

※OS 概念：OS 是控制和管理计算机硬件和软件资源，合理组织计算机工作原理以及方便用户的功能的集合。

※OS 目标：1 有效性；2 方便性；3 可扩充性；4 开放性。

※OS 作用：1 OS 作为用户与计算机硬件系统之间的接口；2 OS 作为计算机系统资源的管理者；3 OS 实现了对计算机资源的抽象。

※OS 特征：并发性、共享性、虚拟性和异步性四个基本特征；最基本的特征是并发性。

※OS 功能：

1 处理机管理：作业和进程调度；

2 存储器管理：内存分配，地址映射，内存保护，内存扩充；

3 设备管理：缓冲区管理，设备分配，设备驱动，设备无关性；

4 文件管理：文件存储空间的管理，文件操作的一般管理，目录管理，文件的存取控制；

5 OS 与用户之间的接口：程序接口，用户接口，联机接口，脱机接口，图形接口；

※单道批处理系统的特征：自动性、顺序性、单道性。

※多道批处理系统：

优点：资源利用率高、系统吞吐量大；

缺点：平均周转时间长、无交互能力。

※实现多道批处理系统需要解决的问题：

1 处理机管理问题；

2 内存管理问题；

3 I/O 设备管理问题；

4 文件管理问题；

5 作业管理问题。

※多道批处理系统形成和发展的主要动力：

1 不断提高计算机资源的利用率；

2 方便用户；

3 器件的不断更新换代；

4 计算机体系结构的不断发展。

※分时系统概念：将一台计算机很好的提供给多个用户同时使用，提高计算机的利用率。

※实时系统概念：是计算机系统可以立即对用户程序要求或者外部信号作出反应的系统，它可以分为硬实时系统和软实时系统。

※分时系统特点：多路性、交互性、及时性、独占性、。

※实时系统特点：有严格的时间限制，每一个信息的接受、分析处理和发送的过程必须在规定的时间内完成。

※分时系统与实时系统的比较(及时性、交互性、可靠性、多路性、独立性)：

1 及时性：实时信息处理系统对实时性的要求与分时系统类似，都是以人所能接受的等待时间来确定；而实时控制系统的及时性，是以控制对象所要求的开始截止时间或完成截止时间来确定的，一般为秒级到毫秒级，甚至有的要低于 100 微妙。

2 交互性：实时信息处理系统具有交互性，但人与系统的交互仅限于访问系统中某些特定的专用服务程序。不像分时系统那样能向终端用户提供数据和资源共享等服务。

3 可靠性：分时系统也要求系统可靠，但相比之下，实时系统则要求系统具有高度的可靠性。因为任何差错都可能带来巨大的经济损失，甚至是灾难性后果，所以在实时系统中，往往都采取了多级容错措施保障系统的安全性及数据的安全性。

※进程的概念：是可并发执行的具有一定功能的程序段，在给定数据机上的一次执行过程。

※进程的引入：为了使程序在多道程序环境下能并发执行，并对并发执行的程序加以控制和描述，在操作系统中引入了进程概念。

※进程的影响：使程序的并发执行得以实行。

※进程的基本属性：

- 1 进程是一个可拥有资源的独立单位；
- 2 进程同时又是一个可独立调度和分派的基本单位。

※线程的属性：

- 1 轻型实体；
- 2 独立调度和分派的基本单位；
- 3 可并发执行；
- 4 共享进程资源。

※进程的基本状态：就绪、执行、阻塞。

※进程在三个基本状态之间转换：

- 1 绪状态→执行状态：进程分配到 CPU 资源；
- 2 执行状态→就绪状态：时间片用完；
- 3 执行状态→阻塞状态：I/O 请求；
- 4 阻塞状态→就绪状态：I/O 完成。

※进程的同步和互斥：是指进程在推进时的相互制约关系，在多道系统中，由于资源共享与进程合作，这种进程间的制约成为了可能。

※进程的同步：主要源于进程合作，是进程间共同完成一项任务时直接发生相互作用的关系。为进程之间的直接制约关系。再多道环境下，这种进程间在执行次序上的协调是必不可少的。

※进程的互斥：主要源于资源共享，市进程之间简介制约关系。再多得到系统中每次只允许一个进程访问的自愿成为临界资源，进程互斥就是保证每一次只有一个进程使用临界资源。

※线程的概念：在操作系统中引入线程，则是为了减少程序在并发执行时所付出的时空开销，使 OS 具有更好的并发性，提高 CPU 的利用率。进程是分配资源的基本单位，而线程则是系统调度的基本单位。

※PCB(Process Control Block 进程控制块的概念)：

PCB 是为了描述和控制进程的运行系统，为每个进程定义的一个数据结构，它是进程实现的一部分，也是操作系统中最重要的记录型数据结构。

※PCB 的作用，为什么 PCB 是进程存在的惟一标志：

PCB 是进程实体的一部分，是操作系统中最重要的记录型数据结构。作用是使一个在多道程序环境下不能独立运行的程序，成为一个能独立运行的基本单位，成为能与其它进程并发执行的进程。OS 是根据 PCB 对并发执行的进程进行控制和管理。

※进程和程序的比较：

- 1 动态性是进程最基本的特性，表现为由创建而产生，由调度而执行，因得不到资源而暂停执行，由撤销而消亡。进程有一定的生命期，而程序只是一组有序的指令集合，是静态实体。
- 2 并发性是进程的重要特征，同时也是 OS 的重要特征。引入进程的目的正是为了使其程序能和其它进程的程序并发执行，而程序是不能并发执行的。
- 3 独立性是指进程实体是一个能独立运行的基本单位，也是系统中独立获得资源和独立调度的基本单位。对于未建立任何进程的程序，不能作为独立单位参加运行。

※进程和线程的比较：

- 1 调度性：线程在 OS 中作为调度和分派的基本单位，进程只作为资源拥有的基本单位。
- 2 并发性：进程可以并发执行，一个进程的多个线程也可并发执行。
- 3 拥有资源：进程始终是拥有资源的基本单位，线程只拥有运行时必不可少的资源，本身基本不拥有系统资源，但可以访问隶属进程的资源。
- 4 系统开销：操作系统在创建、撤消和切换进程时付出的开销显著大于线程。

※信号量的概念：表明资源可以提供给进城使用的量，他是一个整型值。

※计数信号量 S 的物理含义：

$S > 0$: 表示有 n 个可利用的资源。

$S < 0$: 表示有 n 个被阻塞的资源。

$S = 0$: 表示资源正在使用。

※动态重定位: 是指在程序执行过程中, 每当访问指令或数据时, 要将访问的逻辑结构转换为物理地址。

※高级调度的任务: 高级调度的主要任务是根据某种算法, 把外存上处于后备队列中的那些作业调入内存。

※低级调度的任务: 低级调度是保存处理机的现场信息, 按某种算法先取进程, 再把处理器分配给进程。

※低级调度的功能: 1 保存处理机的现场信息; 2 按某种算法选取进程; 3 把处理器分配给进程。

※引入中级调度的目的: 引入中级调度的主要目的是为了提高内存利用率和系统吞吐量。使那些暂时不能运行的进程不再占用内存资源, 将它们调至外存等待, 把进程状态改为就绪驻外存状态或挂起状态。

※周期时间=完成时间-到达时间

※带权周期时间=周转时间/服务时间。

※执行进程调度: 通常是发生某个正在运行的进程或者即应运行完毕、或者隐没种原因进入了等待队列时, CPU 可以为下一个进程提供服务, 另外, 有较高优先级进程进入运行状态, 这种方式成为可剥夺方式。

※作业、作业步和作业流的概念:

1 作业包含通常的程序和数据, 还配有作业说明书。系统根据该说明书对程序的运行进行控制。批处理系统中是以作业为基本单位从外存调入内存。

2 作业步是指每个作业运行期间都必须经过若干个相对独立相互关联的顺序加工的步骤。

3 作业流是指若干个作业进入系统后依次存放在外存上形成的输入作业流; 在操作系统的控制下, 逐个作业进程处理, 于是形成了处理作业流。

※调度算法的采用:

批处理系统的调度算法: 短作业优先、优先权、高响应比优先、多级反馈队列调度算法。

分时系统的调度算法: 时间片轮转法。

※按调度方式实时调度算法分类:

可分为非抢占式和抢占式两种算法。而非抢占式算法又分为非抢占式轮转和优先调度算法; 抢占式调度算法又分为基于时钟中断的抢占式优先权和立即抢占式优先权调度算法。

※临界区是一段程序。

※对信号量的操作分为 P (减) V (加), 这些操作叫做原语。原语是不可再分的操作, 在对信号量的操作中, 与每个信号量相对应的是一个队列, 队列中存储的是排队等待使用这个资源的进程。

※引入信号量、队列、以及 P 、 V 操作的目的是为了解决进程间互斥和同步问题。

※死锁的概念: 死锁是指多个进程在运行过程中因争夺资源而造成的一种僵局, 当进程处于这种僵持状态

时，若无外力作用，它们都将无法再向前推进。

※死锁的原因：1 竞争资源；2 进程间推进顺序非法。

※死锁的必要条件：互斥条件、请求和保持条件、不剥夺条件、环路等待条件。

※死锁避免的处理方法：银行家算法。

※为什么要引入动态重定位？如何实现？

在程序执行过程中，每当访问指令或数据时，将要访问的程序或数据的逻辑地址转换成物理地址，引入了动态重定位；

具体实现方法是在系统中增加一个重定位寄存器，用来装入程序在内存中的起始地址，程序执行时，真正访问的内存地址是相对地址与重定位寄存器中的地址相加之和，从而实现动态重定位。

※抖动：是指页面在内存和外存之间频繁的调入调出引起的。

※连续分配方式：

1 单一连续分配(这是最简单的一种存储管理方式，但只能用于单用户、单任务的操作系统中)；

2 固定分区分配(划分分区的方法有：分区大小相等和不等两种)；

3 动态分区分配(分区分配算法：首次/循环首次/最佳/最坏/快捷适应算法 first/next/best/worst/quick fit)。

※分页(段)存储管理的概念：讲一个进程直接分散的装入到许多不相邻接的分区中，即离散分配的方式，如果离散分配基本单位是页(段)则称为分页(段)存储管理方式。

※引入分段存储管理的目标：

1 方便了编程；

2 实现了分段共享；

3 实现了分段保护；

4 实现了动态链接；

5 实现了动态增长。

※分段和分页存储管理的区别：

页是信息的物理单位，分页是为了实现离散分配方式，以消减内存的外部零头，提高内存利用率。段则是信息的逻辑单位，它含有一组相对完整的信息。

2 页的大小固定且由系统决定，由系统把逻辑地址划分为页号和页内地址两部分，是由机械硬件实现的，因而在系统中只能有一种大小的页面；而段的长度却不固定，决定于用户所编写的程序，通常由编译程序在对原程序进行编译时，根据信息的性质来划分。

3 分页的作业地址空间是一维的，而分段作业地址空间则是二维的。

※虚拟存储器的概念：是指具有请求调入功能和置换功能呢个从逻辑上对内存容量加以扩充的一种存储系

统。

※虚拟存储器的特征：离散性、多次性、对换性和虚拟性。最本质的是离散性，最重要的是虚拟性。

※虚拟存储器的实现方法：

1 在分页请求系统中是在分页的基础上，增加了请求调页功能和页面置换功能所形成的页式虚拟存储系统。允许只装入少数页面的程序(及数据)，便启动运行。

2 在请求分段系统中是在分段系统的基础上，增加了请求调段及分段置换功能后形成的段式虚拟存储系统。允许只装入少数段(而非所有段)的用户程序和数据，即可启动运行。

※虚拟存储请求分页系统的常用哪页面置换算法：最佳(Optimal)、先进先出(FIFO)、最近最久未使用(LRU)、Clock、最少使用(LFU)。

※I/O 控制方式的种类和应用：

1 程序 I/O 方式：早期计算机无中断机构，处理机对 I/O 设备的控制采用程序 I/O 方式或称忙等的方式。

2 中断驱动 I/O 控制方式：适用于有中断机构的计算机系统中。

3 直接存储器访问(DMA)I/O 控制方式：适用于具有 DMA 控制器的计算机系统中。

4 I/O 通道控制方式：具有通道程序的计算机系统中。

※通道又称 I/O 处理机，用于实现（CPU 和外设）之间信息的传输。

※缓冲区：缓冲区是为了协调处理机的高速度和外部设备的低速度之间的区大差距而在内存中开辟的一个区域。

※缓冲技术：缓冲技术是为了协调吞吐速度相差很大的设备之间数据传送的工作，在这两种设备之间不直接进行数据传递，而是在内存中专门开辟的一个存储区域作为中间环节，这种技术就叫做缓冲技术。

※引入缓冲的原因：

1 缓和 CPU 与 I/O 设备间速度不匹配的矛盾；

2 减少对 CPU 的中断频率，放宽对中断响应时间的限制；

3 提高 CPU 与 I/O 设备之间的并行性。

※设备驱动程序的概念：又称为设备处理程序，是 I/O 进程与设备控制器之间的通信程序，因为它总是以进程的形式存在，所以也称为设备驱动进程。

※为何要引入设备独立性？

现代操作系统为了提高系统的可适应性和可扩展性，都实现了设备独立性或设备无关性。基本含义是应用程序独立于具体使用的物理设备，应用程序以逻辑设备名请求使用某类设备。

优点：1 设备分配时的灵活性；2 易于实现 I/O 重定向。

※如何实现设备独立性？

为了实现设备的独立性，应引入逻辑设备和物理设备概念。在应用程序中，使用逻辑设备名请求使用某类设备；系统执行时是使用物理设备名。鉴于驱动程序是与硬件或设备紧密相关的软件，必须在驱动程序之上设置一层设备独立性软件，执行所有设备的公有操作、完成逻辑设备名到物理设备名的转换(为此应设置一张逻辑设备表)并向用户层软件提供统一接口，从而实现设备的独立性。

※设备虚拟：设备虚拟是指把独占设备经过某种技术处理改造成虚拟设备。

※实现设备虚拟所需技术：

可虚拟设备是指一台物理设备在采用虚拟技术后，可变成多台逻辑上的虚拟设备，则可虚拟设备是可共享的设备，将它同时分配给多个进程使用，并对这些访问该物理设备的先后次序进行控制。

※SPOOLing 系统的组成：

SPOOLing 系统由输入井和输出井、输入缓冲区和输出缓冲区、输入进程 SP_i 和输出进程 SP_o 三部分组成。

※SPOOLing 技术的优点：

- 1 提高了 I/O 的速度；
- 2 将独占设备改造为共享设备；
- 3 实现了虚拟设备功能。

※文件：是具有文件名的一组相关信息的集合。

※文件系统：又被管理的文件，操作系统中管理文件的软件和相应的数据结构组成的一个系统。

※逻辑文件：是物理文件中存储的数据的一种视图方式，不包含具体数据，仅包含物理文件中数据的索引。

※物理文件：又称文件存储结构，是指文件在外存上的存储组织形式。

※文件结构：1 逻辑结构(记录式、流式)；2 物理结构(顺序、索引、链接)。

※目录的概念：OS 要求对文件能够实现按名存取，这就需把文件名到文件的物理地址的映射关系存在于文件目录中。为此，系统为每一个文件设置了一个文件控制块。文件目录就是这些控制块的有效集合。

※目录的要求：实现按名存取、提高检索目录的速度、文件共享、允许文件重名。

※目录结构：单级目录、两级目录和多级目录结构。

※磁盘访问时间：

磁盘访问时间由寻道时间 T_s 、旋转延迟时间 T_r 、传输时间 T_t 三部分组成。

1 T_s 是启动磁臂时间 s 与磁头移动 n 条磁道的时间和，即 $T_s = m \times n + s$ 。

2 T_r 是指定扇区移动到磁头下面所经历的时间。硬盘 15000r/min 时 T_r 为 2ms；软盘 300 或 600r/min 时 T_r 为 50~100ms。

3 T_t 是指数据从磁盘读出或向磁盘写入经历的时间。 T_t 的大小与每次读/写的字节数 b 和旋转速度有关： $T_t = b/rN$ 。

※磁盘调度算法的种类和问题：

- 1 先来先服务算法优先考虑进程请求访问磁盘的先后次序；
- 2 最短寻道时间优先算法优先考虑要求访问的磁道与当前磁头所在磁道距离是否最近；
- 3 扫描算法考虑欲访问的磁道与当前磁道间的距离，更优先考虑磁头当前的移动方向。

※对空闲磁盘空间的管理分配方式：空闲表法、空闲链表法、位示图法、成组链接法。UNIX 系统采用的是成组链接法。