**操作系统学习心得体会**

**王珺熠 4042017030**

1. **学到了什么？**

经过近一个学期的学习，我们分从概述，进程管理，内存管理，文件管理，IO管理几个模块对操作系统进行了由浅入深的学习，伴以实验形成了一定的理解，本篇总结将从整体角度与几个模块的Q&A来概述自己的学习体会。

经过学习，直观感觉上，计算机的操作系统是在为应用级编程提供服务，即提供象诸如如何内存寻址，如何管理虚拟存储器、如何进行中断，如何管理磁盘，如何。。。。等等。操作系统为我们做掉了系统级编程中最繁重的一块，所以我们在用应用级编程比如用C++编程的时候才会那么舒服，用int分配一个变量这内存地址就来了，某个函数要调用另一个函数执行的时候这执行行程就能跳到那个函数，内存不够的情况下操作系统自动启动虚拟存储器（只不过慢了些），所以应用级编程是最舒服的，操作系统级编程比较难。

经过一个学期的学习，尤其是同时在学习计算机组成原理，感觉二者的知识上很多融会贯通的地方，相辅相成，考研辅导书上也经常会出现“详见操作系统第XX章”的字眼。。。。。。

**Q:为什么要有操作系统？**

A:操作系统在计算机中起着重要作用，它为所有的应用程序提供一个运行环境，并将应用程序同具体硬件隔离。比如在机器中换一块声卡，只需要重新安装一下声卡驱动程序就完事了，但如果没有操作系统，你就必须为这块新的声卡重新购买所有的应用程序，以便这些应用程序能够认识这块声卡并能够使用它。操作系统就象计算机的大管家，管理着计算机的各种资源，比如内存、磁盘、CPU等。应用程序想使用这些资源，都必须经过操作系统同意（资源申请），并且由操作系统统一安排使用时间（资源分配），应用程序用完后必须将资源还给操作系统（资源回收），以便其他应用程序使用。就这样，计算机系统在操作系统的管理下以最可能高的效率有条不紊地工作着。

**Q:什么叫进程？进程同程序有什么区别？**

A:进程是程序在计算机上的一次执行活动。当你运行一个程序，你就启动了一个进程。显然，程序是死的（静态的），进程是活的（动态的）。

进程可以分为系统进程和用户进程。凡是用于完成操作系统的各种功能的进程就是系统进程，它们就是处于运行状态下的操作系统本身；所有由我们启动的进程都是用户进程。进程是操作系统进行资源分配的单位。在Windows下，进程又被细化为线程，也就是一个进程下有多个能独立运行的

更小的单位。

**Q:什么是多任务？**

A:在同一个时间里，同一个计算机系统中如果允许两个或两个以上的进程处于运行状态，这便是多任务。现代的操作系统几乎都是多任务操作系统，能够同时管理多个进程的运行。

**Q:什么是“并发”？什么是“并行”？**

A:俗话说，一心不能二用，这对计算机也一样，原则上一个CPU只能分配给一个进程，以便运行这个进程。我们通常使用的计算机中只有一个CPU，也就是说只有一颗心，要让它一心多用，同时运行多个进程，就必须使用并发技术。实现并发技术相当复杂，当只有一个CPU时需要通过调度算法来实现，如果一台计算机有多个CPU，情况就不同了，如果进程数小于CPU数，则不同的进程可以分配给不同的CPU来运行，这样，多个进程就是真正同时运行的，这便是并行。但如果进程数大于CPU数，则仍然需 要使用并发技术。在Windows中，进行CPU分配是以线程为单位的，一个进程可能 由多个线程组成，这时情况更加复杂，但简单地说，有如下关系：

总线程数 <= CPU数量：并行运行

总线程数 > CPU数量：并发运行

并行运行的效率显然高于并发运行，所以在多CPU的计算机中，多任务的效率比较高。但是，如果在多CPU计算机中只运行一个进程（线程），就不能发挥多 CPU的优势。

至于内存、文件和IO，感觉复习计组的时候总有一种似曾相识的“跳戏”的感觉，事实上，也是。。。。感觉三者好多原理是相通的，就不做QA了。

1. **建议**

不得不吐槽一下我们发的教材，看起来确实很唬人。。。大部头，印刷精良，名家大作，然而，前几次课的时候翻书快翻到崩溃，里面内容大而全而散，经常是记着记着笔记就找不到讲的内容对应第几页了。。。强烈建议以后要是可以的话，换一下教材吧，考研复习指导讲的就蛮清楚的。

还有就是在线课堂，这个确实挺好的，师生的交互性很强，而且以后看PPT也很方便，要是以后苞米形势不那么严峻了，最好还是能用回来。

1. 操作系统与大数据的联系

我个人感兴趣的方向是网络舆情分析，以谷歌公司的GDELT为例，GDELT每时每刻监控着每个国家的几乎每个角落的100多种语言的新闻媒体—印刷的、广播的和web形式的，识别人员、位置、组织、数量、主题、数据源、情绪、报价、图片和每秒都在推动全球社会的事件，数据每条记录有58个字段，字段间以’\t’分割，每个字段代表着不同的含义，有谷歌自动提取的时间有关的信息，事件有关对象、事件类型等。

之前的实验中，我们曾经基Python编程实现了gdelt数据的定时提取更新、整理与简单的分析功能，但是由于没有在编程中贯彻多线程调度的思想，当数据量大的时候，运算缓慢。

在我看来，舆情分析的一个重要基础就是海量的数据，只有数据样本足够的多才能够实现之后的学习、整理、分析判别，而高效的处理大量数据的分析，这就与操作系统相关知识密不可分，其中涉及到的进程、内存管理等，便是操作系统课程所讲授的原理。

关于GDELT相关的代码和演示PPT已上传至GITHUB：https://github.com/JoeyWANG-IEU/OS-projects