



马哥教育
最专业的Linux培训机构

磁盘管理

- ❖ 磁盘结构
- ❖ 分区类型
- ❖ 管理分区
- ❖ 管理文件系统
- ❖ 挂载设备
- ❖ 管理虚拟内存

马哥教育

www.magedu.com

设备文件

❖ **I/O Ports: I/O设备地址**

❖ 一切皆文件:

`open()`, `read()`, `write()`, `close()`

❖ 设备类型:

块设备: **block**, 存取单位“块”, 磁盘

字符设备: **char**, 存取单位“字符”, 键盘

❖ 设备文件: 关联至一个设备驱动程序, 进而能够跟与之对应硬件设备进行通信

❖ 设备号码:

主设备号: **major number**, 标识设备类型

次设备号: **minor number**, 标识同一类型下的不同设备

马哥教育

www.magedu.com

❖ 并行:

IDE: 133MB/s

SCSI: 640MB/s

❖ 串口:

SATA: 6Gbps

SAS: 6Gbps

USB: 480MB/s

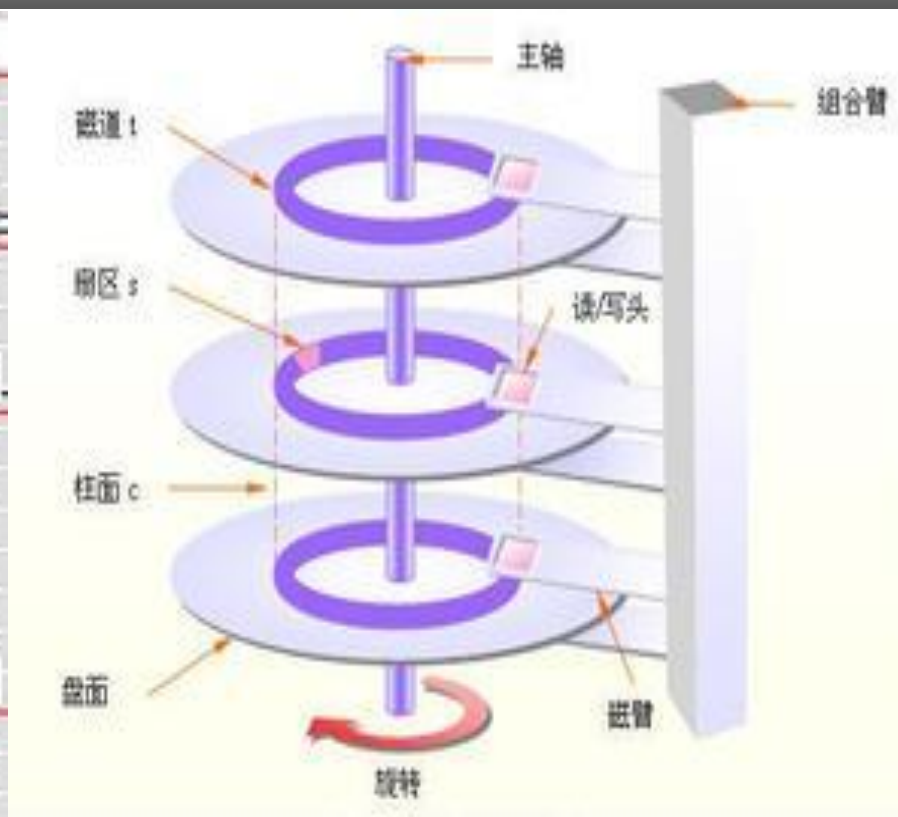
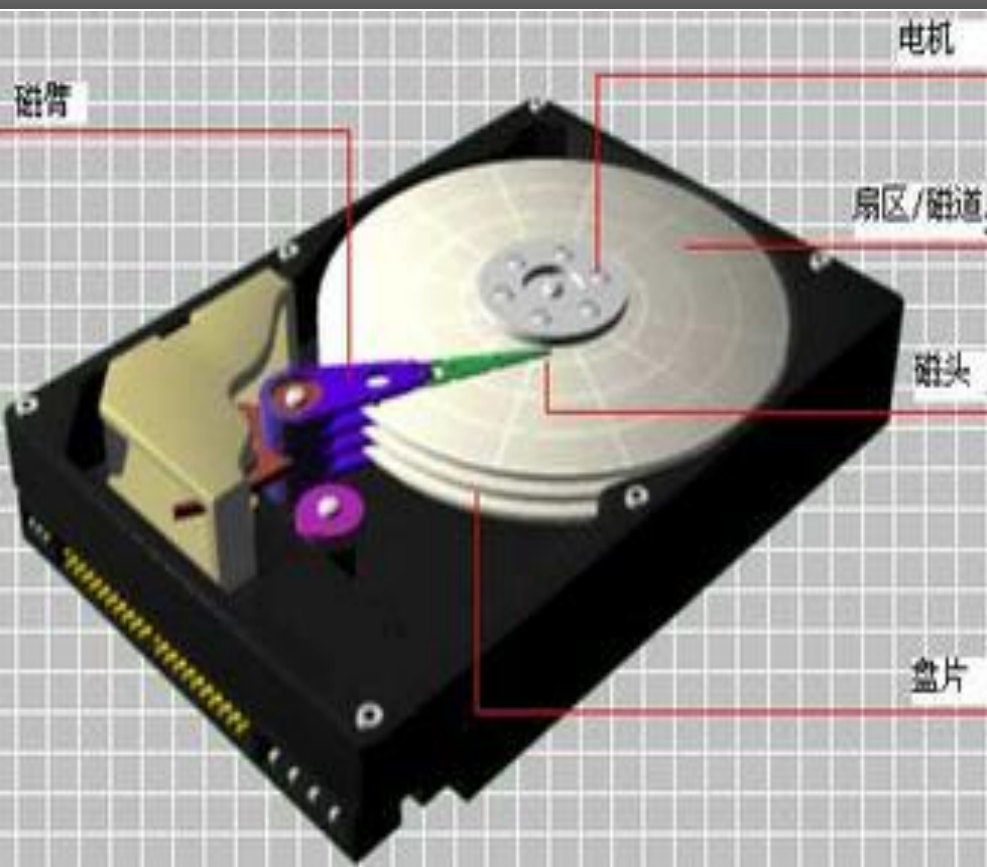
❖ rpm: rotations per minute

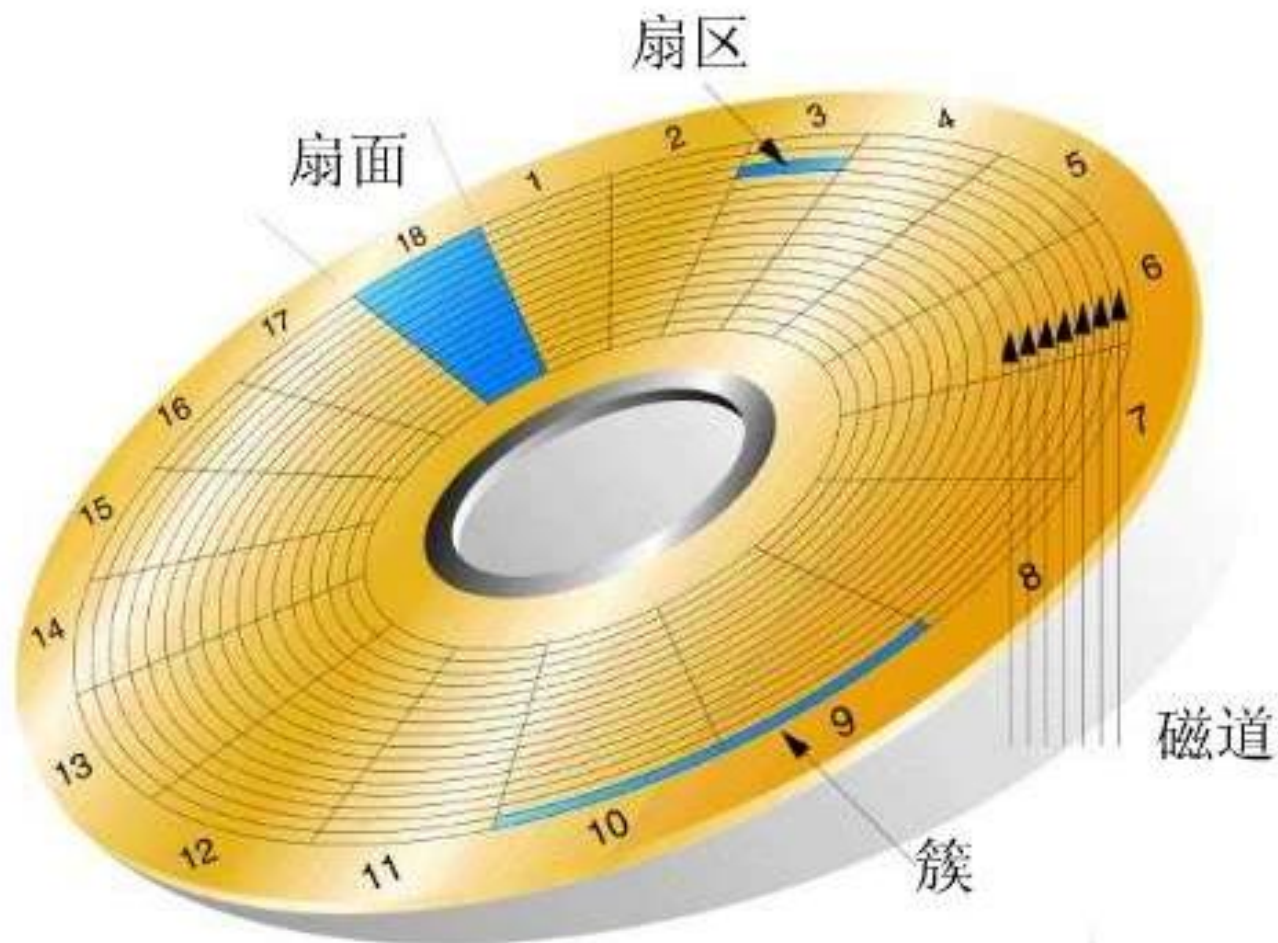


- ❖ 磁盘设备的设备文件命名: `/dev/DEV_FILE`
- ❖ IDE: `/dev/hd`
- ❖ SCSI, SATA, SAS, USB: `/dev/sd`
 不同设备: `a-z`
 `/dev/sda, /dev/sdb, ...`
- ❖ 同一设备上的不同分区: `1,2, ...`
 `/dev/sda1, /dev/sda5`

马哥教育

www.magedu.com





- ❖ head: 磁头
- ❖ track: 磁道
- ❖ cylinder: 柱面
- ❖ secotr: 扇区, 512bytes

马哥教育

www.magedu.com

- ❖ 设备识别
- ❖ 设备分区
- ❖ 创建文件系统
- ❖ 标记文件系统
- ❖ 在`/etc/fstab`文件中创建条目
- ❖ 挂载新的文件系统

马哥教育

www.magedu.com

❖ 为什么是分区？

- 优化**I/O**性能
- 实现磁盘空间配额限制
- 提高修复速度
- 隔离系统和程序
- 安装多个**OS**
- 采用不同文件系统

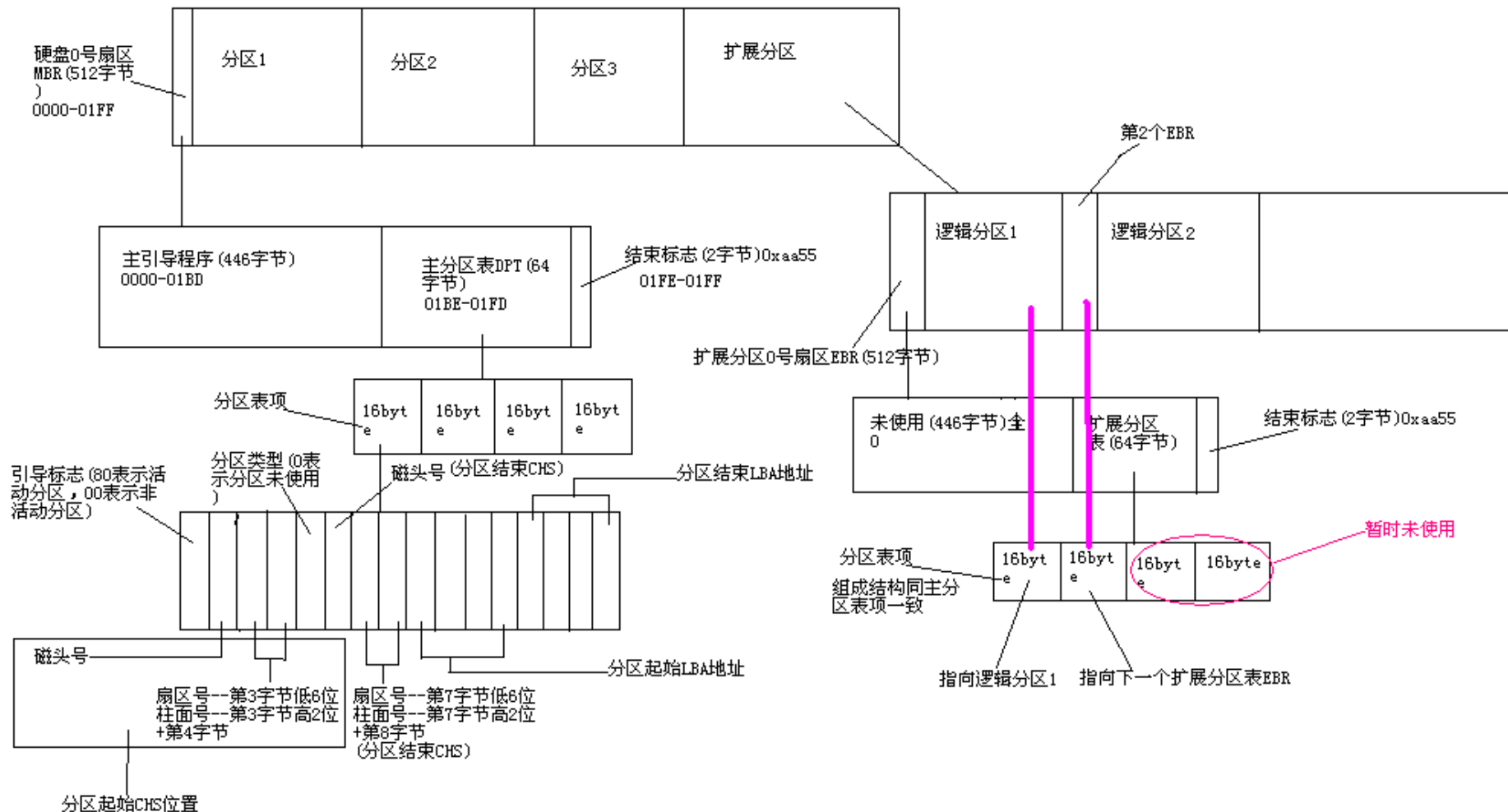
马哥教育

www.magedu.com

- ❖ 两种分区方式: **MBR, GPT**
- ❖ **MBR: Master Boot Record**, 1982年, 使用**32位**表示扇区数, 分区不超过**2T**
- ❖ 如何分区: 按柱面
- ❖ **0磁道0扇区: 512bytes**
 - 446bytes: boot loader**
 - 64bytes: 分区表**
 - 16bytes: 标识一个分区**
 - 2bytes: 55AA**
- ❖ **4个主分区; 3主分区+1扩展(N个逻辑分区)**

MBR分区结构

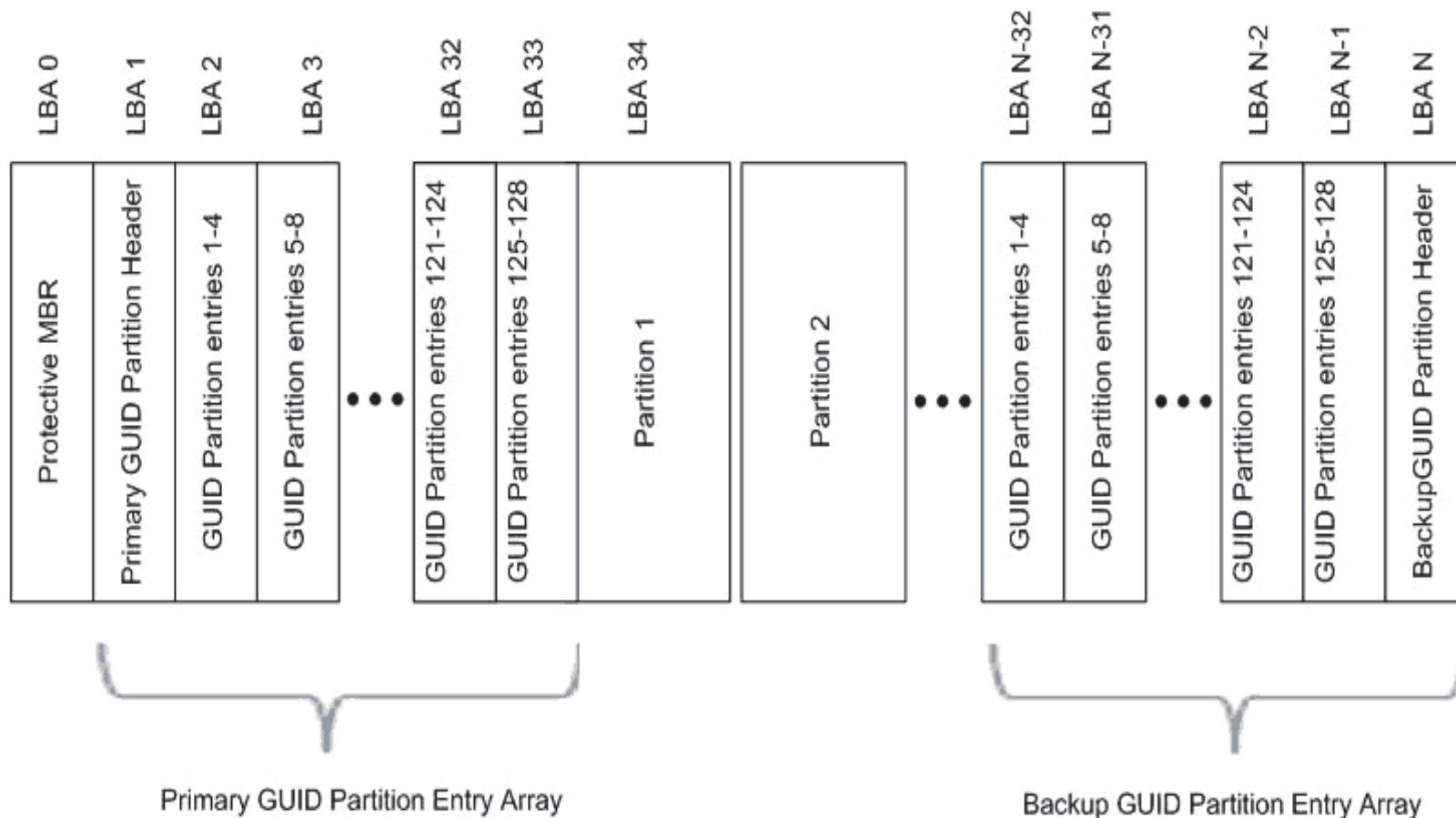
MBR硬盘分区结构图



- ❖ **GPT:GUID (Globals Unique Identifiers) partition table** 支持128个分区，使用64位，支持8Z (512Byte/block) 64Z (4096Byte/block)
- ❖ 使用128位UUID 表示磁盘和分区 GPT分区表自动备份在头和尾两份，并有CRC校验位
- ❖ **UEFI (统一扩展固件接口)**硬件支持GPT，使操作系统启动

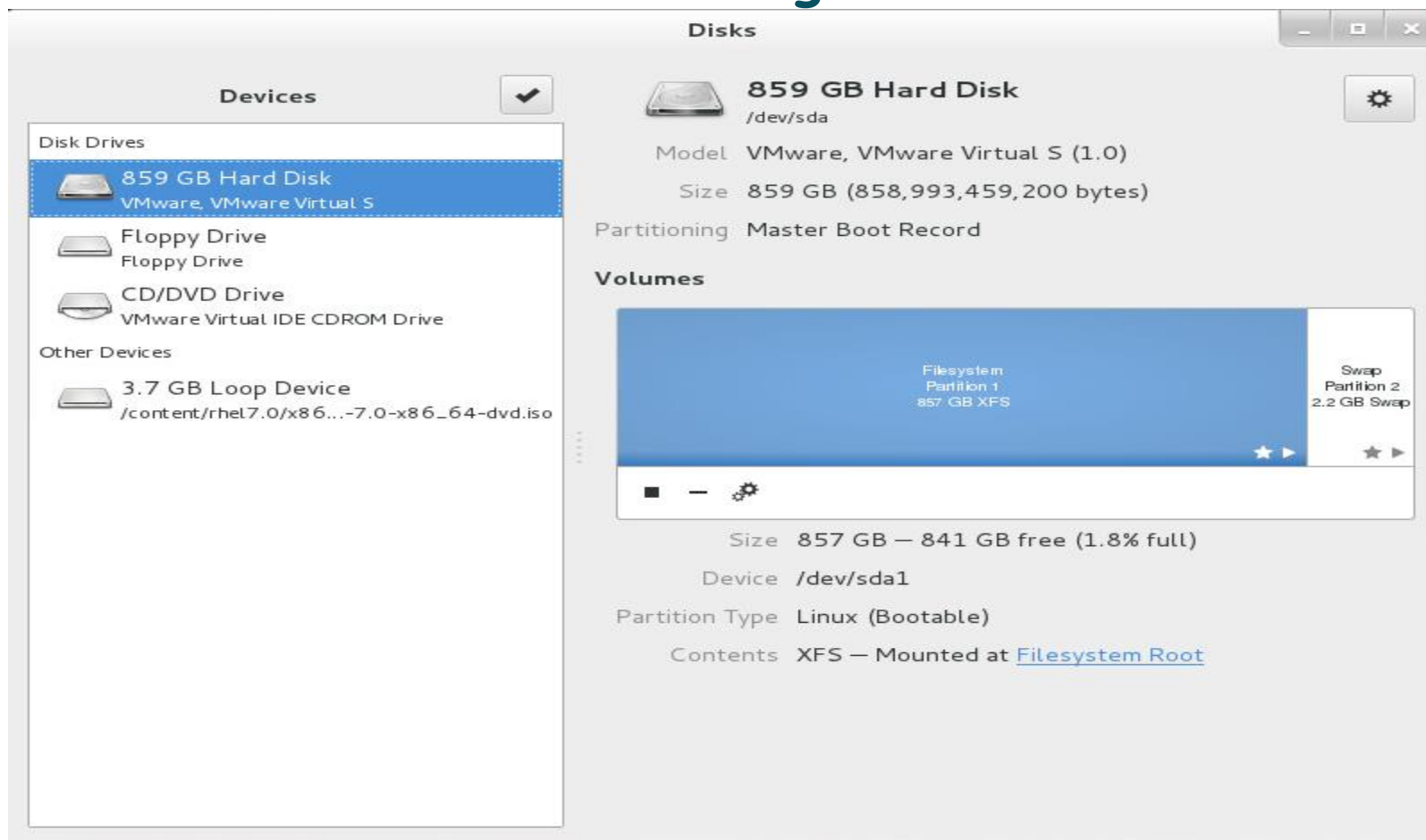
马哥教育

www.magedu.com



EFI部分又可以分为4个区域：EFI信息区(GPT头)、分区表、GPT分区、备份区域

- ❖ 图形化磁盘管理功能工具：点击“应用程序” -> “系统工具” -> “磁盘” 或执行命令 **gnome-disks**



- ❖ 列出块设备
 - **lsblk**
- ❖ 创建分区使用：
 - **fdisk** 创建**MBR**分区
 - **gdisk** 创建**GPT**分区
 - **parted** 高级分区操作（创建、复制、调整大小等等）
- ❖ **partprobe**—重新设置内存中的内核分区表版本

马哥教育

www.magedu.com

- ❖ `fdisk /dev/sdb`
- ❖ `gdisk /dev/sdb` 类fdisk 的GPT分区工具
- ❖ `# fdisk -l [-u] [device...]`
- ❖ 子命令：
 - `p` 分区列表
 - `t` 更改分区类型
 - `n` 创建新分区
 - `d` 删除分区
 - `w` 保存并退出
 - `q` 不保存并退出

马哥教育

www.magedu.com

同步分区表

- ❖ 查看内核是否已经识别新的分区：

```
# cat /proc/partitions
```

- ❖ 通知内核重新读取硬盘分区表

新增分区用

```
partx -a -n M:N /dev/DEVICE
```

```
kpartx -a /dev/DEVICE -f: force
```

删除分区用

```
partx -d -n M:N /dev/DEVICE
```

centos6: -nr N-M

- ❖ CentOS 5, 7: 使用partprobe

- ❖ partprobe [/dev/DEVICE]

- ❖ **parted**的操作都是实时生效的，小心使用
- ❖ 用法: **parted** [选项]... [设备 [命令 [参数]...]...]
parted /dev/sdb mklabel gpt|msdos
parted /dev/sdb print
parted /dev/sdb mkpart primary 1 200 (默认M)
parted /dev/sdb rm 1
parted -l

马哥教育

www.magedu.com

- ❖ 文件系统是操作系统用于明确存储设备或分区上的文件的方法和数据结构；即在存储设备上组织文件的方法。操作系统中负责管理和存储文件信息的软件结构称为文件管理系统，简称文件系统。
- ❖ 从系统角度来看，文件系统是对文件存储设备的空间进行组织和分配，负责文件存储并对存入的文件进行保护和检索的系统。具体地说，它负责为用户建立文件，存入、读出、修改、转储文件，控制文件的存取，安全控制，日志，压缩，加密等。

马哥教育

www.magedu.com

- ❖ Linux文件系统: ext2(Extended file system), ext3, ext4, xfs (SGI) , btrfs (Oracle) , reiserfs, jfs (AIX) , swap
- ❖ 光盘: iso9660
- ❖ Windows: fat32, ntfs
- ❖ Unix: FFS (fast) , UFS (unix) , JFS2
- ❖ 网络文件系统: NFS, CIFS
- ❖ 集群文件系统: GFS2, OCFS2 (oracle)
- ❖ 分布式文件系统: ceph, moosefs, mogilefs, glusterfs, Lustre
- ❖ RAW: 未经处理或者未经格式化产生的文件系统

❖ 根据其是否支持"journal"功能:

日志型文件系统: **ext3, ext4, xfs, ...**

非日志型文件系统: **ext2, vfat**

❖ 文件系统的组成部分:

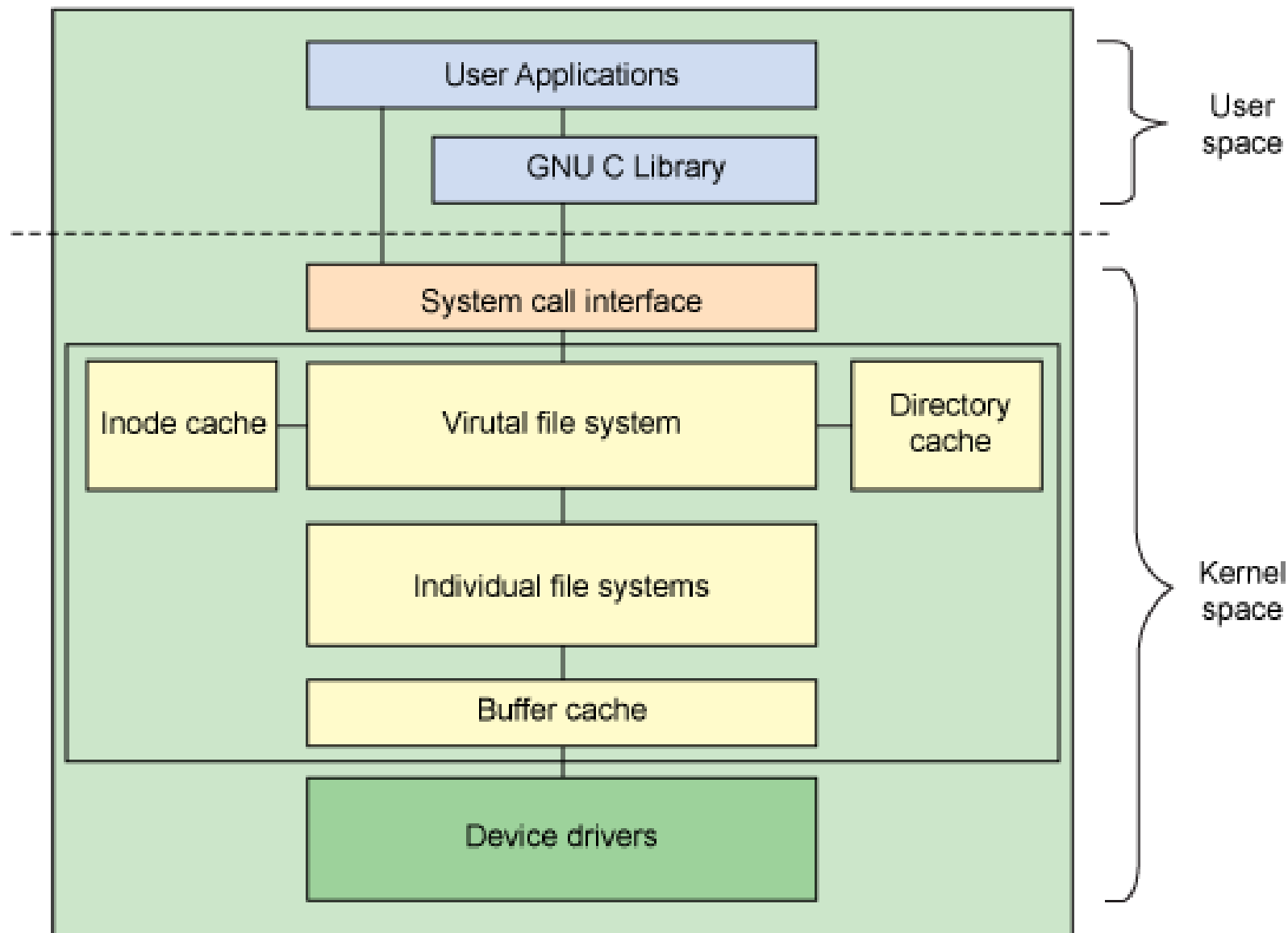
内核中的模块: **ext4, xfs, vfat**

用户空间的管理工具: **mkfs.ext4, mkfs.xfs, mkfs.vfat**

❖ Linux的虚拟文件系统: **VFS**

❖ 查看支持的文件系统: **cat /proc/filesystems**

马哥教育
www.magedu.com



类型	支持限制	Root 分区	Boot分区	注释
单节点				
XFS	500TB	Yes	Yes	默认分区格式
Ext4	50TB	Yes	Yes	兼容 Ext3, Ext2
btrfs	50TB	Yes	Yes	技术预览
网络/多节点				
GFS2	2-16 个节点	Yes	No	集群文件共享存储

❖ mkfs命令:

(1) # mkfs.FS_TYPE /dev/DEVICE

ext4

xfs

btrfs

vfat

(2) # mkfs -t FS_TYPE /dev/DEVICE

-L 'LABEL': 设定卷标

www.magedu.com

❖ **mke2fs**: ext系列文件系统专用管理工具

-t {ext2|ext3|ext4}

-b {1024|2048|4096}

-L 'LABEL'

-j: 相当于 **-t ext3**

mkfs.ext3 = mkfs -t ext3 = mke2fs -j = mke2fs -t ext3

-i #: 为数据空间中每多少个字节创建一个**inode**; 此大小不应该小于**block**的大小

-N #: 为数据空间创建多少个**inode**

-I 一个**inode**记录大小**128---4096**

-m #: 默认**5%**, 为管理人员预留空间占总空间的百分比

-O FEATURE[,...]: 启用指定特性

-O ^FEATURE: 关闭指定特性

❖ 指向设备的另一种方法

❖ 与设备无关

❖ **blkid**: 块设备属性信息查看

blkid [OPTION]... [DEVICE]

-U **UUID**: 根据指定的**UUID**来查找对应的设备

-L **LABEL**: 根据指定的**LABEL**来查找对应的设备

❖ **e2label**: 管理**ext**系列文件系统的**LABEL**

e2label DEVICE [LABEL]

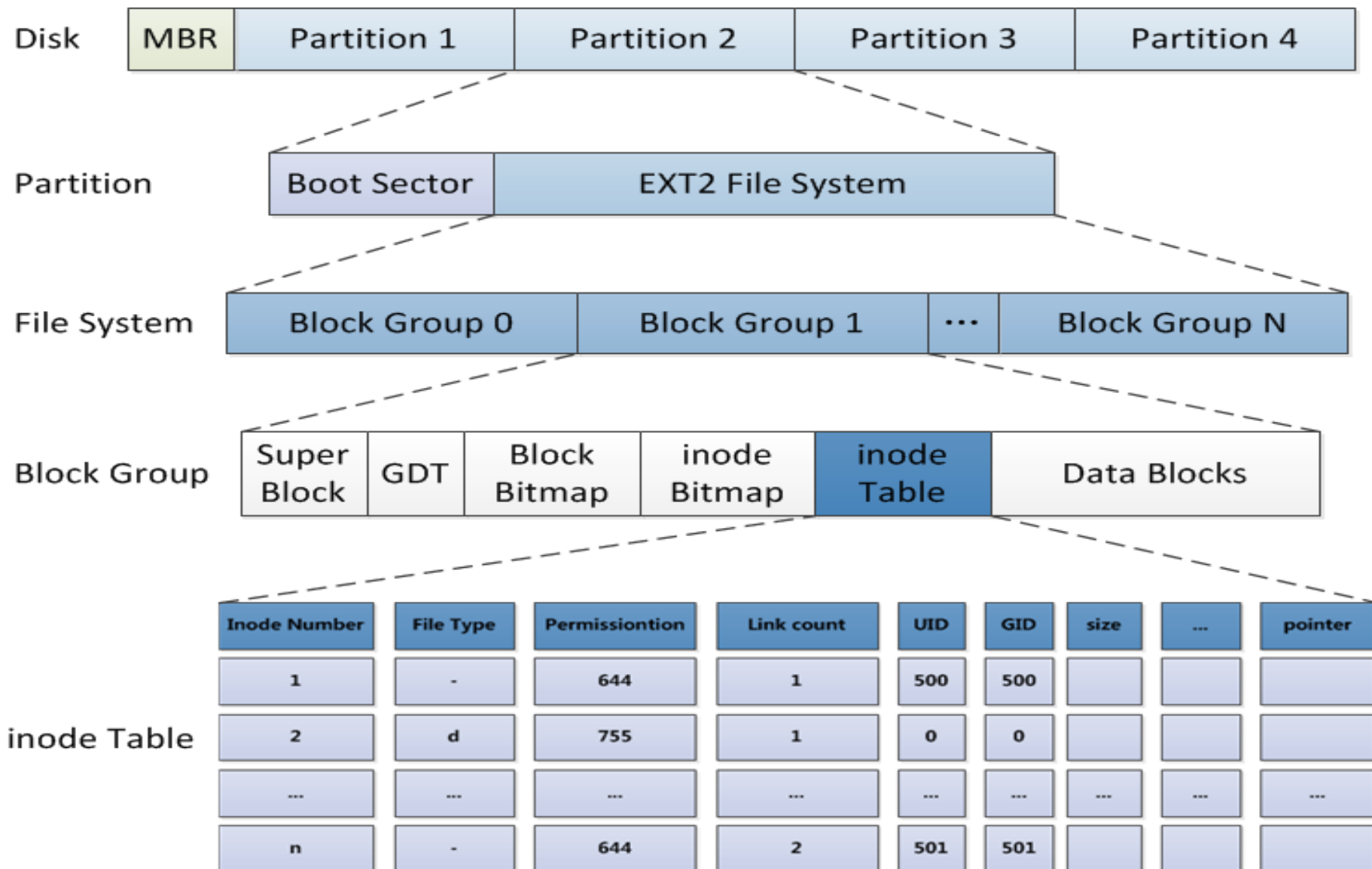
❖ **findfs** : 查找分区

findfs [options] LABEL=<label>

findfs [options] UUID=<uuid>

- ❖ **tune2fs**: 重新设定**ext**系列文件系统可调整参数的值
 - l: 查看指定文件系统超级块信息; **super block**
 - L 'LABEL': 修改卷标
 - m #: 修预留给管理员的空间百分比
 - j: 将**ext2**升级为**ext3**
 - O: 文件系统属性启用或禁用, -O ^has_journal
 - o: 调整文件系统的默认挂载选项, -o ^acl
 - U UUID: 修改**UUID**号
- ❖ **dumpe2fs**:
 - h: 查看超级块信息, 不显示分组信息

超级块和inode table



文件系统检测和修复

- ❖ 常发生于死机或者非正常关机之后
- ❖ 挂载为文件系统标记为“no clean”
- ❖ 注意：一定不要在挂载状态下修复
- ❖ **fsck: File System Check**

fsck.FS_TYPE

fsck -t FS_TYPE

-a: 自动修复错误

-r: 交互式修复错误

FS_TYPE一定要与分区上已经文件类型相同

- ❖ **e2fsck: ext**系列文件专用的检测修复工具

-y: 自动回答为yes

-f: 强制修复

- ❖ 挂载:将额外文件系统与根文件系统某现存的目录建立起关联关系, 进而使得此目录做为其它文件访问入口的行为
- ❖ 卸载:为解除此关联关系的过程
- ❖ 把设备关联挂载点: **mount Point**
mount
- ❖ 卸载时: 可使用设备, 也可以使用挂载点
umount
- ❖ 挂载点下原有文件在挂载完成后会被临时隐藏
- ❖ 挂载点目录一般为空

马哥教育
www.magedu.com

用mount命令挂载文件系统

- ❖ 挂载方法: `mount DEVICE MOUNT_POINT`
- ❖ `mount`: 通过查看`/etc/mtab`文件显示当前已挂载的所有设备
- ❖ `mount [-fnrsvw] [-t vfstype] [-o options] device dir`
 - `device`: 指明要挂载的设备;
 - (1) 设备文件: 例如`/dev/sda5`
 - (2) 卷标: `-L 'LABEL'`, 例如 `-L 'MYDATA'`
 - (3) UUID, `-U 'UUID'`: 例如 `-U '0c50523c-43f1-45e7-85c0-a126711d406e'`
 - (4) 伪文件系统名称: `proc`, `sysfs`, `devtmpfs`, `configfs`
 - `dir`: 挂载点
 - 事先存在; 建议使用空目录
 - 进程正在使用中的设备无法被卸载

- ❖ **-t vsftype**: 指定要挂载的设备上的文件系统类型
- ❖ **-r: readonly**, 只读挂载
- ❖ **-w: read and write**, 读写挂载
- ❖ **-n**: 不更新/etc/mtab, 相当于**#mount**
- ❖ **-a**: 自动挂载所有支持自动挂载的设备(定义在了/etc/fstab文件中, 且挂载选项中有**auto**功能)
- ❖ **-L 'LABEL'**: 以卷标指定挂载设备
- ❖ **-U 'UUID'**: 以UUID指定要挂载的设备
- ❖ **-B, --bind**: 绑定目录到另一个目录上
- ❖ 查看内核追踪到的已挂载的所有设备:

cat /proc/mounts

❖ **-o options:** (挂载文件系统的选项), 多个选项使用逗号分隔

async: 异步模式

sync: 同步模式, 内存更改时, 同时写磁盘

atime/noatime: 包含目录和文件

diratime/nodiratime: 目录的访问时间戳

auto/noauto: 是否支持自动挂载, 是否支持**-a**选项

exec/noexec: 是否支持将文件系统上运行应用程序

dev/nodev: 是否支持在此文件系统上使用设备文件

suid/nosuid: 是否支持**suid**和**sgid**权限

remount: 重新挂载

ro: 只读

rw: 读写

user/nouser: 是否允许普通用户挂载此设备, 默认管理员才能挂载

acl: 启用此文件系统上的**acl**功能

❖ **Defaults:** 相当于**rw, nosuid, dev, exec, auto, nouser, async**

❖ 查看挂载情况:

```
#findmnt MOUNT_POINT
```

❖ 查看正在访问指定文件系统的进程:

```
#lsof MOUNT_POINT
```

```
#fuser -v MOUNT_POINT
```

❖ 终止所有在正访问指定的文件系统的进程:

```
# fuser -km MOUNT_POINT
```

❖ 卸载:

```
# umount DEVICE
```

```
# umount MOUNT_POINT
```


- ❖ 配置文件系统体系
- ❖ 被**mount**、**fsck**和其它程序使用
- ❖ 系统重启时保留文件系统体系
- ❖ 可以在设备栏使用文件系统卷标
- ❖ 使用**mount -a** 命令挂载/etc/fstab中的所有文件系统

马哥教育

www.magedu.com

文件挂载配置文件

❖ `/etc/fstab` 每行定义一个要挂载的文件系统;

- 要挂载的设备或伪文件系统
- 挂载点
- 文件系统类型
- 挂载选项
- 转储频率
- 自检次序

❖ 要挂载的设备或伪文件系统:

设备文件、`LABEL(LABEL="")`、`UUID(UUID="")`、伪文件系统名称(`proc`, `sysfs`)

❖ 挂载选项:

`defaults`

❖ 转储频率:

`0`: 不做备份

`1`: 每天转储

`2`: 每隔一天转储

❖ 自检次序:

`0`: 不自检

`1`: 首先自检; 一般只有`rootfs`才用`1`

- ❖ 交换分区是系统**RAM**的补充
- ❖ 基本设置包括：
 - 创建交换分区或者文件
 - 使用**mkswap**写入特殊签名
 - 在**/etc/fstab**文件中添加适当的条目
 - 使用**swapon -a** 激活交换空间

马哥教育

www.magedu.com

❖ 启用: **swapon**

swapon [OPTION]... [DEVICE]

-a: 激活所有的交换分区;

-p PRIORITY: 指定优先级

/etc/fstab: pri=value

❖ 禁用: **swapoff [OPTION]... [DEVICE]**

马哥教育

www.magedu.com

- ❖ 可以指定**swap**分区**0**到**32767**的优先级，值越大优先级越高
- ❖ 如果用户没有指定，那么核心会自动给**swap**指定一个优先级，这个优先级从**-1**开始，每加入一个新的没有用户指定优先级的**swap**，会给这个优先级减一。
- ❖ 先添加的**swap**的缺省优先级比较高，除非用户自己指定一个优先级，而用户指定的优先级(是正数)永远高于核心缺省指定的优先级(是负数)。
- ❖ 优化性能：分布存放，高性能磁盘存放

马哥教育

www.magedu.com

- ❖ 挂载意味着使外来的文件系统看起来如同是主目录树的一部分
- ❖ 访问前、介质必须被挂载
- ❖ 摘除时，介质必须被卸载
- ❖ 按照默认设置，非根用户只能挂载某些设备（光盘、DVD、软盘、**USB**等等）
- ❖ 挂载点通常在/**media** 或/**mnt**下

马哥教育

www.magedu.com

- ❖ 在图形环境下自动启动挂载/run/media/<user>/<label>
- ❖ 否则就必须被手工挂载

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

- ❖ eject命令卸载或弹出磁盘

- ❖ 创建ISO文件

```
cp /dev/cdrom /root/centos7.iso
```

```
mkisofs -r -o /root/etc.iso /etc
```

- ❖ 刻录光盘

```
wodim -v -eject centos.iso
```

马哥教育
www.magedu.com

- ❖ 被内核探测为SCSI设备
 - /dev/sdaX、/dev/sdbX、或类似的设备文件
- ❖ 在图形环境中自动挂载
 - 图标在[计算机]窗口中创建
 - 挂载在/run/media/<user>/<label>
- 手动挂载
- `mount /dev/sdb1 /mnt`

马哥教育

www.magedu.com

- ❖ 内存空间使用状态:

free [OPTION]

- m: 以MB为单位
- g: 以GB为单位

- ❖ 文件系统空间占用等信息的查看工具:

df [OPTION]... [FILE]...

- H 以1000为单位
- T 文件系统类型
- h: human-readable
- i: inodes instead of blocks
- P: 以Posix兼容的格式输出

- ❖ 查看某目录总体空间占用状态:

du [OPTION]... DIR

- h: human-readable
- s: summary

❖ **dd命令: convert and copy a file**

❖ **用法:**

dd if=/PATH/FROM/SRC of=/PATH/TO/DEST

bs=#: block size, 复制单元大小

count=#: 复制多少个bs

❖ **of=file** 写到所命名的文件而不是到标准输出

if=file 从所命名文件读取而不是从标准输入

bs=size 指定块大小（既是是**ibs**也是**obs**）

ibs=size 一次读**size**个**byte**

obs=size 一次写**size**个**byte**

cbs=size 一次转化**size**个**byte**

skip=blocks 从开头忽略**blocks**个**ibs**大小的块

seek=blocks 从开头忽略**blocks**个**obs**大小的块

count=n 只拷贝**n**个记录

- ❖ **conv=conversion[,conversion...]** 用指定的参数转换文件
- ❖ 转换参数:
- ❖ **ascii** 转换 EBCDIC 为 ASCII
- ❖ **ebcdic** 转换 ASCII 为 EBCDIC
- ❖ **block** 转换一行数据为长度为 **cbs** 的记录，不足部分用空格填充。
- ❖ **unblock** 替代**cbs**长度的每一行尾的空格为新行
- ❖ **lcase** 把大写字符转换为小写字符
- ❖ **ucase** 把小写字符转换为大写字符
- ❖ **nocreat** 不创建输出文件
- ❖ **noerror** 出错时不停止
- ❖ **notrunc** 不截短输出文件
- ❖ **sync** 把每个输入块填充到**ibs**个字节，不足部分用空(NUL)字符补齐

❖ 备份MBR:

```
dd if=/dev/sda of=/tmp/mbr.bak bs=512 count=1
```

❖ 破坏MBR中的bootloader:

```
dd if=/dev/zero of=/dev/sda bs=64 count=1  
seek=446
```

❖ 有一个大与2K的二进制文件fileA。现在想从第64个字节位置开始读取，需要读取的大小是128Byts。又有fileB，想把上面读取到的128Bytes写到第32个字节开始的位置，替换128Bytes，请问如何实现？

```
#dd if=fileA of=fileB bs=1 count=128 skip=63 seek=31  
conv=notrunc
```

❖ 备份:

```
dd if=/dev/sdx of=/dev/sdy
```

将本地的/dev/sdx整盘备份到/dev/sdy

```
dd if=/dev/sdx of=/path/to/image
```

将/dev/sdx全盘数据备份到指定路径的image文件

```
dd if=/dev/sdx | gzip >/path/to/image.gz
```

备份/dev/sdx全盘数据，并利用gzip工具进行压缩，保存到指定路径

❖ 恢复:

```
dd if=/path/to/image of=/dev/sdx
```

将备份文件恢复到指定盘

```
gzip -dc /path/to/image.gz | dd of=/dev/sdx
```

将压缩的备份文件恢复到指定盘

❖ 拷贝内存资料到硬盘

```
dd if=/dev/mem of=/root/mem.bin bs=1024
```

将内存里的数据拷贝到root目录下的mem.bin文件

❖ 从光盘拷贝iso镜像

```
dd if=/dev/cdrom of=/root/cd.iso
```

拷贝光盘数据到root文件夹下，并保存为cd.iso文件

❖ 销毁磁盘数据

```
dd if=/dev/urandom of=/dev/sda1
```

利用随机的数据填充硬盘，在某些必要的场合可以用来销毁数据。执行此操作以后，/dev/sda1将无法挂载，创建和拷贝操作无法执行。

❖ 得到最恰当的block size

```
dd if=/dev/zero bs=1024 count=1000000  
of=/root/1Gb.file
```

```
dd if=/dev/zero bs=2048 count=500000  
of=/root/1Gb.file
```

```
dd if=/dev/zero bs=4096 count=250000  
of=/root/1Gb.file
```

通过比较dd指令输出中所显示的命令执行时间，即可确定系统最佳的block size大小

马哥教育
www.magedu.com

❖ 测试硬盘读写速度

```
dd if=/dev/zero of=/root/1Gb.file bs=1024  
count=1000000
```

```
dd if=/root/1Gb.file bs=64k | dd of=/dev/null
```

通过上两个命令输出的执行时间，可以计算出测试硬盘的写/读
/ 速度

马哥教育

www.magedu.com

❖ 修复硬盘

`dd if=/dev/sda of=/dev/sda`

当硬盘较长时间（比如1,2年）放置不使用后，磁盘上会产生消磁点。当磁头读到这些区域时会遇到困难，并可能导致I/O错误。当这种情况影响到硬盘的第一个扇区时，可能导致硬盘报废。上边的命令有可能使这些数据起死回生,且这个过程是安全高效的。

马哥教育

www.magedu.com

- ❖ 1、创建一个**2G**的文件系统，块大小为**2048byte**，预留**1%**可用空间,文件系统**ext4**，卷标为**TEST**，要求此分区开机后自动挂载至**/test**目录，且默认有**acl**挂载选项
- ❖ 2、写一个脚本，完成如下功能：
 - (1) 列出当前系统识别到的所有磁盘设备
 - (2) 如磁盘数量为**1**，则显示其空间使用信息
否则，则显示最后一个磁盘上的空间使用信息

马哥教育

www.magedu.com

- ❖ 博客: <http://magedu.blog.51cto.com>
- ❖ 主页: <http://www.magedu.com>
- ❖ QQ: 1661815153, 113228115
- ❖ QQ群: 203585050, 279599283

马哥教育
www.magedu.com



马哥教育
最专业的Linux培训机构

Thank You!