

电力终端多核系统的节能调度优化

姚浩^{1,2}, 习伟^{1,2}, 蒋小文³

(1. 南方电网数字电网研究院有限公司, 广东 广州 510670; 2. 南方电网科学研究院有限责任公司, 广东 广州 510663;
3. 浙江大学信息与电子工程学院, 浙江 杭州 310027)

摘要:在嵌入式电力终端系统中应用多核处理器,可以显著提升系统性能,同时也带来实时性和能耗增加的问题。针对电力工控终端系统的运行特性,提出一种基于时间触发的多核系统节能调度方法。提出一种智能调度方案,可以保证多核系统的实时性要求;同时采用动态电压及频率调节(DVFS)和动态电源管理(DPM)这2种常用的系统级能源管理技术,可以降低多核处理器的功耗。

关键词:实时性;智能调度;系统功耗

DOI:10.19781/j.issn.1673-9140.2021.05.001 中图分类号:TM734 文章编号:1673-9140(2021)05-0003-07

Research on energy-saving scheduling optimization method for power terminal multi-core systems

YAO Hao^{1,2}, XI Wei^{1,2}, JIANG Xiaowen³

(1. Digital Grid Research Institute, CSG, Guangzhou 510670, China; 2. Electric Power Research Institute, CSG, Guangzhou 510663, China;
3. College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China)

Abstract: The application of multi-core processors in the embedded power terminal system can significantly improve the system performance, whereas will also bring about the increasing of the reaction time and the energy consumption. Aiming at the operating characteristics of the power industrial control terminal system, this paper proposes a time-trigger-based energy-saving scheduling method for multi-core systems. An intelligent scheduling scheme is established to ensure the requirements of real-time performance for multi-core systems. Also, two common system-level energy management technologies, such as the dynamic voltage and frequency regulation (DVFS) and dynamic power management (DPM), are adopted to reduce the power consumption of multi-core processors.

Key words: real-time performance; intelligent scheduling; system power consumptions

随着集成电路工艺的不断发展和进步,单位面积晶圆上可集成的晶体管数目呈指数级增长,在大幅提升多核处理器性能和满足现代社会高性能复杂计算需要的同时,也带来了功耗剧增的问题。为了推广应用边缘计算技术和容器技术,电力工控终

端日益趋向采用嵌入式多核芯片以满足新技术的计算性能要求^[1]。但对于在高温恶劣工作环境下工作的电力工控终端而言,多核处理器芯片的功耗不但对系统可靠性有突出影响,还可能直接决定依赖电池供电的电力物联网终端的可用时间^[2-4]。因此,结

收稿日期:2019-12-21;修回日期:2020-10-01

基金项目:国家重点研发计划(2018YFB0904900,2018YFB0904902)

通信作者:蒋小文(1990-),男,博士,助理研究员,主要从事多核调度、系统芯片及边缘计算研究;E-mail:xiaowen_jiang@zju.edu.cn

合电力工控业务场景特征,研究能满足性能要求的低功耗系统,是当前亟待解决的问题。

为了最大化发挥多核处理器的优势,同时满足系统实时性的要求,该文采用相应的调度策略解决功耗和实时性问题。其中调度算法主要分为3类,基于数学规划方法;基于启发式方法;基于随机搜索方法。其中整数线性规划(integer linear programming, ILP)和混合整数线性规划(mixed integer linear programming, MILP)在实际生活得到广泛的应用,该方法适合问题规模相对较小的模型,同时能得到问题的最优解^[5-7]。任务复制算法、循环调度算法和任务组合算法是基于启发式方法的3种经典应用^[8-10]。基于随机搜索通常是一类元启发式算法,作为一种新型调度算法,受到了广大研发人员的关注^[11-12]。

多核系统的能耗问题一直是产品能否成功的关键,早期研究主要利用硬件布局优化来减少系统功耗。近年来,研发人员偏向软件层面来优化系统功耗^[13-15]。该文针对中小问题规模,提出一种基于时间触发的面向电力终端多核系统节能调度方法。根据实际应用创建任务模型,并结合 MILP 方法建立一套能效调度机制,得到一个满足约束条件前提下的能耗最小化的最优解。

1 系统模型和问题定义

首先给出研究对象的软硬件模型,包括多核片上系统架构和应用程序的任务模型;然后描述针对任务在硬件上运行所涉及的系统能耗模型;最后给出该文所解决问题的完整定义。

1.1 硬件模型

该文采用经典的多核片上系统架构^[16],如图1所示。该架构支持 n 个处理器,每个处理器的存储区域相对独立,处理器之间通过系统总线进行数据传输,该总线同时支持时间触发和非抢占模式。针对不同的应用场景,该架构支持 M 级电压频率(V/F)等级,则在这 M 个(V/F)等级下系统功耗分别为 $\{P_1, P_2, \dots, P_M\}$ 。

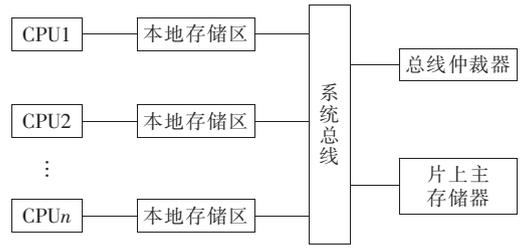


图1 硬件架构

Figure 1 Hardware architecture

1.2 任务模型

该文对周期应用程序进行分析,并引入一个周期任务集合 T ,该任务集合由 K 个应用程序组成,每个应用程序的周期也不相同。对于任何包含于 T 的子集合,其所有任务的完成时间有相同的截止期限。为了表征系统执行完所有任务的时间,用 G 表示公周期,即执行完所有任务的最小时间。系统只有在一个公周期内调度所有任务并执行完毕,这样才被认为是一次可靠的调度结果^[15,17]。

对于任意一个周期应用程序,其要映射到处理器上进行相应的数据操作。数据操作分为计算任务和通信任务,计算任务负责处理多核系统的数据计算工作,通信任务负责处理多核系统的数据传输工作,通常认为计算任务和通信任务可以同时进行。该文用 $T_{i/m}$ 表征处理器 i 在最高(V/F)等级下执行计算任务的最长时间。

在多核片上系统平台上,不同时间触发的实时应用程序启动时间特征不同,根据任务启动的时间特征对任务进行分类,如图2所示。

1) 纯周期任务。应用程序的任务严格周期启动,例如电力工控终端的采集测量业务具有周期性,对周波信号保持定时启动采样,其特点是任意2个相邻任务的时间差是固定的,如图2(a)所示。

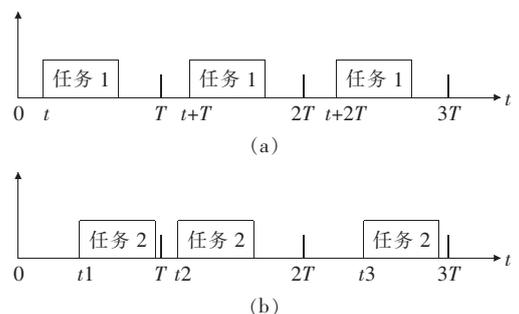


图2 纯周期和非纯周期任务实例

Figure 2 Purely periodic and impurely periodic task instances

2)非纯周期任务。应用程序的任务不按照严格周期启动,例如电力工控终端的保护控制业务,其特点是任务间的时间差不固定,但要保证任务在截止时间前完成,如图 2(b)所示。

1.3 功耗模型

芯片功耗模型具有普遍性,该文采用经典的功耗模型^[18-19]。多核片上系统支持 DVFS 和 DPM 电源管理技术。系统能耗包括计算能耗和通信能耗;CPU 状态包括睡眠模式、空闲模式和运行模式;共享总线状态包括运行模式和空闲模式。该文采用 DVFS 技术,对于任意一个 CPU,其动态功耗、静态功耗和时钟频率为

$$P_D = C_{load} \cdot V_{DD}^2 \cdot f \quad (1)$$

$$P_J = V_{DD} \cdot I_{peak} \quad (2)$$

$$f = K \cdot (V_{DD} - V_e)^2 / V_{DD} \quad (3)$$

式(1)~(3)中 C_{load} 为后级电路等效的电容负载大小; V_{DD} 为电源电压; I_{peak} 为电路的漏电流; K 为一个与电路结构有关的电气常数; V_e 为系统正常运行的极限电压值。因此 CPU 总功耗为

$$P_Z = P_D + P_J + P_B \quad (4)$$

式中 P_B 为额外的功率损耗,其维持处理器的基本运行。根据式(4)得出某一任务在 CPU_{*n*} 内, V/F 等级为 m 下的功耗,即

$$P_A = P_Z \cdot T_{i/m} \quad (5)$$

CPU 会在睡眠模式和工作模式切换,由于睡眠状态中处理器上下文等信息不被保存,所以切换过程中伴随着一定的能量损耗。如果状态切换的能耗开销大于能耗节余,那么显然把处理器转入睡眠状态可能不会节能。为分析处理器是否值得模式切换,引入临界时间,其为基于处理器空闲时功耗所计算的最小空闲周期(或“折中”时间)。通过对临界时间大小的比较,来选择 CPU 的工作模式。如果 CPU 空闲时间大于临界时间,那么 CPU 切换进入睡眠模式更节能;反之,进入空闲模式更节能。临界时间定义为

$$T_c =$$

$$\max\{(E_{switch} - P_s \cdot t_{switch}) / (P_k - P_s), t_{switch}\} \quad (6)$$

式中 t_{switch} 、 E_{switch} 分别为 CPU 在工作模式和睡眠模式之间切换的时间和能量损耗; P_k 、 P_s 分别为 CPU 处于空闲模式和睡眠模式时的功耗。

CPU 在空闲模式和睡眠模式下的能耗为

$$W_k = P_k \cdot t_k \quad (7)$$

$$W_s = P_s \cdot t_s \quad (8)$$

式中 t_k 、 t_s 分别为 CPU 的空闲时间和睡眠时间。根据 CPU 时序特征指定一个调度表 Q ,CPU 总计功耗为

$$E_{all}(Q) = P_A(Q) + W_k(Q) + W_s(Q) + E_{switch}(Q) \quad (9)$$

最后,存在依赖关系的任务间数据传输由 CPU、总线和多核片上系统中的共享存储区域这 3 个组件来完成。

1.4 问题模型

任务分配主要包括调度和任务—CPU 映射 2 部分。该文假设任务—CPU 映射是已知的,主要关注能效调度问题。如图 3 所示,针对 DVFS 和 DPM 触发的多核片上系统,其中包含一系列纯周期和非纯周期的应用程序,以任务—CPU 映射和任务—处理器功耗统计信息作为输入,目标是建立一套基于时间触发的能效调度机制,根据该机制可确定一组最优调度结果和任务分配方案,使系统在公周期内的总能耗最小,同时满足系统的时序约束条件。

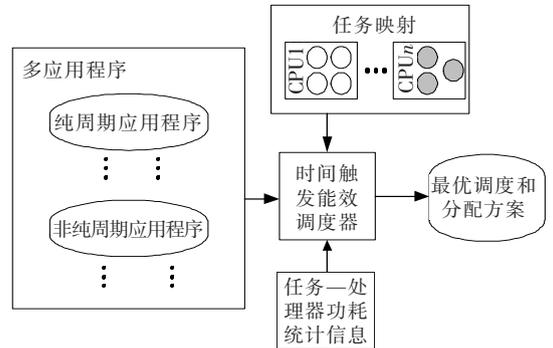


图 3 时间触发多核实时系统的能效调度模型
Figure 3 Energy efficiency scheduling model for time-triggered multi-core real-time systems

2 MILP 方法

考虑由纯周期任务和非纯周期任务组成的多个

应用程序,同时采用 DVFS 和 DPM 技术实现节能调度。该文构建了一个 MILP 模型来解决任务启动是否严格周期的问题,根据该 MILP 模型可以得到一个使系统总能耗最小的调度方案。

针对特定的多核片上系统,结合周期任务集合和任务—CPU 映射情况,考虑具体任务运行的时序特征,找到一个最优调度和任务分配方案,使得系统总能耗最小。根据前文给出的模型,构建符合要求的 MILP 模型。下面定义相关变量。

$X_{n/pm}$:任务 p 映射到处理器 n 上,如果任务 p 对应的 V/F 等级为 m ,则 $X_{n/pm} = 1$,否则 $X_{n/pm} = 0$ 。

$T_{sn/p}$:映射到处理器 n 上的任务 p 的启动时间。

$T_{snN/pq}$:通信任务(任务 p 和任务 q 进行通信)的启动时间(任务 p 映射到处理器 n ,任务 q 映射到处理器 N)。

$C_{sn/pq}$:如果任务 p 的启动时间早于任务 q ,则 $C_{sn/pq} = 1$,否则 $C_{sn/pq} = 0$ 。

最后根据给定的多任务图和任务—CPU 映射,构建的 MILP 模型为

$$E_{\text{all}_G_i} = E_{\text{all}} + E_{\text{cm}} \quad (10)$$

式中 E_{cm} 为一个 G_i 内总通信能耗开销,该模型遵循以下约束条件。

1)任务 V/F 等级约束。DVFS 技术使得每个任务在运行期间有固定的 V/F 等级,直到任务执行完毕才可以变更,因此约束关系为

$$\sum_{m=1}^M X_{n/pm} = 1 \quad (11)$$

2)任务截止时间期限约束。

$$T_{sn/p} + x^n + \sum_{m=1}^M (X_{n/pm} \cdot T_{i/m}) \leq d_i \quad (12)$$

式中 x^n 为任务在 CPU _{n} 上切换的时间开销和 DVFS 切换时间开销; d_i 为任务的截止时间期限。

3)对于非纯周期启动的任务,不存在严格时间约束,而对于纯周期启动的任务,存在如下约束。

$$T_{sn/(p+1)} - T_{sn/p} = p_i \quad (13)$$

式中 p_i 为任务的周期。

4)如果任务 p 和任务 q 映射到同一个处理器

上,则存在以下约束。

$$T_{sn/p} + x^n + \sum_{m=1}^M (X_{n/pm} \cdot T_{i/m}) \leq T_{sn/q} \quad (14)$$

5)如果任务 p 和任务 q 映射到不同的处理器上(任务 p 为源节点任务,任务 q 为目标节点任务),则存在以下约束。

①只有当任务 p 执行完毕,通信任务 $T_{snN/pq}$ 才开始执行,存在约束为

$$T_{sn/p} + x^n + \sum_{m=1}^M (X_{n/pm} \cdot T_{i/m}) \leq T_{snN/pq} \quad (15)$$

②只有当通信任务 $T_{snN/pq}$ 执行完毕,任务 q 才开始启动,存在约束为

$$T_{snN/pq} + x^n + C_{p,q} \leq T_{sn/q} \quad (16)$$

式中 $C_{p,q}$ 为任务 p 和任务 q 的通信时间。

6)非抢占约束。对于没有前后依赖关系的计算任务和通信任务,两者可以并行执行,需要注意任意一个 CPU 不能同时调度映射于相同 CPU 的计算任务实例,该约束同样适合于总线上的通信任务,线性化过程可以参考文献[18]。

根据以上建立的 MILP 模型,多核系统可以通过 ILP 求解器进行相应参数的求解,使系统总能耗 $E_{\text{all}_G_i}$ 达到最优。

3 仿真实验和结果分析

3.1 实验平台简介

为了证明该文提出方法的可行性,该实验硬件平台配置为 Intel i7-6700HQ CPU,主频 2.6 GHz,8 GB 内存,64 位 Windows 10 家庭版操作系统,8 核多处理器架构。该实验采用了同文献[18-20]一致的 70 nm 制程技术的 CPU 功率参数。

3.2 实验方法与实验结果分析

该实验采用的 CPU 功耗模型基于经典的 70 nm 制程技术,其精确度已被 SPICE 仿真模型验证[18-19]。该 CPU 支持 5 个电压等级,电压范围分布为 650~850 mV,每个电压等级相差 50 mV。对应电压等级的频率范围为 1 010~2 100 MHz。根据文献[19-20]提供的工艺常数和 1.3 节的计算公式,

可以计算出对应等级的动态功耗 P_D 和静态功耗 P_J , 汇总以上数据如表 1 所示。

实验包含 15 个不同的应用程序模拟测试示例。TG1~3 来源于工业, 汽车和消费类的应用^[21]; TG4~6 来源于实际应用程序^[22]; TG7~9 来源于电力终端系统应用^[23]; TG10~15 则是由通用随机任务图生成器 TGFF^[24] 提供的软件包生成。将这 15 个任务图按照相应的规则进行组合, 得到 6 组有一定规模任务数的程序基准测试集 (APPG1~6) 结果。同时设置参数 δ ($\delta \in [0, 1]$) 来表示任意一个程序基准测试集中纯周期任务和非纯周期任务的比率。

对于在时间触发系统上的周期应用程序而言, 除启动时间和截止时间是周期性的, 同一个安全关键任务 (例如数据采集任务) 的不同实例的启动时间也必须是周期性的。基于此机制, 时间触发多核系统的能耗优化调度研究同样采用了整数线性规划, 并在建模过程中假设所有任务启动时间间隔等于任务周期, 以保证子任务执行的纯周期特性。实验中选取了文献[18]中的基于纯周期特性的混合整数线性规划节能调度方法 (Pure-MILP) 作为对比基准。

为直观呈现实验效果, 将所提 MILP 方法下调度得到的平均功耗值按传统 Pure-MILP 方法下的平均功耗值进行归一化后, 将 6 种测试集下不同纯周期任务占比的功耗与传统 Pure-MILP 方法下的功耗绘制如图 4 所示。

由图 4 可知, 与 Pure-MILP 相比, 所提 MILP

表 1 处理器功耗模型参数

Table 1 Processor power model parameters

电压等级	$V_{DD} /$ mV	$F /$ MHz	$P_D /$ mW	$P_J /$ mW
1	850	2 100	655.5	462.7
2	800	1 810	489.9	397.6
3	750	1 530	370.4	343.3
4	700	1 260	266.7	290.1
5	650	1 010	184.9	246

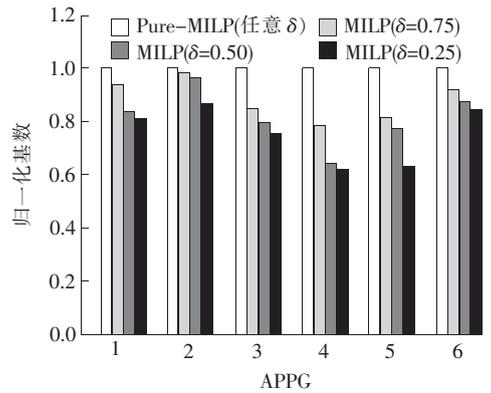


图 4 Pure-MILP 与所提 MILP 方法的平均功耗对比

Figure 4 Comparison of average power consumption between the pure-MILP and the proposed MILP methods

方法的能耗下降显著, 在 6 种程序基准测试集中, 各种不同的周期任务占比下所提 MILP 的总能耗均有不同程度的下降, 非周期任务占比低时表现得尤为明显; 与 Pure-MILP 方法简单假定所有任务严格周期启动不同, 所提方法考虑了具体任务的启动执行特征, 并结合 DVFS 和 DPM 技术, 提供了高效执行策略, 有效地减少了系统能耗开销。

对于不同的应用程序测试集, MILP 和 Pure-MILP 方法平均功耗比也存在较大的差异, APPG1 的平均功耗比最小, APPG5 的平均功耗比最大, 这是因为 APPG5 的任务数较多, 完成任务调度的功耗就相应增加。Pure-MILP 方法在不同的 δ 下平均能耗为一定值; 而 MILP 方法的平均能耗随 δ 变小而减小。这是因为 MILP 方法根据具体任务的严格性与否建立了实际的任务模型, 并且能够执行正确的调度策略。 δ 越小, 则对应的非纯周期任务就越多, 系统的调度灵活性就越强, 系统功耗下降的就越明显。

4 结语

针对实际应用程序是否严格周期启动的特点, 研究了基于时间触发的电力终端多核系统的能耗调度问题, 提出了一个能表征应用内任务是否严格周期启动的模型, 并将该模型和具体的硬件架构和能量模型结合, 描述了一个时间触发多核系统能耗优

化调度问题,提出并构建了 MILP 模型的方法。实验数据表明,所提 MILP 方法可以有效降低系统能耗,同时得到最优节能调度方案。

参考文献:

- [1] 路艳巧,孙翠英,曹红卫,等. 基于边缘计算与深度学习的输电设备异物检测方法[J]. 中国电力,2020,53(6):27-33.
LU Yanqiao, SUN Cuiying, CAO Hongwei, et al. Foreign body detection method for transmission equipment based on edge computing and deep learning[J]. Electric Power,2020,53(6):27-33.
- [2] 唐俊熙,王梓耀,张俊潇,等. 基于文化蚁群算法的高可靠多电源配电网规划模型[J]. 电力系统保护与控制,2020,48(7):99-107.
TANG Junxi, WANG Ziyao, ZHANG Junxiao, et al. High reliability multi-power distribution network planning model based on memetic ant colony algorithm[J]. Power System Protection and Control, 2020, 48(7):99-107.
- [3] 刘远龙,潘筠,王玮,等. 用于泛在电力物联网的配电变压器智能感知终端技术研究[J]. 电力系统保护与控制,2020,48(16):140-146.
LIU Yuanlong, PAN Jun, WANG Wei, et al. Research on intelligent sensing terminal technology of a distribution transformer for ubiquitous power internet of things [J]. Power System Protection and Control, 2020, 48(16):140-146.
- [4] 陈江兴,梁良,付俊峰,等. 基于大数据的智能电网数据调度与快速分发方法研究[J]. 电测与仪表,2020,57(6):88-93.
CHEN Jiangxing, LIANG Liang, FU Junfeng, et al. Research on smart grid data scheduling and fast distribution method based on big data[J]. Electrical Measurement & Instrumentation,2020,57(6):88-93.
- [5] HUANG Kai, YU Min, ZHANG Xiaomeng, et al. ILP based multithreaded code generation for simulink model [J]. Ieice Transactions on Information & Systems, 2014,97(12):3072-3082.
- [6] Walla G, Herkersdorf A, Enger A S, et al. An automotive specific MILP model targeting power-aware function partitioning[C]//International Conference on Embedded Computer Systems: Architectures, Modeling, and Simulation, Agios Konstantinos, Greece, IEEE, 2014.
- [7] Venugopalan S, Sinnen O. ILP formulations for optimal task scheduling with communication delays on parallel systems[J]. IEEE Transactions on Parallel & Distributed Systems, 2014, 26(1):142-151.
- [8] Kwok Y K, Ahmad I. Static scheduling algorithms for allocating directed task graphs to multiprocessors[J]. ACM Computing Surveys, 1999, 31(4):406-471.
- [9] Ahmad I, Kwok Y K. On exploiting task duplication in parallel program scheduling[J]. IEEE Transactions on Parallel & Distributed Systems, 1998, 9(9):872-892.
- [10] Chan W Y, Li C K. Heterogeneous dominant sequence cluster (HDSC): a low complexity heterogeneous scheduling algorithm[C]//IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, Victoria, BC, Canada: IEEE, 1997.
- [11] Zhang C Y, Li P G, Rao Y Q, et al. A very fast TS/SA algorithm for the job shop scheduling problem [J]. Computers & Operations Research, 2008, 35(1):282-294.
- [12] Chan F T, Prakash A, Ma H, et al. A hybrid Tabu sample-sort simulated annealing approach for solving distributed scheduling problem[J]. International Journal of Production Research, 2013, 51(9):2602-2619.
- [13] Liu Y H, Veeravalli B, Viswanathan S. Novel critical-path based low-energy scheduling algorithms for heterogeneous multiprocessor real-time embedded systems [C]//International Conference on Parallel and Distributed Systems, Hsinchu, China, 2007.
- [14] Ghadaksaz E, Safari S. Storage capacity for EDF-ASAP algorithm in energy-harvesting systems with periodic implicit deadline hard real-time tasks[J]. Journal of Systems Architecture, 2018, 89:10-17.
- [15] 陈刚,关楠,吕鸣松,等. 实时多核嵌入式系统研究综述 [J]. 软件学报, 2018, 29(7):2152-2176.
CHEN Gang, GUAN Nan, LV Mingsong, et al. Over-

- view of real-time multi-core embedded system research [J]. *Journal of Software*,2018,29(7):2152-2176.
- [16] Peng D T,Shin K G,Abdelzaher T F. Assignment and scheduling communicating periodic tasks in distributed real-time systems[J]. *IEEE Transactions on Software Engineering*,1997,23(12):745-758.
- [17] 郭锐锋,吴昊天,邓昌义,等. 一种 WCET 比例空闲时间分配的周期任务低功耗算法[J]. *小型微型计算机系统*,2017,38(8):1856-1860.
- GUO Ruifeng,WU Haotian,DENG Changyi, et al. A low-power algorithm for periodic tasks of WCET proportional idle time allocation[J]. *Journal of Chinese Mini-Micro Computer Systems*, 2017, 38 (8): 1856-1860.
- [18] Chen G, Huang K, Knoll A. Energy optimization for real-time multiprocessor system-on-chip with optimal DVFS and DPM combination[J]. *ACM Transactions on Embedded Computing Systems*,2014,13(3S):1-21.
- [19] Wang W, Mishra P. Leakage-aware energy minimization using dynamic voltage scaling and cache reconfiguration in real-time systems[C]//International Conference on VLSI Design,Bangalore,India,2010.
- [20] Martin S M, Flautner K, Mudge T, et al. Combined dynamic voltage scaling and adaptive body biasing for lower power microprocessors under dynamic workloads [C]//IEEE/ACM International Conference on Computer-Aided Design,San Diego,USA,2002.
- [21] Dick R. Embedded system synthesis benchmarks suite (E3S) [EB/OL]. <http://ziyang.eecs.umich.edu/~dickrp/e3s/>,2018-09-30.
- [22] Tobita T,Kasahara H. A standard task graph set for fair evaluation of multiprocessor scheduling algorithms [J]. *Journal of scheduling*,2002,5(5):379-394.
- [23] 陆荣超. 配电自动化终端设备在电力配网自动化的应用研究[J]. *通讯世界*,2018(12):179-180.
- LU Rongchao. Research on application of distribution automation terminal equipment in power distribution automation[J]. *Communication World*,2018(12):179-180.
- [24] Dick R P,Rhodes D L,Wolf W. TGFF:task graphs for free[C]//Hardware/Software Codesign,Seattle,USA:IEEE,1998.