面向同构多核处理器的节能任务调度方法

王颖锋 刘志镜

(西安电子科技大学计算机学院 西安 710071)

摘 要 对于运行在同构多核处理器上的周期性硬实时任务,设计了一个基于动态电压调节的节能调度方法。该方 法首先将计算任务按照周期数降序排序并基于计算任务调度长度最短的原则安排任务映射。然后将各个处理核上具 有最小通讯时间的计算任务设置为最后执行的计算任务而其它计算任务顺序保持不变。在初始映射中所有计算任务 都被分配最高频率的情况下,每个处理核上的计算任务在执行时间扩展过程中确定最佳的计算任务顺序。基于 Intel PXA270 的功耗模型,以几个随机任务集作实验。结果表明提出的方法能够有效地降低多核处理器的能量。 关键词 任务图,节能,实时系统,动态电压调节,多核 中图法分类号 TP316 文献标识码 A

Energy-efficient Task Scheduling Approach for Homogeneous Multi-core Processors

WANG Ying-feng LIU Zhi-jing

(School of Computer Science and Technology, Xidian University, Xi'an 710071, China)

Abstract For periodic hard real-time tasks running on homogeneous multi-core processors, an energy-efficient approach based on dynamic voltage scaling (DVS) was designed. First, computation tasks are ordered by decreasing cycles and task mapping is arranged based on the principle of the shortest scheduling length for computation tasks. Then the computation task with the minimum commun- ication time is set as the last executed computation task while the order of other computation tasks keep unchanged for each processor core. The optimal execution order of computation tasks arranged with the highest frequency in the initial mapping. Experiments were conducted on several random task sets based on the power model of the Intel PXA270. Results show that the proposed approach can decrease energy of multi-core processors effectively.

Keywords Task graph, Energy-efficient, Real-time systems, Dynamic voltage scaling, Multi-core

1 引言

随着移动设备和嵌入式设备平台数目飞速的增长,能量 消耗问题越来越受到关注^[1]。为了满足性能的需要和能量的 限制,多核处理器已经被嵌入式系统广泛采纳。在实际的嵌 入式系统中,许多应用程序存在着严格的时间限制。因此,有 必要以多核处理器为研究对象,研究如何在满足时间限制的 条件下降低能量消耗的问题。

对于单处理机上的节能调度,文献[2]提出了一种动态规 划算法。该算法将节能调度问题转换为一种变异的背包问 题,在满足时间约束的前提下,应用动态电压调节(DVS)技术 降低系统能量。文献[3]提出了一种基于遗传算法的非占先 式能耗感知调度算法。该算法通过构造特定遗传过程来同时 满足能量和时间约束。文献[4]提出了一种同时满足任务同 步和节能要求的调度算法,该算法利用周期性任务运行时剩 余的执行时间调节处理器的频率,以降低系统的动态能量消 耗。

对于多处理机上的节能调度,文献[5]提出了一种基于动态电压调节的改进贪婪算法,该算法考虑的是不相关任务在 多处理器上的调度,通过对静态能量管理进行分析,将动态电 压调节并入任务调度以降低系统的能量消耗。文献[6]提出 了一种针对周期性任务的动态电压缩放的多处理器节能调度 方法。通过静态分析确定最低处理器调度要求,求出满足可 调度性条件下的最低处理器速度。文献[7]提出了一种基于 任务同步与节能的混合调度算法,用静态算法确定任务的静态速度,用动态调度算法在实际运行中固定临界区的运行速 度,并调节处理器的速度,以达到降低能耗并满足实时可调度 性的目的。文献[8]提出了一种两阶段方法来实现异构分布 式系统在时间限制下的能量优化。首先将每个任务的尽可能 晚的开始时间加上一个随机数,将得到的数值按升序排序,得 到的顺序定为任务优先级,然后利用表调度算法实现任务的 分配,采用松弛预算的方法对松弛进行分配。

到稿日期;2010-10-22 返修日期;2011-06-28 本文受陕西省'13115'科技创新工程重大专项(2007ZDKG-57)资助。

王颖锋(1976一),女,博士生,主要研究方向为调度算法、实时系统的功耗压缩及容错技术、高性能计算、并行计算等,E-mail:wyf96126@126. com;**刘志镜**(1957一),男,教授,主要研究方向为网络多媒体技术、视觉计算、虚拟现实技术和数据挖掘等。

然而上述研究都没有考虑如何降低运行依赖任务所产生 的电压转换能量。动态电压调节是在运行时动态调节电压而 降低能量消耗的有效技术。重定时任务图^[9]较有向无环图能 够提高任务的并行性并能为降低能量消耗提供更多的机会。 因此本文在前面研究的基础上^[10],提出了一个将任务顺序调 整并人任务执行时间扩展的系统级节能调度方法。该方法将 重定时任务图作为输入对象。不同于前面的研究,本文提出 的方法是一种非处理器(核)间任务迁移算法,避免了处理器 (核)间任务迁移算法的巨大开销,并针对具有可忽略或常数 电压转换时间的处理器,利用任务重排序降低了电压转换能 量。

2 相关理论

2.1 系统及任务模型

本文研究的系统是存在空闲状态并支持动态电压调节的 同构多核系统,每个处理核都支持 m 个离散的频率/电压模 式。每个处理核能够独立地调节频率/电压模式。假设 $f_1 < f_2 < \cdots < f_m$ 。将频率/电压模式按照频率从低到高的顺序排 列,即 $(f_1,V_1), (f_2,V_2), \cdots, (f_m,V_m)$ 。其中 f_1 所对应的功 耗为空闲功耗。

具有 n 个周期性依赖任务的应用程序可以抽象为有向无 环图,表示为: $G=\{\Gamma, E, \gamma, C, D\}$ 。其中 $\Gamma=\{\tau_1, \tau_2, \dots, \tau_n\}$ 为 节点的集合。E 表示有向边的集合, μ_{e_i} , 表示 τ_i 的执行是 在 τ_i 的执行之后,并且 τ_i 的结果依赖于 τ_i 的结果。 $\gamma(e_{ij})$ 表 示节点 τ_i 和 τ_i 所差的延迟数(迭代次数)。 $c(\tau_i) \in C$ 表示任 务 τ_i 的周期数。借助于 RDAG 算法可以获得有向无环图的 重定时有向无环图。D 表示任务集的时间约束,即截止时间。

2.2 能量模型

本文假定在多核实时系统中,能量消耗主要是由基于 CMOS技术的处理器核产生。能量采用文献[9,11]中的模型。任务 r. 每周期的动态功耗为:

 $P_{AC} = C_{eff} \cdot V_{dd}^{2} \cdot f \tag{1}$

式中,C_{eff}是每周期平均开关电容,V_{dd}是电源电压,f是处理器时钟频率。处理器时钟频率 f 和电源电压 V_{dd}之间存在如下关系:

$$f = k \cdot \frac{(V_{dd} - V_{th})^2}{V_{dd}} \tag{2}$$

式中,k是一个常数,V_#是电路的门限电压。

任务 τi 所消耗的动态能量为:

$$E = P_{AC} \cdot t_u \tag{3}$$

式中,t_u 是执行任务 t_i 所需要的时间,可以由下式得出:

$$t_{u} = \frac{N(u) \cdot V_{dd}}{k \cdot (V_{dd} - V_{ft})} \tag{4}$$

式中,N(u)是执行任务 τ_i 所需要的计算周期数。将式(1)和式(4)代人式(3),得到:

$$E = N(u) \cdot C_{eff} \cdot V_{dd}^{2}$$
(5)

$$\varepsilon_{k,j} = C_r \cdot |V_{ddk} - V_{ddj}|^2 \tag{6}$$

式中,C,是电源导轨电容。

2.3 RDAG 算法

依据重定时原理^[12],文献[9]提出了 RDAG 算法,应用 该算法可以将依赖任务转换为仅存在迭代间数据依赖的任 务。该算法以有向无环图为输入对象,经过重定时值调整,产 生重定时任务图。其具体作法就是通过在每条通讯边上设置 恰当的重定时值,使在原任务图中存在依赖关系的任务分别 处于不同的迭代中,因此消除了迭代内数据依赖。

3 提出的节能调度算法

定理1 在电压转换时间为固定常数或者可以忽略不记 的情况下,当每个任务以一个频率/电压模式执行时,各个处 理器上的任务按电压档降序执行或升序执行,所产生的电压 转换能量最小。

假设某处理核上的任务有 n+1 频率/电压模式,其电压升 序排列为 $V_1, V_1 + a_1, V_1 + a_2, \dots, V_1 + a_{k-1}, V_1 + a_k, \dots, V_1 + a_{m-1}, V_1 + a_m, \dots, V_1 + a_n$ 。其中, $0 \le a_1 \le a_2 \le \dots \le a_k \le \dots \le a_m \le \dots \le a_n \le \dots \le a_n \le n+1$ 个频率/电压模式的电压转换能量为:

 $A+C_{r}[(a_{k-1}-a_{k})^{2}+(a_{k}-a_{k+1})^{2}+\dots+(a_{m-1}-a_{m})^{2}+(a_{m}-a_{m+1})^{2}+\dots+(a_{n-1}-a_{n})^{2}]$ (7) 式中,A为前k个频率/电压模式的电压转换能量。交换V₁+a_k和V₁+a_m的位置,可以得到:

$$A+C_{r}[(a_{k-1}-a_{m})^{2}+(a_{m}-a_{k+1})^{2}+\cdots+(a_{m-1}-a_{k})^{2}+(a_{k}-a_{m+1})^{2}+\cdots+(a_{n-1}-a_{n})^{2}]$$

$$(8)$$
式(7)-式(8),可得:

$$2C_r \cdot (a_k - a_m) \cdot [(a_{k-1} + a_{k+1}) - (a_{m-1} + a_{m+1})] \ge 0$$
(9)

而交换任意两个频率/电压模式的位置都会得到相同的 结果。因此处理核上的任务按电压档升序的顺序执行所产生 的转换能量最小。同理可证处理器核的任务按电压档降序的 顺序执行所产生的转换能量最小。

文中考虑到所应用的处理核从运行状态到空闲状态的转 换时间可能是固定常数且不能忽略。对于这种情况,若给定 的时间限制与某处理核上的最后一个计算任务的完成时间的 差小于从运行状态到空闲状态的转换时间,则剩余的时间以 该处理核上所运行的最后一个计算任务的频率/电压模式运 行。因此为了减少能量消耗,对于从运行状态到空闲状态的 转换时间是固定常数且不能忽略的处理器,其各个处理核上 的任务按电压档降序的顺序执行。而对于从运行状态到空闲 状态的转换时间是可以忽略的处理器,其各个处理核上的任 务按电压档降序或升序的顺序执行均可。

文中提出的方法是用来解决含有依赖任务的应用程序在 多核处理器上的节能调度问题。该算法以从原有向无环图获 得的重定时有向无环图作为输入。除此之外还需要输入的信 息有处理器核所支持的频率/电压模式、模式转换时间、处理 器核的个数,电源导轨电容。具体实现如下:

步骤1 将计算任务按周期数降序排列,按照调度长度 最短处理核(对于计算任务)优先分配的原则进行任务到处理 核的映射。

步骤 2 对于每个处理核,将具有最小通讯时间的计算

任务设置为该处理核最后执行的计算任务,而其它计算任务 的顺序保持不变。

步骤 3 将每个处理核上的计算任务安排以所在处理核 的最低频率(除空闲状态所对应的频率外的最低频率)。如果 产生的调度长度小于等于给定的时间限制,则接受该调度并 返回,否则转步骤 4。

步骤 4 将每个处理核上的计算任务安排以所在处理核的最高频率。

步骤 5 在具有最短调度长度的处理核上,选择一个计 算任务 r_i,当其频率降低一个频率档并按照其变化后,电压档 降序排列该处理核上的计算任务,使总的能量消耗最小,并将 计算任务 r_i进行执行时间扩展,即将该计算任务的频率档降 低一档。

步骤 6 如果该处理核的新调度长度小于等于给定的时间限制,则接受这一变化,计算任务 τ_i 以变化后的频率档执行,该处理器核上的计算任务按降电压档的顺序执行,更新调度和调度长度,转步骤 5,否则转步骤 7。

步骤 7 计算任务 r; 以变换前的频率执行,该处理器核 上的计算任务以此次变化前的顺序执行,即接受上一次变化, 返回。

该方法首先按照计算周期数的大小确定计算任务映射的 优先级,根据对于计算任务调度长度最短的原则实现初步任 务映射。然后考虑到通讯时间对调度长度的影响,对每个处 理核在保持其它计算任务顺序不变的情况下,选择具有最小 通讯时间的计算任务(对不同处理核上的子孙任务有最小通 讯时间或不存在子孙任务)作为各个处理核最后执行的计算 任务,从而为进一步缩短调度长度提供了可能,降低了每个任 务以最低频率运行时背离时间限制的可能性。若每个计算任 务初始被安排以最高频率,则在具有最短调度长度的处理核 上选择一个候选任务,将其频率降低一个频率档而它所在处 理核上的其它任务的频率保持不变。在它频率变化后,将它 所在处理核上的任务按降电压档排序,计算所产生的总能量 消耗,将该总能量消耗与相同处理核上其它任务做相同操作 所产生的总能量消耗进行比较。如果该任务降低一个频率档 后并将它所在处理核上的任务按降电压档顺序执行所产生的 总能量消耗最小,则将它的频率降低一个频率档并安排它所 在处理核上的任务的执行顺序为降电压档的顺序。如此反复 进行计算任务频率调节和计算任务顺序调整。一旦进行频率 调节后且上一次操作后具有最短调度长度的处理器核的调度 长度超过时间限制,就不进行此次降频操作,而保持上一次的 调度结果并返回。

4 实验结果

为了验证提出的节能调度方法的节能效果,本文将提出 的方法与文献[8]中提出的方法进行对比分析。在实验中采 用文献[13]的方法生成了 5 个随机任务图,分别命名为任务 集 1一任务集 5。其计算任务个数在 27 至 68 之间变动,而通 讯边数在 45 至 94 之间变动。任务周期数满足 10 至 40 之间 的均匀分布,具体任务特征见表 1。原有向无环图中存在依 赖关系的两个计算任务如果被分配到不同的处理核,则它们 之间的通讯周期数为从1到4随机选择的数值。这两个计算 任务1如果被分配到相同的处理核,则它们之间的通讯周期 数为0。在所研究的系统中,总线上的频率为固定的,设置为 208MHz。将生成的任务集分别在具有2和4个处理核的多 核系统上进行模拟实验。处理器采用 Intel PXA270^[14]的功 耗模型,处理器可运行在7个频率/电压模式下。每个频率/ 电压模式及其所对应的功耗见表 2。对于 Intel PXA270,其 模式转换时间是可以忽略的[15]。应用式(6)计算电压转换能 量,其中C,设置为5pF。对于每个任务集采用6个不同的时 间限制,初始时间限制为文献[8]中的方法在两处理核系统的 初始调度长度,然后以该调度长度的10%作为步长,设置时 间限制,将得到的能量消耗取均值。为了简单起见,将文献 [8]中的方法和本文提出的方法分别命名为方法1和方法2。 将方法1和方法2的模拟结果从平均能量消耗方面进行比 较。图1是方法1和方法2的平均能量消耗比较结果。

表1 任务集特征表

任务集名称	节点个数	边数	周期数
任务集1	27	45	681
任务集2	35	54	880
任务集3	46	67	1233
任务集4	57	74	1520
任务集5	68	94	1808

表 2	Intel PXA255 的频率及其相应的电压和功料			叻耗
_	频率(MHz)	电压(V)	功耗(mW)	

13(空闲)	0.85	44.2
104	0.9	115
208	1.15	279
312	1.25	390
416	1.35	570
520	1.45	747
624	1.55	925



图 1 两种方法的平均能量消耗比较

从图1可以看出,本文提出的方法的节能效果优于文献 [8]中的方法。这是因为采用重定时有向无环图作为输入,使 原有向无环图中存在依赖关系的任务不再存在迭代内数据依 赖,而只存在迭代间数据依赖,从而提高了任务的并行能力, 进而为任务的执行提供了更多的频率/电压模式调节机会。 又由于在频率/电压模式的选择过程中根据每个处理器核上 任务的频率/电压模式确定任务执行顺序,因此能够同时降低 动态能量消耗和电压转换能量消耗。而方法1由于受到迭代 内数据依赖的影响,降低了动态电压调节能力和多核并行能 力。

结束语 本文对多核处理器上运行的周期性依赖任务在 时间限制下的节能调度问题进行了研究,以同构多核处理器 为研究对象提出了一个节能调度方法。该方法考虑到任务的

• 296 •

依赖关系对频率/电压模式调节的制约作用,采用 RDAG 算 法获得的重定时任务图作为输入对象,将任务顺序调整和频 率/电压模式选择同时进行。实验表明,提出的方法在满足任 务集时间限制的条件下达到了良好的节能效果。由于多核处 理器目前正向异构方向发展,因此下一步的工作将研究如何 在满足时间限制的条件下,降低异构多核处理器上运行的依 赖任务集的能量消耗。

参考文献

- [1] 钟虓,齐勇,侯迪,等.基于 DVS 的多核实时系统节能调度[J].
 电子学报,2006,34(12):2481-2484
- [2] 刘惠,陈平,杜军朝.可变电压处理器的最优动态电压选择算法 [J]. 西安电子科技大学学报,2009,36(6):486-490,540
- [3] 解玉凤,魏少军.实时周期任务的非占先式能耗感知调度[J]. 计 算机辅助设计与图形学学报,2006,18(2):245-250
- [4] 韩建军,刘同涛,李庆华,等.基于任务同步及节能的单机系统实时动态调度算法[J]. 高技术通讯,2008,18(11):1180-1186
- [5] 邢静宇,张立臣. 动态电压调整多处理器实时系统任务调度[J].微电子学与计算机,2006,23(2):51-57,61
- [6] 桑楠,李保宇,马红.多处理器的节能调度算法[J]. 电子科技大 学学报,2008,37(1):116-119
- [7] 韩建军,吴晓东,李庆华,等.硬实时系统中基于任务同步及节能 的动态调度算法[J].通信学报,2009,30(11):15-26
- [8] Luo J, Jha N K. Power-efficient scheduling for heterogeneous distributed real-time embedded systems [J]. IEEE Transaction

(上接第 281 页)

将上述异步 FIFO 的 verilog 模型和异步 FIFO 的关键属 性编码成 SMV 语言源代码 FIFO. smv,输入到 Cadence SMV 工具中运行调试,运行结果没有反例,报出属性定义都是 true。

4.4 实验结果

结构检查、基于断言的模拟验证、形式化验证,采用层次 化验证方法对我们的 CDC 设计进行了检查,验证时间逐步增 长。采用 spyglass 工具进行结构检查,发现我们设计的异步 FIFO 设计有错误。我们基于断言的模拟检查,形式化验证都 没有报告违反。查找原因是该工具的模板库缺少该类型的描 述。设计的芯片已流片回来,经测试,芯片的功能正确,说明 我们的设计和验证方法是有效的。

结束语 随着高性能、低功耗芯片的发展,多时钟域和跨时钟域(CDC)设计越来越多,CDC设计和验证越来越重要, 传统 RTL设计流程不能发现 CDC的问题。我们提出了层次 化验证流程:结构化检查,基于断言的验证(ABV),对关键模 块进行模型检验。设计的芯片已流片回来,经测试芯片的功 能正确,说明我们的设计和验证方法是基本完备的。

参考文献

- [1] Shaker S. Critical clock-domain-crossing bugs[EB/OL]. http:// www. highbeam. com/doc/1G1-177439329. html, 2008-04-03
- [2] Feng Yi, Zhou Zheng, Tong Dong, et al. Clock Domain Crossing

on Computer-aided Design of Integrated Circuits and System, 2007, 26(6):1161-1170

- [9] Liu H, Shao Z, Wang M, et al. Overhead-Aware System-Level Joint Energy and Performance Optimization for Streaming Applications on Multiprocessor Systems-on-Chip[C] // Proc. Euromicro Conference on Real-Time Systems (ECRTS 08). Prague, Czech, IEEE Press, July 2008, 92-101
- [10] 王颖锋,刘志镜.一种变电压多核处理器上的有效节能方法[J]. 计算机科学,2010,35(10):294-296
- [11] Kang J, Ranka S. Assignment algorithm for energy minimization on parallel machines [C] // 2009 International Conference on Parallel Processing Workshops. Vienna, Austria: IEEE Press, Sep. 2009:484-491
- [12] Leiserson C E, Saxe J B. Retiming synchronous circuitry[J]. Algorithmica, 1986, 13(8): 5-35
- [13] http://www.kasahara.elec.waseda.ac.jp/schedule/
- [14] Yang C C, Wang K C, Lin M H, et al. Energy efficient intra-task dynamic voltage scaling for realistic cpus of mobile devices [J]. Journal of Information Science and Engineering, 2009, 25(1): 251-272
- [15] Zeng G, Tomiyama H, Takada H. Power Optimization for Embedded System Idle Time in the Presence of Periodic Interrupt Services[C] // Proc. IESS 2007. Irvine, CA, USA: IEEE Press, May 30- June 1,2007;241-254

Fault Model and Coverage Metric for Validation of SoC Design [C]//Design, Automation& Test in Europe Conference& Exhibition. 2007;1385-1390

- [3] Feng Yi, Yi Jiang-fang, Liu Dan, et al. Property Generation Method for Model Checking on Clock Domain Crossing Design
 [J]. Acta Elecironica Sinica, 2008, 36(5): 886-892
- [4] Li Bing, Kwok C K-K. Automatic Formal Verification of Clock Domain Crossing Signals[C] // Design Automation Conference in Asia and South Pacific, ASP-DAC, 2009:654-659
- [5] Atrenta, Inc. SpyGlass
 Clock-Reset Rules Reference [EB/ OL]. http://www.atrenta.com,2009-03
- [6] Jones I W, Yang Su-wen, Greenstreet M. Synchronizer Behavior and Analysis[C] // IEEE Symposium on Asynchronous Circuits and Systems, 2009:117-126
- [7] Salomon B, Ran G, Michael P, et al. The Devolution of Synchronizers[C] // IEEE Symposium on Asynchronous Circuits and Systems. 2010:94-103
- [8] McMillan K L. Getting started with SMV[DB/OL]
- [9] Cummings C E, Clock Domain Crossing(CDC) Design & Verification Techniques Using SystemVerilog[R]. Boston : SNUG-2008
- Ly T, Hand N, Kwok C K. Formally Verifying Clock Domain Crossing Jitter Using Assertion-based Verification[C]//Design & Verification Conference & Exhibition. DVcon, 2004:1-5