

9.3 操作系统内核I/O子系统

内核 I/O 子系统以系统调用方式提供若干服务，其中有 I/O 调度、缓冲、假脱机、出错处理和设备预留。

内核 I/O 子系统负责：文件和设备命名空间的管理，文件和设备访问控制，操作控制，文件系统空间的分配，设备分配，缓冲、高速缓存、假脱机，I/O 调度，设备状态监控、错误处理、失败恢复，设备驱动程序的配置和初始化等。

1. I/O 调度概念

I/O 调度是按照某种策略来为请求 I/O 的进程分配设备。通过 I/O 调度能改善系统整体性能，能在进程之间公平的共享设备，能减少因 I/O 所需要的等待时间。

操作系统通过为每个设备维护一个请求队列来实现调度。当一个应用程序执行系统调用请求 I/O 时，该请求就加到相应设备的队列上。I/O 调度重新安排队列顺序以改善系统总体效率和应用程序的平均响应时间。

2. 缓冲区与高速缓存

为了解决 CPU 与 I/O 设备间速度不匹配的矛盾，提高的 I/O 速度和设备利用率，在所有的 I/O 设备与处理机之间，操作系统都使用了缓冲区来交换数据。

缓冲技术实现基本思想如下：当一个进程执行操作输入数据时，先向系统申请一个主存区域——缓冲区，系统将一个物理记录的内容读到缓冲区域中，然后根据进程要求，把当前需要的逻辑记录从缓冲区中选出并传送给进程。当一个进程执行写操作输出数据时，先向系统申请一个主存区域——缓冲区，然后，将数据高速送到缓冲区。若为顺序写请求，则不断把数据填到缓冲区，直到它被装满为止。此后，进程可以继续它的计算，同时，系统将缓冲区内容写到 I/O 设备上。

在操作系统管理下，常常辟出许多专用主存区域的缓冲区用来服务于各种设备，支持 I/O 管理功能。常用的缓冲技术有：单缓冲、双缓冲、多缓冲。

高速缓冲（cache）是可以存储数据拷贝的高速内存。对高速缓冲中的数据访问要比普通内存数据访问更为高效。缓冲和高速缓冲的区别是缓冲只保留数据仅有的一个现存拷贝，而高速缓冲只是提供了一个存储在其他地方的数据的一个高速拷贝。

3. 设备分配

在多道程序环境下，设备必须由系统分配。每当进程向系统提出 I/O 请求时，设备分配程序按照一定的策略，把其所需的设备及其有关资源（如缓冲区、控制器和通道）分配给该进程。在分配设备时还必须考虑系统的安全性，避免发生死锁现象。

为了使系统能够安全高效地工作，系统在进行设备分配时应考虑的因素有：设备的固有属性、设备的分配算法、设备分配的安全性和设备的独立性。

（1）根据设备的固有属性而采取的策略。

在分配设备时，首先应考虑设备的属性。根据设备的固有属性采取以下三种策略：

1) 独享方式

独享方式是指将一个设备分配给某进程后，便一直由它独占，直至该进程完成或释放该设备为止，系统才能将该设备分配给其它进程使用。这种分配方式是对独占设备采用的分配策略。它不仅往往造成设备利用率低，而且还会引起系统死锁。

2) 共享方式

共享方式是指将共享设备（磁盘）同时分配给多个进程使用。但是这些进程对设备的访问需进行合理的调度。

3) 虚拟方式

虚拟方式是指通过高速的共享设备，把一台慢速的以独占方式工作的物理设备改造成若干台虚拟的同类逻辑设备，这就需要引入 SPOOLing 技术。虚拟设备属于逻辑设备。

(2) 设备分配算法

与进程的调度算法相似，通常只采用以下两种分配算法：

1) 先来先服务

当多个进程同时向某一设备提出 I/O 请求时，该算法就根据对该设备提出请求的先后次序将这些进程排列成一个设备请求队列，设备分配程序把设备首先分配给队首进程。

2) 优先级高者优先

对优先权高的进程所提出的 I/O 请求赋予高优先级，在形成设备队列时，将优先级高的进程排在设备队列前面，先得到分配。而对于优先权相同的 I/O 请求，则按先来先服务原则排队分配。

(3) 设备分配中的安全性

1) 安全分配方式

每当进程发出一个 I/O 请求后，便进入阻塞状态，直到其 I/O 操作完成时才被唤醒。当它运行时不保持任何设备资源，打破了产生死锁一个必要条件——“请求和保持”，所以这种分配方式是安全的。但是这种分配算法使得 CPU 与 I/O 设备串行工作，设备的利用率比较低。

2) 不安全分配方式

进程发出一个 I/O 请求后仍可以继续运行，需要时还可以发第二个 I/O 请求、第三个 I/O 请求。只有当进程所请求的设备已被另一个进程占用时，进程才进入阻塞状态。这种分配方式是不安全，因为它可能具备“请求和保持”条件，从而可能造成系统死锁。所以，在设备分配程序中应该增加安全性检查的功能。

(4) 设备独立性

为了提高操作系统的可适应性和可扩展性，目前几乎所有的操作系统都实现了设备的独立性（也称为设备无关性）。其基本思想是：用户程序不直接使用物理设备名（或设备的物理地址），而只能使用逻辑设备名；而系统在实际执行时，将逻辑设备名转换为某个具体的物理设备名，实施 I/O 操作。

逻辑设备是实际物理设备属性的抽象，它并不限于某个具体设备。

4. 假脱机技术（SPOOLing）

独占设备每次只能分配给一个进程使用，这种使用特性隐含着死锁的必要条件，所以在考虑独占设备的分配时，一定要结合有关防止和避免死锁的安全算法。

系统中的独占设备是有限的，往往不能满足诸多进程的要求，会引起大量进程由于等待某些独占设备而阻塞，成为系统中的“瓶颈”。另一方面，申请到独占设备的进程在其整个运行期间虽然占有设备，利用率却常常很低，设备还是经常处于空闲状态。为了解决这种矛盾，最常用的方法就是用共享设备来模拟独占设备的操作，从而提高系统效率和设备利用率。这种技术就称为虚拟设备技术，实现这一技术的软、硬件系统被称为假脱机（Simultaneous Peripheral Operation On Line, SPOOL）系统，又叫 SPOOLing 系统。

SPOOLing 系统通常分为输入 SPOOLing 和输出 SPOOLing，两者工作原理类似。下面就以常见的共享打印机为例，说明输出 SPOOLing 的基本原理。

打印机是一种典型的独占设备，引入 SPOOLing 技术后，用户的打印请求传递给

SPOOLing 系统，而不是真正把打印机分配给用户。SPOOLing 系统的输出进程在磁盘上申请一个空闲区，把需要打印的数据传送到里面，再把用户的打印请求挂到打印队列上。如果打印机空闲，就会从打印队列中取出一个请求，再从磁盘上的指定区域取出数据，执行打印操作。由于磁盘是共享的，SPOOLing 系统可以随时响应打印请求并把数据缓存起来，这样就把独占设备改造成了共享设备，从而提高了设备的利用率和系统效率。

5. 出错处理

设备和 I/O 传输的出错有多种形式，有的短暂如网络过载，有的永久如磁盘控制器坏了。操作系统可以对短暂错误进行弥补，但是如果某个重要系统组件出现了永久错误，那么操作系统就不可能从中恢复。

采用内存保护可以预防许多硬件和应用程序的错误，这样就不会因为小的机械问题导致完全的系统崩溃。