



Ubuntu

SAN hosts and cloud clients

NetApp
October 30, 2024

目錄

Ubuntu	1
適用於 ONTAP 的 Ubuntu 22.04 的 NVMe 主機組態	1

Ubuntu

適用於 ONTAP 的 Ubuntu 22.04 的 NVMe 主機組態

含非對稱命名空間存取（ANA）的 Ubuntu 22.04 支援 NVMe over Fabrics（NVMe over Fabric、NVMe over Fibre Channel、NVMe / FC）和其他傳輸。在 NVMe 環境中、ANA 等同於 iSCSI 和 FC 環境中的 ALUA 多重路徑、並以核心內建 NVMe 多重路徑來實作。

下列支援適用於 ONTAP 的 Ubuntu 22.04 NVMe 主機組態：

- 原生 NVMe - CLI 套件中的 NetApp 外掛程式會顯示 NVMe / FC 命名空間的 ONTAP 詳細資料。
- 在指定主機匯流排介面卡（HBA）上的同一主機上使用 NVMe 和 SCSI 共存流量、而不使用明確的 dm-multipath 設定、以避免使用 NVMe 命名空間。

如需支援組態的詳細資訊、請參閱 ["NetApp 互通性對照表工具"](#)。

功能

根據預設、Ubuntu 22.04 已啟用 NVMe 命名空間的核心內建 NVMe 多重路徑。因此、不需要明確的設定。

已知限制

目前不支援使用 NVMe 型傳輸協定進行 SAN 開機。

驗證軟體版本

您可以使用下列程序來驗證最低支援的 Ubuntu 22.04 軟體版本。

步驟

1. 在伺服器上安裝 Ubuntu 22.04。安裝完成後、請確認您正在執行指定的 Ubuntu 22.04 核心：

```
# uname -r
```

◦ 輸出範例：*

```
5.15.0-101-generic
```

2. 安裝「NVMe-CLI (NVMe - CLI) 套件：

```
# apt list | grep nvme
```

◦ 輸出範例：*

```
nvme-cli/jammy-updates,now 1.16-3ubuntu0.1 amd64
```

3. 在 Ubuntu 22.04 主機上、請在中檢查 hostnqn 字串 /etc/nvme/hostnqn :

```
# cat /etc/nvme/hostnqn
```

◦ 輸出範例 *

```
nqn.2014-08.org.nvmexpress:uuid:063a9fa0-438a-4737-b9b4-95a21c66d041
```

4. 確認 hostnqn 字串符合 hostnqn ONTAP 陣列上對應子系統的字串 :

```
::> vserver nvme subsystem host show -vserver vs_106_fc_nvme
```

◦ 輸出範例： *

```
Vserver      Subsystem      Host NQN
-----
vs_106_fc_nvme ub_106 nqn.2014-08.org.nvmexpress:uuid:c04702c8-e91e-4353-9995-ba4536214631
```

+



如果是 hostnqn 字串不相符、請使用 `vserver modify` 命令來更新 hostnqn 對應 ONTAP 陣列子系統上的字串、以符合 hostnqn 字串來源 /etc/nvme/hostnqn 在主機上。

設定NVMe/FC

您可以為 Broadcom / Emulex 或 Marvell/Qlogic 介面卡設定 NVMe / FC 。

Broadcom / Emulex

步驟

1. 確認您使用的是支援的介面卡機型。

```
# cat /sys/class/scsi_host/host*/modelname
```

- 輸出範例：*

```
LPe36002-M64  
LPe36002-M64
```

```
# cat /sys/class/scsi_host/host*/modeldesc
```

- 輸出範例：*

```
Emulex LPe36002-M64 2-Port 64Gb Fibre Channel Adapter  
Emulex LPe36002-M64 2-Port 64Gb Fibre Channel Adapter
```

2. 驗證您使用的是建議的Broadcom lpfc 韌體與內建驅動程式。

```
# cat /sys/class/scsi_host/host*/fwrev  
  
14.2.673.40, sli-4:6:d  
14.2.673.40, sli-4:6:d  
  
# cat /sys/module/lpfc/version  
0: 14.0.0.4
```

如需支援的介面卡驅動程式和韌體版本的最新清單、請參閱 ["NetApp 互通性對照表工具"](#)。

3. 請確認 lpfc_enable_fc4_type 設為 3：

```
# cat /sys/module/lpfc/parameters/lpfc_enable_fc4_type  
3
```

4. 確認啟動器連接埠已啟動並正在執行、而且您可以看到目標生命體：

```

# cat /sys/class/fc_host/host*/port_name
0x100000109bf0447c
0x100000109bf0447b
# cat /sys/class/fc_host/host*/port_state
Online
Online
# cat /sys/class/scsi_host/host*/nvme_info
    NVME Initiator Enabled
XRI Dist lpfc1 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc1 WWPN x100000109bf0447c WWNN x200000109bf0447c DID
x022300 ONLINE
NVME RPORT          WWPN x200cd039eaa8138b WWNN x200ad039eaa8138b DID
x021509 TARGET DISCSRVC ONLINE
NVME RPORT          WWPN x2010d039eaa8138b WWNN x200ad039eaa8138b DID
x021108 TARGET DISCSRVC ONLINE

NVME Statistics
LS: Xmt 000000000e Cmpl 000000000e Abort 00000000
LS XMIT: Err 00000000  Cmpl: xb 00000000 Err 00000000
Total FCP Cmpl 0000000000005238 Issue 000000000000523a OutIO
00000000000000002
    abort 00000000 noxri 00000000 nondlp 00000000 qdepth
00000000 wqerr 00000000 err 00000000
FCP Cmpl: xb 00000000 Err 00000000

NVME Initiator Enabled
XRI Dist lpfc0 Total 6144 IO 5894 ELS 250
NVME LPORT lpfc0 WWPN x100000109bf0447b WWNN x200000109bf0447b DID
x022600 ONLINE
NVME RPORT          WWPN x200bd039eaa8138b WWNN x200ad039eaa8138b DID
x021409 TARGET DISCSRVC ONLINE
NVME RPORT          WWPN x200fd039eaa8138b WWNN x200ad039eaa8138b DID
x021008 TARGET DISCSRVC ONLINE

NVME Statistics
LS: Xmt 000000000e Cmpl 000000000e Abort 00000000
LS XMIT: Err 00000000  Cmpl: xb 00000000 Err 00000000
Total FCP Cmpl 000000000000523c Issue 000000000000523e OutIO
00000000000000002
    abort 00000000 noxri 00000000 nondlp 00000000 qdepth
00000000 wqerr 00000000 err 00000000
FCP Cmpl: xb 00000000 Err 00000000

```

適用於 NVMe / FC 的 Marvell/QLogic FC 介面卡

Ubuntu 22.04 GA 核心中隨附的原生收件匣 qla2xxx 驅動程式具有最新的上游修正程式。這些修正對於 ONTAP 支援至關重要。

步驟

1. 確認您執行的是支援的介面卡驅動程式和韌體版本：

```
# cat /sys/class/fc_host/host*/symbolic_name
```

◦ 輸出範例 *

```
QLE2872 FW: v9.14.02 DVR: v10.02.06.200-k  
QLE2872 FW: v9.14.02 DVR: v10.02.06.200-k
```

2. 請確認 `ql2xnvmeenable` 已設定。這可讓 Marvell 介面卡作為 NVMe / FC 啟動器運作：

```
# cat /sys/module/qla2xxx/parameters/ql2xnvmeenable  
1
```

啟用 1MB I/O (選用)

ONTAP 會在識別控制器資料中報告 8 的 MDTs (MAX Data 傳輸大小)。這表示最大 I/O 要求大小最多可達 1MB。若要針對 Broadcom NVMe / FC 主機發出大小為 1 MB 的 I/O 要求、您必須將 `lpfc`lpfc_sg_seg_cnt`` 參數值從預設值 64 增加至 256。

步驟

1. 將 `lpfc_sg_seg_cnt`` 參數設定為 256：

```
# cat /etc/modprobe.d/lpfc.conf  
options lpfc lpfc_sg_seg_cnt=256
```

2. 執行 `dracut -f`` 命令、然後重新啟動主機：
3. 確認 `lpfc_sg_seg_cnt`` 為 256：

```
# cat /sys/module/lpfc/parameters/lpfc_sg_seg_cnt  
256
```



這不適用於 Qlogic NVMe / FC 主機。

設定NVMe/TCP

NVMe / TCP 沒有自動連線功能。因此、如果某個路徑發生故障、且在 10 分鐘的預設逾時期間內未恢復、則 NVMe / TCP 無法自動重新連線。若要避免逾時、您應該將容錯移轉事件的重試期間設為至少 30 分鐘。

步驟

1. 確認啟動器連接埠可在支援的NVMe/TCP LIF中擷取探索記錄頁面資料：

```
nvme discover -t tcp -w host-traddr -a traddr
```

◦ 輸出範例：*

```
# nvme discover -t tcp -w 10.10.11.47-a 10.10.10.122

Discovery Log Number of Records 8, Generation counter 10
=====Discovery Log Entry 0=====
trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem
treq: not specified
portid: 0
trsvcid: 8009
subnqn: nqn.1992-
08.com.netapp:sn.bbf4ee8dfb611edbd07d039ea165590:discovery
traddr: 10.10.10.122
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 1=====
trtype: tcp
adrfam: ipv4
subtype: current discovery subsystem
treq: not specified
portid: 1
trsvcid: 8009
subnqn: nqn.1992
08.com.netapp:sn.bbf4ee8dfb611edbd07d039ea165590:discovery
traddr: 10.10.10.124
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 2=====
trtype: tcp
```

2. 確認其他的 NVMe / TCP 啟動器目標 LIF 組合能夠成功擷取探索記錄頁面資料：


```
nvme discover -t tcp -w host-traddr -a traddr
```

◦ 輸出範例：*

```
#nvme discover -t tcp -w 10.10.10.47 -a 10.10.10.122
#nvme discover -t tcp -w 10.10.10.47 -a 10.10.10.124
#nvme discover -t tcp -w 10.10.11.47 -a 10.10.11.122
#nvme discover -t tcp -w 10.10.11.47 -a 10.10.11.
```

3. 在節點上所有支援的 NVMe / TCP 啟動器目標生命週期上執行 NVMe connect -all 命令、並將控制器遺失逾時時間設定至少 30 分鐘或 1800 秒：

```
nvme connect-all -t tcp -w host-traddr -a traddr -l 1800
```

◦ 輸出範例：*

```
# nvme connect-all -t tcp -w 10.10.10.47 -a 10.10.10.122 -l 1800
# nvme connect-all -t tcp -w 10.10.10.47 -a 10.10.10.124 -l 1800
# nvme connect-all -t tcp -w 10.10.11.47 -a 10.10.11.122 -l 1800
# nvme connect-all -t tcp -w 10.10.11.47 -a 10.10.11.124 -l 1800
```

驗證 NVMe

您可以使用下列程序來驗證 NVMe。

步驟

1. 確認已啟用核心內建 NVMe 多重路徑：

```
# cat /sys/module/nvme_core/parameters/multipath
Y
```

2. 驗證個別 ONTAP 命名空間的適當 NVMe 設定（例如、模型設定為 NetApp ONTAP 控制器、負載平衡 iopolicing 設定為循環）是否正確反映在主機上：

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/model
NetApp ONTAP Controller
NetApp ONTAP Controller
```

```
# cat /sys/class/nvme-subsystem/nvme-subsys*/iopolicy
round-robin
round-robin
```

3. 確認已在主機上建立並正確探索命名空間：

```
# nvme list
```

◦ 輸出範例：*

```
Node          SN                      Model
-----
/dev/nvme0n1  81CZ5BQuUNfGAAAAAAB   NetApp ONTAP Controller

Namespace Usage    Format                      FW                      Rev
-----
1                 21.47 GB / 21.47 GB    4 KiB + 0 B          FFFFFFFF
```

4. 確認每個路徑的控制器狀態均為有效、且具有正確的ANA狀態：

NVMe / FC

```
# nvme list-subsys /dev/nvme0n1
```

◦ 輸出範例： *

```
nvme-subsys4 - NQN=nqn.1992-08.com.netapp:sn.8763d311b2ac11ed950ed039ea951c46:subsystem. ub_106
\
+- nvme1 fc traddr=nn-0x20a6d039ea954d17:pn-0x20a7d039ea954d17,host_traddr=nn-0x200000109b1b95ef:pn-0x100000109b1b95ef live optimized
+- nvme2 fc traddr=nn-0x20a6d039ea954d17:pn-0x20a8d039ea954d17,host_traddr=nn-0x200000109b1b95f0:pn-0x100000109b1b95f0 live optimized
+- nvme3 fc traddr=nn-0x20a6d039ea954d17:pn-0x20aad039ea954d17,host_traddr=nn-0x200000109b1b95f0:pn-0x100000109b1b95f0 live non-optimized
+- nvme5 fc traddr=nn-0x20a6d039ea954d17:pn-0x20a9d039ea954d17,host_traddr=nn-0x200000109b1b95ef:pn-0x100000109b1b95ef live non-optimized
```

NVMe / TCP

```
# nvme list-subsys /dev/nvme1n1
```

◦ 輸出範例： *

```
nvme-subsys1 - NQN=nqn.1992-08.com.netapp:sn.bbfb4ee8dfb611edbd07d039ea165590:subsystem.rhel_tcp_95
+- nvme1 tcp
traddr=10.10.10.122,trsvcid=4420,host_traddr=10.10.10.47,src_addr=10.10.10.47 live
+- nvme2 tcp
traddr=10.10.10.124,trsvcid=4420,host_traddr=10.10.10.47,src_addr=10.10.10.47 live
+- nvme3 tcp
traddr=10.10.11.122,trsvcid=4420,host_traddr=10.10.11.47,src_addr=10.10.11.47 live
+- nvme4 tcp
traddr=10.10.11.124,trsvcid=4420,host_traddr=10.10.11.47,src_addr=10.10.11.47 live
```

5. 驗證NetApp外掛程式是否顯示每ONTAP 個版本名稱空間裝置的正確值：

欄位

```
# nvme netapp ontapdevices -o column
```

◦ 輸出範例：*

```
Device          Vserver    Namespace Path
-----
/dev/nvme0n1   co_iscsi_tcp_ubuntu /vol/vol1/ns1

NSID           UUID                                           Size
-----
1              79c2c569-b7fa-42d5-b870-d9d6d7e5fa84 21.47GB
```

JSON

```
# nvme netapp ontapdevices -o json
```

◦ 輸出範例 *

```
{
  "ONTAPdevices" : [
    {
      "Device" : "/dev/nvme0n1",
      "Vserver" : "co_iscsi_tcp_ubuntu",
      "Namespace_Path" : "/vol/nvmevol1/ns1",
      "NSID" : 1,
      "UUID" : "79c2c569-b7fa-42d5-b870-d9d6d7e5fa84",
      "Size" : "21.47GB",
      "LBA_Data_Size" : 4096,
      "Namespace_Size" : 5242880
    },
  ]
}
```

已知問題

含 ONTAP 版本的 Ubuntu 22.04 的 NVMe 主機組態有下列已知問題：

NetApp錯誤ID	標題	說明
CONTAPEXT-2037	Ubuntu 22.04 NVMe 主機建立重複的持續探索控制器	在 NVMe 主機上、您可以使用「NVMe 探索 -p」命令來建立持續探索控制器（LC）。此命令只能為每個啟動器目標組合建立一個 PDC。不過、如果您在 NVMe 主機上執行 Ubuntu 22.04、則每次執行「NVMe 探索 -p」時都會建立重複的 PDC。這會導致主機和目標上的資源使用不必要。

版權資訊

Copyright © 2024 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。