

# 文件管理

- 文件与文件系统
- 文件的存储介质
- 磁盘空间管理
- 文件控制块及文件目录
- 文件的物理结构
- 文件系统的实现
- 文件系统实例——UNIX



文件的分类、文件的逻辑结构

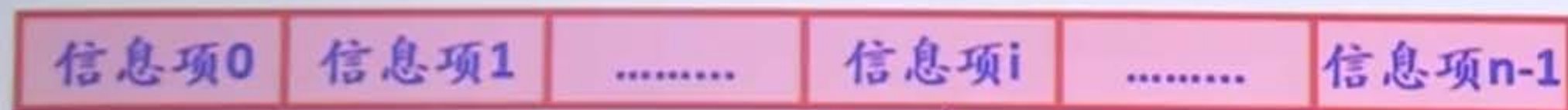
# 文件与文件系统





# 文件是什么？

- 文件 是 对磁盘的 抽象
- 所谓文件 是指 一组带标识（标识即为文件名）的、在逻辑上有完整意义的信息项的序列
- 信息项：构成文件内容的基本单位（单个字节，或多个字节），各信息项之间具有顺序关系
- 文件内容的意义：由文件建立者和使用者解释



读写指针





# 如何设计一个文件系统

## 操作系统角度：

怎样组织、管理文件？

- ✓ 文件的描述、分类
- ✓ 文件目录的实现
- ✓ 存储空间的管理
- ✓ 文件的物理地址
- ✓ 磁盘实际运作方式(与设备管理的接口)
- ✓ 文件系统性能

## 用户角度：

文件系统如何呈现在用户面前：

- ✓ 一个文件的组织
  - ✓ 如何命名？
  - ✓ 如何保护文件？
  - ✓ 可以实施的操作？
- 等等

文件管理的需求分析





# 文件系统

- ◎ 操作系统中统一管理信息资源的一种软件，管理文件的存储、检索、更新，提供安全可靠的共享和保护手段，并且方便用户使用

- 统一管理磁盘空间，实施磁盘空间的分配与回收
- 实现文件的按名存取
  - 名字空间  $\xrightarrow{\text{映射}}$  磁盘空间
- 实现文件信息的共享，并提供文件的保护、保密手段
- 向用户提供一个方便使用、易于维护的接口，并向用户提供有关统计信息
- 提高文件系统的性能
- 提供与I/O系统的统一接口





# 文件的分类

## 按文件性质和用途分类 (UNIX)

普通文件；目录文件；特殊文件(设备文件)；管道文件；套接字

### 普通文件(regular)

包含了用户的信息，一般为ASCII或二进制文件

### 目录文件(directory)

管理文件系统的系统文件

### 特殊文件(special file)

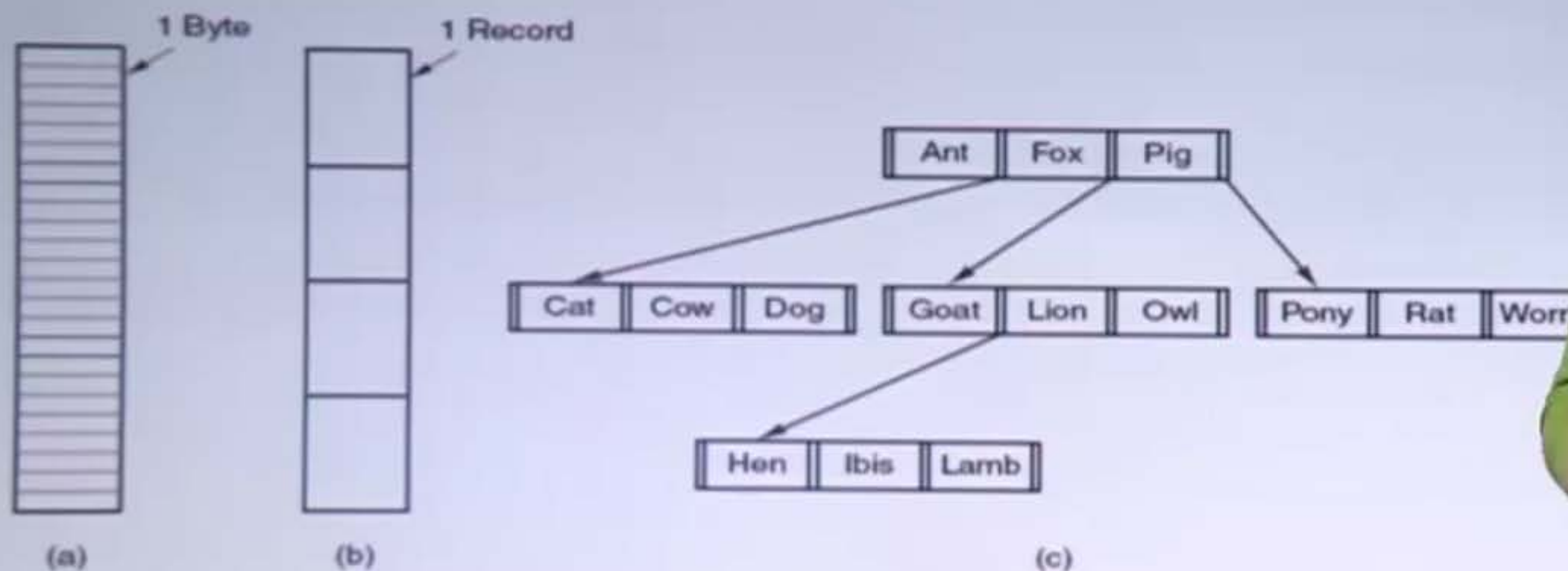
字符设备文件：和输入输出有关，用于模仿串行I/O设备，例如终端，打印机，网卡等

块设备文件：磁盘



# 文件的逻辑结构

从用户角度看文件，由用户的访问方式确定



还可以组织成堆、顺序、索引、索引顺序、散列等结构



# 典型的文件逻辑结构与文件存取

- ◎ **流式文件**：构成文件的基本单位是字符  
文件是有逻辑意义、无结构的一串字符的集合
- ◎ **记录式文件**：文件由若干个记录组成，可以按记录进行读、写、查找等操作  
每条记录有其内部结构

- ◆ 顺序存取(访问)

- ◆ 随机存取(访问)

提供读写位置(当前位置)

例如：UNIX的seek操作





大家好！今天我给大家带来的是操作系统原理课的第九讲 文件系统。在文件系统里头，我们主要介绍文件与文件系统的基本概念 文件存放的存储介质 磁盘空间的管理，文件控制块及文件目录的概念 文件的物理结构，文件系统的实现 以及以 UNIX 为例，介绍文件系统的实例 首先我们来介绍文件与文件系统的概念 什么是文件呢？就如同进程 是对 CPU 的抽象，地址空间 是对内存的抽象一样，文件 是对磁盘的抽象。所谓文件呢 实际上是指一组带标识的 在逻辑上有完整意义的信息项的序列 所谓标识呢，实际上就是我们说的文件名 那么一个文件 它逻辑上有它自己的含义 它有它完整的意义，但是呢这是对于使用者决定的。使用者把这个文件看成是什么样的文件 对它实施什么样的操作，那么这是从使用者的角度来谈它的意义 那么对于操作系统而言，那么文件其实就是信息项的一个序列 那么什么是信息项呢 信息项实际上就是构成文件内容的一个基本单位 可能这个基本单位是单个字节 也可以是多个字节。我们强调的是各信息项之间呢是具有一种顺序关系 那么文件的内容正如刚才我们所谈 逻辑上有完整意义，文件的内容实际上是 文件的建立者和使用者来 对文件的内容进行解释，进行使用的 那么对操作系统而言，我们就把文件 看成是一个信息项的序列，我们可以看一下它是有一个顺序关系 信息项 0，信息项 1，依次排着 那么我们看一下，如果这个信息项是 单个字节，那就是单个字节，那么信息项也可以是什么呢？是多个字节。在这里头我们知道 要去读文件去写文件，需要定位 读写指针，因此呢对于这样一个信息项的序列 我们会有一个读写指针指到某一个具体的信息项 我们这一讲主要是介绍 如何设计一个文件系统，也就是说我们作为一个开发人员 对文件管理的需求进行相应的分析 我们从两个角度来看文件系统应该具备什么样的功能 我们先看从用户的角度 从用户的角度，也就是说文件系统 呈现在用户面前的是个什么样子，用户在使用文件系统时候他所看到的文件系统 通常呢我们会考虑文件的组织 文件的命名 如果需要对文件进行相应的保护，如何进行 在文件上可以实施哪些操作 读、写还有其他的一些操作，等等 就是从用户的角度我们需要对文件 提供这样的一些功能 但是呢，由于我们现在站在的是操作系统的角度 也就是从系统的设计 角度来看怎么样去组织、管理文件 首先我们要考虑对文件 如何描述一个文件，如何能够 对文件的各个属性，刻画它的属性 建立相应的数据结构，文件有哪些分类 然后一个非常重要的概念就是文件目录 它在文件系统里头起着非常重要的作用，那么文件目录是怎么实现的？文件总是存放在某些 存储介质上，那么存储空间如何管理？文件怎么存放在存储介质上，它的物理地址 怎么样去计算？当然了文件系

统呢 因为通常文件系统都存放在磁盘上 所以我们会说文件系统它和设备 这个功能的接口，那么就是磁盘的 实际运作方式也是我们文件系统要考虑的 当然了文件系统呢，因为是呈现在用户面前，用户跟操作系统打交道，非常多的是和文件系统打交道 因此，文件系统的性能至关重要 用户频繁地使用文件系统，那么文件系统的 用户体验好不好，那么取决于它的性能 什么是文件系统？文件系统是操作系统中 统一管理信息资源的一种软件 文件系统呢，为文件的存储 检索、更新提供相应的支持 同时为文件的安全、可靠、共享、保护提供相应的手段 文件系统呢，还要方便用户的使用 下面我们看一下文件系统 主要应该完成哪些重要的任务 第一项任务 文件系统要把磁盘空间管起来 对磁盘空间的分配、回收都是由文件系统统一管理的 第二项任务呢 文件系统是实现了文件的按名存取 这里头也有两个空间 第一个空间呢是名字空间，用户提供给系统 一些文件名。第二个空间呢是磁盘空间 那么文件呢就放在磁盘空间里头 那么文件系统实际上，就完成了从名字空间到磁盘空间的一个映射 它的作用就像我们在存储 管理里头一样，就相当于页表的作用一样，它完成 从文件名到文件在磁盘上具体的存放位置的一个映射 第三项任务呢，对于文件系统而言，应该能够提供文件的共享 并且呢还能够提供保护还有保密这样的一些手段 那么由于文件系统呢 管理各种各样的文件，因此希望给用户提供一个 非常方便的使用或者是易于维护的这样一个接口 同时文件系统还要向用户提供各种各样的统计信息 文件系统的性能影响了整个操作系统的性能，所以呢提高文件系统性能也是文件系统的一个重要任务 最后呢，文件系统要把所有的 文件从磁盘读入内存的时候，或者是写入磁盘的时候 需要和 I/O 系统有统一的接口 在文件系统中 有各种各样对文件的分类的方法 在这里头我们介绍一种 UNIX 的文件分类 它是按照文件的性质和用途来划分文件的 通常分为这样几种类型的文件 普通文件、目录文件、特殊文件 管道文件和套接字文件 下面我们介绍前面三种分类 普通文件呢实际上就是用户 自己建立的文件，是用户自己建立使用的文件 而目录文件呢，是操作系统为了管理文件系统而建立的 系统文件。那么在 UNIX 操作系统中，把各种各样的设备 当作文件来处理，那么就形成了我们现在说的特殊文件，也叫设备文件 为什么这么做呢？UNIX 最早提出了这样一种设计思想 它的主要的出发点是 所有的设备，各种各样的设备，品种繁多 那么比如说打印，我们有打印操作 我们要说扫描，扫描操作 那么从键盘读数据，然后呢对鼠标的这个操作进行相应的处理 因此呢各种各样的设备我



们把它抽象成文件之后呢 那么对于设备的操作就转换为对文件的操作 比如说打印文件, 那么实际上就变成了一个写操作 那么读键盘实际上就变成了一个读操作, 因此 把各种各样的设备抽象成一个文件 有利于提供一个统一的操作的接口 那么特殊文件或者叫做设备文件呢, 又可以分为两类 一类叫做字符设备文件, 那么它和 输入输出有关, 比如说我们的串口的 I/O 设备 和终端啊, 打印机啊或者网卡等等, 那么这些都是 按字符来进行相应的操作的, 因此呢我们把它称之为字符设备文件 而磁盘呢通常是由若干个字节来进行相应的 存储、检索还有传输, 因此呢磁盘呢我们把它称之为块设备文件 下面我们来看一下文件的逻辑结构 所谓文件的逻辑结构呢就是指从用户的角度来怎么看这个文件 它是由用户的访问方式来确定的 对于文件的逻辑结构呢, 我们这里给出了三种典型的结构 那么还可以把文件的逻辑结构组织成堆 顺序文件、索引顺序文件或者是散列文件 有各种各样的文件的逻辑结构。在我们给出的三个例子当中 最左边的呢是以一个字节为单位 把文件看成是一个字节的序列 那么这是一个非常典型的这么一个流式结构文件 中间的呢是以记录为单位 一个记录呢由若干个字节组成, 因此呢 我们这个文件呢, 信息项的单位就是记录, 那么这个文件就是由若干记录组成 那么这个呢是把文件组织成树形结构 好, 那么我们就来看一下典型的一般操作系统所采用的文件的逻辑结构有哪些呢? 第一种就是流式文件 也就是构成文件的基本单位是字符 因此, 文件就是一个有逻辑意义的 没有结构的一个一串字符的集合 第二种 常用的文件逻辑结构 呢就是记录式文件 那么文件呢由若干记录组成。那么对于文件 提供的操作呢就可以提供一些按记录进行读、写、查找的操作 如果是流式文件 或者叫做字符文件的话, 那么我们每次呢 提供读一个字符、写一个字符等等这样的操作 而每个记录呢 内部是有结构的, 它是由若干不同的字段组成 文件的逻辑结构和文件的存取有什么关系呢? 我们来看一下对文件的存取往往是两种 一种呢是顺序存取, 顺序访问 另一种是随机存取, 随机访问 顺序访问很容易理解, 我一个 字符或者一个记录一个记录地从头往后顺序读 而随机存取访问呢指的是我从文件的某个位置开始 读若干个长度的文件内容。因此 如果要提出, 提供随机访问, 就要提供 当前的一个读写的位置。要把这个 读写的指针定位在一个你需要的位置处 比如说 UNIX 就提供了一个 seek 操作, 那它就是 把指针定位在某一个位置。好, 那么文件的逻辑结构呢通常是这样两种 我们说的目录文件呢实际上就是记录式文件的一种典型代表