

## 操作系统作业

阅读第一章《Introduction》，总结操作系统概述。

### 1. 什么是操作系统？

在学习操作系统之前，我们要知道什么是操作系统。

计算机的结构十分复杂，让一个人去管理一台计算机的方方面面是不可能的。因此，需要一个程序把硬件和用户运行的软件管理起来。经过操作系统的管理，用户可以直接运行程序来完成想要的工作，而不必担心硬件连接、资源调配等问题。因此操作系统提供了人与计算机系统的接口。同时，因为这种管理在计算机运行过程中是无时无刻不在进行的，所以操作系统是一直运行在计算机上的程序。

### 2. 操作系统管理哪些方面？

知道了操作系统的定义，我们开始关心操作系统是如何管理计算机的。这时，我们的问题是：操作系统都管理哪些方面？操作系统是如何管理的？接下来先探讨操作系统管理哪些方面的问题。

根据冯诺依曼体系结构，计算机分为存储器，运算器，控制器以及 I/O 设备四个部分。而在一般的计算机中，CPU 提供了运算器和控制器的功能。因此，操作系统在硬件方面要控制的有 CPU，存储器，以及 I/O 设备。此外，操作系统还要对运行的程序进行管理，即进程管理。

### 3. 操作系统是如何管理内存的？

已经知道了管理什么，接下来就要知道如何管理了。本章只是概述，具体内容需要在之后的章节中详细讨论。

我们先来讨论计算机如何管理内存的问题。

内存是处理器能直接访问的惟一的大容量存储区域。计算机要处理 I/O 设备中的任何数据，都要将数据读入内存中才能开始处理。计算机执行的指令也存储在内存中。操作系统要做的是记录内存使用程度，合理分配内存空间，控制数据进出内存。

### 4. 操作系统是如何进行存储管理的？

内存的大小是有限的，而且内存具有易失性，数据无法大量长期保存。所以计算机需要外存来扩充其存储空间。此外，内存访问的速度还不够快，可以使用缓存来加快其访问速度。接下来要讨论的就是不同存储设备的管理。

数据在外存中以文件形式存在。用户需要通过操作系统来访问或修改文件，也需要对文件进行分类整理，以方便查找和再次访问。所以操作系统需要负责以下方面：创建和删除文件或目录，提供文件或目录的路径，提供文件与二级存储的映射以及备份。

文件一般存储在磁盘（二级存储）上，有些很少使用的文件存储在磁带等更慢但是容量更大的设备（三级存储）上。由用户来决定文件放在硬盘的哪些扇区对用户是极不友好的，因此操作系统要负责管理以下方面：空闲空间管理，存储空间分配和硬盘调度。有些操作系统还要负责管理二级存储到三级存储的备份。

内存的速度较硬盘已经很快了，但是对于 CPU 还是不够。寄存器的空间是十分宝贵的。这时需要一个读取速度介于寄存器与内存之间的存储器——高速缓存。对高速缓存进行合理的管理可以大大提高计算机的数据处理速度。类似地，还可以把硬盘内的数据部分复制到内存里以提高硬盘数据的访问速度。一个典型的例子就是微软的 Superfetch。

### 5. 操作系统是如何进行 I/O 设备管理的？

接下来讨论的是对 I/O 设备的管理。硬件设备的具体实现往往对用户并不友好，因此操作系统需要对用户隐藏具体硬件设备的特性。操作系统需要提供设备驱动器接口，需要为 I/O 设备提供内存管理，有时还需要给 I/O 设备提供驱动程序。

## 6. 操作系统是如何进行 CPU 管理的？

最后讨论对 CPU 的管理。

由于程序的执行在很多场合下需要等待，因此单一作业不可能保持 CPU 时时刻刻都忙。这时操作系统可以在一个作业需要等待的时候切换到另一个作业执行，使 CPU 总是有一个作业可以执行，以提高 CPU 的使用率。

这种思想的延伸是分时系统。在某一作业需要等待用户响应的时候，操作系统执行其他的作业，等用户响应完毕之后再切换回去，以充分利用 CPU。

分时系统需要在内存中载入多个程序。如果内存不够，操作系统需要决定哪些程序可以载入内存；如果多个任务需要同时执行，操作系统需要决定它们执行的先后。

CPU 的管理不仅要保证速度，还要保证准确。为了防止用户的非法操作对计算机的稳定运行造成损害，操作系统需要对用户和系统的操作进行区分。分为两种模式：用户模式和内核模式。在用户模式下，用户的进程不能进行超出权限的行为。用户模式与内核模式由模式位区分。这相当于操作系统给用户的程序提供了一个接口，用户通过这个接口来进行系统调用，间接控制计算机。当然，接口可以提供的操作是有限但安全的。

不光那些对计算机可能有有害的操作，一些会引起系统进入死循环的操作也是不正确的，操作系统应当予以禁止。为了鉴别这些操作，操作系统使用定时器对执行时间进行限制，一旦超出规定时间，则产生中断，由系统判断操作是错误还是正常的指令等待，并进行下一步处理。

## 7. 操作系统是如何进行进程管理的？

已经讨论了操作系统对硬件的管理，接下来介绍的是操作系统对进程的管理。进程的执行需要资源，进程出现了错误应当被终止。进程之间的通信需要操作系统的协作，操作系统需要对进程进行同步。操作系统需要对这些进行合理的管理。

## 8. 操作系统还有哪些需要管理的方面？

如果一个操作系统只具备上述功能还是不够的。计算机有受到攻击的风险。因此操作系统需要保护用户数据的安全，控制进程访问计算机资源，防御内部与外部的攻击。

计算机在不同的计算环境下需要有不同的功能以适应计算环境的需要。一些分布式系统是通过网络进行协同工作的。在这种分布式系统上操作系统需要合理控制，使集群的子集出错并不会影响整体的工作。

在一些嵌入式设备上，操作系统还有其他的工作需求，比如进程的实时性要求进程只能在定义的时间内执行。多媒体系统涉及多媒体数据的传输。在一些硬件性能受到限制的设备（如 PDA）上，操作系统还需要在续航和性能之间做出权衡。

## 9. 总结

本章介绍了操作系统控制计算机的概况，介绍了一些专有名词和大体的控制过程，当然具体的细节还需要在之后的章节内进行详细的介绍。