

5 CPU Scheduling

CPU-I/O Burst Cycle

进程的执行包含一个 CPU 执行和 I/O 等待的循环，也就是 CPU burst followed by I/O burst；我们主要进行调度分配的是 CPU burst，也就是这段 CPU 执行应该分配给哪个进程。其中，**CPU-Burst** 由大量的 short burst 和少量的 longer burst 构成。

CPU调度器 (CPU Scheduler) 从准备队列中选择进程，并且向其分配 CPU

- **non-preemptive** (非抢占式)：按照进程的顺序执行进程，有次序的调度方式，需要当前进程执行完任务后在执行下一个进程；
- **preemptive** (抢占式)：进程可以“插队”优先执行。

切换操作 (Dispatch)：从一个进程切换到另一个进程，会有切换延时（基本就是上下文切换的开销，需要保存原有的进程 PCB，导入新的进程的 PCB）。包含：

- 上下文切换；
- 内核态切换到用户态；
- 跳转到下一个进程的指定位置，继续执行进程。

调度效率衡量指标：一般来说用后三个指标评价效率。

- **CPU 利用率** (CPU utilization)；CPU 利用效率；容易受硬件影响；
- **吞吐量** (Throughput)：每个时间单元完成的进程执行的数量；容易受硬件影响；
- **周转时间** (Turnaround time)：从一个进程创建到终止的总时间，执行一个进程的总时间。
- **等待时间** (Waiting time)：进程等待的时间（所有等待时间之和）。
- **响应时间** (Response time)：任务从提交到**第一次执行**的总等待时间。

FCFS / FIFO (先来先服务 / 先进先出调度算法)：优先调度先到来的算法，*非抢占式*。

SJF (shortest-job-first) (最短任务优先算法)：优先调度任务时间最短的算法，*非抢占式*，**最小平均等待时间**。

SRTF (shortest-remaining-time-first) (最短剩余时间优先算法)：优先调度剩余时间最短的进程，*抢占式*。

RR (Round-Robin) (轮转 / 轮询算法)：（对应于操作系统的分时系统）每个进程能在 CPU 能获得一个短的时间片 q （一般 q 在 10ms 到 100ms 之间），时间片执行完成后，则将当前进程保存后放到准备队列的末尾，等待下一次执行。

- 假设有 n 个进程在准备队列中，时间片长度为 q ，那么每个进程的最长等待时间为 $(n - 1)q$ ；
- 如果时间片 q 设置过大，那么转化为 FCFS / FIFO 算法；
- 如果时间片 q 设置过小，那么花费在上下文切换的时间多；
- 时间片 q 设置应该至少大于 80% 的任务执行时间。

Priority Scheduling (优先级调度算法)：给每一个进程分配一个**优先级号** (priority number)（一般为整数，优先级号越小优先级越高 (windows 反之)），优先级高的先执行，默认为抢占式。

- 问题：饿死现象 (starvation)，低优先级的进程可能永远执行不了。
- 解决方式：资历法 (aging)，随着进程的等待时间增加增加其优先级。
- 可能采用周期性优先级调度（每个周期优先级可能不同）；
- 可能赋优先级值为执行时间的倒数。

多级队列：由于优先级数量不多，把优先级一样的放入同一个队列中进行调度（可以将优先级调度算法与轮转/轮询算法结合，相同优先级按照轮转算法进行轮转即可）。

如何划分优先级？（以下从高优先级到低优先级）

- 实时系统的进程 (real-time process): 所有处理必须做实时响应;
- 系统进程 (system process): 内核进程;
- 交互进程 (interactive process);
- 批处理进程 (batch process)。

多种调度算法组合：如轮询算法+优先级调度算法结合，先根据优先级调度，相同优先级再轮询调度。

线程调度：分成两个部分，内核态的系统范围调度 (*System Contention Scope, SCS*) 和用户态的进程范围调度 (*Process Contention Scope, PCS*)。

多核 / 多处理器调度：对于多个 CPU 核心或处理器的调度方法。略。

负载均衡技术：将任务平均分配在每一个核上；

- **在线迁移：**将一个核的任务迁移到另一个空闲的核执行；
- **推迁移** (push migration): 将自己核的任务推给其他空闲的核执行；
- **拉迁移** (pull migration): 当前核空闲，可以接收其他核推来的任务。

实时CPU调度：分软实时调度和硬实时调度。

EDF (Earliest Deadline First): 截止时间靠前的进程优先级高。