1. **本章目标**

通过分析主机规划与磁盘分区掌握规划和分区技巧。

1. **Linux是什么**
2. **Linux与硬件的搭配**
   1. 如果是游戏机/工作机的考虑

用来玩游戏的“游戏电脑”所需要的配置一定比办公室用的“工作电脑”的配置高。主要体现在显卡和CPU及主板芯片组上，因为很多3D游戏需要3D图形运算太多。注：当前（2018年）市场中Linux下的游戏没有Windows下面的多。

* 1. 性价比

除了考虑配件自身的性价比，也要考虑能源消耗。特别是服务器集群的的消耗，不容忽视。

* 1. 支持度

需要考虑配件对Linux系统的驱动程序的支持。

1. **具体配件的选择方式**
   1. CPU

根据实际使用率（主频）来决定是否可以使用。一般系统版本会有最低配置需求。尽量不使用低于推荐配置的等级。

* 1. 内存

越大越好。

* 1. 硬盘

需要考虑两个问题。1是容量，2是安全。如果是服务器应用的话，优先考虑RAID（之后会讲到）。

* 1. 显卡

一般不需要独立显卡，板载显卡足矣。

* 1. 网卡

网卡是服务器上重要的组件之一。因为提供服务时需要超大的吞吐量。所以要考虑使用比较好的网卡，而比较好的网卡通常Linux的驱动程序也做的比较好，用起来会比较顺畅。好一点的Intel或Broadcom等公司的网卡应该是比较合适的选择。

* 1. 光盘、软盘、键盘和鼠标

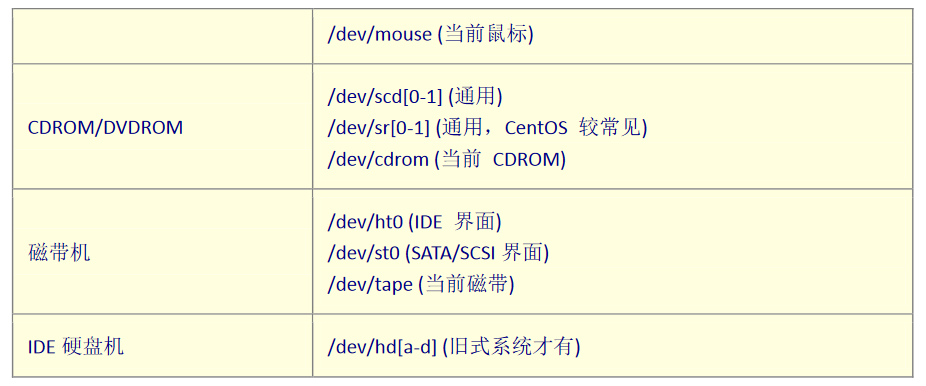
只要不是旧的计算机不支持就可以，因为这这些设备都不是必要配件，而其中的软盘早已经完全退出市场了，而光盘近几年也已经淡出市场。

* 1. 企业用计算机

企业用计算机建议购买商用服务器，因为商用服务器已经通过制造商的散热、稳定性等测试，对于企业来说，会是一个比较好的选择。组装的服务器终究有不确定因素，单如果必须要考虑成本的话，也可以组装使用。而如果是面对可以接受宕机的业务应用时，也可以直接选择组装的服务器。

1. **各硬件在Linux中的文件名**
   1. 每个设备都被当成一个文件来对待
   2. 时至今日，由于IDE接口的磁盘驱动器几乎已经被淘汰，因此现在IDE接口的磁盘文件名也被模拟为/dev/sd[a-p].
   3. 如果使用的是云服务器，可能会得到的是虚拟机。而虚拟机为了加速，在其内部使用的磁盘是使用模拟器产生的，该模拟器产生的磁盘文件名可能为/dev/vd[a-p]系列的文件名。



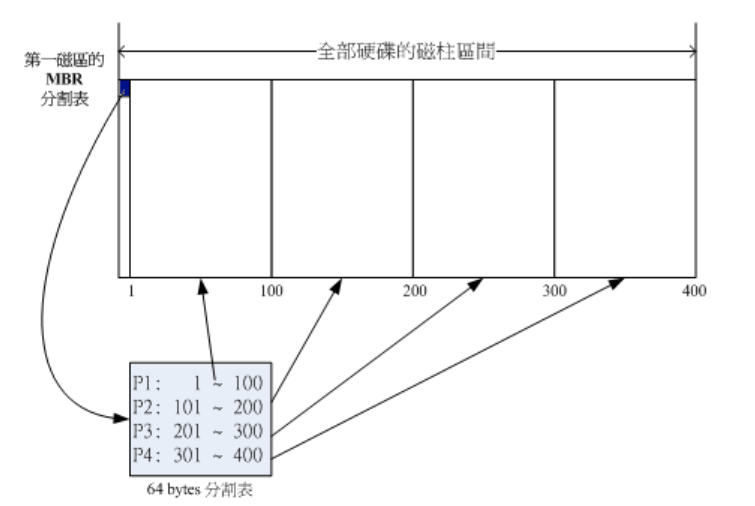


1. **使用虚拟机**
   1. Windows
      1. VirtualBox（免费，简约）
      2. Vmware Workstation（商用，功能强大）
   2. Linux
      1. KVM（开源）
2. **磁盘分区**
3. **磁盘连接方式与设备文件名的关系**
   1. 正常的物理机器大概使用的都是/dev/sd[a-p]的磁盘文件名，至于虚拟机环境中，为了加速可能就会使用/dev/vd[a-p]这种设备文件名。
   2. 根据Linux内核检测到磁盘的顺序来命名，而与实际插槽顺序无关。例：PC上有两个SATA磁盘已经一个USB磁盘，而主板上面有六个SATA的插槽。这两个SATA分别安插在SATA1、SATA5插槽上。最终设备的文件名如下：
      1. SATA1插槽上的文件名：dev/sda
      2. SATA5插槽上的文件名：dev/sdb
      3. USB磁盘（系统启动完成后才被系统识别）：/dev/sdc
4. **MBR与GPT磁盘分区表**
   1. 第一扇区特别重要，因为它记录了整块磁盘的重要信息。早期磁盘第一个扇区里面含有的重要信息被称作“MBR”，但是由于近年来磁盘的容量不断扩大，造成读写上的一切困扰，甚至有些2TB以上的磁盘分区已经让某些操作系统无法存取。因此后来又多了一种新的磁盘分区格式，称为“GPT”（GUID Partition [pɑːrˈtɪʃn] table），这两种分区格式与限制不太相同。分区表：一块新的硬盘就像一根原木，需要在这根原木上进行分割出想要的区段，这个区段才是能够再制作家具的部分。如果没有进行切割，那么原木就不能被有效的使用。硬盘亦如此。
   2. 通常文件系统的最小单位是柱面。因为GPT这个可以达到64位记录功能的分区表，可以使用扇区号码来作为分区单位。所以，综上所述，我们就是**利用参考对照柱面或扇区号码的方式来处理**。
   3. MBR（MS-DOS）分区表格式与限制

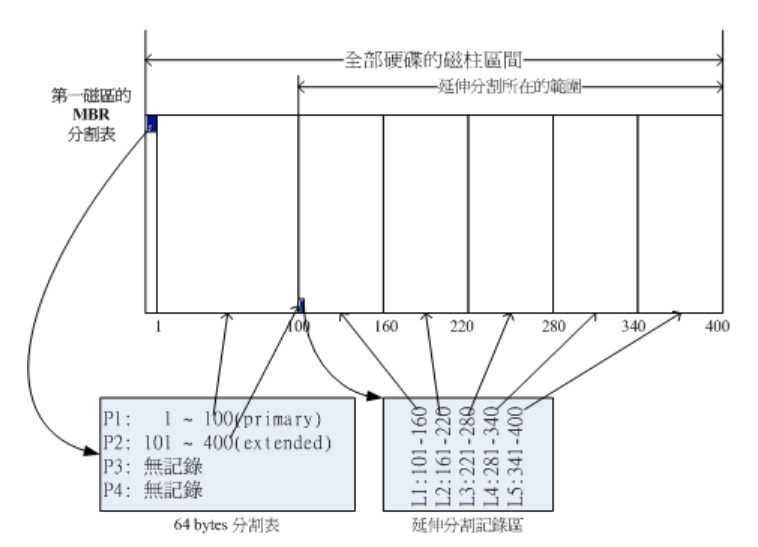
早期Linux为了兼容Windows，因此使用的是支持Windows的MBR（Master Boot Record[ˈrekərd , rɪˈkɔːrd]主引导记录）。

第一扇区的512字节主要会有两个东西：

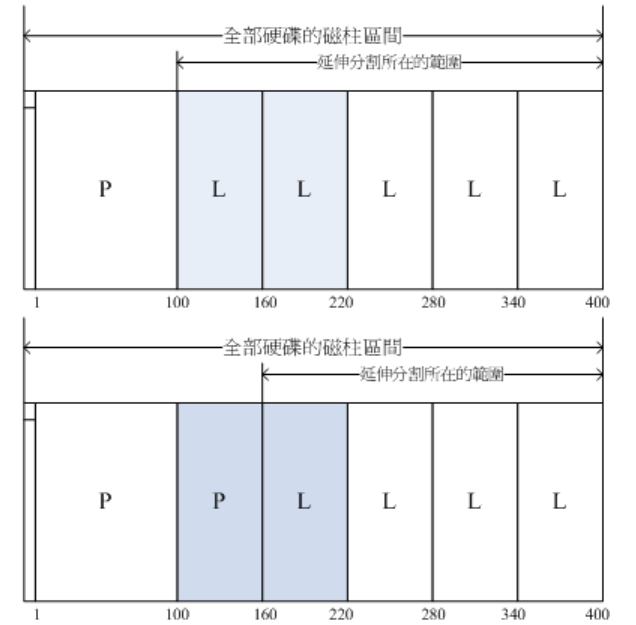
* + 1. 主引导记录MBR（Master Boot Record）：可以安装启动引导程序的地方，有446个字节。
    2. 分区表（partition table）：记录整块硬盘分区的状态，有64个字节。
    3. 由于分区表所在的区块只有64个字节的容量，因此最多仅能有四组记录区，每组记录区记录了该区段的启始与结束的柱面号码。



* + 1. 所谓的分区其实只是针对64个字节的分区表进行设置而已。
    2. 硬盘默认的分区表仅能写入四组分区信息。
    3. 这四组划分信息被称为主要（Primary[ˈpraɪmeri]）或扩展（Extended[ɪkˈstendɪd] ）分区。
    4. 分区的最小单位通常为柱面。
    5. 当系统要写入磁盘时，一定会参考磁盘分区表，才能针对某个分区进行数据的处理。
    6. 分区的好处：a数据的安全性（可以单独格式化）；b系统的性能（数据区段集中）
  1. 利用扩展分区划分多个分区

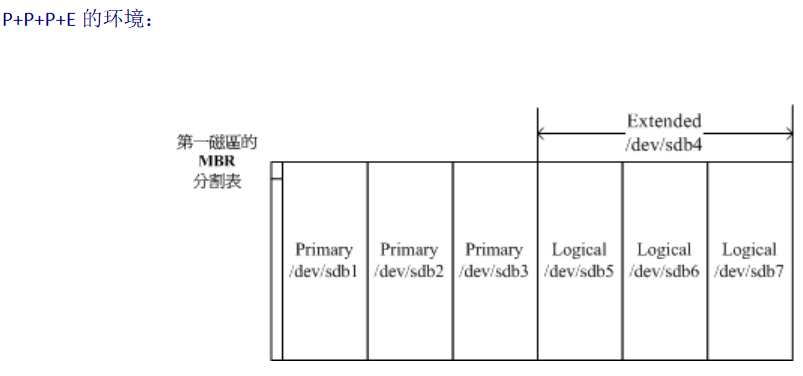


* + 1. P1为主要分区。
    2. P2为扩展分区。
    3. 扩展分区的目的是使用额外的扇区来记录分区信息。
    4. 前四个分区号码会被预留给主要分区或扩展分区使用。所以逻辑分区的设备名称号码由5号开始。这在MBR方式的分区表中是很重要的特性。
  1. 主要分区与扩展分区最多可以有4个。（硬盘的限制）
  2. 扩展分区最多只能有1个。（操作系统的限制）
  3. 逻辑分区是由扩展分区持续划分出来的分区。
  4. 能够被格式化后作为数据存取的分区是主要分区与逻辑分区，扩展分区无法格式化。
  5. 逻辑分区的数量依操作系统而不同，在Linux系统中SATA硬盘已经可以突破63个以上的分区限制。
  6. 分区以柱面为单位的连续磁盘空间，且扩展分区又是个类似独立的磁盘空间。所以分区的时候要注意一下情况。

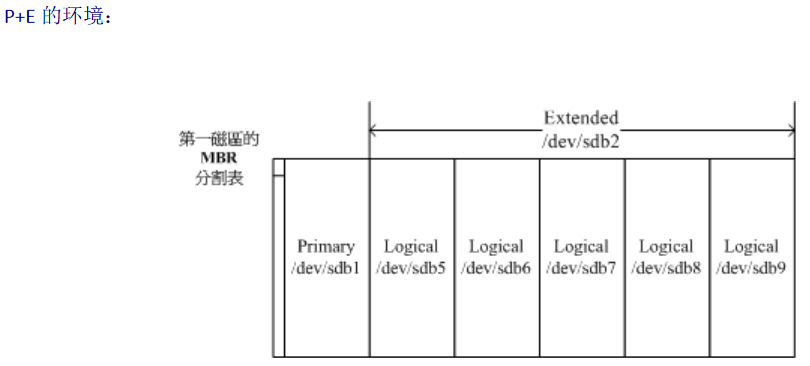


上图二三分区可以重新整合为一个分区，下图不可以。因为下图一个是主要分区，一个是逻辑分区。而如果扩展分区被破坏，所有逻辑分区将会被删除。

* 1. 分区例子

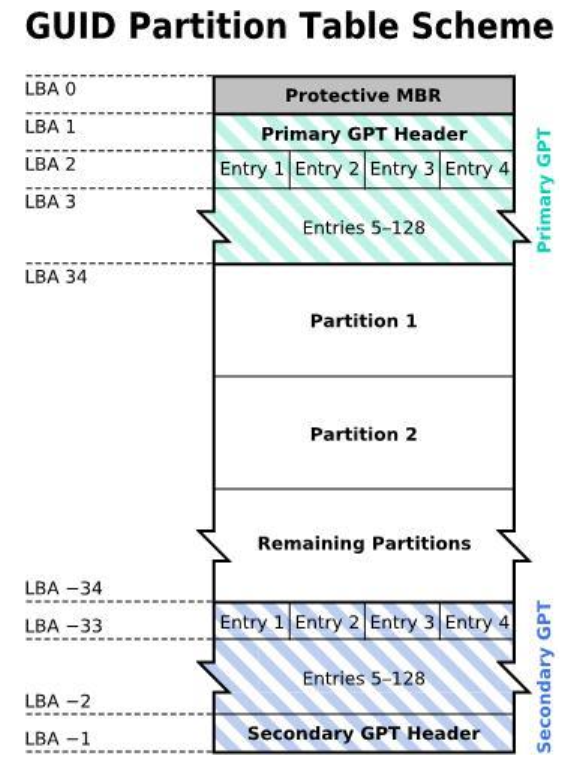


实际可用的是/dev/sdb1, /dev/sdb2, /dev/sdb3, /dev/sdb5, /dev/sdb6, /dev/sdb7这六个，至于/dev/sdb4这个延伸分区本身仅是提供来给逻辑分区建立之用。



因为1~4号是保留给主要/扩展分区槽的，因此第一个逻辑分区一定是由5号开始的！所以/dev/sdb3, /dev/sdb4就会被保留下来没有用到了。

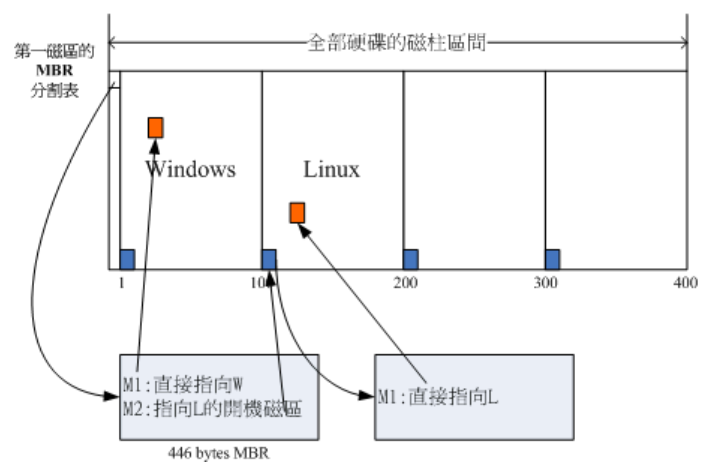
* 1. MBR分区表每组分区表仅有16个字节而已，所以MBR有以下限制：
     1. 操作系统无法使用2.2TB以上的磁盘容量。
     2. MBR仅有一个区块，若被破坏后，经常无法或很难恢复。
     3. MBR内的存放启动引导程序的区块仅有446个字节，无法存储较多的程序代码。
  2. GPT磁盘分区表



为解决各种MBR限制，出现了GPT。

* + 1. 过去扇区大小是512字节，目前已经有4K的扇区设计出现。为了兼容所有的磁盘，因此大多会使用所谓的逻辑区块地址（LBA）来处理。GPT将磁盘所有的区块以此LBA（默认为512字节）来规划。而第一个LBA称为LBA0（从0开始编号）
    2. GPT使用34个LBA区块来记录分区信息。
    3. GPT除了前面34个LBA之外，整个磁盘的最后34个LBA也拿来做为另一个备份。这样可以安全一些。
    4. LBA0作为MBR兼容区块。这个兼容区块也会分为两部分。一个就是跟之前446字节相似的区块，存储了第一阶段的启动引导程序。而在原本的分区表的记录区内，这个兼容模式仅放入一个特殊标识符，用来表示此磁盘为GPT格式。
    5. LBA1作为GPT表头记录。记录了分区表本身的位置和大小，同时记录了备份用的GPT分区，分区表的效验码（CRC32），操作系统可以通过这个效验码来判断GPT是否正确，若有错误，可以通过这个记录区来获取备份的GPT来恢复GPT的正常运行。
    6. LBA2-LBA33存储实际记录分区信息。每个LBA都可以记录4组分区记录，所以在默认的情况下，一共可以有4\*32=128组分区记录。因为每个LBA有512个字节，因此每组记录用到128字节的空间，除了每组记录所需要的标识符与相关记录外，GPT在**每组记录中分别提供了64位来记载开始/结束的扇区号码。**因此，GPT分区表对于单一分区来说，它的最大容量限制就会是1ZB。1ZB=2^30TB=1,073,741,824TB。
  1. 现在GPT分区默认可以提供128组记录，在Linux内核设备记录中，针对单一磁盘，过去最多只能到达15个分区，不过由于udev （udev 是Linux kernel 2.6系列的[设备管理器](https://baike.baidu.com/item/%E8%AE%BE%E5%A4%87%E7%AE%A1%E7%90%86%E5%99%A8" \t "https://baike.baidu.com/item/udev/_blank)。它主要的功能是管理/dev目录底下的设备节点。）等方式的处理，现在Linux已经没有这个限制了。此外，GPT分区没有主、扩展、逻辑分区的概念，每个都可以当作主要分区，每个分区都可以拿来格式化使用。

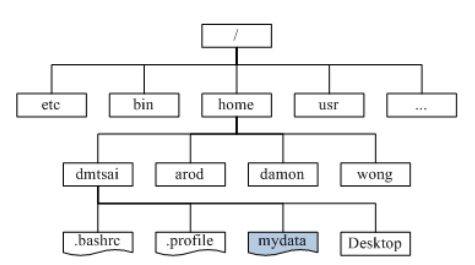
1. **启动流程中的BIOS与UEFI启动检测程序**
   1. 并不是所有的操作系统都可以读取到GPT的磁盘分区格式，也不是所有的硬件都可以支持GPT格式，是否能够读写GPT格式又与启动的检测程序有关。BIOS和UEFI即是启动的检测程序。
   2. BIOS（Basic Input Output System）搭配MBR/GPT的启动流程
      1. CMOS是记录各项硬件参数且嵌入在主板上的存储器。（CMOS电池）
      2. BIOS是写入到主板上的一个固件（固件就是写入到硬件上的一个软件程序）。
      3. BIOS可以修改CMOS当中的参数。
      4. BIOS就是在启动的时候，计算机系统会主动执行的第一个程序。它在执行时第一时间加载CMOS当中的参数。
      5. 接下来BIOS会分析计算机里面的存储设备。以硬盘为例，BIOS根据用户的设置去取得能够启动的硬盘，并且到该硬盘里面去读取第一个扇区的MBR位置。MBR这个仅有446个字节的硬盘容量里面会放置最基本的启动引导程序。此时BIOS的使命已完成。而接下来就是MBR的工作了。
      6. MBR这个启动引导程序的目的是加载（load）内核文件。由于启动引导程序是操作系统在安装的时候提供的，所以它会认识硬盘内的文件系统格式，因此就能够读取内核文件，此时MBR的使命完成。之后的工作就交给了操作系统。
      7. BIOS->MBR->启动引导程序（boot loader）->内核文件
      8. 如果分区表是GPT格式的话，那么BIOS也能够从LBA0的MBR兼容区块读取第一阶段的启动引导程序代码，如果启动引动程序能够支持GPT的话，那么BIOS同样可以读取到正确的操作系统内核。而如果启动引导程序不懂GPT，比如Windows XP的环境，那自然无法读取内核文件，也就无法启动操作系统了。
      9. BIOS与MBR都是硬件本身会支持的功能。
      10. Boot loader是从系统安装在MBR上面的软件。由于MBR只有446字节，所以启动引导程序非常小而高效。Boot loader的主要任务有：
          1. 提供选项。可以选择不同的启动选项。这也是多重引导的重要功能。
          2. 加载内核文件。直接指向可使用的程序区段来启动操作系统。
          3. 转交其他启动引导程序。将启动管理功能转交给其他启动引导程序。
      11. 启动引导程序除了可以安装在MBR之外，还可以安装在每个分区的启动扇区。



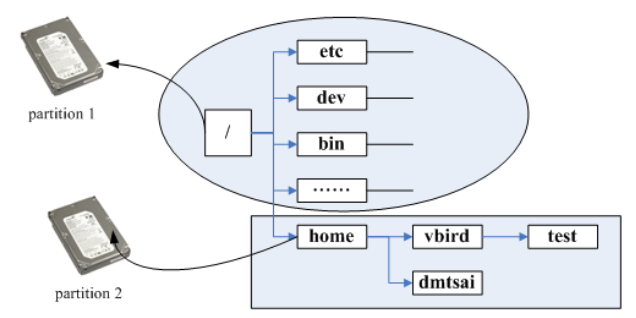
* + - 1. 每个分区都拥有自己的启动扇区。（Boot sector）
      2. 图中系统分区为第一及第二分区。
      3. 实际可启动的内核文件是放置到各分区中的。
      4. 启动引导程序只会认识自己的系统分区内的可启动的内核文件以及其他启动引导程序。
      5. 启动引导程序可直接指向或间接将管理权交给另一个管理程序。
    1. 如果要安装多重引导，最好先安装Windows再安装Linux。因为Linux再安装的时候可以选择启动引导程序安装的位置及设置选项，Windows则是直接覆盖掉MBR。如果先安装Linux，后安装的Windows也无需再重新安装Linux，可以用Linux的恢复模式来修复MBR。
  1. UEFI BIOS搭配GPT启动的流程
  2. UEFI是统一可扩展固件接口。
  3. UEFI主要解决BIOS实际不懂GPT的问题。
  4. UEFI使用C编写的。比起汇编开发的BIOS更容易开发，也拥有更多功能。
  5. UEFI在开机速度上比BIOS快。
  6. UEFI硬件资源使用轮询的方式管理，比BIOS直接使用CPU以中断的方式来管理效率要稍微低一些。
  7. UEFI不能提供完整的缓存功能，因此执行效率也没有办法提升。
  8. 由于过去骇客（Cracker）出于安全考虑，UEFI提供一个安全启动（secure boot）的功能，这个功能代表着即将启动的操作系统必须要被UEFI验证，否则无法顺利启动。微软用了很多这样的功能来管理硬件。但是许多的操作系统，包括Linux，很有可能就无法顺利启动。所以，**某时某刻，可能需要将UEFI的secure （[sɪˈkjʊr]保护） boot功能关闭才能顺利进入Linux。**
  9. 虽然UEFI可以直接获取GPT的分区表，不过最好保留BIOS boot的分区支持，同时，为了与Windows兼容，并且提供其他第三方厂商所使用的UEFI应用程序存储的空间，必须要格式化一个FAT格式的文件系统分区，大概提供512MB到1GB左右的大小，让其它UEFI执行较为方便。
  10. 扇区0（MBR，主引导分区）存放一段代码，这段代码指定了BIOS从哪里启动系统。在计算机启动的时候，BIOS从扇区0读取数据，跳到相应地址，读取系统内核，开始启动。对于磁盘地址的表示，BIOS（BIOS调用INT13）能识别两种模式。一种 linear（线性），另外一种是 triple(CHS,cylinder,head,sector)。但BIOS出现的时间比LINEAR要早，所以在处理LINEAR的时候，也是转化成CHS来访问磁盘的。CHS使用3个字节（BYTE，24bit)来表示磁盘地址，10位表示柱面，6位表示头，8位表示扇区。因为表示柱面的只有10位，那么柱面的个数最多只能有2的10次方也就是1024个（0~1023）。因为一个扇区512字节，那么使用CHS模式能表示的磁盘地址=1024\*64\*256\*512=8589934592字节=8.4GB,这就是著名的1024柱面限制和8.4GB限制。为了表示比8.4GB更大的磁盘，后来出现了一个扩展BIOS接口（Extended BIOS Interface),使用16字节来表示磁盘地址。LILO21.4版本之前缺省使用LINEAR模式，所以不能突破限制，在更新的版本中，使用了LBA32。使用LBA32的有两个意思：1、使用LINEAR模式，2、使用扩展BIOS（也就是使用16字节的磁盘地址）所以，只要在lilo.conf中加上一行“LBA32”或者“lba32”就可以突破1024柱面和8.4GB的限制了（也可以在1024柱面之前分出一个区，用来mount存放内核文件的/boot）
  11. UEFI已经解决了BIOS的1024柱面问题，因此启动引导程序与内核可以放置在硬盘开始的前2TB位置内。加上之前提到的BIOS以及UEFI支持的分区，基本上/boot目录几乎都是/dev/sda3之后的号码。



1. **Linux安装模式下，磁盘分区的选择**
   1. 目录树结构。整个目录树架构最重要的就是根目录（root directory），这个根目录的表示方法为一条斜线“/”。所有的文件都是由根目录衍生来的。



* 1. 文件系统与目录树的关系（挂载）。挂载就是利用一个目录当成进入点，将磁盘分区的数据放置在该目录下。也就是进入该目录就可以读取改分区的内容。这个操作称为“挂载”，进入点的目录称为“挂载点”。由于整个Linux系统最重要的是根目录，因此根目录一定需要挂载到某个分区，而其他的目录则可以根据需要挂载到不同的分区。



上图中硬盘分为两个分区，分别挂载根目录和/home

* 1. Windows也是由挂载的概念的。比如U盘连接后会出现盘符，而且盘符是自动分配的。
  2. 判断某个文件在哪个分区的方法是反向追踪。上图为例：/home/fish/test这个文件，由test->fish->home->/,看哪个“进入点”先被查到那就是使用的进入点。所以test使用的是/home。
  3. 挂载点与磁盘分区的规划
     1. 初次接触Linux，可以只划分“/”和“交换分区”，“交换分区”（SWAP）就是LINUX下的虚拟内存分区,它的作用是在物理内存使用完之后,将磁盘空间(也就是SWAP分区)虚拟成内存来使用。
     2. 预留一个备用的剩余磁盘容量。

1. **安装Linux前的规划**
2. **选择适当的Linux发行版**
3. **主机的服务规划与硬件的关系**
   1. 之前提过，由于主机的服务目的不同，所需要的配件等级与配置自然也就不同。所以需要从服务需求来分析硬件需求。
   2. NAT（网络地址转换）需要好的网卡。
   3. SAMBA（网络邻居）需要好的网卡和大存储量硬盘。
   4. Mail（邮件服务器）需要好的网卡和大存储量硬盘。
   5. Web（WWW服务器）需要好的CPU和大内存。
   6. DHCP（提供客户端自动获取IP功能）无特别硬件需求。
   7. FTP（文件传输服务器）需要好的网卡和大存储量硬盘。
4. **主机硬盘的主要规划**

数据的分类与数据安全性的考虑（非骇客破坏）。当主机系统的硬件出现问题时（比如坏道出现在非用户文件分区），文件数据能否安全的保存。即将用户分区独立挂载，系统出错无需迁移与备份即可重新操作其他分区，比如重装或修复系统。

虽然挂载多个分区会略显麻烦，但是提高了灵活性和安全性，还是值得的。比如上边的/home被独立挂载。

1. **本章作业**
2. **分析自己的计算机如果全年运行的电费消耗是多少钱。**
3. **一部计算机主机是否只要CPU够快，整体速度就会提高？**
4. **Linux对于硬件的要求需要的考虑什么？是否一定要很高的配备才能安装？**
5. **请写下下列设备在Linux的设备文件名：SATA硬盘，CDROM，打印机**
6. **目前个人计算机上面常见的硬盘与主板的简介接口有哪些？**
7. **一台好的主机在组装前，有哪些需要注意的主机规划？**
8. **如果磁盘分区是MBR，且设置了4各主分区，但是还有剩余空间，请问还能不能使用这些空间？如果使用的话需要怎么做？**