

升級叢集有兩種建議方法：滾動升級或完全關閉叢集。每種方法都有自己的優缺點。您的升級程序可能會因心律調整器版本而異。請參閱 ClusterLabs 的"升級 Pacemaker 叢集"說明文件，以判斷要使用哪種方法。在遵循升級方法之前，請確認：

- NetApp BeeGFS 解決方案支援全新的節律器和電暈器同步套件。
- BeeGFS 檔案系統和 Pacemaker 叢集組態有有效的備份。
- 叢集處於正常狀態。

循環升級

此方法涉及從叢集中移除每個節點，將其升級，然後將其重新引入叢集，直到所有節點都執行新版本為止。這種

方法可讓叢集持續運作，這是較大型 HA 叢集的理想選擇，但在處理過程中會有執行混合版本的風險。在雙節點叢集中，應避免使用此方法。

1. 確認叢集處於最佳狀態，且每個 BeeGFS 服務都在其偏好的節點上執行。如 "檢查叢集的狀態" 需詳細資訊、請參閱。
2. 若要升級節點，請將其置入待命模式，以耗盡（或移動）所有 BeeGFS 服務：

```
pcs node standby <HOSTNAME>
```

3. 執行下列步驟，確認節點的服務已耗盡：

```
pcs status
```

請確定未將任何服務回報為待命節點上的服務 `Started`。

> 視叢集大小而定，服務可能需要幾秒鐘或幾分鐘才能移至姊妹節點。如果 BeeGFS 服務無法在姊妹節點上啟動"疑難排解指南"，請參閱。

4. 關閉節點上的叢集：

```
pcs cluster stop <HOSTNAME>
```

5. 升級節點上的 Pacemaker ， corossync 和 PCS 套件：

> 套件管理員命令會因作業系統而異。下列命令適用於執行 RHEL 8 及後續版本的系統。

```
dnf update pacemaker-<version>
```

```
dnf update corosync-<version>
```

```
dnf update pcs-<version>
```

6. 在節點上啟動 Pacemaker 叢集服務：

```
pcs cluster start <HOSTNAME>
```

7. 如果 `pcs`套件已更新，請使用叢集重新驗證節點：

```
pcs host auth <HOSTNAME>
```

8. 確認此工具的節律器組態仍然有效 `crm_verify`。

> 只需在叢集升級期間驗證一次即可。

```
crm_verify -L -V
```

9. 將節點從待機狀態移出：

```
pcs node unstandby <HOSTNAME>
```

10. 將所有 BeeGFS 服務重新部署回其偏好的節點：

```
pcs resource relocate run
```

11. 針對叢集中的每個節點重複上述步驟，直到所有節點都執行所需的節律器，電量器同步和 PCS 版本為止。
12. 最後，請執行 `pcs status`並確認叢集狀況良好，並 `Current DC`回報所需的節律器版本。

> 如果 `Current DC`報告為「 ixed 版本」，則叢集中的某個節點仍在使用舊版 Pacemaker 執行，需要升級。如果任何升級的節點無法重新加入叢集，或資源無法啟動，請查看叢集記錄，並參閱 Pacemaker 版本說明或使用者指南，以瞭解已知的升級問題。

完成叢集關機

在此方法中，所有叢集節點和資源都會關閉，節點會升級，然後重新啟動叢集。如果節律器和電量器同步版本不支援混合版本的組態，則必須使用此方法。

1. 確認叢集處於最佳狀態，且每個 BeeGFS 服務都在其偏好的節點上執行。如 "檢查叢集的狀態" 需詳細資訊、請參閱。
2. 關閉所有節點上的叢集軟體（ Pacemaker 和 corosync ）。

> 視叢集大小而定，整個叢集可能需要幾秒鐘或幾分鐘才能停止。

```
pcs cluster stop --all
```

3. 當所有節點上的叢集服務都關閉之後，請根據您的需求，升級每個節點上的 Pacemaker ， corosync 和 PCS 套件。

> 套件管理員命令會因作業系統而異。下列命令適用於執行 RHEL 8 及後續版本的系統。

```
dnf update pacemaker-<version>
```

```
dnf update corosync-<version>
```

```
dnf update pcs-<version>
```

4. 升級所有節點之後，請在所有節點上啟動叢集軟體：

```
pcs cluster start --all
```

5. 如果 `pcs`套件已更新，請重新驗證叢集中的每個節點：

```
pcs host auth <HOSTNAME>
```

6. 最後，請執行 `pcs status`並確認叢集狀況良好，並 `Current DC`報告正確的節律器版本。

> 如果 `Current DC`報告為「 ixed 版本」，則叢集中的某個節點仍在使用舊版 Pacemaker 執行，需要升級。

更新檔案節點介面卡韌體

請依照下列步驟，將檔案節點的 ConnectX-7 介面卡更新為最新的韌體。

總覽

可能需要更新 ConnectX-7 介面卡韌體，才能支援新的 MLNX_OFED 驅動程式，啟用新功能或修正錯誤。本指南將使用 NVIDIA 的 `mlxfwmanager`公用程式進行介面卡更新，因為它易於使用且效率高。

升級考量

本指南涵蓋兩種更新 ConnectX-7 介面卡韌體的方法：滾動更新和雙節點叢集更新。根據叢集的大小，選擇適當的更新方法。執行韌體更新之前，請確認：

- 已安裝支援的 MLNX_OFED 驅動程式，請參閱"技術需求"。
- BeeGFS 檔案系統和 Pacemaker 叢集組態有有效的備份。
- 叢集處於正常狀態。

韌體更新準備

建議您使用 NVIDIA `mlxfwmanager` 公用程式來更新節點的介面卡韌體，此韌體與 NVIDIA 的 MLNx_OFED 驅動程式一起提供。開始更新之前，請先從下載介面卡的韌體映像"NVIDIA 的支援網站"，並將其儲存在每個檔案

節點上。

> 對於 Lenovo ConnectX-7 介面卡，請使用 `mlxfwmanager_LES` NVIDIA 頁面上的工具"OEM 韌體"。

滾動更新方法

建議任何具有兩個以上節點的 HA 叢集使用此方法。這種方法涉及一次在一個檔案節點上更新介面卡韌體，讓 HA 叢集能夠保留服務要求，不過建議您在此期間避免服務 I/O 。

1. 確認叢集處於最佳狀態，且每個 BeeGFS 服務都在其偏好的節點上執行。如 "檢查叢集的狀態" 需詳細資訊、請參閱。
2. 選擇要更新的檔案節點，並將其置於待命模式，以從該節點移除（或移動）所有 BeeGFS 服務：

```
pcs node standby <HOSTNAME>
```

3. 執行下列步驟，確認節點的服務已耗盡：

```
pcs status
```

驗證沒有任何服務報告為待命節點上的服務 `Started`。

> 視叢集大小而定， BeeGFS 服務可能需要幾秒鐘或幾分鐘才能移至姊妹節點。如果 BeeGFS 服務無法在姊妹節點上啟動"疑難排解指南"，請參閱。

4. 使用更新介面卡韌體 `mlxfwmanager`。

```
mlxfwmanager -i <path/to/firmware.bin> -u
```

記下 `PCI Device Name`接收韌體更新的每個介面卡的。

5. 使用公用程式重設每個介面卡 `mlxfwreset`以套用新韌體。

> 某些韌體更新可能需要重新開機才能套用更新。請參閱"NVIDIA 的 mlxfwreset 限制"以取得指引。如果需要重新開機，請執行重新開機，而非重設介面卡。

a. 停止 opensm 服務：

```
systemctl stop opensm
```

b. 針對先前註明的每個項目執行下列命令 `PCI Device Name`。

```
mlxfwreset -d <pci_device_name> reset -y
```

c. 啟動 opensm 服務：

```
systemctl start opensm
```

d. 重新啟動 `eseries_nvme_ib.service` 。

```
systemctl restart eseries_nvme_ib.service
```

e. 驗證 E 系列儲存陣列的磁碟區是否存在。

```
multipath -ll
```

1. 執行 `ibstat`並驗證所有介面卡是否以所需的韌體版本執行：

```
ibstat
```

2. 在節點上啟動 Pacemaker 叢集服務：

```
pcs cluster start <HOSTNAME>
```

3. 將節點從待機狀態移出：

```
pcs node unstandby <HOSTNAME>
```

4. 將所有 BeeGFS 服務重新部署回其偏好的節點：

```
pcs resource relocate run
```

對叢集中的每個檔案節點重複這些步驟，直到所有介面卡都已更新為止。

雙節點叢集更新方法

建議只有兩個節點的 HA 叢集採用此方法。這種方法類似於滾動更新，但包含其他步驟，可在某個節點的叢集服務停止時，避免服務停機。

1. 確認叢集處於最佳狀態，且每個 BeeGFS 服務都在其偏好的節點上執行。如 "檢查叢集的狀態" 需詳細資訊、請參閱。

2. 選擇要更新的檔案節點，並將節點置於待命模式，以從該節點移除（或移動）所有 BeeGFS 服務：

```
pcs node standby <HOSTNAME>
```

3. 執行以下步驟，確認節點的資源已耗盡：

```
pcs status
```

驗證沒有任何服務報告為待命節點上的服務 `Started`。

> 視叢集大小而定， BeeGFS 服務可能需要幾秒鐘或幾分鐘的時間，才能在姊妹節點上報告為 `Started`。如果 BeeGFS 服務無法啟動，請"疑難排解指南"參閱。

4. 將叢集置於維護模式。

```
pcs property set maintenance-mode=true
```

5. 使用更新介面卡韌體 `mlxfwmanager`。

```
mlxfwmanager -i <path/to/firmware.bin> -u
```

記下 `PCI Device Name`接收韌體更新的每個介面卡的。

6. 使用公用程式重設每個介面卡 `mlxfwreset`以套用新韌體。

> 某些韌體更新可能需要重新開機才能套用更新。請參閱"NVIDIA 的 mlxfwreset 限制"以取得指引。如果需要重新開機，請執行重新開機，而非重設介面卡。

a. 停止 opensm 服務：

```
systemctl stop opensm
```

b. 針對先前註明的每個項目執行下列命令 `PCI Device Name`。

```
mlxfwreset -d <pci_device_name> reset -y
```

c. 啟動 opensm 服務：

```
systemctl start opensm
```

7. 執行 `ibstat`並驗證所有介面卡是否以所需的韌體版本執行：

```
ibstat
```

8. 在節點上啟動 Pacemaker 叢集服務：

```
pcs cluster start <HOSTNAME>
```

9. 將節點從待機狀態移出：

```
pcs node unstandby <HOSTNAME>
```

10. 將叢集移出維護模式。

```
pcs property set maintenance-mode=false
```

11. 將所有 BeeGFS 服務重新部署回其偏好的節點：

```
pcs resource relocate run
```

對叢集中的每個檔案節點重複這些步驟，直到所有介面卡都已更新為止。

升級 **E-Series** 儲存陣列

請依照下列步驟升級 HA 叢集的 E 系列儲存陣列元件。

總覽

使用最新的韌體、讓 HA 叢集的 NetApp E-Series 儲存陣列保持在最新狀態、以確保最佳效能和更高的安全性。儲存陣列的韌體更新是透過 SANtricity OS ， NVSRAM 和磁碟機韌體檔案來套用。

> ⓘ 雖然儲存陣列可以在 HA 叢集上線時進行升級、但建議您將叢集置於維護模式、以便進行所有升級。

區塊節點升級步驟

下列步驟概述如何使用 `Netapp_Eseries.Santricity` Ansible 集合更新儲存陣列的韌體。在繼續之前、請檢閱"升級考量"以更新 E-Series 系統。

> ⓘ 只能從 11.70.5P1 升級至 SANtricity OS 11.80 或更新版本。在套用進一步升級之前、必須先將儲存陣列升級至 11.70.5P1 。

1. 驗證您的 Ansible 控制節點是否使用最新的 SANtricity Ansible Collection 。
   - 可存取的集合升級 "Ansible Galaxy"，執行下列命令：

     ```
     ansible-galaxy collection install netapp_eseries.santricity --upgrade
     ```

   - 若要進行離線升級"Ansible Galaxy"、請從下載集合 tarball 、將其傳輸至您的控制節點、然後執行：

     ```
     ansible-galaxy collection install netapp_eseries-santricity-
     <VERSION>.tar.gz --upgrade
     ```

     請參閱 "安裝集合" 以取得更多資訊。

2. 取得儲存陣列和磁碟機的最新韌體。
   a. 下載韌體檔案。
      - * SANtricity OS 和 NVSRAS:* 瀏覽至"NetApp 支援網站"、並下載適用於您儲存陣列機型的最新版 SANtricity OS 和 NVSRAS* 。
      - * 磁碟機韌體： * 瀏覽"E-Series 磁碟韌體站台"並下載每個儲存陣列磁碟機機型的最新韌體。
   b. 將 SANtricity OS 、 NVSRAM 和磁碟機韌體檔案儲存在 Ansible 控制節點的 `<inventory_directory>/packages`目錄中。
3. 如有必要、請更新叢集的 Ansible 庫存檔案、以納入所有需要更新的儲存陣列（區塊節點）。如需指引、請參閱"Ansible Inventory Overview"一節。
4. 確保叢集處於最佳狀態，且其慣用節點上的每個 BeeGFS 服務都處於最佳狀態。如 "檢查叢集的狀態" 需詳細資訊、請參閱。
5. 按照中的說明將叢集置於維護模式"將叢集置於維護模式"。
6. 建立名為的新 Ansible 教 `update_block_node_playbook.yml`戰手冊。請將下列內容填入教戰手冊、將 SANtricity OS 、 NVSRAM 和磁碟機韌體版本取代為您想要的升級路徑：

```
- hosts: eseries_storage_systems
  gather_facts: false
  any_errors_fatal: true
  collections:
    - netapp_eseries.santricity
  vars:
    eseries_firmware_firmware: "packages/<SantricityOS>.dlp"
    eseries_firmware_nvsram: "packages/<NVSRAM>.dlp"
    eseries_drive_firmware_firmware_list:
      - "packages/<drive_firmware>.dlp"
    eseries_drive_firmware_upgrade_drives_online: true

  tasks:
    - name: Configure NetApp E-Series block nodes.
      import_role:
        name: nar_santricity_management
```

7. 若要啟動更新、請從 Ansible 控制節點執行下列命令：

```
ansible-playbook -i inventory.yml update_block_node_playbook.yml
```

8. 完成教戰手冊後、請確認每個儲存陣列都處於最佳狀態。
9. 將叢集移出維護模式、並驗證叢集處於最佳狀態、每項 BeeGFS 服務都位於其偏好的節點上。

服務與維護

容錯移轉和容錯回復服務

在叢集節點之間移動BeeGFS服務。

總覽

BeeGFS服務可在叢集中的節點之間進行容錯移轉、以確保當節點發生故障或需要執行計畫性維護時、用戶端能夠繼續存取檔案系統。本節說明系統管理員在從故障中恢復後、或在節點之間手動移動服務時、如何修復叢集。

步驟

容錯移轉與容錯回復

容錯移轉（計畫性）

一般而言、當您需要將單一檔案節點離線以進行維護時、您會想要從該節點移動（或耗盡）所有BeeGFS服務。您可以先將節點置於待命狀態、以達成此目標：

```
pcs node standby <HOSTNAME>
```

使用驗證之後 `pcs status` 所有資源都已在替代檔案節點上重新啟動、您可以視需要關機或對節點進行其他變更。

容錯回復（在計畫性容錯移轉之後）

當您準備好將BeeGFS服務還原至首選節點時、請先執行 `pcs status` 並在「Node List（節點清單）」中驗證狀態是否為「standby（待命）」。如果節點重新開機、則會顯示為離線、直到叢集服務上線為止：

```
pcs cluster start <HOSTNAME>
```

節點上線後、請使用以下功能將其從待命模式中移出：

```
pcs node unstandby <HOSTNAME>
```

最後、將所有BeeGFS服務重新部署回其偏好的節點：

```
pcs resource relocate run
```

容錯回復（非計畫性容錯移轉之後）

如果某個節點發生硬體或其他故障、HA叢集應自動回應並將其服務移至正常節點、讓系統管理員有時間採取修正行動。在繼續之前、"疑難排解"請先參閱一節、以判斷容錯移轉的原因、並解決任何未解決的問題。節點重新開機且正常運作後、您就可以繼續進行容錯回復。

當節點在非計畫性（或計畫性）重新開機之後開機時、叢集服務不會設定為自動啟動、因此您必須先使用以下項目使節點上線：

```
pcs cluster start <HOSTNAME>
```

接下來清除任何資源故障並重設節點的隔離記錄：

```
pcs resource cleanup node=<HOSTNAME>
pcs stonith history cleanup <HOSTNAME>
```

請在中驗證 `pcs status` 節點處於線上且健全狀態。根據預設、BeeGFS服務不會自動容錯回復、以避免意外將資源移回不正常的節點。當您準備好時、將叢集中的所有資源、以下列方式傳回其偏好的節點：

```
pcs resource relocate run
```

將個別BeeGFS服務移至替代檔案節點

將BeeGFS服務永久移至新的檔案節點

如果您想要永久變更個別BeeGFS服務的偏好檔案節點、請調整「Ansible」（可執行）資源清冊、使偏好的節點列在第一位、然後重新執行「Ansible」（可執行）資源清冊。

例如、在此範例檔案中 `inventory.yml` 、 beegfs_01 是執行 BeeGFS 管理服務的慣用檔案節點：

```
mgmt:
  hosts:
    beegfs_01:
    beegfs_02:
```

反轉訂單會使beegfs_02上的管理服務更受歡迎：

```
mgmt:
  hosts:
    beegfs_02:
    beegfs_01:
```

暫時將BeeGFS服務移至替代檔案節點

一般而言、如果某個節點正在進行維護、您會想要使用[Failover and failover countures]（#Failover與容錯回復步驟）（#容 錯移轉與容錯回復）將所有服務移出該節點。

如果由於某些原因、您確實需要將個別服務移至不同的檔案節點執行：

```
pcs resource move <SERVICE>-monitor <HOSTNAME>
```

> 請勿指定個別資源或資源群組。請務必指定您要重新部署BeeGFS服務的監視器名稱。例如，要將 BeeGFS 管理服務移至 beegfs_02 ，請執行 `pcs resource move mgmt-monitor beegfs_02`：。您可以重複此程序、將一或多個服務移出偏好的節點。確認 `pcs status` 新節點上的服務已重新定位 / 啟動。

若要將BeeGFS服務移回其慣用節點、請先清除暫用資源限制（視多項服務需要重複此步驟）：

```
pcs resource clear <SERVICE>-monitor
```

當您準備好將服務實際移回偏好的節點時、請執行：

```
pcs resource relocate run
```

請注意、此命令會重新部署任何不再具有暫用資源限制的服務、而這些服務並未位於偏好的節點上。

將叢集置於維護模式

避免HA叢集意外回應環境中的預期變更。

總覽

將叢集置於維護模式會停用所有資源監控、並防止心臟起搏器移動或以其他方式管理叢集中的資源。所有資源都會繼續在其原始節點上執行、無論是否有暫時性的故障情況、都無法存取這些資源。建議/實用的案例包括：

- 可能會暫時中斷檔案節點與BeeGFS服務之間的連線的網路維護。
- 區塊節點升級。
- 檔案節點作業系統、核心或其他套件更新。

一般而言、將叢集手動置於維護模式的唯一理由、是避免它對環境中的外部變更做出回應。如果叢集中的個別節點需要實體修復、請勿使用維護模式、只要依照上述程序將該節點置於待命狀態即可。請注意、重新執行Anunible會自動將叢集置於維護模式、以利進行大部分的軟體維護、包括升級和組態變更。

步驟

若要檢查叢集是否處於維護模式、請執行：

```
pcs property config
```

`maintenance-mode`如果叢集正常運作，則不會顯示內容。如果叢集目前處於維護模式，則內容會報告為`true`。若要啟用維護模式執行：

```
pcs property set maintenance-mode=true
```

您可以執行PCS狀態並確保所有資源都顯示為「（Unmanaged）」、以進行驗證。若要使叢集離開維護模式執行：

```
pcs property set maintenance-mode=false
```

停止並啟動叢集

正常停止及啟動HA叢集。

總覽

本節說明如何正常關機並重新啟動BeeGFS叢集。可能需要的範例案例包括電力維護或資料中心或機架之間的移轉。

步驟

如果您基於任何原因需要停止整個BeeGFS叢集並關閉所有服務執行：

```
pcs cluster stop --all
```

您也可以在個別節點上停止叢集（這會自動將服務容錯移轉至另一個節點）、不過建議您先將節點置於待命狀態（請參閱"容錯移轉"一節）：

```
pcs cluster stop <HOSTNAME>
```

若要在所有節點上啟動叢集服務和資源、請執行下列步驟：

```
pcs cluster start --all
```

或在特定節點上啟動服務：

```
pcs cluster start <HOSTNAME>
```

此時請執行 `pcs status` 並驗證叢集和BeeGFS服務是否在所有節點上啟動、以及服務是否在您預期的節點上執行。

> 視叢集大小而定，整個叢集可能需要幾秒鐘或幾分鐘的時間才能停止，或顯示為「已啟動 `pcs status`」。如果當機超過五分鐘，則 `pcs cluster <COMMAND>`在執行「 Ctrl+C 」取消命令之前，請先登入叢集的每個節點，然後使用 `pcs status`查看叢集服務（電暈器同步 / 節律器）是否仍在該節點上執行。從叢集仍在作用中的任何節點、您都可以檢查封鎖叢集的資源為何。手動解決此問題、命令應已完成或可重新執行、以停止任何剩餘的服務。

取代檔案節點

如果原始伺服器故障、請更換檔案節點。

總覽

這是更換叢集中檔案節點所需步驟的總覽。這些步驟假設檔案節點因為硬體問題而失敗、並以新的相同檔案節點取代。

步驟：

1. 實體更換檔案節點、並將所有纜線恢復至區塊節點和儲存網路。
2. 在檔案節點上重新安裝作業系統、包括新增Red Hat訂閱。
3. 在檔案節點上設定管理和BMC網路功能。
4. 如果主機名稱、IP、PCIe對邏輯介面對應、或是新檔案節點的其他任何變更、請更新Ansible詳細目錄。如

果節點被相同的伺服器硬體所取代、而且您使用的是原始網路組態、則通常不需要這麼做。

a. 例如、如果主機名稱已變更、請建立（或重新命名）節點的庫存檔案 (`host_vars/<NEW_NODE>.yml`` ）、然後在Ansible庫存檔案中 (`inventory.yml`）、將舊節點的名稱改為新節點名稱：

```
all:
    ...
    children:
    ha_cluster:
        children:
        mgmt:
            hosts:
            node_h1_new:   # Replaced "node_h1" with "node_h1_new"
            node_h2:
```

5. 從叢集中的其他節點之一移除舊節點：`pcs cluster node remove <HOSTNAME>`。

執行此步驟之前、請勿繼續。

6. 在Ansible控制節點上：

a. 移除舊的SSH金鑰：

```
`ssh-keygen -R <HOSTNAME_OR_IP>`
```

b. 將無密碼SSH設定為取代節點：

```
ssh-copy-id <USER>@<HOSTNAME_OR_IP>
```

7. 重新執行Ansible playbook以設定節點並將其新增至叢集：

```
ansible-playbook -i <inventory>.yml <playbook>.yml
```

8. 此時、請執行 `pcs status` 並確認已列出更換的節點、並正在執行服務。

擴充或縮小叢集

新增或移除叢集的建置區塊。

總覽

本節說明各種考量事項和選項、以調整BeeGFS HA叢集的大小。一般而言、叢集大小是透過新增或移除建置區塊來調整、而建置區塊通常是以HA配對形式設定兩個檔案節點。如果需要、也可以新增或移除個別的檔案節點（或其他類型的叢集節點）。

將建置區塊新增至叢集

考量

透過新增額外的建置區塊來擴充叢集、是一項簡單的程序。開始之前、請記住每個HA叢集中叢集節點的最小和最大數量限制、並決定是否要將節點新增至現有HA叢集、或是建立新的HA叢集。通常每個建置區塊都由兩個檔案節點組成、但每個叢集的節點數最少為三個節點（以建立仲裁）、建議（測試）最多為十個節點。在進階案例中、您可以新增單一「tiebreaker」節點、而在部署雙節點叢集時、該節點不會執行任何BeeGFS服務。如果您正在考慮進行這類部署、請聯絡NetApp支援部門。

在決定如何擴充叢集時、請謹記這些限制和任何預期的未來叢集成長。例如、如果您有六個節點叢集、需要再新增四個節點、建議您只是啟動一個新的HA叢集。

請記住、單一BeeGFS檔案系統可由多個獨立的HA叢集組成。如此一來、檔案系統就能繼續擴充、遠遠超過基礎HA叢集元件的建議/硬限制。

步驟

將建置區塊新增至叢集時、您需要 `host_vars`為每個新的檔案節點和區塊節點（ E-Series 陣列）建立檔案。這些主機的名稱需要新增至庫存、以及要建立的新資源。 `group_vars`需要為每個新資源建立對應的檔案。如"使用自訂架構"需詳細資訊、請參閱一節。

建立正確的檔案之後、只需要使用以下命令重新執行自動化作業：

```
ansible-playbook -i <inventory>.yml <playbook>.yml
```

從叢集移除建置區塊

當您需要淘汰建置區塊時、請謹記下列幾點考量事項：

- 此建置區塊中執行哪些BeeGFS服務？
- 是否只有檔案節點即將淘汰、而且區塊節點應附加至新的檔案節點？
- 如果整個建置區塊都已淘汰、資料是否應該移至新的建置區塊、分散到叢集中的現有節點、或移至新的BeeGFS檔案系統或其他儲存系統？
- 這種情況是否會在停電期間發生、還是應該在不中斷營運的情況下進行？
- 建置區塊是否正在使用中、或主要包含不再使用的資料？

由於可能的起點和所需的終端狀態各不相同、請聯絡NetApp支援部門、以便我們根據您的環境和需求、找出並協助實作最佳策略。

疑難排解

BeeGFS HA叢集疑難排解。

總覽

本節將說明如何調查及疑難排解在操作BeeGFS HA叢集時可能發生的各種故障和其他情況。

疑難排解指南

正在調查意外的容錯移轉

當節點被意外隔離、其服務移至另一個節點時、第一步應該是查看叢集是否在底部顯示任何資源故障 `pcs status`。通常、如果成功完成隔離並在另一個節點上重新啟動資源、則不會出現任何情況。

一般而言、下一步是使用來搜尋整個系統記錄 `journalctl` 在其餘任一檔案節點上（所有節點上的心臟起搏器記錄都會同步）。如果您知道故障發生的時間、可以在故障發生前立即開始搜尋（建議至少提前10分鐘）：

```
journalctl --since "<YYYY-MM-DD HH:MM:SS>"
```

下列各節顯示您可以在記錄中加入的一般文字、以進一步縮小調查範圍。

調查/解決的步驟

步驟1：檢查BeeGFS監視器是否偵測到故障：

如果容錯移轉是由BeeGFS監控器觸發、您應該會看到錯誤（如果沒有、請繼續下一步）。

```
journalctl --since "<YYYY-MM-DD HH:MM:SS>" | grep -i unexpected
[...]
Jul 01 15:51:03 beegfs_01 pacemaker-schedulerd[9246]:  warning: Unexpected
result (error: BeeGFS service is not active!) was recorded for monitor of
meta_08-monitor on beegfs_02 at Jul  1 15:51:03 2022
```

在此例中、BeeGFS服務meta_08因為某些原因而停止。若要繼續疑難排解、我們應該開機 beegfs_02 、並檢閱服務的記錄、網址為： `/var/log/beegfs-meta-meta_08_tgt_0801.log`。例如、BeeGFS服務可能因為內部問題或節點問題而發生應用程式錯誤。

> 不同於來自心臟起搏器的記錄、BeeGFS服務的記錄不會分散到叢集中的所有節點。若要調查這些故障類型、必須從發生故障的原始節點取得記錄。

監視器可能回報的問題包括：

- 無法存取目標！
  - 說明：表示無法存取區塊磁碟區。
  - 疑難排解：
    - 如果服務也無法在替代檔案節點上啟動、請確認區塊節點正常運作。
    - 請檢查是否有任何實體問題、以免從此檔案節點存取區塊節點、例如發生故障的InfiniBand介面卡或纜線。
- 無法連線到網路！
  - 說明：用戶端用來連線至此BeeGFS服務的介面卡均未連線。
  - 疑難排解：

    - 如果有多個/所有檔案節點受到影響、請檢查網路上是否有故障用於連接BeeGFS用戶端和檔案系統。
    - 請檢查是否有任何實體問題、例如會使此檔案節點無法存取用戶端、例如發生故障的InfiniBand介面卡或纜線。

- BeeGFS服務未啟用！
  - 說明：BeeGFS服務意外停止。
  - 疑難排解：
    - 在報告錯誤的檔案節點上、檢查受影響BeeGFS服務的記錄、以查看是否報告當機。如果發生這種情況、請利用NetApp支援開啟案例、以便調查當機事件。
    - 如果BeeGFS記錄中沒有報告錯誤、請檢查日誌記錄、查看系統是否記錄服務停止的原因。在某些情況下、BeeGFS服務可能沒有機會在程序終止之前記錄任何訊息（例如有人執行） `kill -9 <PID>` ）。

步驟**2**：檢查節點是否意外離開叢集

如果節點發生災難性硬體故障（例如主機板當機）、或發生核心異常或類似軟體問題、BeeGFS監視器將不會報告錯誤。請改為尋找主機名稱、您應該會看到來自心臟起搏器的訊息、指出節點意外遺失：

```
journalctl --since "<YYYY-MM-DD HH:MM:SS>" | grep -i <HOSTNAME>
[...]
Jul 01 16:18:01 beegfs_01 pacemaker-attrd[9245]:  notice: Node beegfs_02
state is now lost
Jul 01 16:18:01 beegfs_01 pacemaker-controld[9247]:  warning:
Stonith/shutdown of node beegfs_02 was not expected
```

步驟**3**：驗證起搏器是否能夠隔離節點

在所有情況下、您都應該看到心臟起搏器試圖將節點隔離、以驗證它實際上是否離線（確切的訊息可能會因隔離原因而異）：

```
Jul 01 16:18:02 beegfs_01 pacemaker-schedulerd[9246]:  warning: Cluster
node beegfs_02 will be fenced: peer is no longer part of the cluster
Jul 01 16:18:02 beegfs_01 pacemaker-schedulerd[9246]:  warning: Node
beegfs_02 is unclean
Jul 01 16:18:02 beegfs_01 pacemaker-schedulerd[9246]:  warning: Scheduling
Node beegfs_02 for STONITH
```

如果隔離動作成功完成、您會看到如下訊息：

```
Jul 01 16:18:14 beegfs_01 pacemaker-fenced[9243]:  notice: Operation 'off'
[2214070] (call 27 from pacemaker-controld.9247) for host 'beegfs_02' with
device 'fence_redfish_2' returned: 0 (OK)
Jul 01 16:18:14 beegfs_01 pacemaker-fenced[9243]:  notice: Operation 'off'
targeting beegfs_02 on beegfs_01 for pacemaker-
controld.9247@beegfs_01.786df3a1: OK
Jul 01 16:18:14 beegfs_01 pacemaker-controld[9247]:  notice: Peer
beegfs_02 was terminated (off) by beegfs_01 on behalf of pacemaker-
controld.9247: OK
```

如果隔離動作因為某種原因而失敗、BeeGFS服務將無法在另一個節點上重新啟動、以避免資料毀損的風險。如果隔離設備（PDU或BMC）無法存取或設定錯誤、則這是需要個別調查的問題。

處理失敗的資源動作（可在**PCS**狀態的底部找到）

如果執行BeeGFS服務所需的資源失敗、BeeGFS監視器會觸發容錯移轉。如果發生這種情況，則可能在底部沒有列出“故障資源操作” `pcs status`，您應該參考有關如何操作的步驟"非計畫性容錯移轉之後的容錯回復"。

否則、通常只會有兩種情況出現「失敗的資源動作」。

調查/解決的步驟

案例**1**：使用隔離代理程式偵測到暫時性或永久性問題、然後重新啟動或移至其他節點。

有些屏障代理程式比其他代理程式更可靠、而且每個代理程式都會實作自己的監控方法、以確保屏障設備已就緒。尤其是RedfISH屏障代理程式已被視為回報失敗的資源動作、例如下列動作、即使它仍會顯示為「已啟動」：

```
  * fence_redfish_2_monitor_60000 on beegfs_01 'not running' (7):
call=2248, status='complete', exitreason='', last-rc-change='2022-07-26
08:12:59 -05:00', queued=0ms, exec=0ms
```

回報特定節點上失敗資源動作的隔離代理程式、預期不會觸發該節點上執行BeeGFS服務的容錯移轉。只需在相同或不同的節點上自動重新啟動即可。

解決步驟：

1. 如果隔離代理程式持續拒絕在所有或一部分節點上執行、請檢查這些節點是否能夠連線至隔離代理程式、並確認隔離代理程式已在「Ansible」（可隔離）資源清冊中正確設定。
   a. 例如、如果RedfISH（BMC）屏障代理程式與負責隔離的節點相同、而且OS管理和BMC IP位於同一個實體介面上、則某些網路交換器組態將不允許兩個介面之間進行通訊（以防止網路迴圈）。根據預設、HA叢集會嘗試避免在其負責隔離的節點上放置隔離代理程式、但在某些情況/組態中可能會發生這種情況。
2. 一旦所有問題都解決（或問題似乎是暫時性的）、請執行 `pcs resource cleanup` 以重設失敗的資源動作。

案例**2**：**BeeGFS**監視器偵測到問題並觸發容錯移轉、但由於某些原因、資源無法在次要節點上啟動。

如果啟用了隔離功能、且資源未被封鎖、無法在原始節點上停止（請參閱「待命（故障時）」的疑難排解一節）、最可能的原因包括在次要節點上啟動資源時發生問題、因為：

- 次要節點已離線。
- 實體或邏輯組態問題使次要實體無法存取做為BeeGFS目標的區塊磁碟區。

解決步驟：

1. 針對失敗資源動作中的每個項目：
   a. 確認失敗的資源動作是啟動作業。
   b. 根據指示的資源和故障資源動作中指定的節點：
      i. 尋找並修正任何會使節點無法啟動指定資源的外部問題。例如、如果BeeGFS IP位址（浮動IP）無法啟動、請確認至少有一個必要的介面已連線/連線、並連接至正確的網路交換器。如果BeeGFS目標（區塊裝置/ E系列磁碟區）故障、請驗證與後端區塊節點的實體連線是否如預期連接、並驗證區塊節點是否正常。
   c. 如果沒有明顯的外部問題、而且您想找出造成此事件的根本原因、建議您先向NetApp支援部門提出案例進行調查、然後再繼續進行、因為下列步驟可能會使根本原因分析（RCA）變得具有挑戰性/不可能。
2. 解決任何外部問題之後：
   a. 從Ansible inventory.yml檔案中註解任何無法運作的節點、然後重新執行完整的可執行教戰手冊、以確保所有邏輯組態都已在次要節點上正確設定。
      i. 附註：當節點正常運作且您已準備好容錯回復時、請別忘了取消註釋這些節點、然後重新執行教戰手冊。
   b. 或者、您也可以嘗試手動恢復叢集：
      i. 使用下列方法將任何離線節點重新連線： `pcs cluster start <HOSTNAME>`
      ii. 使用下列方法清除所有失敗的資源動作： `pcs resource cleanup`
      iii. 執行PCS狀態、並驗證所有服務是否如預期啟動。
      iv. 如有需要、請執行 `pcs resource relocate run` 可將資源移回其首選節點（如果可用）。

常見問題

BeeGFS服務不會在要求時進行容錯移轉或容錯回復

可能的問題： The `pcs resource relocate` 執行命令已執行、但從未成功完成。

*如何檢查：*執行 `pcs constraint --full` 並使用ID檢查任何位置限制 `pcs-relocate-<RESOURCE>`。

*如何解決：*執行 `pcs resource relocate clear` 然後重新執行 `pcs constraint --full` 以驗證是否移除額外的限制。

當隔離功能停用時、**PC**狀態中的一個節點會顯示「待命（故障時）」

*可能的問題：*起搏器無法成功確認故障節點上的所有資源均已停止。

如何解決：

1. 執行 `pcs status` 並檢查是否有任何未「啟動」的資源、或是在輸出底部顯示錯誤、並解決任何問題。
2. 可使節點恢復聯機運行 `pcs resource cleanup --node=<HOSTNAME>`。

在發生非預期的容錯移轉之後、啟用隔離功能時、資源會在**PCS**狀態中顯示「已啟動（故障時）」

*可能的問題：*發生觸發容錯移轉的問題、但心臟起搏器無法驗證節點是否已被隔離。這可能是因為屏障設定錯誤、或屏障代理程式發生問題（例如：PDU已從網路中斷連線）。

如何解決：

1. 驗證節點是否確實關機。

> 如果您指定的節點實際上並未關閉、而是執行叢集服務或資源、則會發生資料毀損/叢集故障。

2. 手動確認隔離： `pcs stonith confirm <NODE>`

此時、服務應完成容錯移轉、並在另一個正常節點上重新啟動。

常見疑難排解工作

重新啟動個別**BeeGFS**服務

通常、如果需要重新啟動BeeGFS服務（例如為了協助變更組態）、則應更新「Ansible」（可存取）清單並重新執行播放手冊。在某些情況下、可能需要重新啟動個別服務、以加快疑難排解的速度、例如變更記錄層級、而不需要等待整個方針執行。

> 除非在Ansible庫存中也新增任何手動變更、否則下次執行Ansible教戰手冊時將會還原這些變更。

選項**1**：系統**d**控制的重新啟動

如果BeeGFS服務可能無法以新組態正確重新啟動、請先將叢集置於維護模式、以防止BeeGFS監視器偵測到服務停止、並觸發不想要的容錯移轉：

```
pcs property set maintenance-mode=true
```

如有需要、請在進行任何服務組態變更 `/mnt/<SERVICE_ID>/_config/beegfs-.conf` （範例：`/mnt/meta_01_tgt_0101/metadata_config/beegfs-meta.conf`）然後使用systemd重新啟動：

```
systemctl restart beegfs-*@<SERVICE_ID>.service
```

範例： `systemctl restart beegfs-meta@meta_01_tgt_0101.service`

選項**2**：心律調整器控制的重新啟動

如果您不擔心新的組態可能會導致服務意外停止（例如、只是變更記錄層級）、或是處於維護期間、不擔心停機、您只需重新啟動BeeGFS監控器、即可取得您要重新啟動的服務：

```
pcs resource restart <SERVICE>-monitor
```

例如、若要重新啟動BeeGFS管理服務： `pcs resource restart mgmt-monitor`

版權資訊

Copyright © 2025 NetApp, Inc. 版權所有。台灣印製。非經版權所有人事先書面同意，不得將本受版權保護文件的任何部分以任何形式或任何方法（圖形、電子或機械）重製，包括影印、錄影、錄音或儲存至電子檢索系統中。

由 NetApp 版權資料衍伸之軟體必須遵守下列授權和免責聲明：

此軟體以 NETAPP「原樣」提供，不含任何明示或暗示的擔保，包括但不限於有關適售性或特定目的適用性之擔保，特此聲明。於任何情況下，就任何已造成或基於任何理論上責任之直接性、間接性、附隨性、特殊性、懲罰性或衍生性損害（包括但不限於替代商品或服務之採購；使用、資料或利潤上的損失；或企業營運中斷），無論是在使用此軟體時以任何方式所產生的契約、嚴格責任或侵權行為（包括疏忽或其他）等方面，NetApp 概不負責，即使已被告知有前述損害存在之可能性亦然。

NetApp 保留隨時變更本文所述之任何產品的權利，恕不另行通知。NetApp 不承擔因使用本文所述之產品而產生的責任或義務，除非明確經過 NetApp 書面同意。使用或購買此產品並不會在依據任何專利權、商標權或任何其他 NetApp 智慧財產權的情況下轉讓授權。

本手冊所述之產品受到一項（含）以上的美國專利、國外專利或申請中專利所保障。

有限權利說明：政府機關的使用、複製或公開揭露須受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中的「技術資料權利 - 非商業項目」條款 (b)(3) 小段所述之限制。

此處所含屬於商業產品和 / 或商業服務（如 FAR 2.101 所定義）的資料均為 NetApp, Inc. 所有。根據本協議提供的所有 NetApp 技術資料和電腦軟體皆屬於商業性質，並且完全由私人出資開發。 美國政府對於該資料具有非專屬、非轉讓、非轉授權、全球性、有限且不可撤銷的使用權限，僅限於美國政府為傳輸此資料所訂合約所允許之範圍，並基於履行該合約之目的方可使用。除非本文另有規定，否則未經 NetApp Inc. 事前書面許可，不得逕行使用、揭露、重製、修改、履行或展示該資料。美國政府授予國防部之許可權利，僅適用於 DFARS 條款 252.227-7015(b)（2014 年 2 月）所述權利。

商標資訊

NETAPP、NETAPP 標誌及 http://www.netapp.com/TM 所列之標章均為 NetApp, Inc. 的商標。文中所涉及的所有其他公司或產品名稱，均為其各自所有者的商標，不得侵犯。