

Unix 系统中进程调度算法的分析与评价

刘晓娇

(石嘴山市第三中学 宁夏 石嘴山 753000)

[摘要] 本文在阐述 Unix 系统中先来先服务算法、最高优先权优先调度算法、时间片轮转调度算法、短进程优先调度算法基本思想的基础上, 对其进行了详细的分析和评价。

[关键词] 进程调度; 算法; 分析; 评价

In Unix System Process Dispatching Algorithm Analysis and Appraisal

LIU Xiao-jiao

(Shizuishan Third Middle School, Shizuishan Ningxia, 753000)

[Abstract] This article based on elaborating the Unix system first come first serve algorithm, the highest priority scheduling algorithm, rotation scheduling algorithm for time slice, short process priority scheduling algorithm, which has carried on the detailed analysis and the appraisal to it.

[Key words] Process dispatching; Algorithm; Analysis; Appraisal

0 引言

Unix 系统中的调度有高级调度、中级调度和低级调度。高级调度属于作业调度, 中级调度是从系统内存中将暂时中止运行的作业调度到外存, 进程调度是 Unix 系统内部的低级调度。在 Unix 系统中进程调度的频率是最高的, 远远高于高级调度和中级调度, 因此, 进程调度算法对整个系统的吞吐量有直接的影响。

Unix 系统中的进程调度的算法有先来先服务算法、时间片轮转算法、最高优先权优先调度算法、最短进程优先调度算法等。衡量进程调度性能通常需要从定性和定量两个方面来综合考虑。

1 进程调度算法评价依据

进程调度算法性能的衡量方法可以分为定性和定量两种, 在定性衡量方面, 首先是调度算法的安全性。比如, 一次进程调度是否可能引起数据结构的破坏等。这要求对调度时机的选择和保存 CPU 现场十分小心。另外, 系统开销也是衡量进程调度的一个重要指标, 进程调度程序的执行涉及到多个进程的上下文切换, 如果调度策略过度繁琐和复杂, 将会耗去较大的系统开销。这在用户进程调度系统调用较多的情况下, 将会造成响应时间大幅度增加。

进程调度的定量评价包括 CPU 的利用率评价、进程在就绪队列中的等待时间与执行时间之比等。实际上, 由于进程进入就绪队列的随机模型很难确定, 而且进程上下文切换等也将影响进程的执行效率, 从而对进程调度进行解析是很困难的, 一般情况下, 大多利用模拟或测试系统响应时间的办法来评价进程调度的性能。

2 Unix 系统中四种常用进程调度算法的分析与评价

2.1 先来先服务算法

2.1.1 算法思想

该算法思想是按照进入就绪队列的先后次序来分配处理机。FCFS 采用非剥夺调度方式, 即一旦某个进程拥有处理机, 就一直运行下去, 直到该进程完成其操作或因等待某一事件而不能继续执行时才释放处理机。

2.1.2 算法实现原理如图 1 所示。

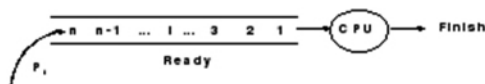


图1

说明: Ready 表示就绪队列, P_i 表示新进入队列的进程, Finish 表示进程运行完毕退出。

2.1.3 算法分析与评价

- 1) 该算法原理简单, 易于实现。
- 2) 各进程平等竞争。
- 3) 由于各进程执行的时间不一样, 从而导致相对不公平现象的产生。
- 4) 该算法有利于长进程, 不利于短进程。

5) 该算法很少用来作为主调度策略, 常常用作辅助调度算法使用。

2.2 最高优先权优先调度算法

2.2.1 算法思想

该算法的基本思想是进程优先权高者优先调度, 是一种最常用的进程调度算法。该算法的关键是如何确定优先数。

通常确定优先数的方法有两种, 即静态法和动态法。

静态优先权是在创建进程时确定的, 其运行特征是优先数确定之后在整个运行期间不再改变。确定静态优先权的依据是进程的类型、进程所使用的资源、进程的估计运行时间等因素。进程所申请的资源越多, 估计的运行时间越长, 进程的优先权越低。进程类型不同, 优先权也不同, 如系统进程的优先权高于用户进程的优先权, I/O 型进程高于 CPU 型进程。

动态优先权是指在创建进程时, 其运行特征是根据系统资源的使用情况和进程的当前状态确定一个优先权, 在进程运行过程中再根据情况的变化调整优先权。动态优先权一般根据进程占有 CPU 时间的长短、进程等待 CPU 时间的长短等因素确定。占有处理机的时间越长, 则优先权越低, 等待时间越长, 则优先权越高。

2.2.2 算法分析与评价

1) 静态优先权调度算法实现较为简单, 但不能反映系统以及进程在运行过程中发生的各种变化。而动态优先权法可以满足这个方面的需要。

2) 动态优先权调度算法的性能介于时间片轮转算法和先来先服务算法之间。

2.3 时间片轮转调度算法

2.3.1 算法思想

该算法的基本思想是使每个进程在就绪队列中的等待时间与享受服务的时间成比例。即将 CPU 的处理时间分成固定大小的时间片, 如果执行进程把分给它的时间片用完了, 但该进程的执行还没有完成, 则它也只能停止下来, 释放它所占的 CPU 资源, 然后排在相应的就绪队列的后面去, 等待进程调度程序的下一次调度, 也就是等待下一个时间片轮到它。

2.3.2 算法实现原理图

该算法实现原理图如下图 2 所示。

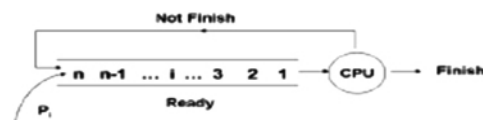


图2

说明: Ready 表示就绪队列, P_i 表示新进入队列的进程, Finish 表示进程运行完毕退出。NotFinish 表示分配给某进程的时间片已用完但任务还没有完成, 从而插入到 Ready 队列尾部。

2.3.3 算法分析与评价

1) 时间片的大小选择难以把握

时间片的大小选择比较困难是因为时间片长度的选择直接关系

到系统开销和进程的响应时间。如果时间片长度过短→导致调度程序剥夺处理器的次数增加→进程的上下文切换的次数增加→系统的开销也加重。如果时间片长度过长,长到能使就绪队列中所需要执行时间最长的进程执行完毕→轮转法就变成了 FCFS 算法→FCFS 短处不足就显示出来了。又因为 CPU 的整个执行时间=各进程执行时间之和+系统开销(各进程切换所花费的 CPU 时间之和,假定存储开销忽略不计)。因此,时间片的长短通常由以下因素确定:

- (1)系统的响应时间。
- (2)就绪队列中的进程数。
- (3)进程的切换时间。
- (4)计算机的处理能力,计算机的速度越高,时间片就可越短。
- 2)时间片长度选择的动态性。

以上仅仅作了静态分析,通常情况下,就绪队列里的进程个数是不断变化的。因此,每一次的调度都需要计算新一轮的时间片长度,不能用固定的时间片长度来进行所有进程的轮转执行。

3)该算法的扩充—多级反馈轮转法

在上面的算法中,未对就绪队列中的进程加以条件分析(即进入就绪队列的因素),由于进入就绪队列的原因不一样,要求占用处理机的紧急程度也不一样。进入就绪队列的进程来自:

- (1)分配给该进程的时间片用完,但进程还未完成。

(2)分给其时间片未用完而发生 I/O 事件等请求后再由阻塞态转变成就绪态。

- (3)新进入的进程。

因此,根据进入就绪队列的进程来源不同,建立多个就绪队列,同时赋予不同的优先级,优先权高的就绪队列优先执行,同一就绪队列中,优先级相同,按照先来先服务进行调度,运行一个给定的时间片,如果没有执行完成则转入到相应的就绪队列中去(运行一次,优先级降低一个等级,等待一个时间片,则优先级升高一个等级)。其实现原理图如图 3 所示。

2.4 短进程优先调度算法

2.4.1 算法思想

该算法的基本思想是从就绪队列(内存)中选择一个估计运行时间最短的进程,将处理机分配给它。

2.4.2 算法分析与评价

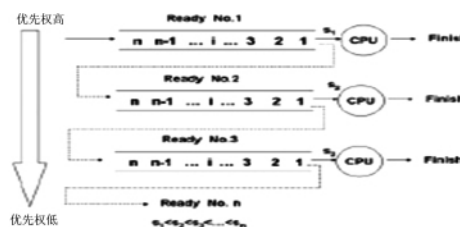


图 3

- 1)该算法能有效降低作业的平均等待时间,提高系统吞吐量。
- 2)对长进程不利,甚至导致长进程长期得不到处理。
- 3)该算法完全未考虑进程的紧迫程度。
- 4)进程的长短通常由某种策略估计提供,不能做到真正的短进程优先。

3 结语

综上所述,本文从算法思想、算法的实现原理、算法的优缺点等几个方面对先来先服务算法、时间片轮转算法、最高优先权优先调度算法、短进程优先调度算法等四种进程调度算法进行详细的论述。

【参考文献】

- [1]张尧学,史美林.计算机操作系统教程[M].北京:清华大学出版社,2001.
- [2]汤子赢,等.操作系统教程.西安:西安电子科技大学出版社,2000.
- [3]周长久,等.计算机操作系统教程.北京:高等教育出版社,2001.
- [4]徐同甲,等.计算机操作系统教程.西安:西安电子科技大学出版社,2004.
- [5][美]Robin Rark.UNIX 技术大全.前导工作室,译.北京:机械工业出版社,1998.
- [6][美]Dave Taylo.UNIX 教程.第三版.李际,译.北京:人民邮电出版社,2002.

作者简介:刘晓娟(1982.6—),女,同济大学软件学院软件工程专业,现工作单位为宁夏石嘴山市第三中学,研究方向为电子商务、软件工程。

[责任编辑:王静]

(上接第 93 页)马克思主义从人类发展的角度来认识人性问题,突破了以往人性论者的思维局限。人在认识世界和改造世界的同时,也在发展社会关系并创造社会。从某种意义上说,人的生成与发展实际上是诸社会关系的生成与发展,人类个体的一切都是以这样或那样的方式与他人进行相互作用。从这个角度出发,马克思主义认为,“人的本质并不是单个人所固有的抽象物,在其现实性上,它是一切社会关系的总和。”^[10]实际上,人总是在一定的社会关系中生存和发展,在人身上本来就凝聚着、表现着一切社会关系,脱离社会关系去考察人,只能陷入对人的抽象研究。而人在社会生产中结成的生产关系和社会关系既规定了他在社会坐标系中的位置,也决定了他在生产劳动中的实际发展程度。换句话说,社会关系是人的限定条件。这里的社会关系体现了人的社会属性。

3.3 主体属性——构建完整人性的核心 与众多人性论者不同,马克思主义人性论的科学性就表现在将人性建立在社会实践活动基础上,突出人在实践活动中的主体性。人与动物的根本分野是人在社会实践中表现出来的主体性或能动性,即“自由自觉的实践”。在这样的活动中,人把自己的生命活动本身变成自己的意志和意识的对象,开展有意识的阶级斗争、生产劳动和科学实验,以此来改造自然、改造社会、改造自身,并最终实现自身发展的自由。本质上说,人的主体性活动过程是人的生命的表现和证实及人的天然禀赋和精神目的的实现过程,也即人的本质力量的对象化过程。人不仅是一个现实存在的客体,也是一个客观能动的主体,人的主体性地位决定了人既是认识和改造世界的主体,也是认识和发展自身的主体。

4 一个结论性评价

纵观现有的研究文献,唯有马克思主义人性理论从人类历史发展

的角度考察人性问题,以辩证唯物主义与历史唯物主义作为分析问题的基本框架,将人性界定为自然性、社会性和主体性的有机统一,科学地阐明了人性的基础、本质和核心,为思想政治教育指明了方向。同时,已有的中西方人性理论尽管不够科学,不够完整,但不乏睿智,我们应吸取其中的合理成分,以便为思想政治教育服务。综合中西方人性理论,有以下几点值得深思:第一,人的需求是分析人性问题的出发点;第二,人性受到多种因素的影响和制约;第三,教育、环境、努力对人的存在和发展具有重要作用。思想政治教育是一项社会化、系统化工程,需调动各方面的积极因素,全面理解和把握马克思主义人性理论,主动借鉴其它人性理论的合理成分,不断增强思想政治教育的实际效果,这是当前教育工作者急需认真思考的重要问题。

【参考文献】

- [1]亚当·斯密.国民财富的性质和原因的研究.上[M].郭大力,王亚南,译.北京:商务印书馆,1972 版:27.
- [2]郭咸纲.西方管理学说史[M].北京:中国经济出版社,2003 版:6.
- [3]李维.认知心理学研究[M].杭州:浙江人民出版社,1998 版:1-145.
- [4]沈善洪,王凤贤.中国伦理思想史.北京:人民出版社,2005 年版:175.
- [5]马克思,恩格斯.马克思恩格斯全集.第 42 卷[M].北京:人民出版社,1979 版:167.
- [6]马克思,恩格斯.马克思恩格斯选集.第 1 卷[M].北京:人民出版社,1972 版:18.

作者简介:陈访贤,男,硕士,主要从事大学生思想政治教育研究。
祖克拉·塔来提,女,学士,从事大学生思想政治教育研究。

[责任编辑:翟成梁]