



计算机科学

COMPUTER SCIENCE

ARM 架构服务器的 CPU 功耗模型研究

金育妍, 余天豪, 王松波, 林伟伟, 潘宇聪

引用本文

金育妍, 余天豪, 王松波, 林伟伟, 潘宇聪. [ARM 架构服务器的 CPU 功耗模型研究](#)[J]. 计算机科学, 2022, 49(10): 59-65.

JIN Yu-yan, YU Tian-hao, WANG Song-bo, LIN Wei-wei, PAN Yu-cong. [CPU Power Model for ARM Architecture Cloud Servers](#)[J]. Computer Science, 2022, 49(10): 59-65.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[移动云计算中面向能耗优化的资源管理](#)

Energy Optimization Oriented Resource Management in Mobile Cloud Computing

计算机科学, 2020, 47(6): 247-251. <https://doi.org/10.11896/jsjx.190400020>

[移动边缘计算中具有能耗优化的任务迁移策略](#)

Task Migration Strategy with Energy Optimization in Mobile Edge Computing

计算机科学, 2020, 47(6): 260-265. <https://doi.org/10.11896/jsjx.190400074>

[区块链共识算法效能优化研究进展](#)

Research Advance on Efficiency Optimization of Blockchain Consensus Algorithms

计算机科学, 2020, 47(12): 296-303. <https://doi.org/10.11896/jsjx.200700020>

[强化学习下能耗优化的虚拟机放置策略](#)

Virtual Machine Placement Strategy with Energy Consumption Optimization under Reinforcement Learning

计算机科学, 2019, 46(9): 291-297. <https://doi.org/10.11896/j.issn.1002-137X.2019.09.044>

[一种基于新能源驱动的存储系统的能耗优化方案](#)

Energy Consumption Optimization Scheme for New Energy-driven Storage System

计算机科学, 2018, 45(7): 66-72. <https://doi.org/10.11896/j.issn.1002-137X.2018.07.010>

ARM 架构云服务器的 CPU 功耗模型研究

金育妍¹ 余天豪¹ 王松波¹ 林伟伟^{1,3} 潘宇聪²

1 华南理工大学计算机科学与工程学院 广州 510641

2 广东省高性能计算重点实验室 广州 510033

3 鹏程实验室 广东 深圳 518066

(jingyy.kk@outlook.com)

摘要 云服务器的功耗模型是云数据中心能耗优化研究的重要内容之一。CPU 功耗模型是云服务器功耗模型的重要组成部分,然而现有 CPU 功耗模型没有考虑 CPU 的异构性,如缺乏对 ARM 架构服务器 CPU 功耗模型的研究。在调研分析现有的 ARM 架构 CPU 功耗模型的基础上,提出了一种面向 ARM 架构的新 CPU 功耗模型——基于混合建模的 CPU 功耗模型(Hybrid Based Model, HBM)。该功耗模型综合考虑了 CPU 利用率和 CPU 性能事件等建模特征,相比现有的测算精度很高的基于性能计数器的 CPU 功耗模型, HBM 的测算精度与其相近且模型训练成本更低,更适合 ARM 服务器的 CPU 功耗建模。文中使用 Sysbench 负载工具对所提 HBM 进行实验验证,实验结果表明, HBM 的平均相对误差(MRE)在 1% 以内,具有良好的测算精度。此外,还针对 x86 和 ARM 架构服务器进行了交叉实验,实验结果表明不同架构服务器的 CPU 功耗行为相异,应当使用不同的 CPU 功耗建模方法。

关键词: 能耗优化; 功耗模型; CPU 异构性; ARM 架构

中图法分类号 TP393

CPU Power Model for ARM Architecture Cloud Servers

JIN Yu-yan¹, YU Tian-hao¹, WANG Song-bo¹, LIN Wei-wei^{1,3} and PAN Yu-cong²

1 School of Computer Science and Engineering, South China University of Technology, Guangzhou 510641, China

2 Guangdong Science and Technology Infrastructure Center, Guangzhou 510033, China

3 Pengcheng Laboratory, Shenzhen, Guangdong 518066, China

Abstract The power model of cloud server is one of the important contents of the research on the energy consumption optimization of cloud data center. The CPU power model is an important part of the power models of cloud servers. However, the existing CPU power models do not consider the CPU heterogeneity, such as lack of research on the CPU power model of ARM architecture cloud servers. Based on the investigation and analysis of existing ARM architecture CPU power models, this paper proposes a new CPU power model oriented to the ARM architecture, namely the hybrid based model(HBM). HBM comprehensively considers modeling features such as CPU utilization and CPU performance events. Compared with existing PMC based model with high measurement accuracy, HBM has similar measurement accuracy and lower model training cost. Thus, HBM is more suitable for CPU power modeling of ARM servers. This paper uses the Sysbench benchmark to verify HBM, and experimental results show that the mean relative error(MRE) of HBM is within 1%, which means HBM has high measurement accuracy. Cross-experiments are also conducted for x86 and ARM architecture servers, and experimental results show that the CPU power behaviors of servers with different architectures are not the same, thus different CPU power modeling methods should be used.

Keywords Energy consumption optimization, Power model, CPU heterogeneity, ARM architecture

到稿日期:2021-08-12 返修日期:2022-03-02

基金项目:国家自然科学基金(62072187);广东省重点研发计划(2021B0101420002);广东省基础与应用基础研究重大项目(2019B030302002);广州市科技计划(202007040002);广州市开发区国际合作项目(2020GH10);广东省高性能计算重点实验室 2019 年课题资助;鹏程实验室重大任务项目(PCL2012A09)

This work was supported by the National Natural Science Foundation of China(62072187), Key-Area Research and Development Program of Guangdong Province(2021B0101420002), Guangdong Major Project of Basic and Applied Basic Research(2019B030302002), Guangzhou Science and Technology Plan(202007040002), Guangzhou Development Zone International Cooperation Project(2020GH10), Guangdong Provincial Key Laboratory of High Performance Computing and 2019 Project Funding and Major Key Project of PCL(PCL2012A09).

通信作者:林伟伟(linww@scut.edu.cn)

1 引言

近年来,云计算技术的发展异常迅速,云数据中心的规模也在不断扩大,在方便使用云资源的同时也带来了巨大的能量消耗。在此背景下,如何对云数据中心的能耗进行有效

管理成为数据中心管理的重要研究方向,而面向云服务器的功耗模型是数据中心能耗研究的基础,其中 CPU 功耗模型是服务器功耗模型的重要组成部分。

目前服务器 CPU 的主流架构主要有 x86 和 ARM 两种,表 1 列出了两者的主要区别^[1-3]。

表 1 x86 和 ARM 架构的区别
Table1 Difference between x86 and ARM architectures

对比项	x86 架构 CPU	ARM 架构 CPU
指令集	复杂指令集(CISC)	精简指令集(RISC)
动态-静态接口(DSI)	DSI 放置位置更高,更多的工作交由硬件处理	DSI 放置位置更低,更多的工作交由软件处理
Load/Store 工作是否由专门的指令执行	其他指令可以执行 Load/Store 工作	专门指令执行 Load/Store 工作
指令长度	可变量指令长度	定长指令长度
寻址模式数量	更多的寻址模式	有限的寻址模式
通用寄存器数量	有限的通用寄存器	大量通用寄存器
效率	采用 CISC 流水线指令集,完成简单任务的效率较低,完成复杂任务的效率较高	采用 RISC 流水线指令集,完成简单任务的效率较高,完成复杂任务的效率较低
性能	明显优于 ARM	较好
功耗	明显高于 ARM	很低

从表 1 可以看出 x86 和 ARM 在指令集体系架构上有着明显区别,例如 x86 会将更多的工作交由硬件实现,ARM 则是将 x86 中微体系架构完成的部分工作交给编译器来完成,从而降低了硬件的复杂度^[3]。另一个差异表现为,RISC 架构处理器可以使用每周指令数(Instruction Per Cycle,IPC)来衡量处理器的运行速度;但 CISC 架构处理器并不适合用 IPC 来反映运行速度,因为 CISC 架构的一条指令可能需要处理器执行大量工作,而 RISC 架构的一条指令通常只对应一个功能或者操作^[4]。

ARM 架构有时会使用 big. LITTLE^[5]异构核架构,这与 x86 架构 CPU 使用的同构核架构有很大区别。另一方面,ARM 架构 CPU 的核心数通常比 x86 架构 CPU 的核心数多,而且 CPU 的 cache 容量也有所不同^[6],不同的核心数和 cache 会带来不同等级的执行资源冲突。Blem 等^[7-8]通过实验验证了 x86 和 ARM 架构的性能和功耗差异。

当前 ARM 架构服务器在云数据中心占有越来越重要的位置,基于 x86 和 ARM 架构的异构云环境是未来的主要趋势^[9],因此,研究针对不同架构的 CPU 功耗测算方法具有重要意义。

指令集是 CPU 的重要性能指标。Vasilakis^[10]分析了 ARM 架构 CPU 的指令集的特性,然后建立了指令级别的 CPU 功耗模型。该模型拥有很高的测算精度,但是指令集数据的获取难度极大,指令级别的功耗粒度太细,模型在实际应用中缺乏易用性。与 x86 架构 CPU 不同,ARM 架构 CPU 使用 big. LITTLE 异构核架构时,不同核心,可能会具有不同的 CPU 性能特征。在此基础上,Obukhova 等^[11]针对某个特定的 CPU 内核进行功耗建模,整体的 CPU 功耗为所有 CPU 核心的功耗总和。该模型利用 CPU 核心利用率这类易于获取的指标进行建模,模型易用性较好,可以应用于异构核环境。但是该模型只适用于特定 CPU 频率下的功耗测算,不同 CPU 运行频率需要重新训练模型,模型通用性不强。为此,Walker 等^[12]提出了结合 CPU 利用率和 CPU 运行频率的功耗模型,其可以测算任何运行频率区间的 CPU 功耗。但是仅靠 CPU 利用率可能达不到理想的测算效果,因为 CPU 功耗

还与 CPU 执行的具体运算有关,例如浮点运算消耗的功耗明显大于整数运算^[13]。Sankaran 等^[14]利用基于性能计数器的线性建模方法来测算 ARM 架构 CPU 的功耗,实验结果表明该方法具有很高的测算精度,但是模型选取的性能事件较多,整个建模过程需要较大的时间成本。Sagi 等^[15]还提出在性能事件采样率很高的情况下,CPU 功耗和 CPU 性能事件不是严格的线性关系,因此他们增加了非线性因素以提高 CPU 功耗模型的测算精度,但该模型复杂度很高且使用场景具有局限性。

另一方面,现代 CPU 中应用了很多节能技术,尤其是以低功耗为设计目标的 ARM 架构 CPU。Zhang 等^[16]提出基于 CPU C-states 的功耗模型,该模型考虑了不同 CPU C-states 状态下的性能特征,结合 CPU 利用率可以较为全面地刻画 CPU 的性能活动,具有良好的测算精度。文献^[17]指出 CPU 功耗和 CPU 运行的线程数有一定的线性关系,且利用基于 CPU 运行线程数的线性模型近似表示 CPU 功耗,中该模型考虑的因素过于单一,难以满足日益复杂的现代 CPU 架构。

本文的主要贡献如下:

(1)分析了现有的主流 ARM 架构 CPU 功耗模型,并在此基础上提出了一种新的基于 ARM 架构云服务器的 CPU 功耗模型——基于混合建模的 CPU 功耗模型(HBM)。

(2)使用 Sysbench 负载工具开展模型的验证实验,实验分析了 HBM 的 MRE 和建模时间成本,实验结果表明 HBM 的 MRE 在 1%以内,测算精度表现良好,且建模时间成本远小于基于性能计数器的 CPU 功耗模型。

(3)最后还针对 x86 和 ARM 架构服务器进行了交叉实验,实验结果表明特定平台下的 CPU 功耗模型在其他平台表现不佳,说明不同架构服务器的 CPU 功耗行为明显不同,在进行 CPU 功耗测算时考虑 CPU 异构性是很有必要的。

本文第 2 节主要介绍了 ARM 架构 CPU 功耗建模方法和模型分析,该部分内容是新模型构建的基础;第 3 节阐述了基于 ARM 架构云服务器的 CPU 功耗模型 HBM 的构建过程;第 4 节为 CPU 功耗模型验证实验,主要是 HBM 的验证实验和交叉验证实验;最后总结全文。

2 CPU 功耗建模方法和模型分析

2.1 建模方法分析

ARM 架构的 CPU 功耗建模方法主要分为 3 个层面,从底层到高层分别为基于指令级别建模、基于性能计数器建模和基于 CPU 利用率建模^[13]。

基于指令级别建模^[10]是最底层的建模方法,拥有最好的测算精度,其主要利用单个指令的功耗和指令之间的功耗关系对 CPU 功耗进行建模。该方法需要研究人员对指令集足够熟悉,而且指令级别的功耗难以获取,因此该方法难以应用于实际应用中。基于 CPU 利用率建模^[11-12]是最顶层的建模方法,测算精度欠佳,其主要将 CPU 功耗表示为基于 CPU 利用率的相关函数。该方法易于实现并广泛应用在实际应用中。基于性能计数器建模^[14-15]的难度低于基于指令级别建模,测算精度优于基于 CPU 利用率建模,应用十分广泛。

综上,本文主要研究基于 CPU 利用率建模和基于性能计数器建模这两种实际使用价值较高的建模方法。

2.2 基于 CPU 利用率的 CPU 功耗模型

Walker 等^[12]提出了基于 CPU 利用率的 CPU 功耗模型(Utilization Based Model, UBM),在特定运行频率下, CPU 的功耗可以表示为:

$$P_{UBM} = c_0 + c_1 U + c_2 U^2 \quad (1)$$

其中, c_0 , c_1 和 c_2 代表模型参数, U 代表 CPU 利用率。为了提高模型的测算精度, Walker 等^[12]后续将运行频率也作为拟合变量,此时 UBM 的计算式为:

$$P_{UBM} = (c_0 + c_1 f) + (c_2 + c_3 f)U + (c_4 + c_5 f)U^2 \\ = c_0 + c_1 f + c_2 U + c_3 Uf + c_4 U^2 + c_5 U^2 f \quad (2)$$

其中, c_0 , c_1 , c_2 , c_3 , c_4 , c_5 代表模型训练参数, f 代表 CPU 运行频率, U 代表 CPU 利用率。UBM 有效地考虑了 CPU 利用率和运行频率对 CPU 功耗的影响,可以测算任何运行频率区间的 CPU 功耗。但是仅靠 CPU 利用率可能达不到理想的测算效果,因为 CPU 功耗还与 CPU 执行的具体运算有关,例如浮点运算消耗的功耗明显大于整数运算^[13], CPU 执行的运算类型可以通过性能计数器获取的性能事件来确定。

2.3 基于性能计数器的 CPU 功耗模型

性能计数器(Performance Monitoring Counter, PMC)是一组专用的寄存器,存储着与计算机硬件活动相关的计数。

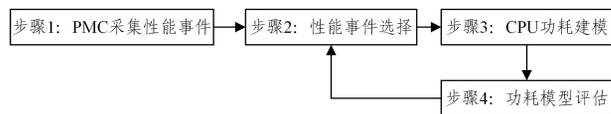


图1 基于 PMC 功耗建模的主要步骤

Fig. 1 Main steps of PMC-based power consumption modeling

由于性能事件数量远多于 PMC 数量,在获取性能事件时,可以采取一个 PMC 负责监控一个事件(One Counter One Event, OCOE)的方式多次运行负载或者多路复用一次运行负载这两种方法。OCOE 多次运行负载测算精确,但是需要运行负载多次,而多路复用一次运行负载只需要一次运行,但是测算误差较大。

PMC 收集性能事件之后,进入性能事件选择阶段。关于如何选择合适的用于建模的性能事件,首先通过式(3)计算

每个性能事件与 CPU 功耗之间的相关系数。

$$r(x, y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3)$$

其中, x 为性能事件, y 为 CPU 功耗, n 为采集的数据数目,相关系数 $r(x, y)$ 介于 $[-1, 1]$ 区间。相关系数越接近 1,表示两者正相关性越好;越接近 -1,则表示两者之间的负相关性越好;而相关系数越接近 0,表示两者的相关性越差。设定一个相关系数阈值,超过该阈值的性能事件都可以归入建模事件集合中。如文献[18]从数百个性能事件中选择约 50 个性能事件组成建模事件集合。

其次利用筛选后的性能事件集合和 CPU 功耗来进行多元线性回归建模。该模型为基于性能计数器的 CPU 功耗模型(PMC Based Model, PBM),模型拟合方法为最小二乘法,模型计算式如式(4)所示:

$$P_{PBM} = c_0 + \sum_{i=1}^n (c_i e_i) \quad (4)$$

其中, c_0 和 c_i 为模型参数, e_i 为性能事件, n 为用于建模的性能事件数目。 e_i 的值普遍较大,为了计算方便可以按比例缩小处理。最后评估 PBM 模型的测算精度,根据精度表现反馈来不断调整性能事件的选择,从而不断提高模型的测算精度。

2.4 基于多线程的 CPU 功耗模型

文献[17,19]指出 CPU 功耗和 CPU 运行的线程数有一定的线性关系。文献[17]利用线性建模近似表示 CPU 功耗,在特定的 CPU 运行频率下,功耗模型简化如式(5)所示,该模型为基于多线程的 CPU 功耗模型(Thread Based Model, TBM)。

$$P_{TBM} = P_{\min} + P_{tb} + n P_t \quad (5)$$

其中, P_{\min} 为 CPU 静止功耗,可以近似为 0; P_{tb} 为多个线程运行时消耗的最小功耗, P_t 为单个线程消耗的功耗, n 为运行线程数。TBM 考虑了 CPU 功耗和多线程的关系,建模方法简单而且模型容易部署,可以快速实现 CPU 的功耗测算。但是 TBM 考虑的因素过于单一,功耗模型过于理想化,较难满足极其复杂的现代 CPU 架构,需要考虑更多的因素来提升模型的测算精度。

3 CPU 功耗模型构建

3.1 基于多线程和 DVFS 的改进 CPU 利用率

现代 CPU 已经加入了多线程技术和 DVFS,这些特征也与 CPU 功耗紧密相关。据本文调研所知,很少有功耗建模方法同时考虑到多线程技术和 DVFS。例如,文献[20]在对多核服务器进行细粒度功耗建模时没有考虑到多线程技术和 DVFS;文献[19]提出的功耗建模方法充分考虑到了多线程技术,但是忽略了 DVFS 对 CPU 功耗的影响。因此,本文在 CPU 功耗建模过程中考虑加入多线程技术和 DVFS,从而实现更精准的功耗测算。

CPU 运行时对应的运行参数和系统参数,运行参数有 CPU 运行线程数 C_{th} , CPU 运行频率 f_r ;系统参数有 CPU 总线程数 T_{th} , CPU 最大运行频率 $f_{r_{\max}}$ 。CPU 通常存在多个逻辑核心,每个逻辑核心都有自己的 CPU 运行频率,因此 f_r 表示所有逻辑核心的平均 CPU 运行频率。本文将 CPU 的运行

参数集合标记为 $O = \{C_{th}, fr\}$, 系统参数集合标记为 $M = \{T_{th}, fr_{max}\}$ 。

文献[19]中提出 CPU 功耗和 CPU 运行线程数呈线性关系, 而文献[21]中提出 CPU 功耗和运行频率呈非线性关系。当 CPU 在特定的电压和运行频率下运行时, 即在特定的 DVFS 状态时, CPU 功耗可以表示为:

$$P_{cpu} = kv^2 f \quad (6)$$

其中, k 为模型参数, v 为 CPU 运行电压, f 为 CPU 运行频率。当 f 升高时, v 也会随之升高。

为了在 CPU 功耗模型中加入多线程技术和 DVFS, 本文提出换算因子:

$$\Delta(O, M) = \left(\frac{C_{th}}{T_{th}} \epsilon\right) \left(\frac{fr}{fr_{max}}\right)^\eta \quad (7)$$

其中, ϵ 和 η 为模型训练参数, ϵ 反映出 CPU 功耗和 CPU 运行线程数的线性关系, η 反映出 CPU 功耗和 CPU 运行频率的非线性关系。将换算因子和 CPU 利用率相结合可以得出改进的 CPU 利用率:

$$U_\Delta = U \cdot \Delta(O, M) = U \cdot \left(\frac{C_{th}}{T_{th}} \epsilon\right) \left(\frac{fr}{fr_{max}}\right)^\eta \quad (8)$$

相比之前的 CPU 利用率 U , 改进的 CPU 利用率 U_Δ 不单考虑了 CPU 的时间片使用率, 而且还考虑了多线程技术和 DVFS, 更加细粒度地反映了 CPU 的使用程度。

3.2 基于混合建模的 CPU 功耗模型 (HBM)

从上述分析结果来看, UBM, PBM 和 TBM 这 3 种建模方法均有不足之处, 而且在适当的范围内, 考虑因素越多, 模型的测算精度越好。基于上述发现, 本文提出一种新型的基于混合建模的 CPU 功耗模型 (HBM), 有效地将 UBM, PBM 和 TBM 这 3 种模型结合起来。

$$P_{HBM} = P_0 + \sum_{i=1}^n (c_i e_i) + \alpha U + \beta f + \gamma C_{th} \quad (9)$$

其中, P_0 为模型参数, e_i 为选中的性能事件, c_i 为对应的性能事件系数, n 为性能事件数目, U 代表 CPU 利用率, f 代表 CPU 运行频率, C_{th} 为 CPU 运行线程数目, α, β 和 γ 分别为对应项的参数。为了提高模型的易用性, HBM 可以进一步表示为式(10)。

$$\begin{aligned} P_{HBM} &= P_0 + \sum_{i=1}^n (c_i e_i) + \alpha U_\Delta \\ &= P_0 + \sum_{i=1}^n (c_i e_i) + \alpha U \cdot \left(\frac{C_{th}}{T_{th}} \cdot \epsilon\right) \left(\frac{fr}{fr_{max}}\right)^\eta \end{aligned} \quad (10)$$

由于性能事件对 CPU 架构有一定的依赖性, 有些性能事件只能在特定的平台上获取。为了提高模型的易用性和通用性, HBM 选取了通用的性能事件集合进行分析, 如表 2 所列。

表 2 HBM 选取的性能事件集合

Table 2 Collection of performance events selected by HBM

性能事件名称	事件类型	描述
cycles	Hardware event	消耗的处理周期数
instructions	Hardware event	执行的指令数
branch-misses	Hardware event	预测错误的分支指令数
L1-dcache-load-misses	Hardware cache event	L1 数据缓存未命中数
L1-icache-load-misses	Hardware cache event	L1 指令缓存未命中数

表 2 列出了 HBM 选取的性能事件集合, 这些事件属于硬件事件或者硬件缓存事件, 几乎在所有的 CPU 平台上都可以获取, 而且和 CPU 功耗有着十分密切的关系。

综上所述, 相比 PBM, HBM 具有以下优点:

(1) 其采取固定的性能事件集合, 可以利用 OCOE 的方法一次运行负载收集性能事件, 比 PBM 的收集方式更方便, 而且选用的性能事件通用性很强, 比 PBM 具有更好的模型易用性和通用性。

(2) 其不需要后期大量的性能事件选择计算, 而且更少的性能事件有利于后续的模式拟合, 减少了线下训练成本。

因此, 使用 HBM 作为 ARM 架构服务器的 CPU 功耗模型更合适。

4 CPU 功耗模型验证实验

4.1 面向不同类型应用负载的调度优化

本节的 CPU 功耗模型验证实验主要包括 HBM 验证实验和交叉验证实验, HBM 验证实验主要使用 ARM 架构服务器验证本文提出的 HBM 的测算精度, 而交叉验证实验分别在 x86 和 ARM 架构服务器中同时验证最新的 x86 架构 CPU 功耗模型和 HBM 的测算精度, 检测特定架构 CPU 功耗模型在不同平台的适用性。实验中使用的服务器配置如表 3 所列。

表 3 服务器配置

Table 3 Server configuration

服务器型号	CPU 型号	CPU 架构	CPU 发行年份	TDP/W	核心/线程	CPU 实际最大睿频频率/GHz
华为 TaiShan 2280 V2 机架式服务器	2 * Kunpeng 920-4826	ARM	2019	158	48/48	2.60
华为 2288H V5 机架式服务器	2 * Intel Xeon Gold 6140	x86	2017	140	18/36	3.0

服务器使用的操作系统为 Centos7.6, 实验使用的负载工具为 Sysbench, 实验利用 Redfish 协议和华为服务器的智能管理系统 (Huawei Intelligent Baseboard Management Controller, iBMC) 进行通信, 可以方便地获取服务器的真实功耗数据。利用 Sysbench 逐渐增加运行线程, 可以实现 CPU 利用率从 0 到 100% 的变化。与 SPECpower_ssj2008 类似, 本文主要收集 CPU 利用率从 0 到 100% 的功耗数据, 间隔为 10%。在每个特定的 CPU 利用率中, 分别使用 top 和 cpupower-monitor 工具收集 CPU 利用率、CPU 运行线程数和 CPU 运行频率数据。

另一方面, 利用 Linux 内核中的性能剖析工具 Perf 获取性能计数器的监控数据。Perf 主要的命令为 perf stat, 可用于收集系统的性能概况。本实验通过该命令的 $-e$ 参数来指定要监控的性能事件, 还通过 $-I$ 和 $-interval-count$ 参数分别确定采集周期时间长度和采集数据数目。性能事件采集和上述功耗采集同步。根据收集的功耗数据和资源利用信息使用最小二乘法进行线下模型训练, 并比较各功耗模型对真实功耗数据的拟合效果。

4.2 HBM 验证实验

为了验证本文针对 ARM 架构服务器 CPU 提出的 HBM 的测算精度, 本文选取了华为 TaiShan 2280 V2 机架式服务器进行实验, 对比上述 UBM, PBM, TBM 和 HBM

这4种模型的测算精度。

为了验证HBM选定的性能事件的可行性,本实验利用收集到的性能事件和CPU功耗数据计算两者的相关系数,如表4所列。从表中可以看出,这5种性能事件和CPU功耗有着很强的相关性,可以很好地反应CPU功耗行为,因此选定这几种性能事件建模是可行的。本实验中的PBM使用的相关系数阈值为0.8,超过该阈值的性能事件都可以列入PBM的建模事件集合。在本实验中PBM采集了约60个性能事件,其中选定用于建模的性能事件接近10个,包括CPU周期数、运行指令数、预测分支数、高速缓存缺失和页表缓存缺失等。

表4 HBM选定的性能事件和CPU功耗的相关系数

Table 4 Correlation coefficients between performance events selected by HBM and CPU power consumption

性能事件名称	相关系数
cycles	0.981
instructions	0.978
branch-misses	0.898
L1-dcache-load-misses	0.877
L1-icache-load-misses	0.847

表5列出了各CPU功耗模型的模型参数,图2给出了HBM和其他模型的拟合效果,图3给出了几种功耗模型的MRE对比,UBM, TBM, PBM和HBM的MRE分别为1.09%, 2.83%, 0.29%和0.73%。可以看出, HBM的测算精度优于UBM和TBM,仅次于最好的模型PBM,但与PBM相差不大,同样满足测算精度要求。

表5 CPU功耗模型参数

Table 5 CPU power model parameters

功耗模型	模型参数
UBM	$c_0:36.42/c_1:-38.61/c_2:232.43/c_3:96.38/c_4:-468.85/c_5:69.83$
TBM	$P_{tb}:13.41/P_t:1.31$
HBM	$P_0:1.77/c_1:-1.06/c_2:8.97/c_3:-1.11/c_4:-2.80/c_5:4.16/\alpha:-3.24/-3.20/-6.50$

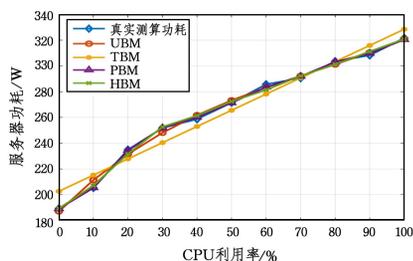


图2 HBM和其他模型的拟合效果对比

Fig. 2 Comparison of fitting effects between HBM and other models

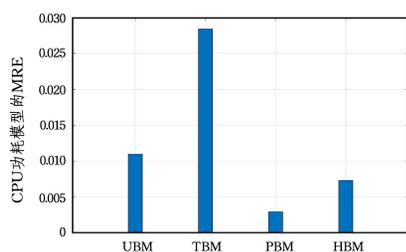


图3 CPU功耗模型MRE对比

Fig. 3 MRE comparison of CPU power consumption model

参照图1所示的基于PMC功耗建模的主要步骤,为了保证性能事件的准确采集,本实验的PBM和HBM都是采用OCOE方式进行性能事件采集。在步骤1(PMC采集性能事件)中,PBM采集了约60个性能事件,由于处理器的PMC数量有限,一次负载运行无法采集完成性能事件,期间需要多次运行负载,数据采集成本很高。在步骤2(性能事件选择)中,还需要对采集到的约60个性能事件进行相关性计算,以便对性能事件进行选择,具有一定的计算成本。最后,在步骤3(CPU功耗建模)中利用步骤2选出的近10个性能事件和功耗数据进行功耗建模。然而,HBM选定的是固定的5个性能事件,监控的性能事件较少。因此在步骤1中HBM仅需要运行一次负载即可采集完毕,比PBM数据采集成本低。HBM选定的性能事件是提前确定的,很明显不需要步骤2中的性能事件选择,且步骤3中HBM仅需对选定的5个性能事件和功耗等数据进行功耗建模,相比PBM,其模型训练成本更低。表6列出了两者的建模过程时间开销对比。

表6 HBM和PBM的建模过程时间开销对比

Table 6 Comparison of time cost between HBM and PBM

modeling processes		
功耗模型	HBM	PBM
步骤1时间开销/s	660	5280
步骤2时间开销/ms	无	236
步骤3时间开销/ms	82	132
整体时间开销/s	660.082	5280.368

本实验在CPU利用率0到100%这11个状态中分别以每秒1次的采集频率进行1min的数据采集,然后取这段时间的数据平均值作为该状态的数据。在步骤1中HBM只需运行负载一次,而PBM需要多达8次负载运行。步骤2和步骤3中的数据均在搭载两颗Intel Core i7-4710MQ处理器的笔记本上进行。从表6可以看出,HBM的整体时间开销明显小于PBM,且该差距会随着PBM采集性能事件的增多而加大。

综上所述,HBM的测算精度虽然稍低于PBM,但是其整体建模成本明显低于PBM,且HBM选定的5个性能事件通用性很强,几乎在所有的CPU平台上都可以获取。因此HBM在满足准确性的前提下具有更好的易用性和通用性,适用于ARM架构服务器CPU的功耗测算。

4.3 交叉验证实验

本文分别对x86和ARM两种架构的CPU进行交叉实验验证,检测特定架构CPU功耗模型在不同平台的适用性。

在x86架构中,本文选择华为2288H V5机架式服务器进行实验。CPU功耗模型为Lin等^[22]提出的一种基于CPU利用率的幂指数CPU功耗模型(Power Exponential Function Model, PEFM)。Lin等在PEFM的基础上还提出了一种基于硬件感知的CPU功耗测算方法(Hardware-aware CPU Power Measurement, CPM)。HCPM在SPECpower_ssj2008数据集中以CPU发行年份为维度分析了几种主流的CPU功耗模型的MRE和模型参数,然后通过获取CPU发行年份来选择合适的CPU功耗模型。他们通过充分的实验证明了PEFM和HCPM是x86架构服务器理想的CPU功耗

模型。在 ARM 架构中,本文选择华为 TaiShan 2280 V2 机架式服务器进行实验,CPU 功耗模型为 HBM。

由于 x86 架构服务器 CPU 的发行年份为 2017 年,因此 HCPM 选择的 CPU 功耗模型为 PEFM,模型参数为 0.64;ARM CPU 的发行年份为 2019 年,HCPM 选择的 CPU 功耗模型为 PEFM,模型参数为 0.61。

图 4 给出了 x86 架构交叉实验拟合效果对比,图 5 给出了 x86 架构交叉实验 CPU 功耗模型 MRE 对比。此时 PEFM,HCPM 和 HBM 的 MRE 分别为 0.45%,1.40% 和 5.14%。可以看出,基于 ARM 架构 CPU 提出的 HBM 应用在 x86 架构时 MRE 较大,明显不适用于 x86 架构。因为 x86 架构比 ARM 架构更为复杂,少量的 CPU 性能事件不能很好地反应 x86 架构的功耗行为。

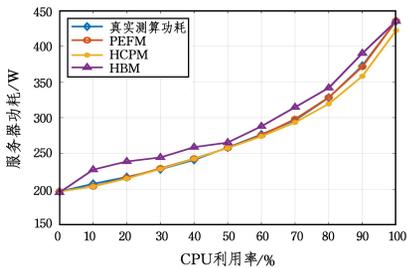


图 4 x86 架构交叉实验拟合效果对比

Fig. 4 Comparison of fitting effects of x86 architecture cross experiment

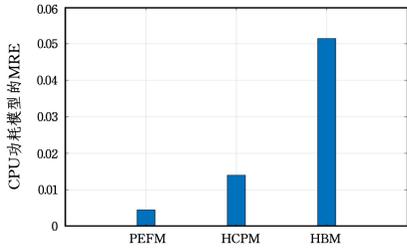


图 5 x86 架构交叉实验 CPU 功耗模型 MRE 对比

Fig. 5 MRE comparison of CPU power model in x86 architecture cross experiment

图 6 给出了 ARM 架构交叉实验拟合效果对比,图 7 给出了 ARM 架构交叉实验 CPU 功耗模型 MRE 对比。此时 PEFM,CPM 和 HBM 的 MRE 分别为 3.70%,12.31% 和 0.73%。可以看出,基于 x86 架构 CPU 提出的 PEFM 和 HCPM 应用在 ARM 架构时 MRE 较大,明显不适用于 ARM 架构。因为 ARM 架构的功耗行为与 x86 架构有一定差异,不一定遵循 PEFM 的函数规律。

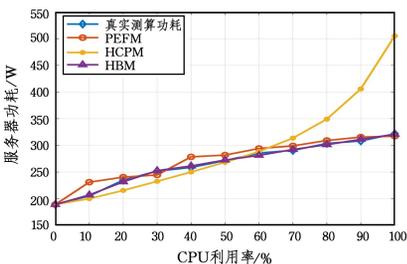


图 6 ARM 架构交叉实验拟合效果对比

Fig. 6 Fitting effect comparison of ARM architecture cross experiment

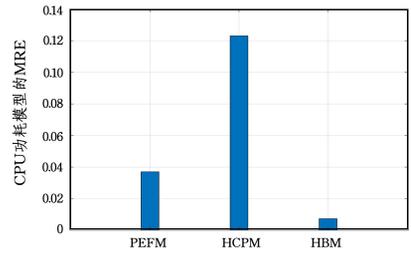


图 7 ARM 架构交叉实验 CPU 功耗模型 MRE 对比

Fig. 7 MRE comparison of CPU power model in ARM architecture cross experiment

综上所述,交叉实验验证了 x86 和 ARM 两种架构的 CPU 拥有不同的功耗行为,应该使用不同的功耗建模方法。特定架构 CPU 功耗模型应用在所属平台时才具有良好的测算精度。

结束语 随着云计算的迅速发展,云数据中心的规模不断扩大,随之而来的是巨大的能源消耗。为了让云计算保持绿色高效,开发精确的功耗模型成为数据中心能耗研究的重要一步。当前 ARM 架构服务器在云数据中心占有越来越重要的位置,但是目前大多数主流的 CPU 功耗模型只专注于 x86 架构。为此本文调研分析了现有的基于 ARM 架构的 CPU 功耗模型,并在此基础上提出了一种新的 CPU 功耗模型,即基于混合建模的 CPU 功耗模型(HBM)。本文对所提 HBM 进行了实验验证,实验结果表明 HBM 具有良好的测算精度。另外,交叉验证实验结果表明不同架构服务器的 CPU 功耗行为相异。

虽然本文提出的 HBM 拥有良好的测算精度,但是未来还有很多需要完善的地方,例如在功耗建模过程中可以引入更多的建模特征,进一步提高模型的测算精度。CPU 功耗模型验证实验也需要在更多型号的服务器或者更大规模的服务器集群中进行,使得实验结果更加可信。

参考文献

- [1] SCHAGAUEV I, KAEGI-TRACHSEL T. Architecture Comparison and Evaluation [M]. Springer International Publishing, 2016.
- [2] AKRAM A. A Study on the Impact of Instruction Set Architectures on Processor's Performance [D]. Davis: University of California, Davis, 2017.
- [3] SHEN J P, LIPASTI M H. Modern Processor Design: Fundamentals of Superscalar Processors [M]. Waveland Press, 2013.
- [4] WANG W. An improved instruction-level power and energy model for RISC microprocessors [D]. Southampton: University of Southampton, 2017.
- [5] GREENHALGH P. ARM big, LITTLE Processing with ARM Cortex-A15 & Cortex-A7 [J/OL]. <https://www.eetimes.com/big-little-processing-with-arm-cortex-a15-cortex-a7>.
- [6] DENG L Y P, CHEN C S, CAI H L. Research on Timing Analysis Based on Cache Characteristics of ARM Processor [J]. Sichuan Ordnance Journal, 2015(11): 118-121, 124.
- [7] BLEME E, MENON J, SANKARALINGAM K. Power Struggles: Revisiting the RISC vs. CISC Debate on Contemporary

- ARM and x86 Architectures [C] // 2013 IEEE 19th International Symposium on High Performance Computer Architecture (HPCA). IEEE, 2013: 1-12.
- [8] AROCA R V, GONALVES L M G. Towards Green Data-Centers: A Comparison of x86 and ARM Architectures Power Efficiency[J]. *Journal of Parallel and Distributed Computing*, 2012, 72(12): 1770-1780.
- [9] LI J W, LUO P. Opportunities and Challenges of Server Based on ARM Architecture [J]. *China New Telecommunications*, 2020, 22(18): 47-48.
- [10] VASILAKIS E. An instruction level energy characterization of arm processors[R]. FORTH-ICS/TR-450, 2015.
- [11] OBUKHOVA K, ZHURAVSKA I, BURENKO V. Diagnostics of Power Consumption of a Mobile Device Multi-core Processor with Detail of Each Core Utilization[C] // 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET). IEEE, 2020: 368-372.
- [12] WALKER M J, DAS A K, MERRETT G V, et al. Run-time Power Estimation for Mobile and Embedded Asymmetric Multi-core CPUs[C] // Hipec Workshop on Energy Efficiency with Heterogenous Computing. 2015.
- [13] CHEN K, KILPATRICK P, NIKOLOPOULOS D S, et al. Cross Architectural Power Modelling[C] // 2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID). IEEE, 2020: 390-399.
- [14] SANKARAN S, SRIDHAR R. Energy Modeling for Mobile Devices Using Performance Counters [C] // 2013 IEEE 56th International Midwest Symposium on Circuits and Systems (MWSCAS). IEEE, 2013: 441-444.
- [15] SAGI M, DOAN N A V, RAPP M, et al. A Lightweight Nonlinear Methodology to Accurately Model Multicore Processor Power[J]. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2020, 39(11): 3152-3164.
- [16] ZHANG Y, LIU Y, LI Z, et al. Accurate CPU Power Modeling for Multicore Smartphones[J/OL]. Microsoft Research, 2015. <https://www.microsoft.com/en-us/research/publication/accurate-cpu-power-modeling-for-multicore-smartphones>.
- [17] KATAOKA H, DUOLIKUN D, ENOKIDO T, et al. Power Consumption and Computation Models of a Server with a Multi-core CPU and Experiments[C] // 2015 IEEE 29th International Conference on Advanced Information Networking and Applications Workshops. IEEE, 2015: 217-222.
- [18] HUANG W, LEFURGY C, KUK W, et al. Accurate Fine-grained Processor Power Proxies[C] // 2012 45th Annual IEEE/ACM International Symposium on Microarchitecture. IEEE, 2012: 224-234.
- [19] ZHAI Y, ZHANG X, ERANIAN S, et al. Happy: Hyperthread-aware Power Profiling Dynamically[C] // 2014 {USENIX} Annual Technical Conference. 2014: 211-217.
- [20] SHEN K, SHRIRAMAN A, DWARKADAS S, et al. Power Containers: An OS Facility for Fine-Grained Power and Energy Management on Multicore Servers[J]. *Computer Architecture News*, 2013, 41(1): 65-76.
- [21] MISHRA S K, PARIDA P P, SAHOO S, et al. Improving Energy Usage in Cloud Computing Using DVFS[M] // *Progress in Advanced Computing and Intelligent Engineering*. Singapore: Springer, 2018: 623-632.
- [22] LIN W, YU T, GAO C, et al. A Hardware-aware CPU Power Measurement Based on the Power-exponent Function Model for Cloud Servers[J]. *Information Sciences*, 2021, 547: 1045-1065.



JIN Yu-yan, born in 1998, postgraduate. Her main research interests include cloud computing, big data technology and resource scheduling.



LIN Wei-wei, born in 1980, Ph.D. professor, is a senior member of China Computer Federation. His main research interests include cloud computing, big data technology and AI application technology.

(责任编辑:杨雪敏)