



定義檔案和區塊節點

BeeGFS on NetApp with E-Series Storage

NetApp
January 27, 2026

目錄

定義檔案和區塊節點	1
設定個別檔案節點	1
總覽	1
步驟	1
設定個別區塊節點	4
總覽	4
步驟	4
指定通用檔案節點組態	8
總覽	8
步驟	8
指定通用區塊節點組態	14
總覽	14
步驟	15

定義檔案和區塊節點

設定個別檔案節點

使用主機變數 (host_vars) 指定個別檔案節點的組態。

總覽

本節將逐步介紹填入的內容 `host_vars/<FILE_NODE_HOSTNAME>.yaml` 叢集中每個檔案節點的檔案。這些檔案應僅包含特定檔案節點專屬的組態。這通常包括：

- 定義IP或Ansible主機名稱應用於連線至節點。
- 設定用於HA叢集服務（起搏器和電量器同步）的其他介面和叢集IP、以便與其他檔案節點通訊。根據預設、這些服務使用的網路與管理介面相同、但應該有額外的介面可供備援。一般做法是在儲存網路上定義額外的IP、避免需要額外的叢集或管理網路。
 - 任何用於叢集通訊的網路效能、對檔案系統效能並不重要。使用預設叢集組態時、通常至少有 1 Gb/s 網路可為叢集作業提供足夠的效能、例如同步節點狀態和協調叢集資源狀態變更。緩慢/忙碌的網路可能會導致資源狀態變更的時間比平常長、而且在極端的情況下、如果無法在合理的時間範圍內傳送訊號、則可能會導致節點從叢集中被逐出。
- 設定介面、用於透過所需的傳輸協定連線至區塊節點（例如：iSCSI/iSER、NVMe/IB、NVMe/RoCE、FCP等）。

步驟

"[規劃檔案系統](#)"針對叢集中的每個檔案節點、參考一節中定義的IP定址方案會建立檔案 `host_vars/<FILE_NODE_HOSTNAME>/yaml`、並填入檔案、如下所示：

1. 在頂端指定Ansible應使用的IP或主機名稱來SSH連接節點並加以管理：

```
ansible_host: "<MANAGEMENT_IP>"
```

2. 設定可用於叢集流量的其他IP：

- a. 如果網路類型為 "InfiniBand (使用IPoIB)"：

```
eseries_ipoib_interfaces:  
- name: <INTERFACE> # Example: ib0 or ilb  
  address: <IP/SUBNET> # Example: 100.127.100.1/16  
- name: <INTERFACE> # Additional interfaces as needed.  
  address: <IP/SUBNET>
```

- b. 如果網路類型為 "融合式乙太網路上的RDMA (RoCE)"：

```
eseries_roce_interfaces:
- name: <INTERFACE> # Example: eth0.
  address: <IP/SUBNET> # Example: 100.127.100.1/16
- name: <INTERFACE> # Additional interfaces as needed.
  address: <IP/SUBNET>
```

c. 如果網路類型為 "乙太網路 (僅TCP、無RDMA) "：

```
eseries_ip_interfaces:
- name: <INTERFACE> # Example: eth0.
  address: <IP/SUBNET> # Example: 100.127.100.1/16
- name: <INTERFACE> # Additional interfaces as needed.
  address: <IP/SUBNET>
```

3. 指出叢集流量應使用哪些IP、優先IP列在較高的位置：

```
beegfs_ha_cluster_node_ips:
- <MANAGEMENT_IP> # Including the management IP is typically but not
  required.
- <IP_ADDRESS> # Ex: 100.127.100.1
- <IP_ADDRESS> # Additional IPs as needed.
```



在步驟2中設定的IPS不會作為叢集IP使用、除非包含在中 `beegfs_ha_cluster_node_ips` 清單。這可讓您使用Ansible來設定其他IP /介面、以便在需要時用於其他用途。

4. 如果檔案節點需要透過IP型傳輸協定來通訊區塊節點、則必須在適當的介面上設定IP、以及安裝/設定該傳輸協定所需的任何套件。

a. 如果使用 "iSCSI"：

```
eseries_iscsi_interfaces:
- name: <INTERFACE> # Example: eth0.
  address: <IP/SUBNET> # Example: 100.127.100.1/16
```

b. 如果使用 "商用"：

```

eseries_ib_iser_interfaces:
- name: <INTERFACE> # Example: ib0.
  address: <IP/SUBNET> # Example: 100.127.100.1/16
  configure: true # If the file node is directly connected to the
block node set to true to setup OpenSM.

```

c. 如果使用 "NVMe / IB" :

```

eseries_nvme_ib_interfaces:
- name: <INTERFACE> # Example: ib0.
  address: <IP/SUBNET> # Example: 100.127.100.1/16
  configure: true # If the file node is directly connected to the
block node set to true to setup OpenSM.

```

d. 如果使用 "NVMe / RoCE" :

```

eseries_nvme_roce_interfaces:
- name: <INTERFACE> # Example: eth0.
  address: <IP/SUBNET> # Example: 100.127.100.1/16

```

e. 其他通訊協定 :

- i. 如果使用 "NVMe / FC" , 不需要設定個別介面。BeeGFS叢集部署會自動偵測傳輸協定、並視需要安裝/設定需求。如果您使用Fabric來連接檔案和區塊節點、請確保交換器已依照NetApp和交換器廠商的最佳實務做法正確分區。
- ii. 使用FCP或SAS不需要安裝或設定其他軟體。如果使用FCP、請確定交換器已正確分區如下 "NetApp" 以及交換器廠商的最佳實務做法。
- iii. 目前不建議使用IB SRP。視E系列區塊節點支援的項目而定、請使用NVMe/IB或iSER。

按一下 "[請按這裡](#)" 例如、代表單一檔案節點的完整庫存檔案。

進階：在乙太網路與InfiniBand模式之間切換NVIDIA ConnectX VPI介面卡

NVIDIA ConnectX-Virtual Protocol Interconnect (VPI) 介面卡支援InfiniBand和乙太網路作為傳輸層。在模式之間切換不會自動協商，而且必須使用中隨附的工具進行設定 `mstconfig mstflint`，這是屬於的開放原始碼套件 "[NVIDIA Firmare 工具 \(MFT\)](#)"。只需變更一次介面卡模式即可。這可以手動完成、也可以納入 Ansible 庫存、做為任何使用庫存區段設定的介面一部分 `eseries-`

`[ib|ib_iser|ipoib|nvme_ib|nvme_roce|roce]_interfaces:`、以自動檢查 / 套用。

例如、將InfiniBand模式中的介面電流變更為乙太網路、以便用於RoCE：

1. 針對您要設定的每個介面指定 `mstconfig` 做為指定的對應 (或字典) `LINK_TYPE_P<N>` 其中 `<N>` 由HCA的介面連接埠號碼決定。◦ `<N>` 值可透過執行來決定 `grep PCI_SLOT_NAME /sys/class/net/<INTERFACE_NAME>/device/uevent` 並從PCI插槽名稱新增1至最後一個數字、然後轉換為十進位。

- a. 例如給定的 `PCI_SLOT_NAME=0000:2f:00.2` (2 + 1 → HCA連接埠3) → `LINK_TYPE_P3: eth` :

```
eseries_roce_interfaces:  
- name: <INTERFACE>  
  address: <IP/SUBNET>  
  mstconfig:  
    LINK_TYPE_P3: eth
```

如需其他詳細資料、請參閱 ["NetApp E系列主機系列文件"](#) 針對您使用的介面類型/傳輸協定。

設定個別區塊節點

使用主機變數 (`host_vars`) 指定個別區塊節點的組態。

總覽

本節將逐步介紹填入的內容 `host_vars/<BLOCK_NODE_HOSTNAME>.yaml` 叢集中每個區塊節點的檔案。這些檔案應僅包含特定區塊節點專屬的組態。這通常包括：

- 系統名稱 (如System Manager所示)。
- 其中一個控制器的HTTPS URL (用於使用REST API管理系統)。
- 用於連線至此區塊節點的儲存傳輸協定檔案節點。
- 設定主機介面卡 (HIC) 連接埠、例如IP位址 (如有需要)。

步驟

"[規劃檔案系統](#)"針對叢集中的每個區塊節點、參考一節中定義的 IP 定址方案 `host_vars/<BLOCK_NODE_HOSTNAME>/yaml`、建立檔案並填入如下內容：

1. 在頂端指定其中一個控制器的系統名稱和HTTPS URL：

```
eseries_system_name: <SYSTEM_NAME>  
eseries_system_api_url:  
https://<MANAGEMENT_HOSTNAME_OR_IP>:8443/devmgr/v2/
```

2. 選取 "[傳輸協定](#)" 檔案節點將用於連線至此區塊節點：

- a. 支援的傳輸協定：`auto`、`iscsi`、`fc`、`sas`、`ib_srp`、`ib_iser`、`nvme_ib`、`nvme_fc`、`nvme_roce`。

```
eseries_initiator_protocol: <PROTOCOL>
```

3. 視使用中的傳輸協定而定、HIC連接埠可能需要額外的組態。必要時、應定義HIC連接埠組態、使每個控制

器組態的頂端項目對應於每個控制器上的實體最左側連接埠、而底部連接埠則對應最右側的連接埠。所有連接埠都需要有效的組態、即使目前未使用。



如果您使用的是具有EF600區塊節點的HDR (200GB) InfiniBand或200GB RoCE、請參閱以下章節。

a. 對於iSCSI：

```
eseries_controller_iscsi_port:
  controller_a:          # Ordered list of controller A channel
  definition.
    - state:             # Whether the port should be enabled.
  Choices: enabled, disabled
    config_method:      # Port configuration method Choices: static,
  dhcp
    address:            # Port IPv4 address
    gateway:            # Port IPv4 gateway
    subnet_mask:        # Port IPv4 subnet_mask
    mtu:                # Port IPv4 mtu
    - (...)             # Additional ports as needed.
  controller_b:         # Ordered list of controller B channel
  definition.
    - (...)             # Same as controller A but for controller B

# Alternatively the following common port configuration can be
# defined for all ports and omitted above:
eseries_controller_iscsi_port_state: enabled          # Generally
specifies whether a controller port definition should be applied
Choices: enabled, disabled
eseries_controller_iscsi_port_config_method: dhcp     # General port
configuration method definition for both controllers. Choices:
static, dhcp
eseries_controller_iscsi_port_gateway:                # General port
IPv4 gateway for both controllers.
eseries_controller_iscsi_port_subnet_mask:           # General port
IPv4 subnet mask for both controllers.
eseries_controller_iscsi_port_mtu: 9000              # General port
maximum transfer units (MTU) for both controllers. Any value greater
than 1500 (bytes).
```

b. 對於iSER：

```
eseries_controller_ib_iser_port:
  controller_a:      # Ordered list of controller A channel address
definition.
  -                 # Port IPv4 address for channel 1
  - (...)           # So on and so forth
  controller_b:      # Ordered list of controller B channel address
definition.
```

c. 適用於NVMe/IB :

```
eseries_controller_nvme_ib_port:
  controller_a:      # Ordered list of controller A channel address
definition.
  -                 # Port IPv4 address for channel 1
  - (...)           # So on and so forth
  controller_b:      # Ordered list of controller B channel address
definition.
```

d. 適用於NVMe / RoCE :

```

eseries_controller_nvme_roce_port:
  controller_a:          # Ordered list of controller A channel
definition.
  - state:              # Whether the port should be enabled.
  config_method:       # Port configuration method Choices: static,
dhcp
  address:             # Port IPv4 address
  subnet_mask:        # Port IPv4 subnet_mask
  gateway:            # Port IPv4 gateway
  mtu:                # Port IPv4 mtu
  speed:              # Port IPv4 speed
  controller_b:       # Ordered list of controller B channel
definition.
  - (...)              # Same as controller A but for controller B

# Alternatively the following common port configuration can be
defined for all ports and omitted above:
eseries_controller_nvme_roce_port_state: enabled          # Generally
specifies whether a controller port definition should be applied
Choices: enabled, disabled
eseries_controller_nvme_roce_port_config_method: dhcp     # General
port configuration method definition for both controllers. Choices:
static, dhcp
eseries_controller_nvme_roce_port_gateway:                # General
port IPv4 gateway for both controllers.
eseries_controller_nvme_roce_port_subnet_mask:           # General
port IPv4 subnet mask for both controllers.
eseries_controller_nvme_roce_port_mtu: 4200              # General
port maximum transfer units (MTU). Any value greater than 1500
(bytes).
eseries_controller_nvme_roce_port_speed: auto            # General
interface speed. Value must be a supported speed or auto for
automatically negotiating the speed with the port.

```

e. FC和SAS傳輸協定不需要額外的組態。不正確建議使用SRP。

如需設定HIC連接埠和主機傳輸協定的其他選項、包括設定iSCSI CHAP的能力、請參閱 "文件" 隨附SANtricity 於此系列產品。注意：部署BeeGFS時、儲存資源池、磁碟區組態及其他資源配置方面將會設定在其他位置、不應在此檔案中定義。

按一下 "請按這裡" 例如、代表單一區塊節點的完整庫存檔案。

在NetApp EF600區塊節點上使用HDR(200Gb) InfiniBand或200GB RoCE：

若要將HDR (200GB) InfiniBand搭配EF600使用、必須為每個實體連接埠設定第二個「虛擬」IP。以下是正確設定配備雙埠 InfiniBand HDR HIC 的 EF600 的範例：

```
eseries_controller_nvme_ib_port:
  controller_a:
    - 192.168.1.101 # Port 2a (virtual)
    - 192.168.2.101 # Port 2b (virtual)
    - 192.168.1.100 # Port 2a (physical)
    - 192.168.2.100 # Port 2b (physical)
  controller_b:
    - 192.168.3.101 # Port 2a (virtual)
    - 192.168.4.101 # Port 2b (virtual)
    - 192.168.3.100 # Port 2a (physical)
    - 192.168.4.100 # Port 2b (physical)
```

指定通用檔案節點組態

使用群組變數（群組_vars）指定通用檔案節點組態。

總覽

所有檔案節點的組態均應定義於 `group_vars/ha_cluster.yml`。通常包括：

- 如何連線及登入每個檔案節點的詳細資料。
- 通用網路組態。
- 是否允許自動重新開機。
- 如何設定防火牆和SELinux狀態。
- 叢集組態、包括警示和隔離。
- 效能調校：
- 通用BeeGFS服務組態。



此檔案中設定的選項也可在個別檔案節點上定義、例如使用混合式硬體模型、或是每個節點的密碼不同。個別檔案節點的組態優先於此檔案中的組態。

步驟

建立檔案 `group_vars/ha_cluster.yml` 並填入如下內容：

1. 指出Ansible Control節點應如何與遠端主機進行驗證：

```
ansible_ssh_user: root
ansible_become_password: <PASSWORD>
```



尤其是在正式作業環境中、請勿以純文字儲存密碼。請改用Ansible Vault (請參閱 "使用Ansible Vault加密內容") 或 `--ask-become-pass` 執行教戰手冊時的選項。如果是 `ansible_ssh_user` 已經是root、您可以選擇省略 `ansible_become_password`。

2. 如果您在乙太網路或InfiniBand介面上設定靜態IP (例如叢集IP) 、且多個介面位於同一個IP子網路 (例如、如果ib0使用192.168.1.10/24、而ib1使用192.168.1.11/24) 、必須設定其他IP路由表和規則、多重主目錄支援才能正常運作。只需啟用提供的網路介面組態掛勾、如下所示：

```
eseries_ip_default_hook_templates:  
  - 99-multihoming.j2
```

3. 部署叢集時、視儲存傳輸協定而定、可能需要重新啟動節點、以協助探索遠端區塊裝置 (E系列磁碟區) 或套用組態的其他層面。依預設、節點會在重新開機前提示、但您可以指定下列項目、讓節點自動重新啟動：

```
eseries_common_allow_host_reboot: true
```

- a. 根據預設、重新開機後、若要確保區塊裝置和其他服務已就緒、Ansible將會等到系統完成 `default.target` 在繼續部署之前就已到達。在使用NVMe / IB的某些情況下、這可能不夠長、無法初始化、探索及連線至遠端裝置。這可能會導致自動部署提早繼續、而且失敗。若要避免這種情況、使用NVMe / IB時也必須定義下列項目：

```
eseries_common_reboot_test_command: "! systemctl status  
eseries_nvme_ib.service || systemctl --state=exited | grep  
eseries_nvme_ib.service"
```

4. BeeGFS和HA叢集服務需要多個防火牆連接埠才能進行通訊。除非您想要手動設定防火牆 (不建議) 、否則請指定下列項目、以建立必要的防火牆區域並自動開啟連接埠：

```
beegfs_ha_firewall_configure: True
```

5. 目前不支援SELinux、建議將狀態設為停用、以避免衝突 (尤其是使用RDMA時) 。請設定下列項目、以確保SELinux已停用：

```
eseries_beegfs_ha_disable_selinux: True  
eseries_selinux_state: disabled
```

6. 設定驗證、讓檔案節點能夠通訊、根據組織原則調整預設值：

```
beegfs_ha_cluster_name: hacluster # BeeGFS HA cluster
name.
beegfs_ha_cluster_username: hacluster # BeeGFS HA cluster
username.
beegfs_ha_cluster_password: hapassword # BeeGFS HA cluster
username's password.
beegfs_ha_cluster_password_sha512_salt: randomSalt # BeeGFS HA cluster
username's password salt.
```

7. 根據"規劃檔案系統"指定此檔案系統的 BeeGFS 管理 IP 一節：

```
beegfs_ha_mgmt_d_floating_ip: <IP ADDRESS>
```



儘管看似冗餘、但當您將BeeGFS檔案系統擴充至單一HA叢集以外的位置時、「beegfs_ha_mgmt_d_浮點_ip」是很重要的。部署後續HA叢集時、不需要額外的BeeGFS管理服務、並指向第一個叢集所提供的管理服務。

8. 視需要啟用電子郵件警示：

```
beegfs_ha_enable_alerts: True
# E-mail recipient list for notifications when BeeGFS HA resources
change or fail.
beegfs_ha_alert_email_list: ["<EMAIL>"]
# This dictionary is used to configure postfix service
(/etc/postfix/main.cf) which is required to set email alerts.
beegfs_ha_alert_conf_ha_group_options:
    # This parameter specifies the local internet domain name. This is
optional when the cluster nodes have fully qualified hostnames (i.e.
host.example.com)
    mydomain: <MY_DOMAIN>
beegfs_ha_alert_verbosity: 3
# 1) high-level node activity
# 3) high-level node activity + fencing action information + resources
(filter on X-monitor)
# 5) high-level node activity + fencing action information + resources
```

9. 強烈建議啟用隔離功能、否則當主要節點故障時、服務可能無法在次要節點上啟動。

- a. 指定下列項目以全域啟用隔離：

```
beegfs_ha_cluster_crm_config_options:
    stonith-enabled: True
```

- i. 附註如有需要、也可在此處指定任何支援 "叢集內容" 的項目。由於 BeeGFS HA 角色隨附許多經過測試的功能，因此通常不需要調整這些 "預設值"功能。
- b. 接下來選取並設定隔離代理程式：
 - i. 選項1：若要使用APC電力分配單元（PDU）啟用隔離功能：

```
beegfs_ha_fencing_agents:
  fence_apc:
    - ipaddr: <PDU_IP_ADDRESS>
      login: <PDU_USERNAME>
      passwd: <PDU_PASSWORD>
      pcmk_host_map:
        "<HOSTNAME>:<PDU_PORT>,<PDU_PORT>;<HOSTNAME>:<PDU_PORT>,<PDU_PORT>"
        "
```

- ii. 選項2：若要使用Lenovo XCC（及其他BMC）提供的Redfish API來啟用屏障：

```
redfish: &redfish
  username: <BMC_USERNAME>
  password: <BMC_PASSWORD>
  ssl_insecure: 1 # If a valid SSL certificate is not available
  specify "1".

beegfs_ha_fencing_agents:
  fence_redfish:
    - pcmk_host_list: <HOSTNAME>
      ip: <BMC_IP>
      <<: *redfish
    - pcmk_host_list: <HOSTNAME>
      ip: <BMC_IP>
      <<: *redfish
```

- iii. 如需設定其他隔離代理程式的詳細資訊，請參閱 "Red Hat 文檔"。

10. BeeGFS HA角色可套用許多不同的調校參數、以協助進一步最佳化效能。其中包括最佳化核心記憶體使用率和區塊裝置I/O、以及其他參數。根據 NetApp E-Series 區塊節點的測試、角色隨附一組合理的 "預設值"、但預設不會套用這些功能、除非您指定：

```
beegfs_ha_enable_performance_tuning: True
```

- a. 如有需要、也可在此處指定預設效能調校的任何變更。如需其他詳細資料、請參閱完整 "效能調校參數" 文件。
11. 為了確保BeeGFS服務所使用的浮動IP位址（有時稱為邏輯介面）可在檔案節點之間容錯移轉、所有網路介面必須一致命名。根據預設、網路介面名稱是由核心產生、因此無法保證產生一致的名稱、即使是安裝在相同PCIe插槽中的網路介面卡、也能在相同的伺服器機型上產生一致的名稱。在部署設備之前建立庫存並已知

產生介面名稱時、這也很有用。根據伺服器或的區塊圖、確保裝置名稱一致 `lshw -class network -businfo` 輸出時、請指定所需的PCIe位址對邏輯介面對應、如下所示：

- a. 對於InfiniBand (IPoIB) 網路介面：

```
eseries_ipoib_udev_rules:
  "<PCIe ADDRESS>": <NAME> # Ex: 0000:01:00.0: i1a
```

- b. 對於乙太網路介面：

```
eseries_ip_udev_rules:
  "<PCIe ADDRESS>": <NAME> # Ex: 0000:01:00.0: e1a
```



為了避免在重新命名介面時發生衝突（避免重新命名）、您不應使用任何可能的預設名稱、例如eth0、ens9f0、ib0或ibs4f0。一般的命名慣例是使用「e」或「i」作為乙太網路或InfiniBand、接著是PCIe插槽編號、以及字母來表示連接埠。例如、安裝在插槽3的InfiniBand介面卡的第二個連接埠為：i3b。



如果您使用已驗證的檔案節點模型、請按一下 ["請按這裡"](#) PCIe位址對邏輯連接埠對應範例。

12. (可選) 指定應套用至叢集中所有BeeGFS服務的組態。可以找到預設組態值 ["請按這裡"](#)、並在其他地方指定個別服務組態：

- a. BeeGFS管理服務：

```
beegfs_ha_beegfs_mgmt_conf_ha_group_options:
  <OPTION>: <VALUE>
```

- b. BeeGFS中繼資料服務：

```
beegfs_ha_beegfs_meta_conf_ha_group_options:
  <OPTION>: <VALUE>
```

- c. BeeGFS儲存服務：

```
beegfs_ha_beegfs_storage_conf_ha_group_options:
  <OPTION>: <VALUE>
```

13. 截至BeeGFS 7.2.7和7.3.1 ["連線驗證"](#) 必須設定或明確停用。您可以使用以Ansible為基礎的部署來設定這項功能：

- a. 根據預設、部署會自動設定連線驗證、並產生 `connauthfile` 將會發佈至所有檔案節點、並搭配BeeGFS服務使用。此檔案也會放置/維護在的Ansible控制節點上

<INVENTORY>/files/beegfs/<sysMgmtHost>_connAuthFile 應將其維護（安全）以供需要存取此檔案系統的用戶端重複使用。

- i. 產生新的金鑰指定 `-e "beegfs_ha_conn_auth_force_new=True` 執行 Ansible 教戰手冊時。請注意、如果是、則會忽略此項 `beegfs_ha_conn_auth_secret` 已定義。
 - ii. 如需進階選項，請參閱隨附的完整預設清單 "[BeeGFS HA 角色](#)"。
- b. 您可以在中定義下列項目、以使用自訂密碼 `ha_cluster.yml`：

```
beegfs_ha_conn_auth_secret: <SECRET>
```

- c. 連線驗證可完全停用（不建議）：

```
beegfs_ha_conn_auth_enabled: false
```

按一下 "[請按這裡](#)" 例如、代表通用檔案節點組態的完整庫存檔案。

使用具有 **NetApp EF600** 區塊節點的 **HDR（200GB）InfiniBand**：

若要將 HDR(200Gb) InfiniBand 搭配 EF600 使用、子網路管理程式必須支援虛擬化。如果使用交換器連接檔案和區塊節點、則必須在整個 Fabric 的子網路管理程式上啟用此功能。

如果使用 InfiniBand 直接連接區塊和檔案節點、`opensm` 則必須在每個檔案節點上為直接連接至區塊節點的每個介面設定執行個體。這是透過指定 `configure: true` 時間來完成 "[設定檔案節點儲存介面](#)" 的。

目前支援的 Linux 套裝作業系統隨附的收件匣版本 `opensm` 不支援虛擬化。而是必須從 NVIDIA OpenFabrics Enterprise Distribution（OFED）安裝和設定的版本 `opensm`。雖然仍支援使用 Ansible 進行部署、但仍需執行幾個額外步驟：

1. 使用 `Curl` 或您想要的工具、將 NVIDIA 網站一節中所列 OpenSM 版本的套件下載 "[技術需求](#)" 到 `<INVENTORY>/packages/` 目錄中。例如：

```
curl -o packages/opensm-5.17.2.MLNX20240610.dc7c2998-0.1.2310322.x86_64.rpm https://linux.mellanox.com/public/repo/mlnx_ofed/23.10-3.2.2.0/rhel9.4/x86_64/opensm-5.17.2.MLNX20240610.dc7c2998-0.1.2310322.x86_64.rpm
curl -o packages/opensm-libs-5.17.2.MLNX20240610.dc7c2998-0.1.2310322.x86_64.rpm https://linux.mellanox.com/public/repo/mlnx_ofed/23.10-3.2.2.0/rhel9.4/x86_64/opensm-libs-5.17.2.MLNX20240610.dc7c2998-0.1.2310322.x86_64.rpm
```

2. 低於 `group_vars/ha_cluster.yml` 定義下列組態：

```

### OpenSM package and configuration information
eseries_ib_opensm_allow_upgrades: true
eseries_ib_opensm_skip_package_validation: true
eseries_ib_opensm_rhel_packages: []
eseries_ib_opensm_custom_packages:
  install:
    - files:
      add:
        "packages/opensm-5.17.2.MLNX20240610.dc7c2998-
0.1.2310322.x86_64.rpm": "/tmp/"
        "packages/opensm-libs-5.17.2.MLNX20240610.dc7c2998-
0.1.2310322.x86_64.rpm": "/tmp/"
    - packages:
      add:
        - /tmp/opensm-5.17.2.MLNX20240610.dc7c2998-
0.1.2310322.x86_64.rpm
        - /tmp/opensm-libs-5.17.2.MLNX20240610.dc7c2998-
0.1.2310322.x86_64.rpm
  uninstall:
    - packages:
      remove:
        - opensm
        - opensm-libs
    files:
      remove:
        - /tmp/opensm-5.17.2.MLNX20240610.dc7c2998-
0.1.2310322.x86_64.rpm
        - /tmp/opensm-libs-5.17.2.MLNX20240610.dc7c2998-
0.1.2310322.x86_64.rpm

eseries_ib_opensm_options:
  virt_enabled: "2"

```

指定通用區塊節點組態

使用群組變數（群組_vars）指定通用區塊節點組態。

總覽

所有區塊節點的組態均定義於 `group_vars/eseries_storage_systems.yml`。通常包括：

- 有關Ansible控制節點應如何連線至用作區塊節點的E系列儲存系統的詳細資料。
- 節點應該執行哪些韌體、NVSRAM/磁碟機韌體版本。

- 全域組態、包括快取設定、主機組態、以及應如何配置磁碟區的設定。



此檔案中設定的選項也可在個別區塊節點上定義、例如使用混合式硬體模型、或是每個節點的密碼不同。個別區塊節點的組態優先於此檔案中的組態。

步驟

建立檔案 `group_vars/eseries_storage_systems.yml` 並填入如下內容：

1. Ansible不會使用SSH連線至區塊節點、而是使用REST API。為了達成此目標、我們必須設定：

```
ansible_connection: local
```

2. 指定用於管理每個節點的使用者名稱和密碼。使用者名稱可以選擇性地省略（預設為admin）、否則您可以指定具有管理員權限的任何帳戶。同時指定是否應驗證或忽略SSL憑證：

```
eseries_system_username: admin
eseries_system_password: <PASSWORD>
eseries_validate_certs: false
```



不建議以純文字列出任何密碼。使用Ansible保存庫或提供 `eseries_system_password` 使用-Extra vars執行Ansible時。

3. (可選) 指定應在節點上安裝哪些控制器韌體、NVSRAM/磁碟機韌體。這些項目必須下載至 `packages/` 執行Ansible之前的目錄。E系列控制器韌體和NVSRAM ["請按這裡"](#) 和磁碟機韌體 ["請按這裡"](#)：

```
eseries_firmware_firmware: "packages/<FILENAME>.dlp" # Ex.
"packages/RCB_11.80GA_6000_64cc0ee3.dlp"
eseries_firmware_nvram: "packages/<FILENAME>.dlp" # Ex.
"packages/N6000-880834-D08.dlp"
eseries_drive_firmware_firmware_list:
  - "packages/<FILENAME>.dlp"
  # Additional firmware versions as needed.
eseries_drive_firmware_upgrade_drives_online: true # Recommended unless
BeeGFS hasn't been deployed yet, as it will disrupt host access if set
to "false".
```



如果指定此組態、Ansible將自動更新所有韌體、包括重新開機控制器（如有必要）、而不會出現其他提示。這對BeeGFS /主機I/O來說是不中斷營運的、但可能會導致效能暫時降低。

4. 調整全域系統組態預設值。此處列出的選項與值通常建議用於NetApp上的BeeGFS、但可視需要調整：

```
eseries_system_cache_block_size: 32768
eseries_system_cache_flush_threshold: 80
eseries_system_default_host_type: linux dm-mp
eseries_system_autoload_balance: disabled
eseries_system_host_connectivity_reporting: disabled
eseries_system_controller_shelf_id: 99 # Required by default.
```

5. 設定全域Volume資源配置預設值。此處列出的選項與值通常建議用於NetApp上的BeeGFS、但可視需要調整：

```
eseries_volume_size_unit: pct # Required by default. This allows volume
capacities to be specified as a percentage, simplifying putting together
the inventory.
eseries_volume_read_cache_enable: true
eseries_volume_read_ahead_enable: false
eseries_volume_write_cache_enable: true
eseries_volume_write_cache_mirror_enable: true
eseries_volume_cache_without_batteries: false
```

6. 如有需要、請依照下列最佳實務做法、調整Ansible選擇儲存資源池和磁碟區群組磁碟機的順序：

- a. 請先列出應用於管理和（或）中繼資料磁碟區的任何（可能較小的）磁碟機、然後列出儲存磁碟區。
- b. 根據磁碟櫃/磁碟機機箱機型、確保在可用磁碟機通道之間平衡磁碟機選擇順序。例如、在EF600不擴充的情況下、磁碟機0-11位於磁碟機通道1、而磁碟機12-23位於磁碟機通道。因此、平衡磁碟機選擇的策略是選擇 `disk shelf:drive 99:0、99:23、99:1、99:22`等如果有多個機箱、第一個數字代表磁碟機櫃ID。

```
# Optimal/recommended order for the EF600 (no expansion):
eseries_storage_pool_usable_drives:
"99:0,99:23,99:1,99:22,99:2,99:21,99:3,99:20,99:4,99:19,99:5,99:18,99
:6,99:17,99:7,99:16,99:8,99:15,99:9,99:14,99:10,99:13,99:11,99:12"
```

按一下 "[請按這裡](#)" 例如、代表通用區塊節點組態的完整庫存檔案。

版權資訊

Copyright © 2026 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。