

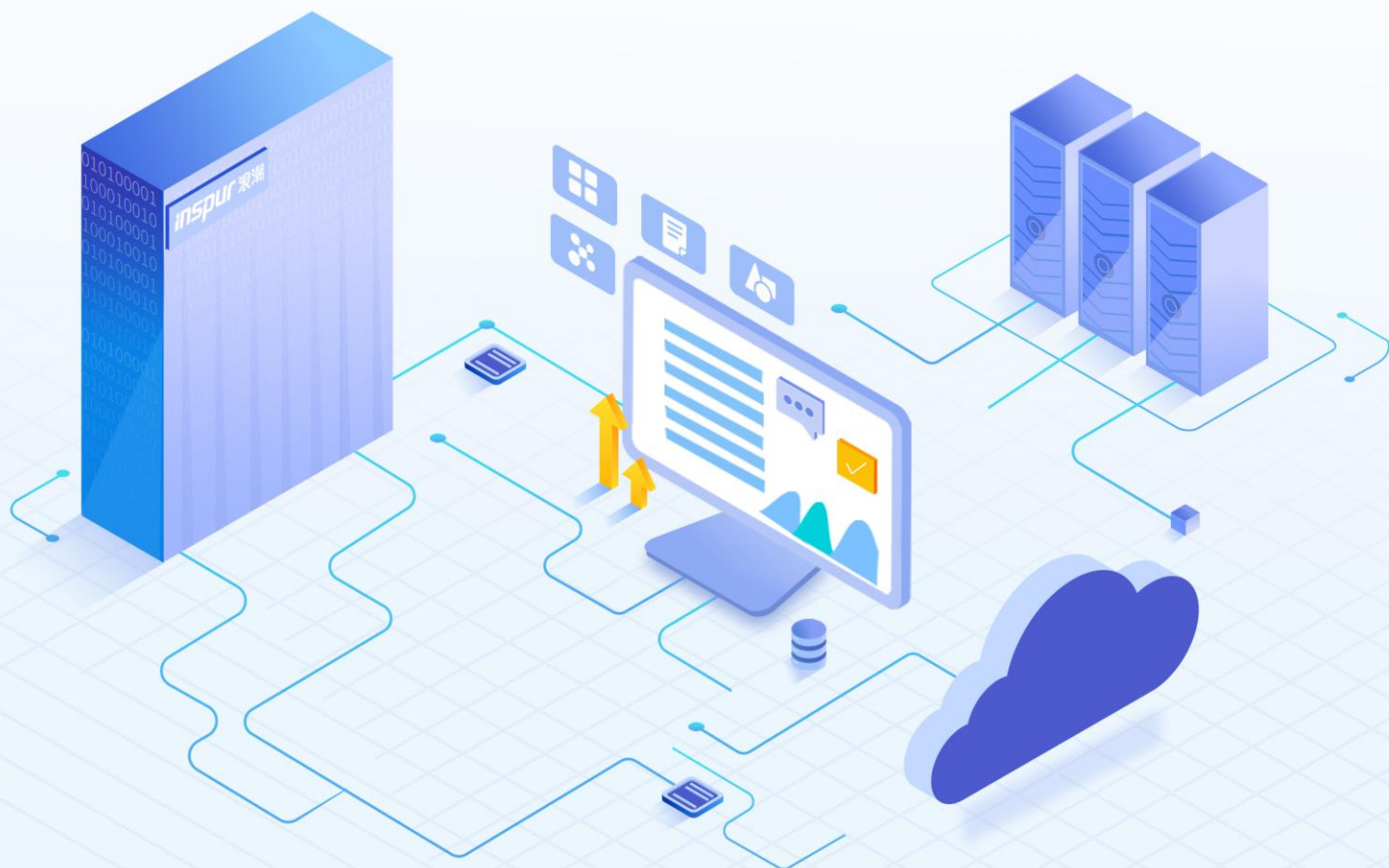
浪潮存储平台

Linux 系统多路径软件用户手册

文档版本 2.5

发布日期 2022-03-25

适用版本 Linux 系统自带多路径的所有版本



尊敬的用户：

衷心感谢您选用浪潮存储！浪潮存储秉承“云存智用 运筹新数据”的新存储之道，致力于为您提供符合新数据时代需求的存储产品和解决方案。

本手册用于帮助您更详细地了解 and 便捷地使用本存储，涉及的截图仅为示例，最终界面请以实际设备显示的界面为准。

由于产品版本升级或其他原因，本手册内容会不定期进行更新，如有变动恕不另行通知。除非另有约定，本手册仅作为使用指导，本手册中的所有陈述、信息和建议不构成任何明示或暗示的担保。

浪潮拥有本手册的版权，保留随时修改本手册的权利。未经浪潮许可，任何单位和个人不得以任何形式复制本手册的内容。

如果您对本手册有任何疑问或建议，请向浪潮电子信息产业股份有限公司垂询。

技术服务电话： 4008600011

地 址： 中国济南市浪潮路 1036 号
浪潮电子信息产业股份有限公司

邮 编： 250101

使用声明

在您正式使用本存储之前，请先阅读以下声明。只有您阅读并且同意以下声明后，方可正式开始使用本存储。如果您对以下声明有任何疑问，请和您的供货商联系或直接与我们联系。如您在开始使用本系统前未就以下声明向我们提出疑问，则默认您已经同意了以下声明。

1. 请不要自行拆卸本存储的机箱及机箱内任何硬件设备。在本存储出现任何硬件故障或您希望对硬件进行任何升级时，请您将机器的详细硬件配置反映给我们的客户服务中心。
2. 请不要将本存储的设备与任何其他型号的相应设备混用。本存储的内存、CPU、CPU 散热片、风扇、硬盘托架、硬盘等都是特殊规格的。
3. 在使用本存储时遇到任何软件问题，请您首先和相应软件的提供商联系。由提供商和我们联系，以方便我们共同沟通和解决您遇到的问题。对于数据库、网络管理软件或其他网络产品的安装、运行问题，我们尤其希望您能够这样处理。
4. 上架安装本存储前，请先仔细阅读相关产品手册中的快速安装指南。我们致力于产品功能和性能的持续提升，部分功能及操作与手册描述可能会有所差异，但不会影响使用。如果您有任何疑问问题，请与我们的客户服务中心联系。
5. **我们特别提醒您：在使用过程中，注意对您的数据进行必要的备份。**
6. 本存储为 A 级产品，在生活环境中可能会造成无线电干扰，需要您对其干扰采取切实可行的措施。
7. 请仔细阅读并遵守本手册的安全声明和安全细则。
8. 本手册中涉及的各项软、硬件产品的标识、名称版权归产品的相应公司拥有。

以上声明中，“我们”指代浪潮电子信息产业股份有限公司；浪潮电子信息产业股份有限公司拥有对以上声明的最终解释权。

安全声明

我们非常重视数据安全和隐私，且一如既往地严密关注产品和解决方案的安全性，为您提供更满意的服务。在您正式使用本存储之前，请先阅读以下安全声明。

1. 为了保护您的数据隐私，在调整存储产品用途或淘汰存储设备时，请您将存储软件恢复固件出厂设置、删除信息、清除日志。同时，建议采用第三方安全擦除工具对存储软件所在的系统盘进行全面安全擦除。
2. 您购买的存储产品业务运营或故障定位的过程中可能会获取或使用用户的某些个人数据（如告警邮件接收地址、IP 地址）。因此，您有义务根据所适用国家或地区的法律法规制定必要的用户隐私政策，并采取足够的措施以确保用户的个人数据受到充分的保护。
3. 如需获取本存储开源软件声明，请直接联系我们的客户服务人员。
4. 本存储的某些安全特性需要您自行配置，如认证、传输加密、存储数据加密等，这些配置操作可能会对存储的性能和使用方便性造成一定影响。您可以根据应用环境，权衡是否进行安全特性配置。
5. 本存储自带了部分用于生产、装备、返厂检测维修的接口、命令及定位故障的高级命令，如使用不当，可能会导致设备异常或者业务中断，不建议您自行使用。如需使用，请联系我们的客户服务人员。
6. 我们已全面建立产品安全漏洞应急和处理机制，确保第一时间处理产品安全问题。若您在存储产品使用过程中发现任何安全问题，或者寻求有关产品安全漏洞的必要支持，请直接联系我们的客户服务人员。

以上声明中，“我们”指代浪潮电子信息产业股份有限公司；浪潮电子信息产业股份有限公司拥有对以上声明的最终解释权。

安全细则

在使用本存储时，若操作不当，可能会危及您的人身安全。为避免发生意外，在正式使用本存储之前，请务必认真阅读以下安全细则，严格按照要求进行操作。

1. 本存储中的电源设备可能会产生高电压和危险电能，从而导致人身伤害。请勿自行卸下主机盖以拆装、更换系统内部的任何组件。除非另外得到我们的通知，否则只有经过我们培训的维修技术人员才有权拆开主机盖及拆装、更换内部组件。
2. 请将设备连接到适当的电源，仅可使用额定输入标签上指明的外部电源为设备供电。为保护您的设备免受电压瞬间升高或降低所导致的损坏，请使用相关的稳压设备或不间断电源设备。
3. 如果必须使用延长线缆，请使用配有正确接地插头的三芯线缆，并查看延长线缆的额定值，确保插入延长线缆的所有产品的额定电流总和不超过延长线缆额定电流限制的百分之八十。
4. 请务必使用随机配备的供电组件，如电源线、电源插座（如果随机配备）等。为了本存储及使用者的安全，切勿随意更换电源线缆或插头。
5. 为防止因系统漏电而造成电击危险，请务必将本存储和外围设备的电源电缆插入已正确接地的电源插座。在未安装接地导线及不确定是否已有适当接地保护的情况下，请勿操作和使用本存储，并及时与电工联系。
6. 切勿将任何物体塞入本存储的开孔处，否则，可能会导致内部组件短路而引起火灾或电击。
7. 请将本存储置于远离散热片和有热源的地方，切勿堵塞通风孔。
8. 切勿在高潮湿、高灰尘的环境中使用本存储，切勿让食物或液体散落在系统内部或其它组件上。
9. 使用错误型号的电池会有爆炸的危险，需要更换电池时，请先向制造商咨询并使用与制造商推荐型号相同或相近的电池。切勿拆开、挤压、刺戳电池或使其外部接点短路。不要将其丢入火中或水中，也不要暴露在温度超过 60 摄氏度的环境中。请勿尝试打开或维修电池，务必合理处置用完的电池，不要将用完的电池及可能包含电池的电路板及其它组件与其它废品放在一起。有关电池回收政策请与当地废品回收处理机构联系。
10. 由于 3U、4U 和 5U 设备过重，建议使用本公司官方选配导轨上架。如果您使用自备导轨，请联系本公司客服，确保设备可以正常上架。在您使用自备导轨上架本设备前，请确认该导轨承载重量大于 150kg，否则，请务必使用本公司官方选配的导轨。若使用非本公司官方选配的导轨上架本设备，可能导致自备导轨的承载重量小于设备重量，造成设备上架失效等风险，对此，本公司将不承担任何责任或义务。

以上内容中，“我们”指代浪潮电子信息产业股份有限公司；浪潮电子信息产业股份有限公司拥有对以上内容的最终解释权。

目录

使用声明.....	ii
安全声明.....	iii
安全细则.....	iv
1 概述.....	1
2 安装 multipath-tools.....	3
2.1 安装方法.....	3
2.2 加载模块与服务.....	5
2.3 检查 multipath-tools 配置文件.....	7
2.4 设置 multipathd 服务开机启动.....	8
2.5 安装 inpath 管理工具.....	9
3 配置管理.....	11
3.1 逻辑盘的创建与映射.....	11
3.1.1 SAN switch 组网.....	11
3.1.2 iSCSI 组网.....	12
3.2 扫描映射的逻辑盘.....	12
3.2.1 SAN switch 组网.....	12
3.2.2 iSCSI 组网.....	13
3.3 配置 multipath-tools 多路径工具.....	13
3.3.1 配置文件介绍.....	13
3.3.2 配置文件修改.....	14
3.4 配置 FC 目标端口设备丢失超时参数.....	16
3.5 配置访问 SCSI block device 的超时时间.....	18
3.6 配置使用 FC-NVMe 逻辑卷.....	18
3.6.1 禁用 nvme_multipathing 多路径.....	18
3.6.2 配置文件修改.....	19
3.6.3 建立主机-存储连接.....	20
3.6.4 聚合多路径.....	21
3.7 Inpath 管理工具使用.....	21
3.7.1 inpath aggregation.....	21
3.7.2 inpath config.....	22
3.7.3 inpath delalias.....	22
3.7.4 inpath delete.....	23
3.7.5 inpath lsalias.....	24

3.7.6 inpath lspathpolicy	24
3.7.7 inpath port.....	25
3.7.8 inpath query.....	26
3.7.9 inpath restart.....	27
3.7.10 inpath scan.....	27
3.7.11 inpath setalias.....	28
3.7.12 inpath setpathpolicy.....	29
3.7.13 inpath setstate	30
3.7.14 inpath start.....	30
3.7.15 inpath status.....	31
3.7.16 inpath stop	31
3.7.17 inpath version	32
3.7.18 inpath -h.....	32
4 故障分析与解决.....	34
4.1 多路径设备聚合问题.....	34
4.1.1 问题描述.....	34
4.1.2 问题解答.....	34
4.2 分区表显示问题.....	35
4.2.1 问题描述.....	35
4.2.2 问题解答.....	36
4.3 Lustre 文件系统 block size 故障.....	39
4.3.1 故障描述.....	39
4.3.2 故障原因.....	39
4.3.3 故障处理.....	39
4.4 统信 UOS V20 多路径设备不聚合问题.....	39
4.4.1 故障描述.....	39
4.4.2 故障原因.....	40
4.4.3 故障处理.....	40
4.5 大量 FC 卷映射引起的多路径设备异常.....	40
4.5.1 故障描述.....	40
4.5.2 故障原因.....	40
4.5.3 故障处理.....	41
4.6 盘符漂移问题.....	41
4.6.1 故障描述.....	41
4.6.2 故障原因.....	41
4.6.3 故障处理.....	41
5 缩略语.....	43

1 概述

Linux 系统平台通常包含多路径模块 multipath-tools，multipath-tools 是 Linux 平台用于配置存储设备多路径支持的工具，通过配置 Linux 内核模块 device-mapper 实现对存储设备的多路径支持。

不同版本的多 multipath-tools 对浪潮统一系列存储的支持情况不同：

- multipath-tools 的版本号为 0.7.5 及以上时，默认支持浪潮统一系列存储，无需修改配置文件即可支持存储设备的多路径功能。
- multipath-tools 的版本号低于 0.7.5 时，需要修改多路径模块的配置文件才能支持浪潮统一系列存储的多路径功能。

不同 Linux 系统发行版本的 multipath-tools 会有差异，需根据具体版本进行安装配置。本文档以 Linux 发行版 RHEL 6.6（Red Hat Enterprise Level 6.6）为例，介绍在 Linux 平台配置多路径工具 multipath-tools，以支持浪潮统一系列存储的多路径功能。

- RHEL 6.6 发行版本基于 Linux 2.6.32 内核版本，其多路径工具 multipath-tools 的版本为 0.4.9。
- 其它 RHEL 发行版本的多路径工具与 RHEL 6.6 类似，请参考文档中对应的说明。
- CentOS、Oracle Linux、Kylin 3.2-4F、Ubuntu 与 RHEL 类似，请参考文档中对应的说明。
- SuSE 与 RHEL 存在差别，请参考文档中对应的说明。

适用存储产品

Linux 系统多路径软件适用的存储产品如表 1-1 所示。

表 1-1 适用存储产品

产品系列	产品型号
G2 系列	AS2150G2&AS2200G2&AS2600G2&AS5300G2&AS5500G2&AS5600G2&AS5800G2&AS6800G2
G2-F 系列	AS2600G2-F&AS5300G2-F&AS5500G2-F&AS5600G2-F&AS5800G2-F&AS6800G2-F
G5 系列	AS5300G5&AS5500G5&AS5600G5&AS5800G5&AS6800G5&HF5000G5&HF6000G5&HF8000G5
G5-I 系列	AS18000G5-I&HF18000G5-I
G6-CG 系列	HF6000G6-CG&HF8000G6-CG&HF18000G6-CG

G6-CGF 系列	HF6000G6-CGF&HF8000G6-CGF&HF18000G6-CGF
其他	HF5500

2 安装 multipath-tools

2.1 安装方法

1. 执行以下命令，检查当前系统中是否已经安装多路径工具：

```
rpm -qa | grep device-mapper-multipath
```

SUSE 系统 multipath 相关 rpm 包是 multipath-tools-*.rpm，检查安装包时使用如下命令：

```
rpm -qa | grep multipath
```

Linux-Rocky4.2 系统多路径工具的相关包名称为 multipath-tools-0.4.8-x86_64-linux-Rocky4.2-TS.pkg.tar.gz，检查安装包时使用如下命令：

```
pkginfo -l | grep multipath
```

Ubuntu/UOSV20 系统 multipath 相关 deb 包是 multipath-tools_*.deb，检查安装包时使用如下命令：

```
dpkg -l | grep multipath
```

2. 如果系统中已经安装了此模块，以上命令会列出模块名及版本。如果没有安装此模块，请执行以下安装步骤，从系统安装光盘中安装多路径工具：

- a. 将系统安装光盘装入光驱，并 mount 到一个挂载目录（例如：/mnt/cd0）。请执行以下命令：

```
mount /dev/cdrom /mnt/cd0
```

- b. 进入系统模块所在目录。如表 2-1 所示查找目录，并执行以下命令，如：

```
cd /mnt/cd0/Packages
```

不同 Linux 发行版 RPM 包目录会有差异，以系统发布版本的实际情况为准。如下表所示。

表 2-1 不同 Linux 发行版的 RPM 包在光盘中的目录差异

发行版	RPM 包目录
RHEL5.x	/mnt/cd0/Server
RHEL6.x	/mnt/cd0/Packages
RHEL7.x	/mnt/cd0/Packages
RHEL8.x	/mnt/cdrom/BaseOS/Packages
CentOS6.x	/mnt/cd0/Packages
CentOS7.x	/mnt/cd0/Packages

CentOS8.x	/mnt/cdrom/BaseOS/Packages
SuSE	/mnt/cd0/suse/x86_64 (选择相应的处理器)
Kylin 3.2-4F	/mnt/cd0/KYLIN
KylinV10	网站下载: http://archive.kylinos.cn/kylin/KYLIN-ALL
Ubuntu	Ubuntu 官方网站下载: https://ubuntu.com/
UOS V20	网站下载: https://uos.deepin.cn/uos

- c. 执行 rpm 包安装命令。如表 2-2 所示查找 rpm 包，并执行以下命令，如：

```
rpm -ivh device-mapper-multipath-libs-0.4.9-80.el6.x86_64.rpm
rpm -ivh device-mapper-multipath-0.4.9-80.el6.x86_64.rpm
```

- d. 不同 Linux 发行版 RPM 包会有差异，以系统发布版本的实际情况为准。例如，不同 Linux 发行版的 DM-Multipath 模块 RPM 包对比如下表所示：

表 2-2 不同 Linux 发行版的 RPM 包版本差异

发行版	device-mapper-multipath-libs	device-mapper-multipath	multipath-tools
RHEL5.1	0.4.7-12	0.4.7-12	-
RHEL5.2	0.4.7-17	0.4.7-17	-
RHEL5.8	0.4.7-48	0.4.7-48	-
RHEL6.2	0.4.9-46	0.4.9-46	-
RHEL6.3	0.4.9-56	0.4.9-56	-
RHEL6.4	0.4.9-64	0.4.9-64	-
RHEL6.5	0.4.9-72	0.4.9-72	-
RHEL6.6	0.4.9-80	0.4.9-80	-
RHEL6.7	0.4.9-87	0.4.9-87	-
RHEL6.8	0.4.9-93	0.4.9-93	-
RHEL7.0	0.4.9-66	0.4.9-66	-
RHEL7.1	0.4.9-77	0.4.9-77	-
RHEL7.2	0.4.9-85	0.4.9-85	-
CentOS7.2	0.4.9.85	0.4.9.85	-

CentOS6.4	0.4.9.64	0.4.9.64	-
CentOS6.5	0.4.9.72	0.4.9.72	-
SuSE-10	-	-	0.4.6-25.8
SuSE-11-SP1	-	-	0.4.8-40.21.1
SuSE-11-SP2	-	-	0.4.9-0.60.1
SuSE-11-SP3	-	-	0.4.9-0.83.2
SuSE-11-SP4	-	-	0.4.9-109.1
SuSE-12-SP1	-	-	0.5.0-46.1
Kylin 3.2-4F	0.4.9-64.ky3	0.4.9-64.ky3	-
Ubuntu16.04	-	-	0.5.0
Ubuntu18.04	-	-	0.7.4
Ubuntu19.04	-	-	0.7.4

3. 完成以上步骤后，需检查多路径工具是否已经安装成功。请参考步骤 1 中的命令。

2.2 加载模块与服务

请逐步执行以下命令加载内核 device-mapper 模块，并启动 multipathd 服务。

1. modprobe dm-multipath
2. modprobe dm-service-time
3. service multipathd start (RHEL6.x、CentOS6.x、SuSE、Kylin 3.2-4F)
或 systemctl start multipathd.service(RHEL7.x、CentOS7.x、Ubuntu)

Linux-Rocky4.2 系统请执行以下步骤：

1. 修改/etc/sysconfig/modules 配置文件，在文件最后一行添加 dm-multipath。
2. 创建 multipath 配置文件
 - a. 制作连接文件 ln -s /lib/udev/scsi_id /lib64/udev/scsi_id。
 - b. 执行 scsi_id -g -u -s /block/sdb,获取 sdb 的 wwid。
 - c. 创建一个/etc/multipath.conf 的配置文件，该文件在安装后不会自动创建，将步骤 b)中查看到的 wwid 写入到配置文件。配置文件内容如下：

```
defaults{
    user_friendly_names yes
}
```

```
blacklist{
devnode "^sda"

#不扫描本地磁盘，这里根据本地磁盘盘符来确定，该示例中假定本地盘符为 sda

}

multipaths {
    multipath {
        wwid 3600508b4000156d70001200000b0000
        alias mpatha
    }
    multipath {
        wwid 3600508b4000156d70001200000b0001
        alias mpathb
    }
}

devices{
    device{
        vendor "INSPUR"
        product "MCS"
        path_grouping_policy group_by_prio
        path_selector "round-robin 0"
        features "1 queue_if_no_path"
        prio alua
        path_checker tur
    }
}
```

```

failback immediate

no_path_retry "60"

rr_min_io 1000

dev_loss_tmo 120

fast_io_fail_tmo 5

}
}

```

3. 启动服务：`/etc/rc.d/init.d/multipathd start`

2.3 检查 multipath-tools 配置文件

查看系统中`/etc/multipath.conf` 文件是否存在。

如果不存在，将`/usr/share/doc/device-mapper-multipath-*/multipath.conf` 文件拷贝一份，文件命名为`multipath.conf` 放置于`/etc/`目录下。

上述斜体部分“`device-mapper-multipath-*`”表示相应 DM-Multipath 版本。不同 Linux 发行版默认的配置文件会有差异，以系统发布版本的实际情况为准：

表 2-3 不同 Linux 发行版的配置文件路径及文件名差异

发行版	配置文件路径及文件名
RHEL5.x	<code>/usr/share/doc/device-mapper-multipath-0.4.7/multipath.conf.synthetic</code>
RHEL6.x	<code>/usr/share/doc/device-mapper-multipath-0.4.9/multipath.conf</code>
RHEL7.x	<code>/usr/share/doc/device-mapper-multipath-0.4.9/multipath.conf</code>
RHEL8.x	<code>/usr/share/doc/device-mapper-multipath/multipath.conf</code>
CentOS6.x	<code>/usr/share/doc/device-mapper-multipath-0.4.9/multipath.conf</code>
CentOS7.x	<code>/usr/share/doc/device-mapper-multipath-0.4.9/multipath.conf</code>
CentOS8.x	<code>/usr/share/doc/device-mapper-multipath/multipath.conf</code>
SuSE10	<code>/usr/share/doc/packages/multipath-tools/multipath.conf.synthetic</code>
SuSE11	<code>/usr/share/doc/packages/multipath-tools/multipath.conf.synthetic</code>

SuSE12	/usr/share/doc/packages/multipath-tools/multipath.conf.synthetic
SuSE15	请参考 SuSE12 系统上的 multipath.conf 文件
Kylin 3.2-4F	/usr/share/doc/device-mapper-multipath-0.4.9/multipath.conf
Kylin V10	请参考 RHEL/CentOS 系统上的 multipath.conf 文件
Ubuntu16.04	/usr/share/doc/multipath-tools/examples/multipath.conf.synthetic
Ubuntu18.04	请参考 RHEL/CentOS 系统上的 multipath.conf 文件
Ubuntu19.04	请参考 RHEL/CentOS 系统上的 multipath.conf 文件
UOS V20	请参考 RHEL/CentOS 系统上的 multipath.conf 文件

2.4 设置 multipathd 服务开机启动

执行如下命令查看 multipathd 服务是否开机启动：

表 2-4 不同 Linux 发行版查看 multipathd 服务是否开机启动命令差异

发行版	命令
RHEL5.x	chkconfig --list multipathd
RHEL6.x	chkconfig --list multipathd
RHEL7.x	systemctl is-enabled multipathd.service
RHEL8.x	systemctl is-enabled multipathd.service
CentOS6.x	chkconfig --list multipathd
CentOS7.x	systemctl is-enabled multipathd.service
CentOS8.x	systemctl is-enabled multipathd.service
SuSE10	chkconfig --list multipathd
SuSE11	chkconfig --list multipathd
SuSE12	chkconfig --list multipathd
SuSE15	systemctl is-enabled multipathd.service
Kylin 3.2-4F	chkconfig --list multipathd
Kylin V10	systemctl is-enabled multipathd.service
Ubuntu16.04	systemctl is-enabled multipathd.service

Ubuntu18.04	systemctl is-enabled multipathd.service
Ubuntu19.04	systemctl is-enabled multipathd.service
UOS V20	systemctl is-enabled multipathd.service

若 multipathd 未配置开机启动，执行如下命令，设置 multipathd 服务开机启动。

表 2-5 不同 Linux 发行版设置 multipathd 服务开机启动命令差异

发行版	命令
RHEL5.x	chkconfig --level 35 multipathd on
RHEL6.x	chkconfig --level 35 multipathd on
RHEL7.x	systemctl enable multipathd.service
RHEL8.x	systemctl enable multipathd.service
CentOS6.x	chkconfig --level 35 multipathd on
CentOS7.x	systemctl enable multipathd.service
CentOS8.x	systemctl enable multipathd.service
SuSE10	chkconfig --level 35 multipathd on
SuSE11	chkconfig --level 35 multipathd on
SuSE12	chkconfig --level 35 multipathd on
SuSE15	systemctl enable multipathd.service
Kylin 3.2-4F	chkconfig --level 35 multipathd on
Kylin V10	systemctl enable multipathd.service
Ubuntu16.04	systemctl enable multipathd.service
Ubuntu18.04	systemctl enable multipathd.service
Ubuntu19.04	systemctl enable multipathd.service
UOS V20	systemctl enable multipathd.service

2.5 安装 inpath 管理工具

在 RedHat 6.3/6.5/6.7/7.3/7.5/8.0/8.1/8.2/8.3、CentOS 6.3/6.5/6.7/7.3/7.5/8.0/8.1/8.2/8.3、中标麒麟 6.x/7.x、SuSE12SP5/15SP1/15SP2、KylinV10 操作系统发行版上，执行以下步骤，安装 inpath 管理工具。

1. 执行以下命令，检查当前系统中是否已经安装 inpath 管理工具：

```
rpm -qa | grep InPath
```

2. 执行 rpm 包安装命令，如：

```
rpm -ivh InPath-tool.x86_64.rpm
```

不同 Linux 发行版 RPM 包会有差异，以系统发布版本的实际情况为准。

在 UOSV20 操作系统发行版上，执行以下步骤，安装 inpath 管理工具。

1. 执行以下命令，检查当前系统中是否已经安装 inpath 管理工具：

```
dpkg -l | grep inpath
```

2. 执行 deb 包安装命令。

```
dpkg -i inpath-tools_3.3.1-0_amd64.deb
```

3 配置管理

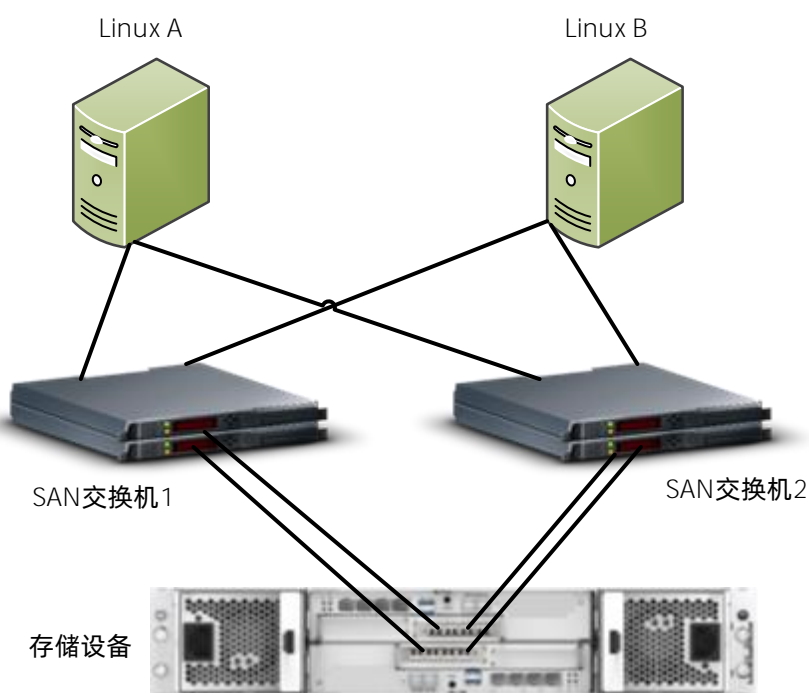
3.1 逻辑盘的创建与映射

使用浪潮统一系列存储的管理工具（GUI 或 CLI），添加新的逻辑盘，该逻辑盘可以称为“LUN”或“Virtual Disk”或“卷”，然后将新创建的逻辑盘映射给主机。

3.1.1 SAN switch 组网

典型的双机应用集群与浪潮统一系列存储通过双冗余 FC 网络组网的方式，如图 3-2 所示。实际应用中，主机与每个 FC 网络的连接数不限于 1 条。

图 3-1 SAN switch 组网



1. 执行以下命令，查看主机上光纤 HBA 的 WWN：

```
cat /sys/class/fc_host/host*/port_name
```

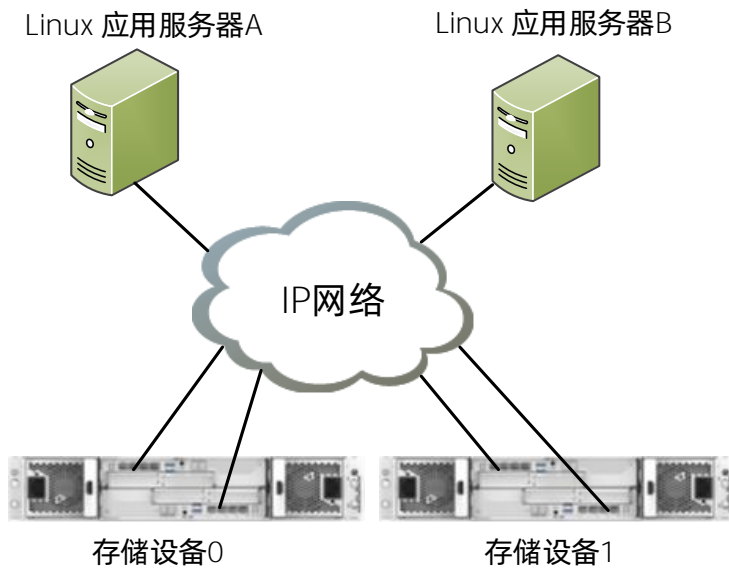
2. 使用浪潮统一系列存储的 GUI 管理页面，利用步骤 1 中获取的光纤 HBA 的 WWN，创建主机（参考《基本功能配置手册》）。

3. 创建新的逻辑盘，然后将新创建的逻辑盘，映射给步骤 2 创建的主机。

3.1.2 iSCSI 组网

应用服务器与存储系统通过 iSCSI 方式组网的典型应用，如图 3-2 所示。这是一种较为灵活的部署方式，可以通过现有的 IP 网络组网。如果考虑性能，也可以组建专门的用于连接应用服务器与存储设备的专用 IP 网络。

图 3-2 iSCSI 组网



1. 执行以下命令，记录 iSCSI initiator name (IQN):

```
cat /etc/iscsi/initiatorname.iscsi
```

2. 使用浪潮统一系列存储的管理工具（图形界面），利用步骤 1 中获取的 IQN，创建服务器主机。
3. 创建新的逻辑盘，然后将新创建的逻辑盘，映射给步骤 2 创建的主机。

3.2 扫描映射的逻辑盘

3.2.1 SAN switch 组网

1. 执行以下命令，查看 FC HBA 卡端口状态：

```
more /sys/class/ fc_host /hostX/state
```

hostX：代表具体的 FC HBA 卡端口。

如果 HBA 卡端口未知，执行以下指令，查看所有 HBA 卡端口状态：

```
for x in `ls /sys/class/ fc_host` ; do more /sys/class/ fc_host /$x/state; done
```

2. 执行以下命令，重新扫描 SCSI 设备：

```
echo "- - -" > /sys/class/scsi_host/hostX/scan
```

hostX:代表具体的 FC HBA 卡端口。

如果 HBA 卡端口未知，执行以下指令，重新扫描所有设备：

```
for x in `ls /sys/class/fc_host`; do echo "- - -" > /sys/class/scsi_host/$x/scan; done
```

3. 以下命令，查看磁盘设备：

```
fdisk -l
```



说明

上面命令中的 `ls /sys/class/fc_host`，“`”不是单引号，而是英文输入法下“~、”键按出来的，一般是在“Esc”键下面的一个按键。

3.2.2 iSCSI 组网

1. 执行以下命令，重新扫描设备：

```
for x in `ls /sys/class/iscsi_host`; do echo "- - -" > /sys/class/scsi_host/$x/scan; done
```

2. 执行以下命令，查看磁盘设备：

```
fdisk -l
```

3.3 配置 multipath-tools 多路径工具

3.3.1 配置文件介绍

multipath-tools 模块的配置文件默认是/etc/multipath.conf，如果此文件不存在，或没有任何有效配置，则采用多路径工具缺省的配置（见 multipath.conf.defaults）。在/usr/share/doc/device-mapper-multipath-0.4.9 目录下，通过以下文件，便于了解配置设备的支持信息：

表 3-1 配置文件

文件名	说明
multipath.conf	配置文件样例
multipath.conf.annotated	配置项取值详细说明
multipath.conf.defaults	多路径工具缺省配置

multipath.conf 包括如下几个配置单元：

- blacklist, 需要过滤屏蔽的设备, 即不支持列表中设备的多路径功能
- defaults, 多路径工具缺省配置, 覆盖 multipath.conf.defaults 中的 defaults 值
- multipaths, 多路径单元 (LUN) 参数配置, 按 WWID 指定每个需要配置的多路径单元
- devices, 存储设备参数配置, 按 vendor, product 指定要配置的存储设备



说明

具体每一配置单元及其配置项的解释, 请参考文件：

/usr/share/doc/device-mapper-multipath-0.4.9/multipath.conf.annotated。

3.3.2 配置文件修改

多路径工具 multipath-tools 0.7.5 及之后的版本, 无需修改配置文件, 默认支持浪潮统一系列存储的最佳配置。

多路径工具 multipath-tools 0.7.5 之前的版本, 默认情况下不支持浪潮统一系列存储的多路径优化配置, 需要修改/etc/multipath.conf 配置文件实现对浪潮统一系列存储的最佳支持。请联系浪潮客服人员获取最新版本的自动化配置脚本 instoragempiocfg.sh, 并在主机端执行此脚本, 即自动完成/etc/multipath.conf 的配置工作。如果没有上述脚本, 或上述脚本不支持当前的系统环境, 请按以下步骤修改多路径配置文件并使之生效。



说明

如果是云平台 (如 OpenStack、ICOS、ICKS、ICS 等), 请确认 multipath.conf 文件中 user_friendly_names 参数为 no。如果不是, 请修改该参数。

1. 请在/etc/multipath.conf 的 devices{}配置单元中, 增加以下配置信息。

```
devices{
    device {
        vendor "INSPUR"
        product "MCS"
        path_grouping_policy group_by_prio
        path_selector "round-robin 0"
        features "1 queue_if_no_path"
        prio alua
    }
}
```

```
path_selector tur
path_checker tur
failback immediate
no_path_retry "60"
rr_min_io 1000
dev_loss_tmo 120
fast_io_fail_tmo 5
}
}
```

以上配置项中，`path_selector` 设置路径选择策略为 `round-robin`，即路径组中轮循使用路径。

在其他配置信息中，若 `path_selector` 设置路径选择策略为 `service time`，则是最优路径组中轮循使用路径。如果需要在所有路径（最优路径组+非最优路径组）中使用负载均衡策略，请修改该配置项的值：

```
path_grouping_policy multibus
```

请确保 `multipath.conf` 配置文件中，`blacklist{}` 配置项没有设置对阵列的屏蔽。例如，在 Red Hat 5.x 系统中，`multipath.conf` 配置文件中，默认屏蔽了所有的 `devnode`，即：

```
blacklist{
devnode "*"
}
```

需要将屏蔽设置注释掉，或者在 `blacklist_exceptions{}` 配置项中设置屏蔽例外。

2. 重启多路径服务，执行以下命令，重新加载配置。

```
service multipathd reload (RHEL6.x、CentOS6.x、SuSE、Kylin 3.2-4F)
```

或

```
systemctl reload multipathd.service (RHEL7.x/8.x、CentOS7.x/8.x、Ubuntu、UOS V20、Kylin V10)
```

3. 检查参数是否生效



说明

除了需要在多路径的配置文件中将 `dev_loss_tmo` 设置为 120，`fast_io_fail_tmo` 设置为 5 外，需要进一步检查设置的值是否确实生效了，详见 3.4 章节。如果未生效，请按 3.4 章节永久生效方法修改。

检验配置是否生效，请执行以下命令：

```
multipath -ll
```

如果结果中输出存储的多路径拓扑，那么表示配置生效；若没有输出存储的多路径拓扑，表示配置没有生效，请检查系统日志文件或 multipath 命令的前台输出信息。

请执行以下命令，显示当前多路径配置信息：

```
multipathd show config
```

3.4 配置 FC 目标端口设备丢失超时参数

如果未设置 fast_io_fail_tmo 为 5 的情况下，FC 目标端口的设备丢失超时参数(dev_loss_tmo)，将影响到多路径工具的 failover 功能。请根据系统中实际的 FC 卡类型进行修改。



说明：双活场景下，请将 dev_loss_tmo 修改为 120。

在双活场景下，为了防止多路径磁盘出现掉盘问题，请务必将 FC 目标端口的设备丢失超时参数(即 dev_loss_tmo)修改为 120，同时 fast_io_fail_tmo 修改为 5。

下表介绍两种常用 FC HBA 卡的参数修改方法。

表 3-2 参数修改方法

HBA 卡	缺省值	修改驱动超时参数方法
emulex	30	<p>方法一：</p> <p>修改 emulex FC HBA 卡驱动模块 lpfc 的驱动超时参数。以修改驱动超时参数 120 秒为例，步骤如下：</p> <ol style="list-style-type: none"> 3. 创建/etc/modprobe.d/lpfc.conf 文件，在文件中增加以下配置，options lpfc lpfc_devloss_tmo=120 4. 备份 initramfs： mv /boot/initramfs-`uname -r`.img /boot/initramfs-`uname -r`.img.bak 或备份 initrd： mv /boot/initrd-`uname -r`.img /boot/initrd-`uname -r`.img.bak 5. 更新 initramfs： dracut -v -f 或更新 initrd： mkinitrd -f -v /boot/initrd-`uname -r`.img `uname -r` 6. 重启操作系统 reboot <p>说明：上面命令中的`uname -r`， ``不是单引号，而是英文输入法下“~、”键按出来的，一般是在“Esc”键下面的一个按键。</p>

		<p>方法二（临时生效，重启后参数回退默认值）：</p> <p>开机启动执行脚本：</p> <pre>for targetport in /sys/class/fc_remote_ports/rport*;do echo 120 > \${targetport}/dev_loss_tmo;done</pre> <p>方法三（永久生效）：</p> <p>新增/etc/udev/rules.d/inspur-tmo.rules 文件内容如下（以 120 秒为例）：</p> <pre>ACTION!="add change", GOTO="tmo_end" KERNELS=="rport-?*",SUBSYSTEM=="fc_remote_ports", ATTR{roles}=="FCP Target*", ATTR{dev_loss_tmo}="120", ATTR{fast_io_fail_tmo}="5" LABEL="tmo_end"</pre> <p>修改后，执行以下命令使其生效：</p> <pre>udevadm control -reload udevadm trigger --subsystem-match=fc_remote_ports</pre>
qllogic	35	<p>方法一（临时生效，重启后参数回退默认值）：</p> <p>开机启动执行脚本：</p> <pre>for targetport in /sys/class/fc_remote_ports/rport*;do echo 120 > \${targetport}/dev_loss_tmo;done</pre> <p>方法二（永久生效）：</p> <p>新增/etc/udev/rules.d/inspur-tmo.rules 文件内容如下（以 120 秒为例）：</p> <pre>ACTION!="add change", GOTO="tmo_end" KERNELS=="rport-?*",SUBSYSTEM=="fc_remote_ports", ATTR{roles}=="FCP Target*", ATTR{dev_loss_tmo}="120", ATTR{fast_io_fail_tmo}="5" LABEL="tmo_end"</pre> <p>修改后，执行以下命令使其生效：</p> <pre>udevadm control -reload udevadm trigger --subsystem-match=fc_remote_ports</pre>

修改之后，参数修改是否成功验证方法：

```
cat /sys/class/fc_remote_ports/rport-*/dev_loss_tmo
```


3.5 配置访问 SCSI block device 的超时时间

请根据实际应用场景对此节介绍的参数进行修改。

通过修改 udev-rules 进行修改，新增规则文件：inspur-storage-mcs.rules：

以修改为 30 秒为例，在/etc/udev/rules.d/ inspur-storage-mcs.rules 文件内容如下：

```
SUBSYSTEM=="block",
ACTION=="add",
ENV{ID_VENDOR}=="INSPUR",
ENV{ID_MODEL}=="MCS",
RUN+="/bin/sh -c 'echo 30 > /sys/block/%k/device/timeout'"
```

修改之后，参数修改是否成功验证方法：

```
cat /sys/block/sd*/device/timeout
```

3.6 配置使用 FC-NVMe 逻辑卷



说明：仅 G5 和 G5-I 系列支持配置 FC-NVMe 逻辑卷。

在 RedHat 8.0/8.1/8.2/8.3、CentOS 8.0/8.1/8.2/8.3、SuSE12SP5/15SP1/15SP2 操作系统发行版上，支持使用系统自带的多路径 DM-multipath 接管 FC-NVMe 逻辑卷。

3.6.1 禁用 nvme_multipathing 多路径

当使用 DM-multipath 接管 FC-NVMe 逻辑卷时，需要将 nvme_multipathing 多路径禁用掉。可以使用如下命令查看 nvme_multipathing 多路径状态：

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

如果查询结果显示 “No such file or directory” 或 “N”，不需要进行禁用 nvme_multipathing 自带多路径操作。如果查询结果显示 “Y”，表示 nvme_multipathing 多路径已启用，禁用方法请参考表 3-3。

表 3-3 禁用 nvme_multipathing 多路径方法

操作系统	禁用 nvme_multipathing 多路径方法
Redhat 8.X CentOS 8.X	1. 编辑/etc/default/grub 文件，增加 nvme_core.multipath=N，例如：

	<pre>[root@localhost ~]# cat /etc/default/grub GRUB_TIMEOUT=5 GRUB_DISTRIBUTOR="\$(sed 's, release .*\$,,g' /etc/system-release)" GRUB_DEFAULT=saved GRUB_DISABLE_SUBMENU=true GRUB_TERMINAL_OUTPUT="console" GRUB_CMDLINE_LINUX="crashkernel=auto resume=/dev/mapper/rhel-swap rd.lvm.lv=rhel /root rd.lvm.lv=rhel/swap nvme_core.multipath=N rhgb quiet" GRUB_DISABLE_RECOVERY="true" GRUB_ENABLE_BLSCFG=true [root@localhost ~]#</pre> <p>2. 配置完成后执行如下命令：</p> <pre>grub2-mkconfig -o /boot/grub2/grub.cfg</pre> <p>3. 重启主机，再次查看确认 NVME 多路径状态已禁用成功：</p> <pre>cat /sys/module/nvme_core/parameters/multipath</pre> <p>4. 命令查询结果为 NO 即可</p>
SLES 15SP1	<p>1. 编辑/boot/grub2/grub.cfg，在如下两处添加 nvme_core.multipath=N</p> <pre>linux-4931:~ # cat /boot/grub2/grub.cfg grep nvme_core linux /boot/vmlinuz-4.12.14-120-default root=UUID=d6a51f53-c9d5-4966-9a 80-c68a6f93552e \${extra cmdline} resume=/dev/sda1 mitigations=auto splash=silent quiet nvme_core.multipath=N showopts crashkernel=178M,high crashkernel=72M,low linux /boot/vmlinuz-4.12.14-120-default root=UUID=d6a51f53-c9d5 -4966-9a80-c68a6f93552e \${extra cmdline} resume=/dev/sda1 mitigations=auto splas h=silent quiet nvme_core.multipath=N showopts crashkernel=178M,high crashkernel=7 2M,low</pre> <p>2. 重启主机，再次查看确认 NVME 多路径状态已禁用成功：</p> <pre>cat /sys/module/nvme_core/parameters/multipath</pre> <p>3. 命令查询结果为 NO 即可</p>

3.6.2 配置文件修改

请在/etc/multipath.conf 的 devices {}配置单元中，增加以下推荐配置信息。

```
devices {
    device {
        vendor "NVME"
        product "INSPURMCS"
        uid_attribute "ID_WWN"
        path_grouping_policy "multibus"
        path_selector "round-robin 0"
```

```
path_checker "directio"

prio "const"

failback immediate

rr_min_io_rq 1

user_friendly_names yes

fast_io_fail_tmo 15

dev_loss_tmo 600

no_path_retry "queue"

}

}
```

3.6.3 建立主机-存储连接

1. 在主机上，获取主机适配器 NQN 信息，执行以下命令：

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

2. 利用主机适配器 NQN 信息，在存储 GUI 上创建 NVMe 类型主机
3. 在存储 GUI 上创建卷，并映射给步骤 2 创建的 NVMe 类型主机
4. 在主机上，查找存储目标器的 WWPN 和 WWNN。

执行命令，获取所有存储目标器 WWPN：

```
cat /sys/class/fc_remote_ports/rport-*/scsi_target_id
cat /sys/class/fc_remote_ports/rport-*/port_name
```

记录结果中满足 scsi_target_id 值为-1，且 WWPN 含字符串"0x56c92bf"的存储器目标端口。

执行命令，获取所有存储目标器 WWNN：

```
cat /sys/class/fc_remote_ports/rport-*/node_name
```

根据 WWPN 和 WWNN 的对应关系，记录存储目标器的 WWNN

5. 在主机上，查找主机适配器的 WWPN 和 WWNN

执行命令：

```
cat /sys/class/fc_host/host*/port_state
```

```
cat /sys/class/fc_host/host*/port_name
cat /sys/class/fc_host/host*/node_name
```

根据上述三个命令执行结果，查找状态为 online 的主机适配器的 WWPN 和 WWNN

6. 在主机上，执行 nvme discover 发现 NVMe 存储目标器

利用前面步骤获取的存储目标器和主机适配器的 WWNN 和 WWPN 信息，执行 nvme discover 命令：

```
nvme discover -transport=fc -traddr=nn-0xWWNN:pn-0xWWPN -host-traddr= nn-0xWWNN:pn-0xWWPN
```

发现成功后，记录 nvme discover 命令响应信息中的存储的 Subsystem_NQN 信息。

7. 在主机上，执行 nvme connect 连接 NVMe 存储目标器

利用步骤 5 获取的存储目标器和主机适配器的 WWNN 和 WWPN 信息和步骤 6 中记录的存储的 Subsystem_NQN 信息，执行 nvme connect 命令：

```
nvme connect -transport=fc -traddr=nn-0xWWNN:pn-0xWWPN -host-traddr= nn-0xWWNN:pn-0xWWPN -n Subsystem_NQN
```

3.6.4 聚合多路径

执行如下命令，聚合多路径。

```
multipath -r
```

3.7 Inpath 管理工具使用



说明：仅 G5 和 G5-I 系列支持 FC-NVMe 相关配置。

提供与多路径设备及路径相关的配置与管理功能。

3.7.1 inpath aggregation

inpath aggregation 命令用于重载聚合多路径设备。

命令行语法格式：

```
inpath aggregation
```

参数：无

示例：

执行 `inpath aggregation`，输出以下信息：

图 3-2 执行 `inpath aggregation` 命令

```
[root@localhost ~]# inpath aggregation
-----
Command executed successfully.
-----
```

输出项解释：

Command executed successfully: 命令执行成功

3.7.2 inpath config

`inpath config` 命令用于自动配置多路径设备参数。



说明

使用此命令会主动修改配置文件中的值，一般即时生效。如果未即时生效，请执行 `inpath aggregation` 重载操作。

命令行语法格式：

```
inpath config
```

参数：无

示例：

执行 `inpath config`，输出以下信息：

图 3-3 执行 `inpath config` 命令

```
[root@localhost ~]# inpath config
-----
Command executed successfully.
-----
```

输出项解释：

Command executed successfully: 命令执行成功

3.7.3 inpath delalias

`inpath delalias` 命令用于删除某一多路径设备别名。

命令行语法格式：

```
inpath delalias -wwid <wwid_val>
```

参数：

wwid_val：多路径设备的 WWID 或 FC-NVMe 设备 EUI 值

示例：

执行 `inpath delalias -wwid 36005076708a20800800000000000094`，输出以下信息：

图 3-4 执行 `inpath delalias` 命令

```
[root@localhost ~]# inpath delalias -wwid 36005076708a20800800000000000094
-----
Command executed successfully.
-----
```

输出项解释：

Command executed successfully：命令执行成功

3.7.4 inpath delete

`inpath delete` 命令用于删除特定多路径设备。

命令行语法格式：

```
inpath delete <path_alias>
```

参数：

path_alias：多路径设备别名

示例：

执行 `inpath delete mpathb`，输出以下信息：

图 3-5 执行 `inpath delete mpathb` 命令

```
[root@localhost ~]# inpath delete mpathb
-----
Command executed successfully.
-----
```

输出项解释：

Command executed successfully：命令执行成功

3.7.5 inpath lsalias

inpath lsalias 命令用于查询所有多路径设备别名名称。

命令行语法格式：

```
inpath lsalias
```

参数：无

示例：

执行 inpath lsalias，输出以下信息：

图 3-6 执行 inpath lsalias 命令

```
[root@localhost ~]# inpath setalias -wwid 36005076708a208008000000000000094 -alias lun1
-----
Command executed successfully.
-----
[root@localhost ~]# inpath lsalias
-----
wwid                alias
36005076708a208008000000000000094  lun1
-----
```

输出项解释：

Command executed successfully: 命令执行成功

3.7.6 inpath lspathpolicy

inpath lspathpolicy 命令用于查询多路径设备路径分组策略和组内负载均衡算法。

命令行语法格式：

```
inpath lspathpolicy
```

参数：无

示例：

执行 inpath lspathpolicy，输出以下信息：

图 3-7 执行 inpath lspathpolicy 命令

```
[root@localhost ~]# inpath lspathpolicy
-----
SCSI DISK:
group_policy: multibus
chooser_type: service-time
NVMe DISK:
group_policy: multibus
chooser_type: service-time
-----
```

输出项解释：

- SCSI DISK: iSCSI 或 FC-SCSI 多路径块设备
- NVMe DISK: FC-NVMe 多路径块设备
- group_policy: 路径分组策略
- chooser_type: 组内负载均衡算法

3.7.7 inpath port

inpath port 命令用于查询主机适配器及存储目标端口基本信息。

命令行语法格式:

```
inpath port
```

参数: 无

示例:

执行 inpath port, 输出以下信息:

图 3-8 执行 inpath port 命令

```
[root@localhost ~]# inpath port
```

host	wwnn	wwpn	state	remote_port	remote_wwnn	remote_wwpn	type	state
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-10	0x5005076803002764	0x50050768032a2764	NVME	OnLine
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-11	0x5005076805002884	0x5005076805272084	SCSI	OnLine
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-12	0x5005076805002884	0x5005076805262084	SCSI	OnLine
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-13	0x5005076805002885	0x5005076805262085	SCSI	OnLine
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-14	0x5005076805002885	0x5005076805262085	NVME	OnLine
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-15	0x5005076803002764	0x5005076803222764	Primary	OnLine
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-16	0x5005076803002764	0x50050768032b2764	NVME	OnLine
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-17	0x5005076805002884	0x50050768052b2884	NVME	OnLine
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-18	0x5005076805002884	0x50050768052a2884	NVME	OnLine
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-19	0x5005076805002885	0x50050768052a2885	NVME	OnLine
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-20	0x5005076805002885	0x5005076805272885	SCSI	OnLine
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-3	0x5005076803002764	0x5005076803262764	SCSI	OnLine
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-4	0x5005076805002884	0x5005076805232884	Primary	OnLine
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-5	0x5005076803002764	0x5005076803232764	Primary	OnLine
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-6	0x5005076805002884	0x5005076805222884	Primary	OnLine
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-7	0x5005076805002885	0x5005076805222885	Primary	OnLine
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-8	0x5005076805002885	0x5005076805232885	Primary	OnLine
host33	0x200000109bc18a3f	0x100000109bc18a3f	OnLine	rport-33:0-9	0x5005076803002764	0x5005076803272764	SCSI	OnLine

输出项解释:

- host: 主机适配器名称
- wwnn: 主机适配器的 WWNN
- wwpn: 主机适配器端口的 WWPNN
- state: 主机适配器端口的状态
- remote_port: 远端存储目标端口名称
- remote_wwnn: 远端存储目标端口 WWNN
- remote_wwpn: 远端存储目标端口 WWPNN
- type: 端口类型
- state: 远端存储目标端口状态

3.7.8 inpath query

inpath query 命令显示系统中多路径设备的相关信息。

命令行语法格式：

```
inpath query
```

参数：无

示例：

执行 inpath query，输出以下信息：

图 3-9 执行 inpath query 命令

```
[root@localhost ~]# inpath query
-----
path_alias: mpatha
wwid: 3600507600d689c0d000000000000002
device_number: 2
vendor: INSPUR
product_type: MCS
lun_size: 4.0G
vdisk_name: centos8.3_fc_160174_vdisk0
host_interface: fc
=====
dev_id dev_name dev_node path_group_policy path_group_priority path_group_status path_status inpath_status target_port
8:80 sdr 33:0:9:0 service-time 0 50 active active running 0x5005076805262085
8:112 sdr 33:0:11:0 service-time 0 50 active active running 0x5005076805272085
8:16 sdb 33:0:7:0 service-time 0 10 enabled active running 0x5005076805272084
8:48 sdd 33:0:8:0 service-time 0 10 enabled active running 0x5005076805262084
-----
path_alias: mpathb
wwid: eui.d00000000000000016005076000d689c06
device_number: 3
vendor: NVME
product_type: INSPUR1813
lun_size: 3.0G
vdisk_name: centos8.1_nvme_160173_vdisk0
host_interface: fc-nvme
=====
dev_id dev_name dev_node path_group_policy path_group_priority path_group_status path_status inpath_status target_port
259:0 nvme0n1 0:4144:1:23 round-robin 0 50 active active running 0x50050768052b2085
259:2 nvme1n1 1:48:1:23 round-robin 0 50 active active running 0x50050768052b2084
259:4 nvme2n1 2:49:1:23 round-robin 0 50 active active running 0x50050768052a2084
259:6 nvme3n1 3:4145:1:23 round-robin 0 50 active active running 0x50050768052a2085
```

输出项解释：

- path_alias: 用户自定义设备别名
- wwid: 全球识别符
 - World Wide Identifier(wwid): SCSI 块设备使用该标识符，唯一的标识一个磁盘
 - Extended Unique Identifier(EUI): FC-NVMe 块设备使用该标识符，唯一的标识一个磁盘
- device_number: 多路径设备编号
- vendor: 厂商编号
- product_type: 产品编号
- lun_size: 多路径存储设备容量
- vdisk_name: 多路径设备存储侧设备名
- host_interface: 指定接口类型，可选的值：iscsi, fc, fc-nvme
- dev_id: 路径 ID 标识
- dev_name: 路径名称
- dev_node: 路径 host: chanel: target: lun 信息

- path_group_policy: 负载均衡策略
- path_group_priority: 路径优先级权重值
- path_group_status: 路径组状态
 - active: 处理 IO 的路径组
 - enabled: 备用路径组, 若处理 IO 路径组异常, 则接管 IO
- path_status: 路径的物理状态
 - active: 路径状态正常, 可处理 IO
 - fail: 路径状态异常, 无法处理 IO
- inpath_status: 多路设备状态
- target_port: 远端存储目标端口 WWPN

3.7.9 inpath restart

inpath restart 命令用于重启多路径服务。

命令行语法格式:

```
inpath restart
```

参数: 无

示例:

执行 inpath restart, 输出以下信息:

图 3-10 执行 inpath restart 命令

```
[root@localhost ~]# inpath restart
-----
Command executed successfully.
-----
```

输出项解释:

Command executed successfully: 命令执行成功

3.7.10 inpath scan

inpath scan 命令用于扫描物理主机上报的逻辑卷, 包含 FC、iSCSI 和 FC-NVMe 链路。

命令行语法格式:

```
inpath scan
```

参数: 无

示例：

执行 inpath scan，输出以下信息：

图 3-11 执行 inpath scan 命令

```
[root@localhost /]# inpath scan
-----
Command executed successfully.
```

输出项解释：

Command executed successfully: 命令执行成功

3.7.11 inpath setalias

inpath setalias 命令用于为多路径设备设置别名。



说明

使用此命令会主动修改配置文件中的值，一般即时生效。如果未即时生效，请执行 inpath aggregation 重载操作。

命令行语法格式：

```
inpath setalias -wwid <wwid_val> -alias <alias_val>
```

参数：

- wwid_val: 多路径设备的 WWID 或 FC-NVMe 设备 EUI 值
- alias: 多路径设备别名名称

示例：

执行 inpath setalias -wwid 36005076708a20800800000000000094 -alias lun1，输出以下信息：

图 3-12 执行 inpath setalias 命令

```
[root@localhost /]# inpath setalias -wwid 36005076708a20800800000000000094 -alias lun1
-----
Command executed successfully.
```

输出项解释：

Command executed successfully: 命令执行成功

3.7.12 inpath setpathpolicy

inpath setpathpolicy 命令用于设置多路径设备的路径分组策略和组内负载均衡路径选择算法。



说明

使用此命令会主动修改配置文件中的值，一般即时生效。如果未即时生效，请执行 inpath aggregation 重载操作。

命令行语法格式：

```
inpath setpathpolicy [-scsi/-nvme]-group <group_type> -chooser <choose_type>
```

参数：

- scsi：设置 SCSI 协议多路径设备
 - nvme：设置 FC-NVMe 协议多路径设备
- 说明：若不指定参数 -scsi 或 -nvme，则为设置 SCSI 多路径块设备
- group：多路径设备路径分组策略
 - failover：每个路径为一个路径组
 - multibus：所有有效路径为一个路径组
 - group_by_serial：每个检测到的序列号为一个路径组
 - group_by_prio：相同优先权重值的路径为一个路径组
 - group_by_node_name：每个目标节点名为一个路径组
 - chooser：组内负载均衡算法
 - round-robin：在路径组内循环向每个路径发送同样数量的 I/O
 - queue-length：将待处理 I/O 发送到未处理 I/O 请求最少的路径上
 - service-time：将待处理 I/O 发送到预计服务时间最短的路径上

示例：

执行 inpath setpathpolicy -group multibus -chooser service-time，输出以下：

图 3-13 执行 inpath setpathpolicy 命令

```
[root@localhost ~]# inpath setpathpolicy -group multibus -chooser service-time
-----
Command executed successfully.
-----
[root@localhost ~]# inpath setpathpolicy -nvme -group multibus -chooser service-time
-----
Command executed successfully.
-----
[root@localhost ~]# inpath setpathpolicy -scsi -group multibus -chooser service-time
-----
Command executed successfully.
-----
[root@localhost ~]# inpath lspathpolicy
-----
SCSI DISK:
group_policy: multibus
chooser_type: service-time
NVMe DISK:
group_policy: multibus
chooser_type: service-time
-----
```

输出项解释：

Command executed successfully: 命令执行成功

3.7.13 inpath setstate

inpath setstate 命令用于设置多路径设备的路径状态。

命令行语法格式：

```
inpath setstate -path_alias <alias_val> -device_name <pathname_val> -state <state_val>
```

参数：

- path_alias: 多路径设备别名
- device_name: 路径设备的名称
- state: 路径状态
- online: 设置路径状态为在线
- offline: 设置路径状态为离线

示例：

执行 inpath setstate -path_alias mpatha -device_name sdb-state offline, 输出以下信息：

图 3-14 执行 inpath setstate 命令

```
[root@localhost ~]# inpath setstate -path_alias mpatha -device_name sdb -state offline
-----
Command executed successfully.
-----
```

输出项解释：

Command executed successfully: 命令执行成功

3.7.14 inpath start

inpath start 命令用于启动多路径服务。

命令行语法格式：

```
inpath start
```

参数：无

示例：

执行 inpath start, 输出以下信息：

图 3-15 执行 inpath start 命令

```
[root@localhost ~]# inpath start
-----
Command executed successfully.
-----
```

输出项解释：

Command executed successfully: 命令执行成功

3.7.15 inpath status

inpath status 命令显示多路径服务的状态。

命令行语法格式：

```
inpath status
```

参数：无

示例：

执行 inpath status，输出以下信息：

图 3-16 执行 inpath status 命令

```
[root@localhost ~]# inpath status
-----
Service status: active
-----
```

输出项解释：

Command executed successfully: 命令执行成功

3.7.16 inpath stop

inpath stop 命令用于停止多路径服务。

命令行语法格式：

```
inpath stop
```

参数：无

示例：

执行 inpath stop，输出以下信息：

图 3-17 执行 inpath stop 命令

```
[root@localhost ~]# inpath stop
-----
Command executed successfully.
-----
```

输出项解释：

Command executed successfully: 命令执行成功

3.7.17 inpath version

inpath version 命令用于查询多路径工具的版本号。

命令行语法格式：

```
inpath version
```

参数：无

示例：

执行 inpath version，输出以下信息：

图 3-18 执行 inpath version 命令

```
[root@localhost ~]# inpath version
-----
Software name: InPath
Vendor: Inspur
Version: 3.3.1
Copyright Inspur Corporation 2019,2021.All Rights Reserved.
-----
```

输出项解释：

Command executed successfully: 命令执行成功

3.7.18 inpath -h

inpath -h 命令用于显示多路径工具的帮助信息。

命令行语法格式：

```
inpath -h
```

参数：无

示例：

执行 inpath -h，输出以下信息：

图 3-19 执行 inpath -h 命令

```
[root@localhost ~]# inpath -h
-----
aggregation - Aggregate the mpio device.
config      - Config mpio service.
delalias    - Delete the alias configure for the specified path.
delete      - Delete the specified mpio device.
lsalias     - Query path alias configure.
lspathpolicy - Query path group policy and path chooser type.
port        - Query the fc connection information.
query       - Query the mpio information.
restart     - Restart the mpio service.
scan        - Scan luns reported by all fc and iscsi hosts.
setalias    - Set alias for the specified path.
setpathpolicy - Set path group policy and path chooser type.
setstate    - Make the path online or offline.
start       - Start the mpio service.
status      - Show the mpio service status.
stop        - Stop the mpio service.
version     - Show mpio software information
-----
```

输出项解释：

显示帮助信息

4 故障分析与解决

4.1 多路径设备聚合问题

4.1.1 问题描述

对于多路径管理过程中，出现多路径卷聚合不正常，如：缺少某一条路径，或者某一条路径显示不正常。此时需要对此多路径卷进行重新聚合的操作。

4.1.2 问题解答

尝试通过以下步骤恢复：

1. 执行以下命令，删除设备文件：

```
multipath -f mpathX  
echo 1 > /sys/block/sdY/device/delete
```

mpathX：某一多路径设备

sdY：多路径设备的某一条路径对应的设备文件名。

例如多路径设备/dev/mapper/mpathb 由/dev/sdc 、/dev/sdd 、/dev/sde 、/dev/sdf 4 个设备聚合而成，则应执行以下命令：

```
multipath -f mpathb  
echo 1 > /sys/block/sdc/device/delete;  
echo 1 > /sys/block/sdd/device/delete;  
echo 1 > /sys/block/sde/device/delete;  
echo 1 > /sys/block/sdf/device/delete
```

2. 执行以下命令，查询系统中的 scsi host：

```
ls /sys/class/scsi_host  
或者
```

```
ls /sys/class/iscsi_host
```

3. 执行以下命令，扫描 scsi host，使上报 LUN：

```
echo "--" > /sys/class/scsi_host/hostX/scan
```

hostX：代表具体的物理 host 端口。

对于 SAN switch 组网，上报所有 LUN，命令如下：

```
for x in `ls /sys/class/fc_host`; do echo "--" > /sys/class/scsi_host/$x/scan; done
```

对于 iSCSI 组网，上报所有 LUN，命令如下：

```
for x in `ls /sys/class/iscsi_host`; do echo "--" > /sys/class/scsi_host/$x/scan; done
```

4. 执行以下命令，重新聚合多路径设备：

```
multipath -r
```

5. 执行以下命令，检验重新聚合结果：

```
multipath -ll
```

4.2 分区表显示问题

4.2.1 问题描述

使用浪潮统一系列存储的 GUI 管理页面，添加新的逻辑盘，然后将新创建的逻辑盘，映射给新创建的主机；在安装有凝思操作系统的主机上扫描识别出分配的逻辑盘后，使用 `fdisk -l`，报错：`Error:/dev/sdX:unrecognised disk lable`。如下所示：

```
root@linx:~# cat /etc/linx-release
6.0.60.4
root@linx:~#
root@linx:~# fdisk -l
Disk /dev/sda: 2398 GB, 2398335367680 bytes
255 heads, 63 sectors/track, 291581 cylinders
```

```
Units = cylinders of 16065 * 512 = 8225280 bytes
```

```
Device Boot      Start      End      Blocks      Id System
/dev/sda1          1         32      257008      82 Linux Swap / Solaris
```

```
Warning: Partition 1 does not end on cylinder boundary.
```

```
/dev/sda2        32        3922     31246425    83 Linux
```

```
Warning: Partition 2 does not end on cylinder boundary.
```

```
/dev/sda3        3922       4895     7815622     83 Linux
```

```
Warning: Partition 3 does not end on cylinder boundary.
```

```
/dev/sda4        4895       5503     4883760     83 Linux
```

```
Warning: Partition 4 does not end on cylinder boundary.
```

```
/dev/sda5        5503       6110     4875727     83 Linux
```

```
Warning: Partition 5 does not end on cylinder boundary.
```

```
Error: /dev/sdb: unrecognised disk label
```

```
Error: /dev/sdc: unrecognised disk label
```

```
Error: /dev/mapper/mpath3: unrecognised disk label
```

4.2.2 问题解答

问题原因：映射到主机的逻辑盘，没有分区表。针对逻辑盘没有分区表这种情况，有的版本的 fdisk 会报错，有的不会，例如：

Linux-6.0.60.4 操作系统（凝思）上运行的 fdisk 的版本为：GNU Fdisk 1.2.4，对于映射给主机的逻辑盘，如果逻辑盘没有分区表会报错。

RHEL7.0 操作系统上运行的 fdisk 的版本为：fdisk from util-linux 2.23.2，对于映射给主机的逻辑盘，如果逻辑没有分区不会报错。二者对比如下：

- 执行 parted /dev/sdb，检查逻辑盘有无分区表：

Linux 发行版	检查逻辑盘有无分区表
Linx-6.0.60.4	<pre> root@linx:~# parted /dev/sdb GNU Parted 2.3 Using /dev/sdb Welcome to GNU Parted! Type 'help' to view a list of commands. (parted) p Error: /dev/sdb: unrecognised disk label (parted) quit root@linx:~# </pre>
RHEL7.0	<pre> [root@localhost ~]# parted /dev/sdb GNU Parted 3.1 Using /dev/sdb Welcome to GNU Parted! Type 'help' to view a list of commands. (parted) p Error: /dev/sdb: unrecognised disk label Model: INSPUR MCS (scsi) Disk /dev/sdb: 17.2GB Sector size (logical/physical): 512B/512B Partition Table: unknown Disk Flags: (parted) quit [root@localhost ~]# </pre>

- 查看 fdisk -l 命令是否报错

Linux 发行版	fdisk -l 打印
Linx-6.0.60.4	<pre> Error: /dev/sdb: unrecognised disk label Error: /dev/sdc: unrecognised disk label Error: /dev/mapper/mpath3: unrecognised disk label </pre>
RHEL7.0	<pre> Disk /dev/sdc: 17.2 GB, 17179869184 bytes, 33554432 sectors Units = sectors of 1 * 512 = 512 bytes Sector size (logical/physical): 512 bytes / 512 bytes I/O size (minimum/optimal): 32768 bytes / 32768 bytes </pre>

	Disk /dev/sdb: 17.2 GB, 17179869184 bytes, 33554432 sectors Units = sectors of 1 * 512 = 512 bytes Sector size (logical/physical): 512 bytes / 512 bytes I/O size (minimum/optimal): 32768 bytes / 32768 bytes
	Disk /dev/mapper/mpatha: 17.2 GB, 17179869184 bytes, 33554432 sectors Units = sectors of 1 * 512 = 512 bytes Sector size (logical/physical): 512 bytes / 512 bytes I/O size (minimum/optimal): 32768 bytes / 32768 bytes

可以通过创建分区表，解决报错 Error:/dev/sdX:unrecognised disk lable 问题。

1. 执行 `parted /dev/sdX`，进入分区交互界面，执行 `mktable msdos`（容量小于 2T 的逻辑盘）或 `mktable gpt`（容量大于 2T 的逻辑盘）创建分区表：

```

root@linx: ~# parted /dev/sdb
GNU Parted 2.3
Using /dev/sdb
Welcome to GNU Parted! Type 'help' to view a list of commands.
(parted) p
Error: /dev/sdb: unrecognised disk label
(parted) mktable msdos
(parted) p
Model: INSPUR MCS (scsi)
Disk /dev/sdb: 27.9GB
Sector size (logical/physical): 512B/512B
Partition Table: msdos

Number  Start  End  Size  Type  File system  Flags

(parted) quit
Information: You may need to update /etc/fstab.

```

2. 输入以下命令，查看有无报错：

```
fdisk -l
```

4.3 Lustre 文件系统 block size 故障

4.3.1 故障描述

挂载 Lustre 文件系统的卷，在进行读写操作时，操作进入长时间等待，且系统日志报以下信息，显示卷路径异常：

```
kernel: blk_cloned_rq_check_limits: over max size limit.  
device-mapper: multipath: Failing path 8:144  
device-mapper: multipath: Failing path 8:176
```

4.3.2 故障原因

其原因是 Linux 内核的 block layer 中有一个函数 blk_cloned_rq_check_limits，被 device-mapper 在 remap block request 时调用。这个函数会将 request 的 block size 与 mapped device 的 max_sectors_kb 这个参数进行比较，若 block size 大于 max_sectors_kb，则系统进行报错。在 dm-multipath 里，若接收到该报错信息，则认为是路径异常。

4.3.3 故障处理

修改系统中卷的参数 max_sectors_kb 为 4MB。参数修改方法为：在 /etc/udev/rules.d 目录下新增一个文件 inspur-instorage-bs.rules，重启主机即可。

inspur-instorage-bs.rules 文件包含的内容如下：

```
SUBSYSTEM=="block", ACTION=="add|change", ATTRS{vendor}=="INSPUR", ATTRS{model}=="MCS",  
RUN+="/bin/sh -c '/bin/echo 4096 > /sys/block/%k/queue/max_sectors_kb'"
```

4.4 统信 UOS V20 多路径设备不聚合问题

4.4.1 故障描述

统信 UOS V20 操作系统上，可以发现路径设备，当配置完多路径配置文件 /etc/multipath.conf，使用命令 multipathd reconfigure 或者 multipath -r 重新配置聚合多路径时，仍然无法聚合成相应的多路径设备。

4.4.2 故障原因

其原因是统信 UOS V20 操作系统跟常见 Linux 发行版相比，其多路径软件在使用上存在差异。

4.4.3 故障处理

1. 在/etc/multipath.conf 配置文件中，在 defaults 段内配置 find_multipaths 为 yes；如下所示：

```
defaults {  
    find_multipaths "yes"  
}
```

2. 如果统信 UOS V20 多路径无法聚合时，需要在/etc/multipath/wwid 文件中添加已经映射磁盘的 WWID。

注意：添加 WWID 时，WWID 必须为小写字母。

以添加 WWID 为 36005076708a2080080000000000000a 的多路径磁盘为例，执行如下命令：

```
multipath -a 36005076708a2080080000000000000a
```

4.5 大量 FC 卷映射引起的多路径设备异常

4.5.1 故障描述

当存储向单一主机映射了大量 FC 卷，在存储设备升级、存储节点热重启、主机重启等场景下，您可能会遇到路径震荡，多路径拓扑信息异常，多路径设备丢失等问题。

4.5.2 故障原因

1. 与 ulimit -n 的设置相关。该参数表示进程能打开的最大文件描述符数目，一般系统里面默认是 1024。如果存储映射大量卷到主机，主机上总路径条数(卷数量*单个卷的路径条数)超过 1024，那么在存储进行升级等场景下就会出现多路径设备丢失问题。
2. 当存储系统或主机系统刚起来时，会有大量的 uevent 事件上来，遭遇事件风暴，从而出现路径震荡问题。

4.5.3 故障处理

1. 调整 ulimit -n 的值，至少使其大于主机上总路径条数(卷数量*单个卷的路径条数)。例如调整为 65535，

```
ulimit -n 65535
```

2. 当出现路径震荡，多路径拓扑信息异常，多路径设备丢失时，需要多次重新启动多路径服务并查看多路径拓扑信息，使其逐步恢复正常。

4.6 盘符漂移问题

4.6.1 故障描述

当存储使用 SAN 存储与主机对接时，由于重新启动主机操作系统或存储系统后盘符发生变化，出现盘符漂移的情况，该情况可能导致业务中断或数据写入错乱。

4.6.2 故障原因

一般情况下，新增磁盘名称在系统中显示为 mpath[n]开头的格式，盘符显示在/dev/mapper/目录下，由于主机系统或存储系统重新启动时磁盘发现顺序发生变化，导致盘符发生漂移。

4.6.3 故障处理

- 方法 1

可以依据 wwid 号对每个多路径设备设定一个唯一的别名，以免造成磁盘盘符漂移的情况出现。编辑配置文件/etc/multipath.conf，增加 multipath 段进行 wwid 和 alias 的绑定，修改完成后重启 multipathd 服务，后续使用绑定的别名进行挂载使用。

```
multipaths {
    multipath {
        wwid 366c92bf7088201e4b80000000000241
        alias inspur_lun1
    }
    ...
    multipath {
        wwid 366c92bf7088201e4b8000000000024X
```



```
alias inspur_lunX  
}  
}
```

- 方法 2

设置 `user_friendly_name` 为 `no`，如果设定为 `no`，即指定该系统应使用 `wwid` 作为该多路径的别名，后续使用 `wwid` 为别名进行挂载使用。

5 缩略语

F		
FC	Fibre Channel	光纤通道
H		
HBA	Host Bus Adapter	主机总线适配器
I		
IO	Input Output	输入输出