

第一章、操作系统引论

1. 操作系统是什么？

1.1 定义

操作系统是配置在计算机上的第一层软件，是对硬件系统的首次扩充；其主要作用是管理好这些设备，提高他们的利用率和系统的吞吐量，并为用户和应用程序提供一个简单的接口。

1.2 操作系统的目标

1. 方便性：配置了操作系统更加方便使用
2. 有效性：改善资源利用率，提高系统吞吐量
3. 可扩充性：现在的微内核结构有良好的可扩充性
4. 开放性：OS遵循世界标准范围，可兼容

1.3 作用

1. OS作为用户与计算机硬件系统之间的接口（API / CUI / GUI）

注意：OS处于用户与计算机硬件系统之间，用户可以通过OS来使用计算机

2. OS作为计算机系统资源的管理者（处理机、存储器、I/O设备、文件（数据和程序））

注意：处理机管理是用于分配和控制处理机

存储器管理是负责内存的分配与回收

I/O设备管理是负责I/O设备的分配（回收）与操纵

文件管理是用于实现文件的存取、共享和保护

3. OS实现对于计算机资源的抽象（OS是扩充机/虚拟机）

2. 操作系统的发展过程

1. 未配置操作系统的计算机系统

1. 手工操作方式：用户独占全机（资源利用率低）、CPU等待人工操作)
2. 脱机输入/输出（Off-Line I/O）方式（磁带）
减少了CPU空闲时间、提高了I/O速度

2. 单道批处理系统

1. 解决问题：

单道批处理系统是在**解决人机矛盾和CPU与I/O设备速度不匹配矛盾的过程中形成的**。批处理系统旨在提高系统资源的利用率和系统的吞吐量。（但单道批处理系统仍不能充分利用资源，故现在已很少用）。

单道批处理分为：联机批处理、脱机批处理

联机批处理：CPU直接控制作业输入输出

脱机批处理：由外围机控制作业输入输出

2. 缺点：

主要缺点是系统资源得不到充分的利用。（CPU和I/O不能同时使用）

3. 特征：

//自动性。正常情况下，一批作业能自动依次运行，无需人工干预。 //顺序性。各道作业是顺序进入内存，FCFS //单道性。内存中任何时候只能有一道作业。
--

3. 多道批处理系统

目的：进一步提高资源利用率和系统吞吐量。

基本过程：该系统中，用户提交的作业都先存在外存中，在作业A在执行I/O请求时，CPU空闲，此时调用作业B，防止CPU空闲。同理按一定的算法调用作业，防止CPU空闲

PS:推动多道批处理系统形成和发展的动力是提高资源利用率和系统吞吐量。

1. 优点:

资源利用率高（资源指CPU、存储器、设备、数据）
系统吞吐量大（系统吞吐量指系统在单位时间完成的总工作量（作业数量））

2. 缺点:

平均周转时间长（作业的周转时间是指从作业进入系统开始，直至其完成并退出系统为止所经历的时间。）
无交互能力

3. 特点:

多道：内存同时存在多个作业
宏观上并行：同时有多道程序在内存运行，某一时间段上，各道程序不同程度地向前推进。
微观上串行：任一时刻最多只有一道作业占用CPU，多道程序交替使用CPU

4. 存在问题:

处理机管理问题
内存管理问题（内存分配和保护）
I/O设备管理问题
文件管理问题（文件的组织和管理）
作业管理问题
用户与系统的接口（OS还应提供用户与OS的接口，方便用户操作）

4.分时系统（解决人机交互问题）

PS:推动分时系统形成和发展的主要动力是用户的需要：交互、共享主机

1. 为实现人机交互（多个用户从键盘输入命令，系统能全部及时接收并处理）的关键问题:

及时接收（多路卡和缓冲区）：作业直接进入内存、不允许一个作业长期占用处理机
即时处理（划分时间片）：多个用户分时使用主机，每一用户分得一个时间片，用完这个时间片后操作系统将处理机分给另一用户，如此循环，每一用户可以周期性地获得CPU使用权，这样每一用户都有一种独占CPU的感觉。

2. 特征:

多路性：允许同一主机联接多台终端。
独立性：每一用户独占一个终端；每个用户感觉不到其他用户的存在。
及时性：用户请求能及时响应。
交互性： 可进行广泛的人机对话。

5. 实时操作系统

实时系统（Real - Time System）是指系统能及时响应外部事件的请求，在规定的时间内完成对该事件的处理，并控制所有实时任务协调一致地运行。

1. 特征：

<p>多路性</p> <p>对实时处理系统，系统按分时原则为多个终端服务；</p> <p>对实时控制系统，系统经常对多路现场信息进行采集；以及对多个对象或多个执行机构进行控制。</p>
<p>独立性</p> <p>实时处理系统，每个终端用户向实时系统提出服务请求时，彼此独立；</p> <p>实时控制系统，对信息的采集和对对象的控制彼此不干扰。</p>
<p>及时性（开始截止时间/完成截止时间）</p> <p>交互性（仅限于访问专用服务程序）</p> <p>可靠性（多级容错措施保障系统和数据安全）</p>

2. 硬实时任务和软实时任务

6. 其他操作系统

微机操作系统、多处理机操作系统、网络操作系统、分布式操作系统、嵌入式操作系统

3. 操作系统的基本特征

并发、共享、虚拟、异步

1. 并发

并发是指在内存中放多道作业，在一个时间段上来看，每一道作业都能不同程度地向前推进。但在任何一个时间点上只能有一道占用CPU。

串行：在内存中每次只能放一道作业，只有它完全执行完后别的作业才能进入内存执行。

并行：存在于有多个CPU的环境中，在内存中放多道作业，在任一时间点上都可能有多道作业在不同的CPU上同时执行。

2. 共享

系统中资源可供多个并发的进程使用

根据资源属性的不同，有两种资源共享方式：

互斥共享方式（临界/独占资源）

同时访问方式（分时共享）（时间片）

PS：并发和共享是OS的两个最基本的特性，二者互为条件！

3. 虚拟

通过某种技术将一个物理实体映射为若干个逻辑上对应物（如CPU；一个屏幕可看成多个屏幕-窗口）。或将多个物理实体映射为一个逻辑实体（如虚拟存储是内存和外存的虚拟）。

可提高资源利用率。

4. 异步

不确定性：什么时候开始，执行多久时间，终止时间。

走走停停，交替进行，结果相同。

4. 操作系统的主要功能

处理机管理功能：进程控制、进程同步、进程通信、调度（作业调度、进程调度）

存储器管理功能：内存分配、内存保护、存储扩充、地址映射

设备管理功能：缓冲管理、设备分配、设备处理

文件管理功能：文件存储空间的管理、目录管理、文件的读/写管理和保护

操作系统与用户之间的接口：命令接口、程序接口、图形接口

现代OS的新功能：

系统安全：为保障数据的保密性、完整性和可用性，常使用的技术：认证技术、密码技术、访问控制技术、反病毒技术

网络功能和服务：网络通信、资源管理、应用互操作

支持多媒体

5. OS结构设计

1. 传统的体系结构

无结构OS

模块化OS结构

分层式OS结构

2. 模块化结构OS

1. 模块接口法

2. 模块独立性

内聚性：各个模块直接联系的紧密程度，越高模块越独立

耦合性：模块间相互影响程度，越低，越独立。

3. 优点

提高OS正确性、可理解性、可阅读性

提高OS的可适应性

加速OS的开发过程

4. 缺点

难以满足实际接口

无序性

3. 分层式OS

原则：每一步建立在可靠的基础上

优点：易保证系统的正确性、易扩充和以维护

缺点：系统效率低

4. 微内核OS

正确性、灵活性、易维护性、可扩充性

1. 足够小的内核

微内核并非一个完整的OS，将操作系统中最基本的部分放入微内核：

1. 与硬件紧密相关的部分
2. 一些较基本的功能
3. 客户与服务端的通信

2. 基于C/S模式

上下文切换

3. 基本功能

1. 线程（进程）管理
2. 低级存储器管理
3. 中断和陷入处理（陷入只能在用户态使用）

4. 优点

可扩展性、可靠性、可移植性、提供了对于分布式系统的支持（C/S）、面向对象技术

5. 存在问题

运行效率有所降低

习题

8、10、11、13、14、15、16、17、19、20、21、23、25