# 概念

Multipath即多路径，是个通用概念。这里要介绍的是开源的存储多路径技术，也就是DM multipath。

# Device Mapper

# 命名

有时看到一串16进制数字(WWID)，有时是以mpath为前缀的名字(user-friendly name), 有时是任意字母串(alia name)。multipath默认用的是WWID，为什么不用好记的名字呢？好记的名字不能工作的一个情景：根文件系统不能在multipath设备上面。好记的名字和 WWID之间的映射是保存在/etc/multipath/bindings文件里的。要访问这个文件，根文件系统必须已经挂载上了，而multipath服务在initrd里就要开始工作，那个时候还没有根系统。因此，默认设置为wwid是为了安全。

# 设备地址

这种以冒号分割4个数字的设备地址如2:0:0:1怎么解释？

2:0:0:1设备地址，数字分别对应：Host:Bus:Target:Lun 。比如我们让iscsi target走了两个IP地址，那么对于同一个设备只有 host字段不同。比如：2:0:0:1和3:0:01。

# 操作

## 基本指令

multipath

modprobe dm-multipath

service multipathd reload （start,stop,restart）

multipath -v2

multipath -v3 格式化路径 -v3 更详细

multipath -F

multipath -ll

multipathd -k (help)

注意：由于配置multipath 后，在新添加的本地磁盘或存储阵列的lun时，需要考虑blacklist的和LVM配置文件对磁盘的过滤问题，不正确的参数可能导致主机无非正常识别和使用新添加的磁盘,需要视情况修改。

两个选项：

file:/etc/lvm/lvm.conf

filter = [ "a/.\*/" ]

file:/etc/multipath.conf

blacklist {

devnode "^sda"

}

## 操作步骤

1. 安装rpm包；
2. 启动multipath：

将多路径软件添加至内核模块中

modprobe dm-multipath

modprobe dm-round-robin

检查内核模块加载状况：

lsmod | grep multipath

将多路径软件multipath设置为开机自启动：

chkconfig –level 2345 multipath on

或

chkconfig multipathd on

chkconfig –list | grep multipathd

启动multipath服务：

service multipathd restart

1. 多路径配置：

停止服务：

service multipathd stop

获取所有存储LUN的WWID：

创建配置文件/etc/multipath.conf：

/sbin/,ultipathconf –enable –find\_multipaths y –with\_module y –with\_chkconfig y

多路径配置文件中添加内容：

blacklist {

devnode "本地磁盘" ##本地磁盘不需要多路径，加入黑名单

}

multipaths{

multipath{

wwid

alias #设置数据盘、日志盘等

}

}

# path group

什么是path group?

path grouping policy and IO scheduling policy?

起初，我对这个概念有混淆：认为一个真实设备对应的所有路径为一个path group，即认为下面是一个path group:

multipath-demo:~ # multipath -l14945540000000000ccb70d0ceeee4280f8450284d6298b59 dm-0 IET,VIRTUAL-DISK

size=10G features='1 retain\_attached\_hw\_handler' hwhandler='0' wp=rw

|-+- policy='service-time 0' prio=0 status=active

| `- 2:0:0:0 sda 8:0  active undef unknown`-+- policy='service-time 0' prio=0 status=enabled

 `- 3:0:0:0 sdc 8:32 active undef unknown

其实，dm-0设备有两个path group，每个PG都只有一个路径（真实环境有多条），状态active的是正在工作的路径，状态enabled处于备用状态，并不下发IO。

path grouping policy 默认是failover, 如martin所说，各设备厂商默认策略不同，主流的在用group\_by\_prio，作用就是把路径分组。IO scheduling policy默认是service time, 负责如何在一个PG的路径中分配IO。

# 应用

在没有企业存储的情况下，怎么玩multipath？

使用虚拟机和iscsi。装一虚拟机，添加块设备，添加两个网卡，再用这个块设备建一个iscsi target。然后在一个想玩multipath的机器上面，用iscsi client去连接iscsi target。至此，用lsblk会查看到原来的块设备有两个设备节点。