



Ubuntu

SAN hosts and cloud clients

NetApp
March 29, 2024

目录

Ubuntu	1
适用于采用ONTAP的Ubuntu 22.04的NVMe-oF主机配置	1

Ubuntu

适用于采用ONTAP的Ubuntu 22.04的NVMe-oF主机配置

采用非对称命名空间访问(AANA)的Ubuntu 22.04支持基于网络结构的NVMe (NVMe-oF)、包括基于光纤通道的NVMe (NVMe/FC)和其他传输。在NVMe-oF环境中、ANA相当于iSCSI和FC环境中的ALUA多路径功能、并可通过内核NVMe多路径实施。

对于采用ONTAP的Ubuntu 22.04的NVMe-oF主机配置、可获得以下支持：

- 本机NVMe-CLI软件包中的NetApp插件可显示NVMe/FC命名库的ONTAP详细信息。
- 在给定主机总线适配器(HBA)的同一主机上同时使用NVMe和SCSI流量、而不使用显式dm-dpath设置、以防止声明NVMe命名空间。

有关支持的配置的其他详细信息、请参见 "[NetApp 互操作性表工具](#)"。

功能

默认情况下、Ubuntu 22.04已为NVMe命名空间启用内核NVMe多路径。因此、无需显式设置。

已知限制

目前不支持使用NVMe-oF协议启动SAN。

验证软件版本

您可以使用以下操作步骤验证支持的最低Ubuntu 22.04软件版本。

步骤

1. 在服务器上安装Ubuntu 22.04。安装完成后、验证是否正在运行指定的Ubuntu 22.04内核：

```
# uname -r
```

示例输出：

```
5.15.0-101-generic
```

2. 安装 NVMe-CLI 软件包：

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

示例输出：

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

3. 在Ubuntu 22.04主机上、检查中的hostnqn字符串 /etc/nvme/hostnqn:

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

示例输出

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

4. 验证是否已 hostnqn 字符串与匹配 hostnqn ONTAP 阵列上对应子系统的字符串:

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

示例输出:

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



如果 hostnqn 字符串不匹配、请使用 `vservers modify` 用于更新的命令 hostnqn 要匹配的相应ONTAP 阵列子系统上的字符串 hostnqn 字符串自 /etc/nvme/hostnqn 在主机上。

配置 NVMe/FC

您可以为Broadcom/Emulex或Marvell/Qlogic适配器配置NVMe/FC。

Broadcom/Emulex

步骤

1. 验证您使用的是支持的适配器型号。

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

示例输出：

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

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

示例输出：

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

2. 确认您使用的是建议的Broadcom lpfc 固件和收件箱驱动程序。

```
# 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
```

有关支持的适配器驱动程序和固件版本的最新列表、请参见 ["NetApp 互操作性表工具"](#)。

3. 请验证 lpfc_enable_fc4_type 设置为 3：

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

4. 验证启动程序端口是否已启动且正在运行、以及您是否可以看到目标生命周期：

```

# 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 DISCSRV ONLINE
NVME RPORT          WWPN x2010d039eaa8138b WWNN x200ad039eaa8138b DID
x021108 TARGET DISCSRV 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 DISCSRV ONLINE
NVME RPORT          WWPN x200fd039eaa8138b WWNN x200ad039eaa8138b DID
x021008 TARGET DISCSRV 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

```

适用于NVMe/FC的Marvell/QLogic FC适配器

步骤

1. Ubuntu 22.04 GA内核中包含的本机内置qla2xxx驱动程序具有ONTAP支持所必需的最新版上游修复程序。验证您是否正在运行受支持的适配器驱动程序和固件版本：

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

示例输出

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

2. 请验证 ql2xnvmeenable 已设置。这样、Marvell适配器便可用作NVMe/FC启动程序：

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

启用1 MB I/O (可选)

ONTAP会在"识别控制器"数据中报告MDTS (MAX Data传输大小)为8、这意味着最大I/O请求大小最多可以为1 MB。但是、要使Broadcom NVMe/FC主机的问题描述I/O请求大小为1 MB、必须增加 lpfc 的值 lpfc_sg_seg_cnt 参数从默认值64更改为256。

步骤

1. 将 lpfc_sg_seg_cnt 参数设置为 256 。

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

2. 运行 dracut -f 命令，然后重新启动主机。
3. 验证 lpfc_sg_seg_cnt 是否为 256 。

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



这不适用于逻辑NVMe/FC主机。

配置 NVMe/TCP

NVMe/TCP没有自动连接功能。因此、如果某个路径发生故障、并且未在默认超时时间10分钟内恢复、

则NVMe/TCP无法自动重新连接。为了防止超时、您应将故障转移事件的重试期限至少设置为30分钟。

步骤

1. 验证启动程序端口是否可以通过受支持的NVMe/TCP LIF提取发现日志页面数据：

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

示例输出：

```
# 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.bbfb4ee8dfb611edbd07d039ea165590: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.bbfb4ee8dfb611edbd07d039ea165590:discovery
traddr: 10.10.10.124
eflags: explicit discovery connections, duplicate discovery information
sectype: none
=====Discovery Log Entry 2=====
trtype: tcp
```

2. 验证其他NVMe/TCP启动程序-目标LIF组合是否能够成功提取发现日志页面数据：

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


示例输出:

```
#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. 对节点中所有受支持的NVMe/TCP启动程序-目标Li运行NVMe Connect-all命令、并将控制器丢失超时期限设置为至少30分钟或1800秒:

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

示例输出:

```
# 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
```

验证 NVMe-oF

您可以使用以下操作步骤验证NVMe-oF。

步骤

1. 验证是否已启用内核NVMe多路径:

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

2. 验证相应ONTAP命名库的适当NVMe-oF设置(例如、型号设置为NetApp ONTAP控制器、负载平衡iopolicy设置为循环)是否正确反映在主机上:

```
# 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. 验证是否已在主机上创建并正确发现命名空间：

```
# nvme list
```

示例输出：

```
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. 验证每个路径的控制器状态是否为活动状态且是否具有正确的ANA状态：

NVMe/FC

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

示例输出:

```
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
```

示例输出:

```
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. 验证NetApp插件是否为每个ONTAP 命名空间设备显示正确的值:

列

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

示例输出:

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
```

示例输出

```
{
  "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
    },
  ]
}
```

已知问题

对于采用ONTAP版本的Ubuntu 22.04的NVMe-oF主机配置、没有已知问题。

版权信息

版权所有 © 2024 NetApp, Inc.。保留所有权利。中国印刷。未经版权所有者事先书面许可，本文档中受版权保护的任何部分不得以任何形式或通过任何手段（图片、电子或机械方式，包括影印、录音、录像或存储在电子检索系统中）进行复制。

从受版权保护的 NetApp 资料派生的软件受以下许可和免责声明的约束：

本软件由 NetApp 按“原样”提供，不含任何明示或暗示担保，包括但不限于适销性以及针对特定用途的适用性的隐含担保，特此声明不承担任何责任。在任何情况下，对于因使用本软件而以任何方式造成的任何直接性、间接性、偶然性、特殊性、惩罚性或后果性损失（包括但不限于购买替代商品或服务；使用、数据或利润方面的损失；或者业务中断），无论原因如何以及基于何种责任理论，无论出于合同、严格责任或侵权行为（包括疏忽或其他行为），NetApp 均不承担责任，即使已被告知存在上述损失的可能性。

NetApp 保留在不另行通知的情况下随时对本文档所述的任何产品进行更改的权利。除非 NetApp 以书面形式明确同意，否则 NetApp 不承担因使用本文档所述产品而产生的任何责任或义务。使用或购买本产品不表示获得 NetApp 的任何专利权、商标权或任何其他知识产权许可。

本手册中描述的产品可能受一项或多项美国专利、外国专利或正在申请的专利的保护。

有限权利说明：政府使用、复制或公开本文档受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中“技术数据权利 — 非商用”条款第 (b)(3) 条规定的限制条件的约束。

本文档中所含数据与商业产品和/或商业服务（定义见 FAR 2.101）相关，属于 NetApp, Inc. 的专有信息。根据本协议提供的所有 NetApp 技术数据和计算机软件具有商业性质，并完全由私人出资开发。美国政府对这些数据的使用权具有非排他性、全球性、受限且不可撤销的许可，该许可既不可转让，也不可再许可，但仅限在与交付数据所依据的美国政府合同有关且受合同支持的情况下使用。除本文档规定的情形外，未经 NetApp, Inc. 事先书面批准，不得使用、披露、复制、修改、操作或显示这些数据。美国政府对国防部的授权仅限于 DFARS 的第 252.227-7015(b)（2014 年 2 月）条款中明确的权利。

商标信息

NetApp、NetApp 标识和 <http://www.netapp.com/TM> 上所列的商标是 NetApp, Inc. 的商标。其他公司和产品名称可能是其各自所有者的商标。