



# 浪潮存储系统 Linux 系统多路径软件用户手册

文档版本 1.6

发布日期 2019-08-23

适用版本 Linux 系统自带多路径的所有版本及以上

## 尊敬的存储系统用户：

衷心感谢您选用了浪潮存储系统！

本手册介绍了本存储系统的多路径软件的安装部署、配置管理等，有助于您更详细地了解  
和便捷地使用本款存储系统。

浪潮拥有本手册的版权。

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。浪潮保留随时修改本手册的  
权利。

本手册中的内容如有变动恕不另行通知。

如果您对本手册有疑问或建议，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

# 声 明

在您正式使用本存储系统之前，请您先阅读以下声明。只有您阅读了以下声明并且同意以下各条款后，方可正式开始使用本存储系统；如果您对以下条款有任何疑问，请和您的供货商联系或直接与我们联系。如您未向我们就以下条款提出疑问并开始使用本系统，则是默认您已经同意了以下各条款。

1. 在您使用的存储系统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；您不要自行拆卸存储系统机箱及机箱内任何硬件设备。
2. 本存储系统的内存、CPU、CPU 散热片、风扇、硬盘托架、硬盘等都是特殊规格的，请您不要将它们和任何其他型号机器的相应设备混用。
3. 您在使用存储系统过程中遇到的任何软件问题，我们希望您首先和相应软件的提供商联系，由他和我们联系，以方便我们沟通、共同解决您遇到的问题。对于如数据库、网络管理软件或其他网络产品等的安装、运行问题，我们尤其希望您能够这样处理。
4. 如果上架安装本存储系统，请先仔细阅读相关产品手册中的快速安装指南。浪潮致力于产品功能和性能的持续提升，这可能导致部分功能及操作与手册描述有所差异，但不会影响使用，如果您有任何使用疑难问题，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. 此为 A 级产品，在生活环境中，该产品可能会造成无线电干扰。在这种情况下，可能需要用户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细则。
8. 本手册中涉及的各软、硬件产品的标识、名称版权归产品的相应公司拥有。
9. 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

1. 本系统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，除非另外得到浪潮的通知，否则只有经过浪潮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，为了设备及使用者的安全，不要随意更换电源线缆或插头。
5. 为防止系统漏电造成电击危险，务必将系统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，请勿操作使用本设备，请与电工联系。
6. 切勿将任何物体塞入系统的开孔处。如果塞入物体，可能会导致内部组件短路而引起火灾或电击。
7. 请将系统置于远离散热片和有热源的地方，切勿堵塞通风孔。
8. 切勿让食物或液体散落在系统内部或其它组件上，不要在高潮湿、高灰尘的环境中使用产品。
9. 用错误型号的电池更换会有爆炸危险，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，有关电池回收请与当地废品回收处理机构联系。

# 目 录

声 明 .....	ii
安全细则 .....	iii
<b>1 概述 .....</b>	<b>5</b>
<b>2 安装 multipath-tools .....</b>	<b>7</b>
2.1 安装方法.....	7
2.2 加载模块与服务.....	9
2.3 检查 multipath-tools 配置文件 .....	10
2.4 设置 multipathd 服务开机启动 .....	11
<b>3 逻辑卷配置 .....</b>	<b>13</b>
3.1 逻辑盘的创建与映射.....	13
3.1.1 SAN switch 组网 .....	13
3.1.2 iSCSI 组网 .....	14
3.2 扫描映射的逻辑盘.....	15
3.2.1 SAN switch 组网 .....	15
3.2.2 iSCSI 组网 .....	15
3.3 配置 multipath-tools 多路径工具 .....	16
3.3.1 配置文件介绍.....	16
3.3.2 配置文件修改.....	16
3.4 配置 FC 目标端口设备丢失超时参数.....	18
3.5 配置访问 SCSI block device 的超时时间 .....	19
<b>4 故障分析与解决 .....</b>	<b>20</b>
4.1 多路径设备聚合问题.....	20
4.1.1 问题描述.....	20
4.1.2 问题解答.....	20
4.2 分区表显示问题.....	21
4.2.1 问题描述.....	21
4.2.2 问题解答.....	22
4.3 Lustre 文件系统 block size 故障 .....	25
4.3.1 故障描述.....	25
4.3.2 故障原因.....	25
4.3.3 故障处理.....	25
<b>5 术语&amp;缩略语 .....</b>	<b>27</b>

# 1 概述

Linux 系统平台通常会包含多路径模块：multipath-tools，根据多路径模块的版本号不同，支持浪潮存储的多路径配置也不同：

- 从版本 0.7.5 开始（包含 0.7.5），multipath-tools 的默认支持设备包含了浪潮 G2、G5 系列存储设备，无需修改配置文件，即可支持 G2、G5 系列存储设备的多路径功能。
- 当 multipath-tools 模块的版本号低于 0.7.5 时，需要修改多路径模块的配置文件，实现对 G2、G5 系列存储设备的多路径功能支持。

本文介绍在 Linux 系统下配置多路径工具 multipath-tools，以支持浪潮 G2、G5 系列存储设备多路径功能。multipath-tools 是 Linux 平台用于配置存储设备多路径支持的工具，其作用是配置 Linux 内核模块 device-mapper，实现对存储设备的多路径支持。

本文以 Linux 发行版 RHEL 6.6（Red Hat Enterprise Level 6.6）为例，介绍在 Linux 平台配置浪潮 G2、G5 系列设备多路径支持。

- RHEL 6.6 发行版本基于 Linux 2.6.32 内核版本，其多路径工具版本为 multipath-tools 0.4.9。
- 其它 RHEL 发行版本的多路径工具，与此例类似，请参考其对应的模块说明。
- 系统 CentOS、Oracle Linux、Kylin 3.2-4F 与 RHEL 系统类似，请参考其对应的模块说明。
- 系统 SuSE 与 RHEL 存在差别，请参考文中备注对应的模块说明。

**本文适用产品型号：**

AS2150G2&AS2200G2&AS2600G2&AS5300G2&AS5500G2&AS5600G2&AS5800  
G2&AS6800G2

AS2600G2-F&AS5300G2-F&AS5500G2-F&AS5600G2-F&AS5800G2-F&AS6800G2  
-F

AS5300G5&AS5500G5&HF5000G5

# 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 -I | 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：

- 1) 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
```

- 2) 进入系统模块所在目录。参考表 2-1 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



CentOS6.x	/mnt/cd0/Packages
CentOS7.x	/mnt/cd0/Packages
SuSE	/mnt/cd0/suse/x86_64（选择相应的处理器）
Kylin 3.2-4F	/mnt/cd0/KYLIN

- 3) 执行 rpm 包安装命令。参考表 2-2 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
```

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	-

3. 完成以上步骤后，需检查多路径工具是否已经安装成功。请执行以下命令：

```
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 -I | grep multipath
```

## 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)

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
  - 1) 制作连接文件 ln -s /lib/udev/scsi\_id /lib64/udev/scsi\_id。

- 2) 执行 `scsi_id -g -u -s /block/sdb`, 获取 `sdb` 的 `wwid`。
- 3) 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 "service-time 0"
        prio alua
    }
}
```

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的会有差异，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>
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>
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	<code>/usr/share/doc/packages/multipath-tools/multipath.conf.synthetic</code>
Kylin 3.2-4F	<code>/usr/share/doc/device-mapper-multipath-0.4.9/multipath.conf</code>

## 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	<code>chkconfig --list multipathd</code>
RHEL6.x	<code>chkconfig --list multipathd</code>
RHEL7.x	<code>systemctl is-enabled multipathd.service</code>
CentOS6.x	<code>chkconfig --list multipathd</code>
CentOS7.x	<code>systemctl is-enabled multipathd.service</code>
SuSE10	<code>chkconfig --list multipathd</code>

SuSE11	chkconfig --list multipathd
SuSE12	chkconfig --list multipathd
Kylin 3.2-4F	chkconfig --list multipathd

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
CentOS6.x	chkconfig --level 35 multipathd on
CentOS7.x	systemctl enable multipathd.service
SuSE10	chkconfig --level 35 multipathd on
SuSE11	chkconfig --level 35 multipathd on
SuSE12	chkconfig --level 35 multipathd on
Kylin 3.2-4F	chkconfig --level 35 multipathd on

# 3 逻辑卷配置

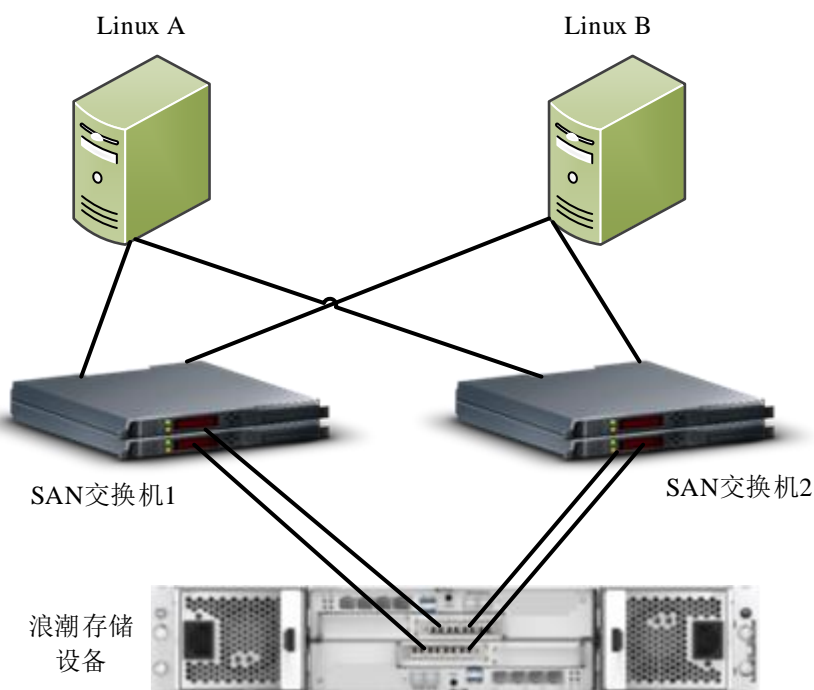
## 3.1 逻辑盘的创建与映射

使用浪潮 G2、G5 系列存储设备的管理工具（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

图 3-1 所示为典型的双机应用集群与浪潮 G2、G5 系列存储设备通过双冗余 FC 网络组网的方式。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



4. 执行以下命令，查看主机上光纤 HBA 的 WWN:

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

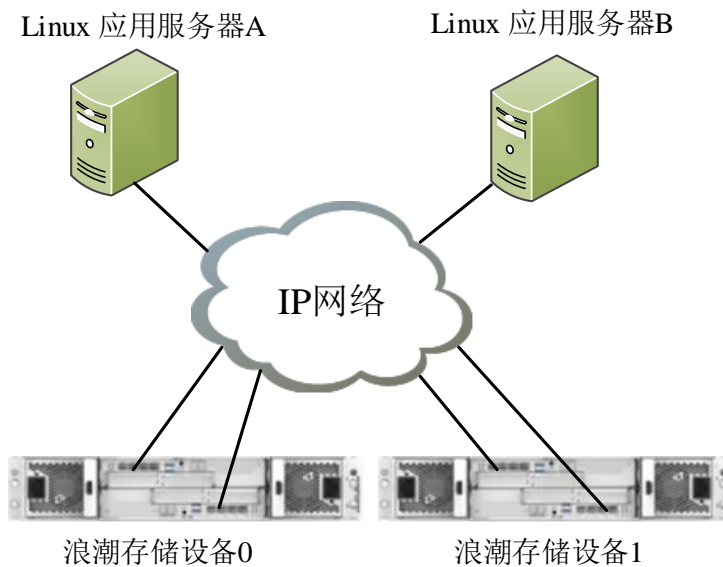
5. 使用浪潮 G2、G5 系列存储设备的 GUI 管理页面，利用步骤 1 中获取的光纤 HBA 的 WWN，创建主机（参考《基本功能配置手册》）。
6. 创建新的逻辑盘，然后将新创建的逻辑盘，映射给步骤 2 创建的主机。

--结束

### 3.1.2 iSCSI 组网

图 3-2 所示为应用服务器与存储系统通过 iSCSI 方式组网的典型应用。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 G2、G5 系列存储设备的的管理工具（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。

--结束

## 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。

--结束

### 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
```

--结束



## 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 G2、G5 系列存储设备的最佳配置。

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 G2、G5 系列存储设备的多路径优化配置，需要修改/etc/multipath.conf 配置文件实现对浪潮 G2、G5 系列存储设备的最佳支持。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。

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

```
device{
    vendor "INSPUR"
    product "MCS"
    path_grouping_policy group_by_prio
    path_selector "service-time 0"
    path_checker tur
    features "1 queue_if_no_path"
    prio alua
    failback immediate
}
```

以上配置项中，`path_selector` 设置路径选择策略为 `service time`，即在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、CentOS7.x)
```

3. 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。

### 3.4 配置 FC 目标端口设备丢失超时参数

FC 目标端口的设备丢失超时参数，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。下表介绍两种常用 FC HBA 卡的参数修改方法。

HBA 卡	缺省值	修改驱动超时参数方法
emulex	30	<p><b>方法一：</b></p> <p>修改 emulex FC HBA 卡驱动模块 lpfc 的驱动超时参数。</p> <p>以修改驱动超时参数 10 秒为例，步骤如下：</p> <ol style="list-style-type: none"> <li>1. 创建/etc/modprobe.d/lpfc.conf 文件，在文件中增加以下配置，options lpfc lpfc_devloss_tmo=10</li> <li>2. 备份 initramfs:           <pre>mv /boot/initramfs-`uname -r`.img /boot/initramfs-`uname -r`.img.bak</pre>           或备份 initrd:           <pre>mv /boot/initrd-`uname -r`.img /boot/initrd-`uname -r`.img.bak</pre> </li> <li>3. 更新 initramfs: <code>dracut -v -f</code> 或更新 initrd: <code>mkinitrd -f -v /boot/initrd-`uname -r`.img `uname -r`</code></li> <li>4. 重启操作系统 <code>reboot</code></li> </ol> <p><b>说明：</b>上面命令中的 <code>`uname -r`</code>，“`”不是单引号，而是英文输入</p>

		<p>法下“~、”键按出来的，一般是在“Esc”键下面的一个按键。</p> <p><b>方法二：</b></p> <p>开机启动执行脚本：</p> <pre>for targetport in /sys/class/fc_remote_ports/rport*;do echo 10 &gt; \${targetport}/dev_loss_tmo;done</pre>
qlogic	35	<p>开机启动执行脚本：</p> <pre>for targetport in /sys/class/fc_remote_ports/rport*;do echo 10 &gt; \${targetport}/dev_loss_tmo;done</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
```

# 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
```

--结束

## 4.2 分区表显示问题

### 4.2.1 问题描述

使用浪潮 G2、G5 系列存储设备的 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/linux-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</pre>

	<pre>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</pre>



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

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'"
```

# 5 术语&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