

操作系统电源管理研究进展

赵霞^{1,2} 陈向群¹ 郭耀¹ 杨芙清¹

¹(北京大学信息科学技术学院软件研究所高可信软件技术教育部重点实验室 北京 100871)

²(北京工商大学计算机学院 北京 100037)

(zhaox@th. btbu. edu. cn)

A Survey on Operating System Power Management

Zhao Xia^{1,2}, Chen Xiangqun¹, Guo Yao¹, and Yang Fuqing¹

¹(Key Laboratory of High Confidence Software Technologies (Ministry of Education), Institute of Software, School of Electronics Engineering and Computer Science, Peking University, Beijing 100871)

²(College of Computer Science and Technology, Beijing Technology and Business University, Beijing 100037)

Abstract With the rapid development of micro-electronics and mobile computing technology, reducing power dissipation of computing systems has become a hot topic for both academia and industry. As the resource manager of systems, operating system power management makes decision based on the characteristics of workload and hardware components, in order to reduce energy consumption while satisfying the performance constraints of applications simultaneously. Because of the diversity of applications and complexity of multi-task operating system concurrency, OS dynamic power management strategy is facing workload uncertainty. One of the key issues becomes the trade-off between performance and energy consumption. In this paper, the OS dynamic power management techniques are surveyed from the directions of optimal control strategy and operating system design. From the perspective of optimal control strategy, an OS power management sub-system model and two types of dynamic power management strategies, namely dynamic power management and dynamic voltage scaling are summarized in detail. From the perspective of operating system design, how to abstract power resources and how to design power management mechanisms and strategies in existing OS resources management framework are also important issues. Some key ideas and recent progresses on these issues are discussed. Finally, several challenges and open issues in OS power management are summarized.

Key words operating system; power management; dynamic power management; dynamic voltage scaling; low energy consumption

摘要 系统级动态电源管理技术的关键问题是在系统性能与能耗之间进行权衡。在硬件提供支持的基础上,操作系统电源管理的优势在于操作系统作为系统资源管理者,根据工作负载和硬件资源电源状态的特征,在满足系统性能约束条件下作出电源管理决策和控制,降低系统能耗。从优化控制策略和操作系统资源管理的角度,综述操作系统电源管理研究现状及其关键技术。对两类动态电源管理策略进行详细分析,讨论了从资源管理角度进行操作系统电源管理的研究思路和问题,对当前面临的主要问题和发展趋势进行了总结和展望。

关键词 操作系统; 电源管理; 动态功耗管理; 动态频率调节; 低能耗

中图分类号 TP316

收稿日期:2007-06-08;修回日期:2008-02-02

基金项目:国家“八六三”高技术研究发展计划基金项目(2004AA1Z2024, 2007AA010304, 2007AA01Z462)

随着微电子技术和移动计算技术的快速发展, 能耗和散热问题日益突出, 特别在采用电池供电的系统中降低能耗已经成为和提高性能、降低成本同等重要的设计目标, 推动人们从能耗角度重新考虑系统设计和实现。

低功耗技术研究出现于硬件电路设计领域^[1], 随着研究的发展, 在体系结构^[2]、编译器^[3-5]、操作系统及网络等各层次展开, 涉及领域包括嵌入式系统、桌面系统和服务器系统等。管理的硬件从磁盘、人机交互设备、网卡等各类外设到处理器、内存和总线等核心部件。

在电能有限的条件下, 电源管理的目标是提高计算机系统的能耗效率, 关键在于系统性能与能耗之间的权衡。在多任务并发的操作系统中, 系统的能耗是工作负载驱动硬件操作的结果。由于应用场景和任务特征的多样性和存在性能约束, 以及操作系统对任务的管理调度作用, 给电源管理决策带来复杂性和不确定性因素; 硬件电源状态转换固有的能耗与时间开销, 也是电源管理决策不可忽视的影响因素。

传统的硬件低功耗技术从电路设计角度降低局部器件功耗, 难以顾及任务运行特征, 降低能耗的程度和范围有限。而操作系统能够在系统运行期间根据系统运行状态, 动态调节系统部件功耗状态, 在系统范围内更有效地发挥低功耗硬件的特性, 降低系统能耗。因此, 如何根据工作负载的复杂性和不确定性进行电源管理决策和控制, 在不降低用户可察觉的系统性能情况下降低系统能耗, 成为移动计算和操作系统研究领域的热点问题。

20世纪90年代以来, 操作系统电源管理技术受到广泛重视, 研究的关注点可以分为策略与算法、资源管理机制、规范与接口等多个方面。

从优化控制角度来看, 操作系统电源管理的主要问题是硬件特性基础上研究电源管理系统模型以及电源管理策略。从操作系统设计角度来看, 操作系统管理电源资源面临的主要问题是如何扩充现有操作系统资源管理框架来支持电源管理, 以及如何把电源当做一种资源进行操作系统资源抽象, 建立高效资源管理机制。

1 操作系统电源管理模型与策略

操作系统电源管理技术的硬件基础是系统部件的低功耗特性。虽然硬件的种类繁多, 但降低能耗

都依据 CMOS 电路的功耗原理。CMOS 电路功耗分为 3 类: 动态功耗、静态功耗和泄漏功耗。泄漏功耗相对较小, 可以忽略。静态功耗由于供电电源和地之间漏电流造成, 相对较小, 但由于芯片特征尺寸缩小而明显提高, 在低功耗设计中逐渐受到重视。动态功耗是门电路跳变的功耗在 CMOS 电路功耗中占主要部分, 是低功耗技术中的主要研究对象。

根据动态功耗和能耗表达式^[6-7], CMOS 器件的动态功耗是供电电压的单调递增函数, 即如果电压降低为原来的 $1/k$, 则动态功耗 P 近似为原来的 $1/k^3$; 能耗近似为原来的 $1/k^2$ 。降低电压会造成 CMOS 电路的延迟, 从而同步降低时钟频率^[6]。

在此基础上, 器件级低功耗设计技术通常采取降低供电电压、降低时钟频率或者减少硬件电路的分布电容等方法来降低器件的功耗。充分发挥可调节功耗的系统部件的低功耗特性成为软件电源管理技术的降低系统能耗的重要途径。以笔记本电脑为例, 显示屏背光、处理器以及硬盘这 3 个部件能耗占系统总能耗的 57%~68%^[8], 通过软件技术控制硬盘空闲时停转, 调低显示屏背光亮度等就可以显著降低系统能耗。

在空闲外设被关闭情况下, 处于运行态的系统主要能耗来自于处理器。因此, 处理器的动态功耗管理和电压调节技术是软硬件低能耗研究者的关注热点。

根据所管理的硬件特性抽象电源管理子系统模型是操作系统电源管理研究的出发点。根据硬件特性的区别, 把操作系统电源管理策略研究内容划分为两类: 动态功耗管理(dynamic power management)策略和动态电压调节(dynamic voltage scaling)策略。

1.1 电源管理子系统模型

Stanford 大学的 Benini 等人^[9]给出动态功耗管理的概念: 根据工作负载的变化, 选择性地设置系统部件到低功耗状态或者关闭空闲系统部件, 以最小的活动部件数目或最小的部件功耗来提供系统所需的服务和性能级别。

电源管理软件和被管理的系统部件被抽象为如图 1 所示的子系统, 由电源管理器(power manager)和系统部件构成。电源管理器监视系统运行, 决策选择何种功耗状态及状态转换时间, 并实施控制过程。可管理部件是系统部件或整个系统的抽象, 用

电源状态机来建模,如图2所示.每个状态代表特定能耗及性能级别,状态间弧线标注是状态间转换时间.

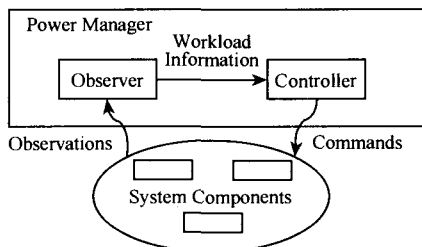


Fig. 1 Abstract model of the power management subsystem.

图1 电源管理子系统抽象模型

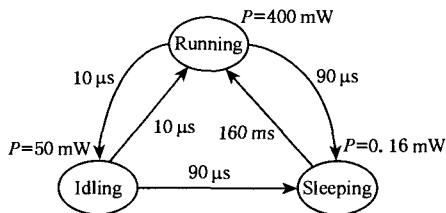


Fig. 2 Power state machine of a processor.

图2 处理器的功耗状态图

在该模型基础上评价电源管理策略的基本指标是策略可利用性.该指标依赖于部件的电源状态特征、工作负载特征以及性能约束定义.其中,部件的电源状态特征通过部件低功耗状态的能耗平衡时间来表达,表示部件能够节能所需要的最短空闲时间.

该电源管理子系统模型从过程控制的角度,强调电源管理器对系统部件的监视和控制作用以及系统部件的电源状态特征,对电源管理策略研究和评估具有重要的参考价值.

1.2 动态功耗管理策略

在系统部件确定的情况下,操作系统电源管理器决策面临的主要问题是对工作负载特征进行建模.在性能约束条件下,对系统部件功耗状态转换时机和选择何种状态作出决策.根据对工作负载特征抽象方法不同,把动态功耗管理策略分为3类:超时策略、基于预测的启发式策略和基于随机过程的优化策略.

1.2.1 超时策略

超时策略^[10]是最简单通用的DPM策略.如果部件空闲时间超过预定的时间,则该策略认为未来的空闲时间将大于部件的能耗平衡时间,从而设置系统为低功耗状态.该策略不关注应用的特征差异

和需求,比较通用.超时策略的主要缺陷是等待超时过程中增加能耗,唤醒部件时会造成系统部件响应延迟.如果超时阈值选择不当会带来性能损失或者降低节能效率.针对非平稳工作负载,可以通过保存阈值、加权平均以及动态自适应等方法^[11-12]优化超时阈值.研究表明,对无线网络传输设备等具有非平稳自相似业务请求的设备,最优的动态功耗管理策略是超时策略(确定性Markov策略)^[13].

1.2.2 基于预测的DPM策略

动态预测策略分为预测关闭^[14]和预测唤醒^[15]策略.预测关闭根据预测结果关闭系统部件;预测唤醒则根据预测的空闲时间提前唤醒部件.

预测方法属于启发式方法,假设系统部件访问在时间上存在关联性,未来的空闲时间可以通过历史信息进行估算,如何提高部件空闲时间预测的准确度是这类研究面临的主要问题.一种采用离线的非线性回归方法^[14]可以较好地拟合系统部件空闲时间特征,但对应用依赖较严重,并需要离线计算,适用于事先确定的工作负载.指数滑动平均方法^[15]易于实现并具有较高的准确性,应用范围较广.基于BP神经网络自适应学习来预测系统空闲时间^[16]的方法无需预先获得工作负载特性,具有传统回归算法不可比拟的优点.

1.2.3 基于随机过程的优化策略

基于随机过程的优化策略是通过较高层次数学抽象建立系统的概率模型,解决启发式方法遇到的最优化问题.

基于随机过程的方法用Markov链对部件能耗、状态转换时间和工作负载的不确定性建模,定义给定性能约束下的全局能耗优化问题,用线性规划方法求解.需要得到系统工作负载的先验信息,但系统工作负载很难提前建模.一种改进静态随机过程的方法^[17]在不同的工作负载下在线地学习,动态调节工作负载的Markov模型参数.

用这种方法得到的性能和能耗是期望值,不能保证对特定工作负载得到最优解.构造的Markov模型只是复杂随机过程的近似,如果模型不准确,优化策略也只是近似解.与启发式方法相比,该方法求解复杂度和计算开销较大,实际中难以直接实现,但可以用于离线系统分析和DPM策略评估.

从前面的分析可以看到,动态功耗管理策略的安全性和效率很大程度上取决于所管理部件电源状

态特性,以及系统中所运行的工作负载的特性。对各种策略的总结与比较见表 1 所示:

Table 1 Comparison of Three Power Management Strategies

表 1 三类动态功耗管理策略的比较

Strategy	Description	Features & Efficiency	Applicability	Workload Features
Timeout Strategy	When the idle period exceeds a fixed threshold, the component is switched to low power state.	Safety can be improved by increasing the timeout threshold, but large threshold may cause energy and performance penalty.	Various kinds of components	Does not consider workload features
Predictive Strategy	Predicts the current idle period based on the recent idle and active periods. If the preset condition is satisfied, the component is switched to low power state as soon as it becomes idle.	Depends on workload features; more efficient than the timeout approach, but less safe. Normally based on heuristics.	Interactive devices, such as keyboard, touch screen, and mouse	Efficient when workload has predictive timing patterns
Stochastic Optimal Strategy	Assumes stochastic models for component state transition, workload, and cost matrix, searches for the optimal states through learning.	Depends on the stochastic model of the components and workloads. The optimal policies are approximate solutions	Hard disk	Markov model for components and workload is required for policy optimization.

1.3 动态电压调节策略

DVS 技术主要针对处理器等电压/频率可调节的系统部件。根据功耗和频率的递增凸函数关系,在完成同样的工作量情况下,当处理器稳定在尽可能低的频率上时能耗最低^[18]。另一方面,应用程序并非总是需要以最快的速度执行,如果能以较低处理器频率满足任务的性能要求,则可以降低系统能耗或者达到期望电池寿命。问题关键在于正确预测工作负载对处理器的需求,合理分配处理器配额^[19]。根据对工作负载特性抽象方法不同,有以下两种 DVS 策略。

1.3.1 基于间隔处理器利用率的 DVS 策略

在这类 DVS 策略中,处理器利用率被用来衡量工作负载对处理器的需求,从而根据过去一段时间的处理器的利用率调节处理器频率^[20]。后来有许多这种策略的变种,其目的是提高预测未来的处理器利用率准确度,包括 Aged-a, Flat-u^[21]、指数滑动平均^[22]、概率分布模型^[23]等方法。

这种抽象方法简单有效,但处理器利用率只是对处理器需求的粗略抽象,不包含应用对系统性能需求的信息,不能很好保证系统的性能需求^[24]。适合应用特征确定或者不强调应用性能的系统。

1.3.2 基于任务处理器需求的 DVS 策略

实时系统和交互式系统对系统性能和能耗都有较高要求,要保证实时任务可调度性或满足延迟时间约束,不应当造成用户的可感知性能下降。为了

满足性能需求,需要在 DVS 策略中引入表达性能约束的抽象,把处理器频率调节与任务调度结合起来,研究带有 DVS 目标的任务调度问题。

在硬实时嵌入式系统中,每个任务有明确的到达时间和截止时间。通过用形式化方法描述实时任务调度条件下的处理器调频问题,可以给出离线的任务集调度算法^[25]。这类问题进一步被扩展为离散电压和基于固定优先级调度算法的 DVS 问题^[26]、DVS 下的任务同步问题^[27]和在线调度 DVS 调度问题^[28]。

与硬实时系统应用不同,交互式应用和软实时流媒体应用的性能约束不严格,允许存在一定程度的延迟,性能评估单位可以是任务或交互事务。同时,这类应用通常没有预知的处理器需求信息,任务截止时间和执行时间等参数通常不能预先确定。因此,针对这类应用的频率设置与预测算法确定应用的性能约束,以及对工作负载处理器需求的自动监测机制等都是需要解决的问题。

一种 DVS 调节方法是从全局任务管理和局部的任务调度两个层次预测任务的处理器需求^[29]。监测工作负载的处理器需求有多种方法,一种是利用线程并行性分析方法^[30]来监测交互响应过程的执行时间,由操作系统内核自动监测任务执行^[31];另一种是用监视 GUI 事件的方法^[32]获得交互事件的发生时机和特征。这类处理器需求监视及预测等功能增大了操作系统的开销,但同时也为操作系统自适应管理系统能耗提供了支持。

1.4 电源管理策略的评估指标

在操作系统电源管理技术研究中,如何评价电源管理策略的有效性和可用性是一个不可忽视的问题,对策略的评估和推广应用有决定作用。

评估原则强调策略对系统整体能耗的作用^[33],同时包括策略本身的开销、能耗、效果(电池寿命、截止时间超时率)等。

策略评估可以采用实际系统测量^[34]、仿真实验^[23]、模拟器估算和验证^[35]等方法。仿真验证收集电源管理前后的系统跟踪数据作为原始数据进行离线计算和比较,易于实现,是大多数研究者采用的方法。

目前,对电源管理策略的量化评估比较困难:一方面,实际测量每个操作的能耗开销很困难,另一方面,策略对系统整体能耗的影响还缺乏系统、有效的评测工具和评价指标。

2 操作系统电源管理机制与框架

操作系统承担着两种角色:管理系统资源和为用户提供接口,通过对被管理资源进行逻辑抽象并建立有效的管理机制来实现。操作系统设计者要把物理资源抽象和管理、逻辑资源的组织与管理以及面向应用的接口等不同层次的软件组成部分统一于整体的资源管理框架之中,为实现操作系统资源管理策略提供支持。因此,如何在已有的操作系统资源管理框架中加入电源管理机制与策略是操作系统电源管理研究的重要内容之一。

2.1 电源资源抽象与资源管理机制

传统操作系统没有考虑电源资源的管理,管理处理器、内存、外设等资源的方法是通过定义资源模型来建立资源抽象,建立资源的记账、分配与回收、共享与保护等的机制和策略。随着操作系统电源管理研究的深入,人们越来越多地从资源管理的角度意识到,应当把电源当做一个关键因素来思考操作系统的设计和实现。

一些研究者提出把电源当做系统资源来管理^[36],把操作系统传统资源(如CPU等)的记账机制用于电源的记账和收费,操作系统和应用程序协同进行电源管理。Vahdat等人^[37]提出把电源当做操作系统中的首要资源来考虑操作系统的设计问题。

用电量(currentcy)作为系统范围内能源的统一抽象,管理所有消耗能源的系统部件的方法^[38]将

电源资源抽象研究推进了一大步。该方法在嵌入式操作系统原型中提供简单接口来手工设置电池使用寿命目标以及任务电量分配比例,基于电量模型实现电源管理目标。这种方法的缺点是造成不成比例的性能损失。如何抽象细微和复杂的系统行为,通过策略研究来改进电源资源的抽象模型还是一个有待进一步研究的问题^[39]。

2.2 操作系统电源管理框架

操作系统电源管理框架起步于桌面操作系统,在嵌入式系统领域得到广泛研究。近年来,工业界和学术界关于操作系统电源管理框架的研究一直很活跃。

操作系统电源管理框架的需求主要有:1)支持系统范围的电源管理策略。除了调节处理器核频率,还要能支持其他系统部件和外设的电源状态以及频率调节,例如支持总线频率动态调节或者相对独立地调节外设的电源状态。2)灵活地支持不同平台的需求和电源管理策略的动态调整。系统部件能耗及有效的电源管理策略与应用类型有很大相关性,应当允许系统开发商根据应用的特点,定制特殊的电源管理策略。3)支持细粒度性能-能耗权衡和设置。在能源有限、性能约束严格的嵌入式应用系统中,细粒度基于任务的动态电源管理策略是保证性能-能耗权衡的必要条件。

ACPI规范(advanced configuration and power interface specification)^[40]面向台式机、服务器和移动电脑等带有BIOS固件的计算机系统,首次提出操作系统电源管理的概念,并给出软件接口、硬件接口和ACPI数据结构的定义以及这些接口之间的语义关系。2006年,ACPI 3.0b改进通用配置规范,增加多处理器以及新外设功耗管理、温度管理等支持。目前支持ACPI规范的操作系统主要有Windows 2000, XP, Vista和Linux等。1997年,微软公司提出面向台式机的电源管理架构OnNow Initiative^[41],2002年提出基本电源管理机制基础上的增强电源管理特性^[42]。从2004年至今的Linux论坛^[43]每年都推出Linux电源管理技术新进展,从支持ACPI框架发展为构建内核电源管理机制,开发可扩展的电源管理策略以及支持多核处理器电源管理等。

面向嵌入式应用的片上系统(system-on-chip)一般都不包括传统的BIOS,系统电源状态控制必须由操作系统来完成。针对嵌入式系统的特点,Brock等人提出了由动态电源管理方案、策略、策略管理器和DPM模块组成的动态电源管理框架^[44]。

嵌入式系统应用领域广泛、系统结构和设备类型复杂,目前尚无统一规范。能够有效支持电源管理策略,并可扩展的操作系统管理机制与框架仍然是操作系统设计和低能耗研究领域尚待研究的问题。

3 总结与展望

从移动计算设备到服务器领域,随着硬件低功耗技术的发展和器件层出不穷,系统级动态电源管理技术的研究也不断发展。操作系统电源管理集成硬件技术与软件技术作为降低计算机系统能耗的关键技术之一已取得了飞速进展与广泛应用,但是还有许多问题需要解决。我们认为以下几个方向会成为操作系统电源管理的研究重点。

1) 从提高操作系统资源管理的自适应控制能力角度,对任务执行、资源利用等情况的自我监测、控制机制的研究。操作系统的资源管理模式从传统的以资源分配回收为核心,向着资源利用的自适应控制方向发展。自适应管理电源资源,提高系统能耗效率是操作系统电源管理的目标。

2) 从提高电源管理策略有效性和可用性角度,对操作系统中工作负载特征、交互式系统中用户行为对工作负载的影响等方面的进一步研究。尤其在嵌入式系统领域,非常需要针对工作负载特征进行优化和可动态调节的电源管理策略。

3) 从操作系统资源管理的角度,电源的资源模型和资源管理机制问题的研究。例如用适当的资源抽象改进现有的资源模型,对系统各部分电源需求和能耗进行量化评估、电源资源的记账、分配等。在严重依赖电池资源的应用领域,这些问题具有非常重要的意义。

4) 从电源管理技术的量化评估角度,对系统部件和软件能耗的估算模型及评测方法的研究。利用模拟器技术进行操作系统能耗估算与优化,可以在软硬件设计早期阶段从软件角度对计算机系统各个组成部分进行能耗分析与评估,分析影响系统能耗的关键软件设计要素,优化操作系统对系统能耗的影响。这方面的研究成为操作系统低能耗研究领域兴起的新热点。

回顾计算机系统设计技术发展历史,如同提高计算机系统性能的需求推动软件性能优化技术和评测工具的快速发展一样,对降低计算机系统能耗的需求也必然推动软件电源管理和评测技术的发展,

我们可以预期未来会产生更多的能耗优化技术和评测工具。

参 考 文 献

- [1] O S Unsal, I Koren. System-level power-aware design techniques in real-time systems [C]. In: Proc of the IEEE, Special Issue on Real-Time Systems. Los Alamitos: IEEE Computer Society Press, 2003. 1055-1069
- [2] Y Guo, S Chheda, I Koren, *et al.* Energy characterization of hardware-based data prefetching [C]. In: Proc of the IEEE Int'l Conf on Computer Design. Los Alamitos: IEEE Computer Society Press, 2004. 518-523
- [3] Y Guo, M Ben-naser, C A Moritz. PARE: A power-aware data prefetching engine [C]. In: Proc of Int'l Symp on Low Power Electronics and Design (ISLPED'05). Los Alamitos: IEEE Computer Society Press, 2005. 339-344
- [4] Lei Ting, Li Xi, Zhou Xuehai. Performance lossless voltage scheduling for low energy software [J]. Journal of Computer Research and Development, 2006, 43(6): 153-158 (in Chinese)
(雷霆, 李曦, 周学海. 低能耗软件设计中的性能无损电压调度技术研究[J]. 计算机研究与发展, 2006, 43(6): 153-158)
- [5] Yi Huizhan, Chen Juan, Yang Xuejun, *et al.* A real-time dynamic voltage scaling algorithm based on syntax tree for low power [J]. Journal of Software, 2005, 16(10): 1726-1734 (in Chinese)
(易会战, 陈娟, 杨学军, 等. 基于语法树的实时动态电压调节低功耗算法[J]. 软件学报, 2005, 16(10): 1726-1734)
- [6] K S Yeo, S S Rofail, W Goh. CMOS/BiCMOS ULSI: Low Voltage, Low Power [M]. Beijing: Publishing House of Electronics Industry and Pearson Education Press, 2003
- [7] A P Chandrakasan, S Sheng, R W Brodersen. Low-power CMOS digital design [J]. IEEE Journal of Solid-State Circuits, 1992, 27(4): 473-484
- [8] J R Lorch. A complete picture of the energy consumption of a portable computer: [Master dissertation] [D]. Berkeley: University of California at Berkeley, 1995
- [9] L Benini, A Bogliolo, G D Micheli. A survey of design techniques for system-level dynamic power management [J]. IEEE Trans on Very Large Scale Integration (VLSI) Systems, 2000, 8(3): 299-316
- [10] D Ramanathan, R Gupta. System level online power management algorithms [C]. Conf on Design, Automation and Test in Europe, Paris, France, 2000
- [11] P Krishnan, P Long, J Vitter. Adaptive disk spindown via optimal rent-to-buy in probabilistic environments [C]. In: Proc of Int'l Conf on Machine Learning. East Sussex: Omni Press, 1995. 322-330

- [12] D Helmbold, D Long, E Sherrod. Dynamic disk spin-down technique for mobile computing [C]. In: Proc of IEEE Conf on Mobile Computing. Los Alamitos; IEEE Computer Society Press, 1996. 130-142
- [13] Wu Qi, Xiong Guangze. Adaptive dynamic power management for non-stationary self-similar requests [J]. Journal of Software, 2005, 16(8): 1499-1505 (in Chinese) (吴琦, 熊光泽. 非平稳自相似业务下自适应动态功耗管理 [J]. 软件学报, 2005, 16(8): 1499-1505)
- [14] M Srivastava, A Chandrakasan, R Brodersen. Predictive system shutdown and other architectural techniques for energy efficient programmable computation [J]. IEEE Trans on VLSI System, Los Alamitos; IEEE Computer Society Press, 1996, 4(1): 42-55
- [15] C - Hwang, A Wu. A predictive system shutdown method for energy saving of event-driven computation [C]. In: Proc of Int'l Conf on Computer-Aided Design. Los Alamitos; IEEE Computer Society Press, 1997. 28-32
- [16] Tang Zhifang, Shi Haitao, Lu Huaxiang, *et al.* System-level power management algorithm based on BP neural network [J]. Computer Engineering, 2006, 32(4): 214-216 (in Chinese) (唐志芳, 时海涛, 鲁华祥, 等. 基于 BP 神经网络的系统级电源管理算法 [J]. 计算机工程, 2006, 32(4): 214-216)
- [17] E Chung, L Benini, A Bogliolo, *et al.* Dynamic power management for nonstationary service requests [C]. In: Proc of Design and Test in Europe Conf. Los Alamitos; IEEE Computer Society Press, 1999. 77-81
- [18] R Gonzalez, B M Gordon, M Horowitz. Supply and threshold voltage scaling for low power CMOS [J]. IEEE Journal of Solid-State Circuits, 1997, 32(8): 1210-1216
- [19] M Weiser, B Welch, A Demers, *et al.* Scheduling for reduced CPU energy [C]. In: Proc of the 1st Symp on Operating Systems Design and Implementation. San Francisco; USENIX Association Press, 1994. 13-23
- [20] M Weiser, B Welch, A Demers, *et al.* Scheduling for reduced CPU energy [C]. The 1st Symp on Operating Systems Design and Implementation (OSDI), Monterey, CA, 1994
- [21] K Govil, E Chan, H Wasserman. Comparing algorithms for dynamic speed-setting of a low-power CPU [C]. In: Proc of the 1st Int'l Conf on Mobile Computing and Networking. New York; ACM Press, 1995
- [22] T Pering, T Burd, R Brodersen. Voltage scheduling in the IpARM microprocessor system [C]. In: Proc of IEEE Int'l Symp on Low Power Electronics and Design. Los Alamitos; IEEE Computer Society Press, 2000. 96-101
- [23] J R Lorch. PACE: A new approach to dynamic voltage scaling. [J]. IEEE Trans on Computers, 2004, 53(7): 856-869
- [24] T Pering, T Burd, R Brodersen. The simulation and evaluation of dynamic voltage scaling algorithms [C]. In: Proc of IEEE Int'l Symp on Low Power Electronics and Design. Los Alamitos; IEEE Computer Society Press, 1998
- [25] F Yao, A Demers, S Shenker. A scheduling model for reduced CPU energy [C]. In: Proc of IEEE Annual Foundations of Computer Science. Los Alamitos; IEEE Computer Society Press, 1995. 374-382
- [26] W Kwon, T Kim. Optimal voltage allocation techniques for dynamically variable voltage processors [C]. In: Proc of the 40th Conf on Design Automation. New York; ACM Press, 2003
- [27] R Jejurikar, R K Gupta. Energy aware EDF scheduling with task synchronization for embedded real time systems [R]. University of California at Irvine, Tech Rep; 02-24, 2002
- [28] P Pillai, K G Shin. Real-time dynamic voltage scaling for low-power embedded operating systems [C]. In: Proc of the 18th Symp on Operating Systems Principles. New York; ACM Press, 2001. 89-102
- [29] W Yuan, K Nahrstedta, S V Advea, *et al.* Design and evaluation of a cross-layer adaptation framework for mobile multimedia systems [C]. In: Proc of SPIE/ACM Multimedia Computing and Networking Conference (MMCN). New York; ACM Press, 2003
- [30] K Flautner, R Uhlig, S Reinhardt, *et al.* Thread-level parallelism and interactive performance of desktop applications [C]. In: Proc of Conf on Architectural Support for Programming Languages and Operating Systems (ASPLOS-IX). New York; ACM Press, 2000. 129-138
- [31] K Flautner, T Mudge. Vertigo: Automatic performance-setting for Linux. [C]. In: Proc of the 1st USENIX Symp on Operating Systems Design and Implementation (OSDI). San Francisco; USENIX Association Press, 2002. 105-116
- [32] J R Lorch. Operating System Modifications for Task-Based Speed and Voltage Scheduling [C]. In: Proc of the 1st Int'l Conf on Mobile System Applications and Services (Mobisys), New York; ACM Press, 2003. 215-229
- [33] J R Lorch, A Smith. Software strategies for portable computer energy management [J]. IEEE Personal Communications, 1998, 5(3): 60-73
- [34] D grunwald, P Levis, K Farkas, *et al.* Policies for dynamic clock scheduling [C]. In: Proc of the 4th Symp on Operating Systems Design & Implementation. San Francisco; USENIX Association Press, 2000. 6-6
- [35] S Kang, Y C Huayong Wang, X Wang, *et al.* Skyeeye: An instruction simulator with energy awareness [C]. In: Proc of the 1st Int'l Conf on Embedded Software and System. Berlin; Springer-Verlag, 2005
- [36] R Neugebauer, D Mcauley. Energy is just another resource; energy accounting and energy pricing in the nemesis OS [C]. In: Proc of the 8th IEEE Workshop on Hot Topics in Operating Systems (HotOS-VIII). Los Alamitos; IEEE Computer Society Press, 2001
- [37] A Vahdat, A Lebeck, C Ellis. Every joule is precious: The case for revisiting operating system design for energy efficiency [C]. In: Proc of the 9th ACM SIGOPS European Workshop. New York; ACM Press, 2000. 31-36

- [38] H Zeng, C S Ellis, A R Lebeck, *et al.* Ecosystem: Managing energy as a first class operating system resource [C]. In: Proc of the 10th Int'l Conf on Architectural Support for Programming Languages and Operating Systems (ASPLOS X). New York: ACM Press, 2002. 123-131
- [39] H Zeng, C S Ellis, A R Lebeck, *et al.* Currentcy: A unifying abstraction for expressing energy management policies [C]. In: Proc of USENIX 2003 Annual Technical Conference. San Francisco: USENIX Association Press, 2003. 43-56
- [40] Hewlett-Packard, Intel, Microsoft, *et al.* Advanced Configuration and Power Interface Specification V3.0 [OL]. <http://www.acpi.info/>, 2003
- [41] Microsoft. OnNow: The Evolution of the PC Platform [OL]. <http://www.microsoft.com/hwdev/pcfuture/OnNow.HTM>, 1997
- [42] P Yao. Power Management Features of Windows CE. NET [OL]. <http://msdn.microsoft.com/library/>, 2002
- [43] P Mochel. The State of Linux Power Management [OL]. <http://www.linuxsymposium.org/2006/proceedings.php>, 2006-11-01
- [44] B Brock, K Rajamani. Dynamic power management for embedded systems [C]. In: Proc of the IEEE Int'l SOC Conference. Los Alamitos: IEEE Computer Society Press, 2003



Zhao Xia, born in 1972. Ph. D. candidate and associate professor. Senior member of China Computer Federation. Her main research interests include low energy software, system software, and software engineering.

赵霞, 1972年生, 博士研究生, 副教授, 中国计算机学会高级会员, 主要研究方向为低功耗软件、系统软件与软件工程等。



Chen Xiangqun, born in 1961. Professor and senior member of China Computer Federation. Her main research interests include system software and software engineering, *etc.*

陈向群, 1961年生, 教授, 中国计算机学会高级会员, 主要研究方向为系统软件与软件工程(cherry@sei.pku.edu.cn)。



Guo Yao, born in 1976. Lecturer. His main research interests include system software, low power system design, compiler technology and software engineering, *etc.*

郭耀, 1976年生, 讲师, 主要研究方向为系统软件、低功耗系统设计、编译技术、软件工程等。



Yang Fuqing, born in 1933. Professor and doctoral supervisor. Member of the Chinese Academy of Sciences. Honorary president and member of Steering Committee of China Computer Federation.

Her main research interests include system software and software engineering, *etc.*

杨芙清, 1933年生, 博士生导师, 教授, 中国科学院院士, 中国计算机学会名誉理事和指导委员会委员, 主要研究方向为系统软件与软件工程。

Research Background

This work is supported in part by the National Natural Science Foundation of China under grant No. 2004AA1Z2024, No. 2007AA010304, and No. 2007AA01Z462 (Research on key techniques of embedded operating system and desktop operating system security).

With the rapid development of mobile computing technology, it becomes critical to reduce power dissipation on battery-powered systems with operating system design techniques. In this review paper, we present the state-of-the-art in this growing field, and introduce the key techniques from two directions: optimal control strategy and operating system resource management. Finally, we discuss the challenges and open issues in this field.