第47卷 第14期 2023年7月25日
DOI：10.7500/AEPS20220905004
电力系统自动化
Automation of Electric Power Systems
Vol. 47 No. 14 Jul. 25，2023

基于CPU工作状态进行温度估算的数据中心服务器功耗模型

朱介北[1]，刘东昊[1]，李炳森[1,2]，俞露杰[1]，田　娜[3]，贾宏杰[1]
（1. 天津大学电气自动化与信息工程学院，天津市 300072；2. 国网信息通信产业集团有限公司，北京市 102211；
3. 国网河北省电力有限公司经济技术研究院，河北省石家庄市 300171）

摘要：为实现数据中心节能降耗，数据中心服务器功耗模型的精确构建至关重要。传统仅基于CPU利用率的服务器功耗模型，忽略了受温度变量影响的服务器功耗延迟变化动态特性，无法区分在CPU波动或阶跃不同工作状态下的服务器功耗分层现象，导致较大的模型误差。文中提出一种基于CPU工作状态进行温度估算的数据中心服务器功耗模型(PMTC)。PMTC通过在模型建立过程中计及温度变量，捕捉受CPU核心温度延迟变化影响的服务器功耗延迟变化动态特性，消除基于CPU工作状态的服务器功耗分层导致的模型误差；在功耗计算过程中，通过判断CPU工作状态进行CPU核心温度的精确估算，避免了对温度变量的额外测量。通过搭建的功耗测试平台对比了PMTC与传统服务器功耗模型，验证了所提PMTC在不增加额外温度测量设备和输入数据维度的前提下，还原了服务器功耗延迟变化动态特性，显著降低了模型计算误差。

关键词：负荷建模；数据中心；服务器；功耗模型；热电耦合

0　引言

近年来，随着国家大数据战略加速落地，物联网、云计算和5G等技术得到了高速发展[1]。作为基础支撑设施的数据中心，其数量和规模持续增长，能源消耗比重逐年增加[2]。据国家能源局数据显示，2020年中国数据中心耗电量已突破200 TW·h，占全国总用电量的2.7%[3]；预计到2025年，其占比将达到4.1%[4]。数据中心的巨大能耗问题已成为制约其运营和发展的重要因素[5]。如何降低数据中心能耗已经成为国内外工业界和学术界广泛关注的问题。

在数据中心各类用电设备中，IT设备和冷却设备能耗最多，分别约占据总能耗的50%和40%。其中，服务器作为执行运算、存储、通信等基础功能的核心设备，其能耗约占IT设备能耗的70%～80%[5]。由于数据中心的强热电耦合特性，服务器能耗降低亦能减少冷却设备能耗，服务器的节能降耗是数据中心节能的关键。当前，服务器节能降耗的方法主要包含面向节点和集群的合理设计和优化配置，均需以精确高效的服务器功耗模型为基础。因此，准确建立服务器功耗模型对于数据中心节能降耗具有重要意义[6-7]。

目前，国内外相关研究主要基于CPU利用率对服务器功耗展开建模，主要原因在于CPU功耗所占比重最大，且服务器内其他组件活动与CPU活动耦合程度高[8]。文献[9]利用空载和满载参数建立了服务器线性模型，该模型简单直观，但随着近年来超线程技术等CPU技术的广泛应用，功耗模型的非线性特征愈发明显，线性模型误差较大[10]。针对服务器功耗的非线性特征，文献[11]采用分段线性模型对服务器功耗进行建模；文献[12-13]分别提出基于幂指数和基于多种多项式拟合的功耗模型，并指出三次多项式拟合方法较为精确。以上模型虽还原了服务器功耗的非线性规律，但缺乏对非线性现象的物理解释。

随着服务器功率密度越来越高，其热电耦合现象日益显著。除了CPU利用率，温度对服务器功耗的影响也不容忽视。由于温度变化的延迟特性，功耗变化随CPU利用率变化亦呈现延迟特性。实测结果表明，服务器功耗在CPU不同工作状态下呈现不同的变化规律[7,13]，其主要体现在：服务器功耗在CPU利用率短时波动状态下变化不明显，但在CPU利用率阶跃前后发生显著变化。因此，服务器功耗与CPU利用率的对应关系呈现“分层”离散分布现象，本文称之为基于CPU工作状态的服务器功耗分层（CPU working states based server power consumption layering，CPL）现象。

收稿日期：2022-09-05；修回日期：2023-01-31。
上网日期：2023-06-26。
国家重点研发计划资助项目(2018YFA0702200)。

已有研究引入温度实测值作为模型输入，可解决CPL现象导致的误差问题。文献[14-15]在估算服务器功耗时引入温度实测值，建立了更为准确的功耗模型，但其同样基于数据驱动的建模方法并未探究温度影响功耗的内在机理。在服务器内部与温度耦合密切的耗能部件中，CPU以及散热风扇耗能较大[6,16]。文献[17-18]将CPU功耗分为动态功耗和泄漏功耗，阐述了温度影响CPU功耗的机理；文献[19]对目前服务器常用的智能温控调速型散热风扇展开研究，说明了其转速随测量温度的变化规律。然而，相比于CPU利用率，即时温度不易获取，在使用模型进行能耗监测或预测时需要增加额外测量设备，且易受测量误差影响。

针对以上问题，本文提出一种基于CPU工作状态进行温度估算的数据中心服务器功耗模型（power consumption model based on temperature estimation considering CPU working state，PMTC）。该模型在建模过程中综合考虑CPU核心温度对服务器功耗影响的实际物理过程，可捕捉由温度延迟变化引起的服务器CPL现象，消除由CPL现象导致的模型误差；但在功耗计算时不需要对温度直接测量，而是基于CPU工作状态对温度进行估算，避免了温度测量设备的使用以及相应的温度测量误差。

1　计及温度变量的服务器功耗模型

如前所述，计及温度变量的服务器功耗模型可有效提高模型准确度。本章首先介绍计及温度变量的包括CPU功耗子模型和风扇功耗子模型在内的服务器功耗模型。

1.1　服务器总体功耗模型

在t_n时刻，服务器总体功耗$P_{\mathrm{srv}}^{t_n}$主要由CPU功耗$P_{\mathrm{CPU}}^{t_n}$、散热风扇功耗$P_{\mathrm{fan}}^{t_n}$以及内存磁盘等其余部分功耗P_{other}三部分组成，如式(1)所示[6]。

$$P_{\mathrm{srv}}^{t_n}=P_{\mathrm{CPU}}^{t_n}+P_{\mathrm{fan}}^{t_n}+P_{\mathrm{other}} \tag{1}$$

式中：P_{other}为其余功耗，由于占比较小且波动不大，其通常视为恒定值。

1.2　计及温度变量的CPU功耗子模型

CPU芯片功耗$P_{\mathrm{CPU}}^{t_n}$包括空闲功耗P_{idle}、动态功耗$P_{\mathrm{dyn}}^{t_n}$和泄漏功耗$P_{\mathrm{leak}}^{t_n}$[17,20]，如式(2)所示。

$$P_{\mathrm{CPU}}^{t_n}=P_{\mathrm{idle}}+P_{\mathrm{dyn}}^{t_n}+P_{\mathrm{leak}}^{t_n} \tag{2}$$

式中：P_{idle}为固定常数。

CPU动态功耗$P_{\mathrm{dyn}}^{t_n}$由组成CPU的晶体管信号切换过程产生，与CPU执行计算任务相关，可由CPU利用率$U_{\mathrm{CPU}}^{t_n}$表示为[15]：

$$P_{\mathrm{dyn}}^{t_n}=aU_{\mathrm{CPU}}^{t_n} \tag{3}$$

式中：a为通过测量结果拟合所得系数。

CPU泄漏功耗$P_{\mathrm{leak}}^{t_n}$由晶体管泄漏电流产生，与内部载流子运动有关，受温度影响，可表示为与CPU核心温度$T_{\mathrm{CPU}}^{t_n}$相关的指数形式[21]：

$$P_{\mathrm{leak}}^{t_n}=b_0\exp\left(-\frac{b_1}{T_{\mathrm{CPU}}^{t_n}}\right) \tag{4}$$

式中：b_0和b_1为通过测量结果拟合所得系数。

将式(3)和式(4)代入式(2)，可得计及温度变量的CPU功耗$P_{\mathrm{CPU}}^{t_n}$为：

$$P_{\mathrm{CPU}}^{t_n}=P_{\mathrm{idle}}+aU_{\mathrm{CPU}}^{t_n}+b_0\exp\left(-\frac{b_1}{T_{\mathrm{CPU}}^{t_n}}\right) \tag{5}$$

由式(5)可见，CPU核心温度主要通过改变泄漏功耗来影响CPU的功耗。

1.3　计及温度变量的风扇功耗子模型

服务器中CPU等电子元件消耗的电能以热能形式耗散，从而导致服务器温升。在散热方面，目前服务器常采用温控调速型散热风扇。其转速$f_{\mathrm{fan}}^{t_n}$可由CPU核心温度分段表示为[19]：

$$f_{\mathrm{fan}}^{t_n}=\begin{cases}f_{\min} & T_{\mathrm{CPU}}^{t_n}\leqslant T_{\min}\\ f_{\min}+k_{\mathrm{fan}}(T_{\mathrm{CPU}}^{t_n}-T_{\min}) & T_{\min}<T_{\mathrm{CPU}}^{t_n}\leqslant T_{\max}\\ f_{\max} & T_{\mathrm{CPU}}^{t_n}>T_{\max}\end{cases} \tag{6}$$

式中：$f_{\min}$和$f_{\max}$分别为服务器低负荷和高负荷状态下的风扇转速；$T_{\min}$和$T_{\max}$为风扇转速在恒速状态与变速状态间切换时的温度阈值；k_{fan}为变速状态下转速随温度改变的变化率。

根据测量数据可拟合获得风扇功耗$P_{\mathrm{fan}}^{t_n}$与风扇转速$f_{\mathrm{fan}}^{t_n}$对应的三次多项式关系[17]，如式(7)所示。

$$P_{\mathrm{fan}}^{t_n}=c_0+c_1f_{\mathrm{fan}}^{t_n}+c_2(f_{\mathrm{fan}}^{t_n})^2+c_3(f_{\mathrm{fan}}^{t_n})^3 \tag{7}$$

式中：c_0至c_3为实验结果拟合所得系数。

基于式(6)和式(7)，可得计及温度变量的风扇功耗模型。

2　CPU不同工作状态下服务器功耗变化规律

第1章所述计及温度变量的服务器功耗模型须使用实测温度作为模型输入。然而，相比于CPU利用率，CPU核心温度数据的获取依赖于额外测量设备，且受制于测量误差。

为避免对温度实际测量，须考虑CPU利用率变化通过CPU核心温度对服务器功耗的影响。本章首先阐述服务器温度变量的延迟变化动态特性，并在此基础上分析计及温度变量的服务器功耗延迟变化动态特性。最后，基于上述变化特性揭示CPU不

同工作状态下的服务器功耗变化规律。

2.1　服务器温度变量的延迟变化动态特性

由于服务器本质上可看作是具有一定温度差和热阻的热量传输体系[22]，服务器运行过程中的温度变化可使用热力学动态分析常用的等效热参数模型进行描述。假设在采样时刻 t_{n-1}，CPU 利用率为 $U_{\mathrm{CPU}}^{t_{n-1}}$，且在采样间隔 $[t_{n-1}, t_n)$ 内保持不变，则采样时刻 t_n 的 CPU 核心温度 $T_{\mathrm{CPU}}^{t_n}$ 可由式(8)计算得到[23-24]。

$$T_{\mathrm{CPU}}^{t_n}=T_{\mathrm{CPU,s}}^{t_{n-1}}+\left(T_{\mathrm{CPU}}^{t_{n-1}}-T_{\mathrm{CPU,s}}^{t_{n-1}}\right)\exp\left(\frac{t_{n-1}-t_n}{\tau_{\mathrm{CPU}}^{t_{n-1}}}\right) \tag{8}$$

式中：$\tau_{\mathrm{CPU}}^{t_{n-1}}$ 为CPU利用率等于 $U_{\mathrm{CPU}}^{t_{n-1}}$ 状态下的温度变化时间常数(根据本文实验测量结果平均在150 s左右)；$T_{\mathrm{CPU,s}}^{t_{n-1}}$ 为CPU利用率等于 $U_{\mathrm{CPU}}^{t_{n-1}}$ 状态下的CPU核心温度稳态解，其与 $U_{\mathrm{CPU}}^{t_{n-1}}$ 的对应关系如式(9)所示(具体求解过程详见附录A)，有

$$T_{\mathrm{CPU,s}}^{t_{n-1}}=\mu_0 U_{\mathrm{CPU}}^{t_{n-1}}+\mu_1+\sqrt{\left(\mu_0 U_{\mathrm{CPU}}^{t_{n-1}}+\mu_1\right)^2-\mu_2} \tag{9}$$

式中：参数 μ_0、μ_1、μ_2 均为定值，可由实测数据拟合得到。

图1(a)和(c)展示了式(8)所述的服务器温度变量延迟变化动态特性。当CPU利用率由 $U_{\mathrm{CPU},0}$ 阶跃至 $U_{\mathrm{CPU},1}$ 后(如图1(a)所示)，CPU核心温度 $T_{\mathrm{CPU}}^{t_n}$ 将按照式(8)呈现延迟变化，如图1(c)所示。图中：$T_{\mathrm{CPU},0}$、$T_{\mathrm{CPU,s}}$ 分别为初始和稳态CPU核心温度；$P_{\mathrm{srv},0}$、$P_{\mathrm{srv,s}}$ 分别为初始和稳态服务器功耗。

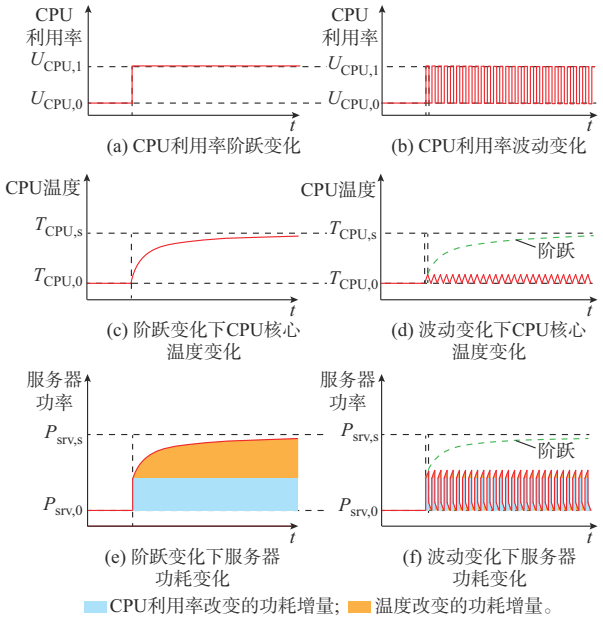

图1　服务器不同工作状态下参数变化规律示意图
Fig. 1　Schematic diagram of parameter change law under different server working states

2.2　计及温度变量的服务器功耗延迟变化动态特性

如2.1节所述，CPU核心温度 $T_{\mathrm{CPU}}^{t_n}$ 随CPU利用率实时变化呈现如式(8)所示的动态延迟特性。将式(8)的CPU核心温度动态模型代入式(5)，获得计及温度变量的CPU功耗动态特性，如式(10)所示。

$$P_{\mathrm{CPU}}^{t_n}=P_{\mathrm{idle}}+aU_{\mathrm{CPU}}^{t_n}+b_0\exp\left(\frac{-b_1}{T_{\mathrm{CPU,s}}^{t_{n-1}}+\left(T_{\mathrm{CPU}}^{t_{n-1}}-T_{\mathrm{CPU,s}}^{t_{n-1}}\right)\exp\left(\frac{t_{n-1}-t_n}{\tau_{\mathrm{CPU}}^{t_{n-1}}}\right)}\right) \tag{10}$$

同样地，将式(8)代入式(6)与式(7)，获得计及温度变量的风扇功耗动态特性。由于风扇工作在恒速区时其功耗恒定，故此处仅展示风扇变速区($T_{\min}<T_{\mathrm{CPU}}^{t_n}\leqslant T_{\max}$)功耗动态特性：

$$P_{\mathrm{fan}}^{t_n}=\lambda_0+\lambda_1\exp\left(\frac{t_{n-1}-t_n}{\tau_{\mathrm{CPU}}^{t_{n-1}}}\right)+\lambda_2\exp\left(2\frac{t_{n-1}-t_n}{\tau_{\mathrm{CPU}}^{t_{n-1}}}\right)+\lambda_3\exp\left(3\frac{t_{n-1}-t_n}{\tau_{\mathrm{CPU}}^{t_{n-1}}}\right) \tag{11}$$

式中：参数 λ_0 至 λ_3 详细表达式见附录B，在2.1节所述假设条件下均为常系数。

最后，将式(10)与式(11)代入式(1)，可得计及温度变量的服务器功耗动态特性(仍以风扇工作在变速区为例)：

$$P_{\mathrm{srv}}^{t_n}=P_{\mathrm{other}}+P_{\mathrm{idle}}+aU_{\mathrm{CPU}}^{t_n}+\lambda_0+\lambda_1\exp\left(\frac{t_{n-1}-t_n}{\tau_{\mathrm{CPU}}^{t_{n-1}}}\right)+\lambda_2\exp\left(2\frac{t_{n-1}-t_n}{\tau_{\mathrm{CPU}}^{t_{n-1}}}\right)+\lambda_3\exp\left(3\frac{t_{n-1}-t_n}{\tau_{\mathrm{CPU}}^{t_{n-1}}}\right)+b_0\exp\left(\frac{-b_1}{T_{\mathrm{CPU,s}}^{t_{n-1}}+\left(T_{\mathrm{CPU}}^{t_{n-1}}-T_{\mathrm{CPU,s}}^{t_{n-1}}\right)\exp\left(\frac{t_{n-1}-t_n}{\tau_{\mathrm{CPU}}^{t_{n-1}}}\right)}\right) \tag{12}$$

图1(e)展示了当CPU利用率按照图1(a)产生阶跃变化后服务器功耗的动态变化。由图1可见，服务器功耗在CPU利用率变化后呈现如式(12)所示的延迟变化动态特性。

2.3　CPU不同工作状态下服务器功耗变化规律

在服务器的实际运行中，CPU有“波动”和“阶跃”两种工作状态。波动状态为：当CPU运行固定任务时，CPU利用率稳定在某一数值附近并伴有不定幅度的短时波动。阶跃状态为：当出现负载迁移、任务增减等变化时，CPU利用率呈现大幅度阶跃变

化。图1简要展示了CPU利用率由$U_{\mathrm{CPU},0}$阶跃至$U_{\mathrm{CPU},1}$和在$U_{\mathrm{CPU},0}$与$U_{\mathrm{CPU},1}$之间波动两种工作状态下，CPU核心温度和服务器功耗的不同变化情况。

1)在波动状态时(如图1(b)所示)，由于CPU利用率$U_{\mathrm{CPU}}^{t_n}$变化周期远小于式(8)和式(12)中动态变化时间常数$\tau_{\mathrm{CPU}}^{t_{n-1}}$，故无论$U_{\mathrm{CPU}}^{t_n}$变化幅度大小，CPU核心温度$T_{\mathrm{CPU}}^{t_n}$与服务器功耗$P_{\mathrm{srv}}^{t_n}$仅小幅变化，如图1(b)、(f)所示。同时，$P_{\mathrm{srv}}^{t_n}$的变化主要由动态功耗$aU_{\mathrm{CPU}}^{t_n}$部分贡献，如图1(f)的蓝色部分所示。

2)在阶跃状态时(如图1(a)所示)，$U_{\mathrm{CPU}}^{t_n}$呈现大幅度阶跃变化，此时式(8)和式(12)有足够时间达到稳态，故$T_{\mathrm{CPU}}^{t_n}$与$P_{\mathrm{srv}}^{t_n}$均发生大幅变化，如图1(c)、(e)所示。

由于服务器功耗在CPU不同工作状态下具有上述不同变化规律，仅以CPU利用率作为功耗模型输入的传统方法存在误差。下文将在服务器功耗模型中考虑温度与功耗的延迟变化动态特性，以实现更为准确的功耗计算。

3　基于CPU工作状态进行温度估算的服务器功耗模型

本章提出一种基于CPU工作状态的温度估算方法，将该方法与第1章所述计及温度变量的服务器功耗模型相结合，提出一种新型基于CPU工作状态进行温度估算的服务器功耗模型。

3.1　基于CPU工作状态的温度估算方法

由于CPU核心温度在CPU波动或阶跃工作状态下具有不同变化规律(如2.3节所述)，故对CPU工作状态的判断是保障温度估算准确的前提。为估算t_n时刻的CPU核心温度，首先提取t_n时刻前Δt时间段内CPU利用率时间序列$U=\{U_{\mathrm{CPU}}^{t_n-\Delta t},U_{\mathrm{CPU}}^{t_n-\Delta t+1},\cdots,U_{\mathrm{CPU}}^{t_n}\}$，再计算其平均值离散度(mean value dispersion，MVD)，进而判断CPU工作状态类型。平均值离散度$R_{\mathrm{MVD}}^{t_n}$为CPU利用率时间序列U前后两部分(等分)的平均值之差，计算式为：

$$R_{\mathrm{MVD}}^{t_n}=\left|\frac{2}{\Delta t}\sum_{t=t_n-\Delta t/2}^{t_n}U_{\mathrm{CPU}}^{t}-\frac{2}{\Delta t}\sum_{t=t_n-\Delta t}^{t_n-\Delta t/2}U_{\mathrm{CPU}}^{t}\right| \tag{13}$$

式中：Δt为所取时间序列长度，可通过参数优化获得(具体见3.2节)。

附录C图C1展示了实测CPU利用率阶跃变化前后$R_{\mathrm{MVD}}^{t_n}$随$U_{\mathrm{CPU}}^{t_n}$的变化情况。当CPU处于短时波动状态时，$R_{\mathrm{MVD}}^{t_n}$维持在较低水平；当CPU处于阶跃状态时，$R_{\mathrm{MVD}}^{t_n}$在短时间内升至较高数值，而后又恢复至较低水平。

因此，通过检测$R_{\mathrm{MVD}}^{t_n}$变化，可对CPU工作状态进行判定，进而确定当前时刻CPU核心温度估算值$T_{\mathrm{pre}}^{t_n}$。当$R_{\mathrm{MVD}}^{t_n}$小于状态判定阈值$R_{\mathrm{MVD,th}}$时，判断为CPU处于波动状态，温度维持在原有估算值$T_{\mathrm{pre}}^{t_{n-1}}$不变；当$R_{\mathrm{MVD}}^{t_n}$大于阈值$R_{\mathrm{MVD,th}}$时，判断为CPU处于阶跃状态，利用式(9)的CPU核心温度稳态解$T_{\mathrm{CPU,s}}^{t_n}$与CPU利用率$U_{\mathrm{CPU}}^{t_n}$的对应关系式对$T_{\mathrm{pre}}^{t_n}$进行更新，有

$$T_{\mathrm{pre}}^{t_n}=\begin{cases}T_{\mathrm{pre}}^{t_{n-1}} & R_{\mathrm{MVD}}^{t_n}\leqslant R_{\mathrm{MVD,th}}\\ \dfrac{\mu_0}{\delta t}\displaystyle\sum_{t=t_n-\delta t}^{t_n}U_{\mathrm{CPU}}^{t}+\mu_1+\\ \qquad\sqrt{\left(\dfrac{\mu_0}{\delta t}\displaystyle\sum_{t=t_n-\delta t}^{t_n}U_{\mathrm{CPU}}^{t}+\mu_1\right)^2-\mu_2} & R_{\mathrm{MVD}}^{t_n}>R_{\mathrm{MVD,th}}\end{cases} \tag{14}$$

式中：δt为消除暂态误差设置的矫正序列长度(本文取8 s)。

3.2　PMTC服务器功耗模型建立流程

通过式(1)和式(5)至式(7)可获得计及温度变量的服务器功耗数学模型。其中，温度变量输入可按照式(13)和式(14)估算，进而避免实际功耗计算过程中对CPU核心温度的测量。

为获取式(5)中参数a、b_0、b_1和P_{idle}，式(6)中参数$f_{\min}$、$f_{\max}$、$T_{\min}$、$T_{\max}$和k_{fan}，式(7)中参数c_0至c_3，式(14)中参数μ_0至μ_2，模型的建立首先需要在测试服务器上运行可生成特定CPU利用率动态的测试程序。同时，利用软硬件工具记录服务器$U_{\mathrm{CPU}}^{t_n}$、$T_{\mathrm{CPU}}^{t_n}$、$P_{\mathrm{CPU}}^{t_n}$、$f_{\mathrm{fan}}^{t_n}$性能指标以及服务器整体功耗$P_{\mathrm{srv}}^{t_n}$。最后，基于最小二乘法对测量数据进行拟合，得到上述参数准确数值。

除上述待拟合参数，式(13)中的CPU利用率时间序列长度Δt和式(14)中的状态判定阈值$R_{\mathrm{MVD,th}}$对CPU工作状态的准确判断具有重要影响。为获得合理的Δt和$R_{\mathrm{MVD,th}}$，本文以模型功耗计算误差最小为优化目标，对Δt和$R_{\mathrm{MVD,th}}$进行优化，并采用均方根误差R_{RMSE}来衡量实测值与模型计算值之间的误差。均方根误差R_{RMSE}计算公式如式(15)所示。

$$R_{\mathrm{RMSE}}=\sqrt{\frac{1}{t_{\mathrm{N}}}\sum_{t=1}^{t_{\mathrm{N}}}(P_{\mathrm{srv,m}}^{t}-P_{\mathrm{srv}}^{t})^2} \tag{15}$$

式中：t_{N}为测试过程总时间；$P_{\mathrm{srv,m}}^{t}$为t时刻服务器实测功耗数据；P_{srv}^{t}为t时刻模型计算服务器功耗数据。优化Δt和$R_{\mathrm{MVD,th}}$参数的流程伪代码如附录C表C1所示。

基于上述建立的服务器功耗模型，服务器功耗计算流程如图2右侧所示。

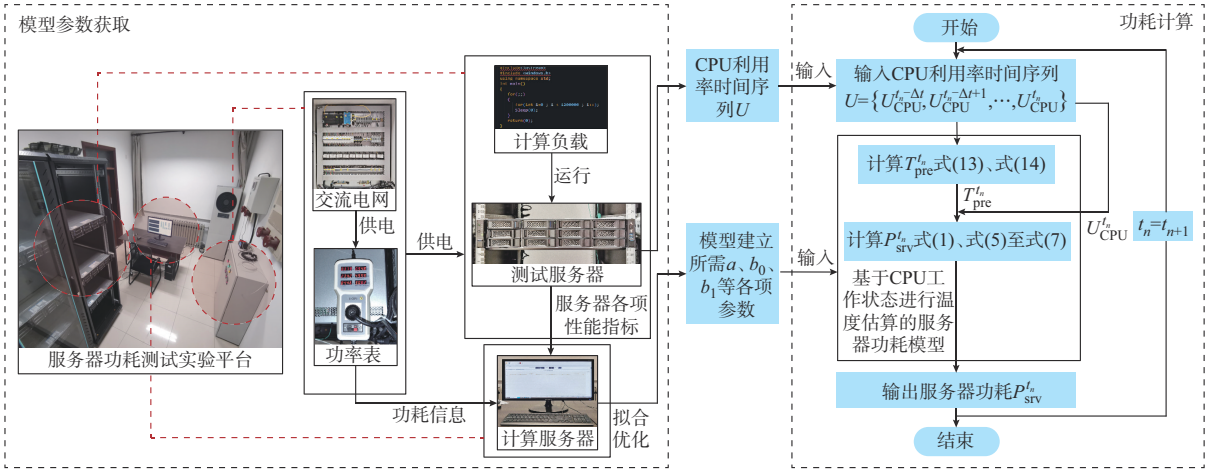

图2　PMTC建立流程

Fig. 2　Establishment process of PMTC

为求得服务器 t_n 时刻的功耗，首先以 t_n 时刻CPU利用率时间序列 U 作为输入，通过式(13)、式(14)估算CPU核心温度 $T_{\mathrm{pre}}^{t_n}$，然后将 $T_{\mathrm{pre}}^{t_n}$ 和当前 $U_{\mathrm{CPU}}^{t_n}$ 输入式(1)和式(5)至式(7)，获得最终输出服务器功耗 $P_{\mathrm{srv}}^{t_n}$。通过实时获取时间序列 U，可实现服务器运行过程中功耗的实时计算与监测。此外，该模型还可根据任务运行CPU利用率占用历史数据，实现服务器功耗的精准预测。

4　算例演示与验证

为验证第3章所提PMTC的准确性，搭建了服务器功耗测试平台。首先，利用实测数据验证了2.3节所述变化规律导致的CPL现象及其误差；然后，依据3.2节所提流程，对测试服务器建立了PMTC，并与5种传统服务器功耗模型进行了准确度对比。

为充分验证PMTC的准确性和通用性，本文选取了3台异构服务器进行测试和对比，其配置如下：服务器Ⅰ搭载Xeon Silver 4210处理器，满载功耗为220 W左右；服务器Ⅱ搭载Core i7-8700处理器，满载功耗为120 W左右；服务器Ⅲ搭载Core i7-9700处理器，满载功耗为110 W左右。计算负载模拟采用C语言循环测试程序，以占用处理器特定数量内核方式实现CPU利用率灵活变化，从 $U_{\mathrm{CPU}}^{t_n}\approx5\%$ 的初始状态每1 000 s以8%～12%左右的幅度梯度递增。使用AIDA64检测引擎获取CPU利用率、CPU功率和散热风扇转速等信息[25]；使用功率表测量服务器整体功耗信息；所有采样记录间隔均为1 s。

4.1　服务器CPL现象

图3通过实测数据展示了服务器Ⅰ的CPL现象。

由图3(a)可见，CPU利用率 $U_{\mathrm{CPU}}^{t_n}$ 包括两次阶跃变化。图3(a)中选取 t=650 s附近CPU工作状态进行展示。由图可见，$U_{\mathrm{CPU}}^{t_n}$ 从点①的33%变化至

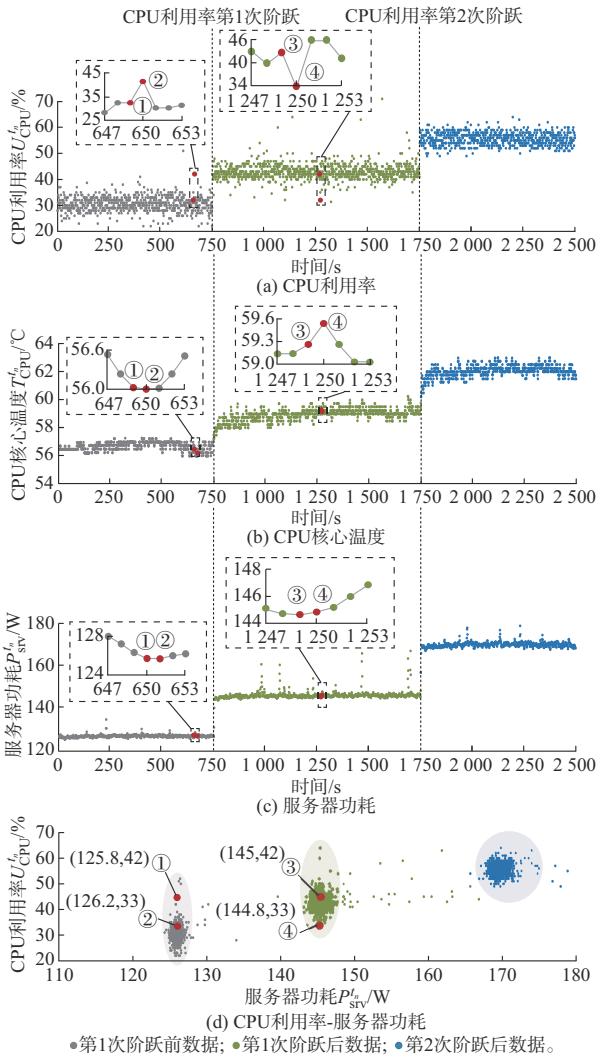

图3　服务器CPL现象示意图

Fig. 3　Schematic diagram of server CPL phenomenon

点②的42%又恢复至31%。依据2.1和2.2节所述延迟变化的动态特性，由于该过程中波动周期远小于式(8)所得CPU核心温度$T_{\mathrm{CPU}}^{t_n}$和式(12)所得服务器功耗$P_{\mathrm{srv}}^{t_n}$暂态变化的时间常数，故$T_{\mathrm{CPU}}^{t_n}$和$P_{\mathrm{srv}}^{t_n}$仅小幅变化，如图3(b)、(c)所示。

在t=750 s，设置服务器负载增加使$U_{\mathrm{CPU}}^{t_n}$产生第1次阶跃，$T_{\mathrm{CPU}}^{t_n}$遵循式(8)所述的温度延迟变化特性，从56.8 ℃左右缓慢升至59.1 ℃左右，如图3(b)所示；$P_{\mathrm{srv}}^{t_n}$亦遵循式(12)从128 W左右升至145.8 W左右，如图3(c)所示。

通过比较阶跃变化前后的运行点①②和运行点③④可以观测到，虽然运行点①④的$U_{\mathrm{CPU}}^{t_n}$同为34%、运行点②③的$U_{\mathrm{CPU}}^{t_n}$同为42%，但对应服务器功耗$P_{\mathrm{srv}}^{t_n}$却相差19 W(运行点①②功耗分别为126.2 W、125.8 W，运行点③④功耗分别为145 W、144.8 W)，如图3(d)所示。

由实验结果可见，$U_{\mathrm{CPU}}^{t_n}$阶跃变化前后，服务器功耗与CPU利用率数据的对应分布出现分层现象，连续的CPU利用率数值无法对应连续的服务器功耗数值，如图3(d)所示，即CPL现象。CPL现象导致传统仅基于CPU利用率的服务器功耗模型产生显著误差，影响计算准确度。本文所提PMTC通过计及温度变量，并根据CPU不同工作状态对温度变量准确估算，解决了CPL现象对功耗计算的影响。

4.2　PMTC的参数拟合与优化

本节按照3.2节所述建模流程，通过拟合优化方法获得服务器Ⅰ(环境温度为20 ℃左右)的PMTC参数。首先，依据在测试服务器上运行测试程序记录的服务器$U_{\mathrm{CPU}}^{t_n}$、$T_{\mathrm{CPU}}^{t_n}$、$P_{\mathrm{CPU}}^{t_n}$、$f_{\mathrm{fan}}^{t_n}$性能指标以及服务器功耗$P_{\mathrm{srv}}^{t_n}$数据，通过最小二乘法拟合获得式(5)至式(7)和式(14)中的主要参数。针对式(5)，对$P_{\mathrm{CPU}}^{t_n}$的拟合如附录C图C2(a)所示，从中获取a、b_0、b_1和P_{idle}的准确参数；由于式(1)中P_{other}占比较小且较为恒定，因此，将P_{fan}视为P_{srv}与P_{CPU}之差，则针对式(7)对P_{fan}的拟合如图C2(b)所示，从中获取c_0至c_3的准确参数；针对式(6)，对f_{fan}的拟合如图C2(c)所示，从中获取$f_{\min}$、$f_{\max}$、$T_{\min}$、$T_{\max}$和k_{fan}的准确参数；针对式(14)，对CPU核心温度稳态解$T_{\mathrm{CPU,s}}^{t_n}$的拟合如图C2(d)所示，从中获取μ_0、μ_1、μ_2的准确参数。服务器Ⅰ的PMTC拟合参数结果如表1所示。

表1　PMTC拟合参数测试结果

Table 1　Test results of PMTC fitting parameters

模型构成	模型参数	拟合参数值			
		服务器Ⅰ(20 ℃)	服务器Ⅰ(25 ℃)	服务器Ⅱ(25 ℃)	服务器Ⅲ(25 ℃)
式(5)	a	0.43	0.42	0.21	0.26
	b_0	247.00	245.00	72.30	40.45
	b_1	35.90	35.00	46.01	28.50
	P_{idle}	46.20	46.00	20.89	17.90
式(6)	c_0	−184.10	−175.90	−219.90	−602.00
	c_1	0.51	0.50	14.98	1.64
	c_2	-3.46×10^{-4}	-3.22×10^{-4}	-0.31×10^{-2}	-1.39×10^{-3}
	c_3	7.96×10^{-8}	7.91×10^{-8}	0.21×10^{-2}	3.93×10^{-7}
式(7)	$f_{\min}$	911.40	913.00	920.90	875.00
	$f_{\max}$				
	$T_{\min}$	56.15	55.59	40.20	52.70
	$T_{\max}$				
	k_{fan}	75.20	75.20	67.40	78.51
式(14)	μ_0	0.21	0.21	0.13	0.16
	μ_1	10.91	13.71	99.63	61.20
	μ_2	9.42	9.19	152.30	81.77

另外，使用式(15)的误差指标和附录C表C1的求解流程，通过寻找R_{RMSE}最小值以获得式(13)、式(14)中Δt和$R_{\mathrm{MVD,th}}$的最优值。附录C图C3展示了本文测试平台R_{RMSE}随Δt和$R_{\mathrm{MVD,th}}$的变化。最终确定误差值最小(R_{RMSE}=3.71 W，如图C3中红点所示)对应的最优模型参数Δt=12，$R_{\mathrm{MVD,th}}$=3.8。

4.3　PMTC的准确度验证

为验证所建立的PMTC的准确度，将其与仅以CPU利用率作为模型输入的服务器功耗模型以及引入温度实测值的服务器功耗模型进行准确度对比：线性模型[9](linear model，LM)、分段线性模型[11](piecewise linear model，PLM)、指数模型[12]

(exponential model, EM)、多项式模型[13](polynomial model, PM)和考虑温度的功耗模型[14](temperature considered power consumption model, TCM)。其模型算法详见附录D。

图4展示了每一种模型对于服务器Ⅰ的功耗计算值与实测功耗值及其误差百分比(相比于满载功耗)。可以看到,与传统模型的误差(±9%以内)相比,PMTC对实测值的拟合更为精确,误差更小(±3%以内)。

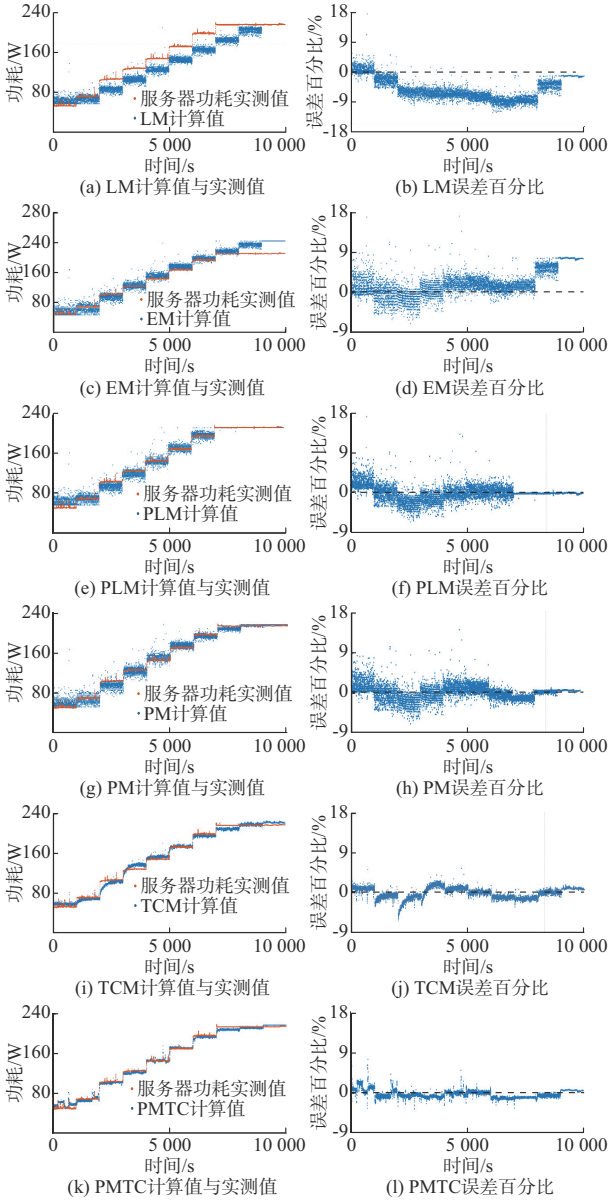

图4　服务器Ⅰ模型计算值与实测值对比及误差变化
Fig. 4　Comparison between calculated value and measured value and error change of server Ⅰ model

选取PMTC和同样仅以CPU利用率作为模型输入的功耗模型中准确度最高的PM(以三次多项式为例)进行比较,并选取t=4 000 s处CPU利用率阶跃前后的输出数据,绘制服务器功耗-CPU利用率散点图,如图5所示。可以看出,相比于PM中CPU利用率和服务器功耗的直接对应关系,PMTC由于引入温度变量及CPU核心温度估算方法,使其计算结果更加符合图3(d)所示实际情况下服务器CPL现象,降低了模型计算误差。

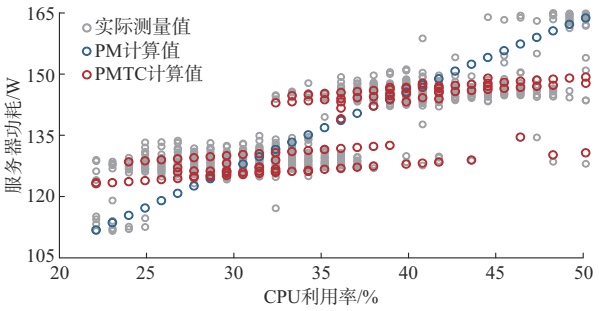

图5　PMTC对CPL现象造成误差的改善情况
Fig. 5　Improvement of errors of CPL phenomenon caused by PMTC

最后,本文根据式(15)对由图4所示模型测试全流程数据计算均方根误差R_{RMSE},以评估模型准确度,如附录C图C4所示。与其他传统模型相比,本文所提的PMTC准确度更高,其R_{RMSE}(3.71 W)比LM(R_{RMSE}=11.92 W)、PLM(R_{RMSE}=7.85 W)、EM(R_{RMSE}=4.87 W)、PM(R_{RMSE}=4.73 W)、TCM(R_{RMSE}=4.15 W)分别降低了69%、53%、22%、21%、11%左右。

除R_{RMSE}外,本文也使用以下几种常用的拟合度评价指标对模型进行准确度评估:平均绝对误差R_{MAE}、平均绝对百分比误差R_{MAPE}、R方值R_{RSV}和矫正决定系数R_{ARSV},其计算方法详见附录E。其中,R_{RSV}和R_{ARSV}用来衡量模型对实测值的拟合能力,其值越大模型越精确;其余指标用来衡量误差大小,其值越小模型越精确。用以上评价指标对本文所提模型和传统功耗模型进行准确度分析的结果如附录C图C4所示。可见,PMTC的各方面指标均优于其他传统模型。

4.4　PMTC通用性验证

为验证PMTC的通用性,本节使用前述3台异构服务器,在相同实验温度(25 ℃)下使用3.2节功耗模型建立方法分别搭建其PMTC。由表1可知,3台异构服务器的各项PMTC参数均有所不同。附录C图C5展示了3台服务器的功耗实测值与PMTC功耗计算值及其误差情况。由图可知,PMTC计算结果均有较高的准确度(误差百分比在±3%以内),其准确度亦明显优于图4所示的5种传统服务器功耗建模方法。

除此之外,对比表2中服务器Ⅰ的PMTC拟合

参数可知，相同服务器在不同温度下式(14)中的参数μ_1存在明显差别(20 ℃和25 ℃时μ_1分别为10.91和13.71)。这是由于包含参数μ_1的式(14)与环境温度T_{amb}有较强关联性(如附录A式(A5)和式(A6)所示)，参数μ_1的获取结果受环境温度影响。如图4(f)以及附录C图C5(a)所示，在上述两种不同环境温度下，其PMTC误差百分比均在±3%以内，具有较高准确度，亦体现了PMTC建模过程充分计及服务器运行环境温度影响的优势。

5　结语

由于温度变量延迟变化动态特性对服务器功耗的影响，服务器在CPU波动或阶跃不同工作状态下呈现功耗分层现象，造成模型的计算误差。本文提出一种利用CPU工作状态进行温度估算的服务器功耗模型PMTC，具有以下主要特点：

1)PMTC通过运行服务器测试程序并获得建模所需数据，对模型待求参数进行拟合优化完成模型建立，其建模方法具有较高通用性；

2)PMTC在模型建立过程中计及温度变量，可有效捕捉受CPU核心温度影响的服务器功耗延迟变化动态特性及CPL现象，降低模型误差；

3)PMTC根据CPU利用率时间序列数据的平均值离散度，判断CPU工作状态并准确估算CPU核心温度，可避免在使用过程中对温度变量的额外测量。

本文搭建了功耗测试平台并对PMTC和传统功耗模型进行准确度对比，对异构服务器进行模型通用性分析，验证了PMTC在不增加输入数据维度和额外温度测量设备的前提下有效消除了由于CPL现象产生的模型误差，显著提高了功耗计算准确度。后续研究将在本文基础上，进一步考虑如何将所提PMTC应用在服务器调度策略中，以实现数据中心节能降耗的目标。

附录见本刊网络版(http://www.aeps-info.com/aeps/ch/index.aspx)，扫英文摘要后二维码可以阅读网络全文。

参考文献

[1] 丁肇豪，曹雨洁，张素芳，等.能源互联网背景下数据中心与电力系统协同优化(一)：数据中心能耗模型[J].中国电机工程学报，2022，42(9)：3161-3177.
DING Zhaohao, CAO Yujie, ZHANG Sufang, et al. Coordinated operation for data center and power system in the context of energy Internet (Ⅰ): energy demand management model of data center[J]. Proceedings of the CSEE, 2022, 42(9): 3161-3177.

[2] JIN C, BAI X, YANG C, et al. A review of power consumption models of servers in data centers[J]. Applied Energy, 2020, 265: 114806.

[3] 王扬，李鹏，冀浩然，等.考虑多类型资源的数据中心园区供电协调规划[J].电力系统自动化，2022，46(14)：19-28.
WANG Yang, LI Peng, JI Haoran, et al. Coordination planning of power supply in data center park considering multiple resources[J]. Automation of Electric Power Systems, 2022, 46(14): 19-28.

[4] 高晋坤，余娟，刘珏麟，等.考虑多时段设备耦合的数据中心能效优化方法[J].电力系统自动化，2022，46(15)：153-161.
GAO Jinkun, YU Juan, LIU Juelin, et al. Optimization method for energy efficiency of data center considering multi-period equipment coupling[J]. Automation of Electric Power Systems, 2022, 46(15): 153-161.

[5] 高赐威，曹晓峻，闫华光，等.数据中心电能管理及参与需求侧资源调度的展望[J].电力系统自动化，2017，41(23)：1-7.
GAO Ciwei, CAO Xiaojun, YAN Huaguang, et al. Energy management of data center and prospect for participation in demand side resource scheduling[J]. Automation of Electric Power Systems, 2017, 41(23): 1-7.

[6] DAYARATHNA M, WEN Y G, FAN R. Data center energy consumption modeling: a survey[J]. IEEE Communications Surveys & Tutorials, 2015, 18(1): 732-794.

[7] 罗亮，吴文峻，张飞.面向云计算数据中心的能耗建模方法[J].软件学报，2014，25(7)：1371-1387.
LUO Liang, WU Wenjun, ZHANG Fei. Energy modeling based on cloud data center[J]. Journal of Software, 2014, 25(7): 1371-1387.

[8] 林伟伟，吴文泰.面向云计算环境的能耗测量和管理方法[J].软件学报，2016，27(4)：1026-1041.
LIN Weiwei, WU Wentai. Energy consumption measurement and management in cloud computing environment[J]. Journal of Software, 2016, 27(4): 1026-1041.

[9] FAN X B, WEBER W D, BARROSO L A. Power provisioning for a warehouse-sized computer[C]// Proceedings of the 34th Annual International Symposium on Computer Architecture, June 9-13, 2007, San Diego, USA: 13-23.

[10] LIN Weiwei. A cloud server energy consumption measurement system for heterogeneous cloud environments[J]. Information Sciences, 2018, 468: 47-62.

[11] 肖丁，贾亚璞，石川.云计算环境下的非线性能耗建模方法[J].北京邮电大学学报，2016，39(1)：107-111.
XIAO Ding, JIA Yapu, SHI Chuan. Non-liner energy consumption model for cloud computing[J]. Journal of Beijing University of Posts and Telecommunications, 2016, 39(1): 107-111.

[12] LIN Weiwei. A hardware-aware CPU power measurement based on the power-exponent function model for cloud servers[J]. Information Sciences, 2021, 547: 1045-1065.

[13] ZHANG Xiao. A high-level energy consumption model for heterogeneous data centers[J]. Simulation Modelling Practice and Theory, 2013, 39: 41-55.

[14] BAI Y, GU L J, QI X. Comparative study of energy performance between chip and inlet temperature-aware workload allocation in air-cooled data center[J]. Energies, 2018, 11(3): 669.

[15] WANG Y, NÖRTERSHÄUSER D, LE MASSON S, et al. An empirical study of power characterization approaches for

servers [C]// ENERGY 2019-The Ninth International Conference on Smart Grids, Green Communications and IT Energy-aware Technologies, June 2-6, 2019, Athens, Greece: 1-6.
[16] WANG Z K, BASH C, TOLIA N, et al. Optimal fan speed control for thermal management of servers[C]// ASME 2009 InterPACK Conference, July 19-23, 2009, San Francisco, USA.
[17] HAM S W. Simplified server model to simulate data center cooling energy consumption[J]. Energy and Buildings, 2015, 86: 328-339.
[18] ZAPATER M, TUNCER O, AYALA J L, et al. Leakage-aware cooling management for improving server energy efficiency[J]. IEEE Transactions on Parallel and Distributed Systems, 2015, 26(10): 2764-2777.
[19] 洪世辉.机架式存储服务器散热及实验研究[D].上海:上海交通大学,2016.
HONG Shihui. Cooling design and experimental study on storage server[D]. Shanghai: Shanghai Jiao Tong University, 2016.
[20] SHIN D, CHUNG S W, CHUNG E Y, et al. Energy-optimal dynamic thermal management: computation and cooling power co-optimization [J]. IEEE Transactions on Industrial Informatics, 2010, 6(3): 340-351.
[21] REDA S, NOWROZ A N. Power modeling and characterization of computing devices: a survey [J]. Foundations and Trends in Electronic Design Automation, 2012, 6(2): 121-216.
[22] LI X, GARRAGHAN P, JIANG X H, et al. Holistic virtual machine scheduling in cloud datacenters towards minimizing total energy[J]. IEEE Transactions on Parallel and Distributed Systems, 2018, 29(6): 1317-1331.
[23] 戚野白,王丹,贾宏杰,等.基于归一化温度延伸裕度控制策略的温控设备需求响应方法研究[J].中国电机工程学报,2015, 35(21):5455-5464.
QI Yebai, WANG Dan, JIA Hongjie, et al. Research on demand response for thermostatically controlled appliances based on normalized temperature extension margin control strategy[J]. Proceedings of the CSEE, 2015, 35(21): 5455-5464.
[24] SONG M, GAO C W, YAN H G, et al. Thermal battery modeling of inverter air conditioning for demand response[J]. IEEE Transactions on Smart Grid, 2018, 9(6): 5522-5534.
[25] Benchmarking[EB/OL]. [2022-07-01]. https://www.aida64.com/products/features/benchmarking.

朱介北(1987—),男,博士,教授,博士生导师,主要研究方向:分布式信息能源系统、电力系统稳定分析与控制、新能源控制。E-mail:jiebei.zhu@tju.edu.cn
刘东昊(1998—),男,硕士研究生,主要研究方向:数据中心节能技术。E-mail:liudonghao@tju.edu.cn
李炳森(1975—),男,通信作者,博士研究生,高级工程师,主要研究方向:信息化管理、数字化转型、电网智能化、能源智慧化等。E-mail:libingsen@sgitg.sgcc.com.cn

(编辑 代长振)

Data Center Server Power Consumption Model Based on Temperature Estimation Considering CPU Working State

ZHU Jiebei[1], *LIU Donghao*[1], *LI Bingsen*[1,2], *YU Lujie*[1], *TIAN Na*[3], *JIA Hongjie*[1]
(1. School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China;
2. State Grid Information & Telecommunication Group Co., Ltd., Beijing 102211, China;
3. Eco-Tech Research Institute of State Grid Hebei Electric Power Co., Ltd., Shijiazhuang 300171, China)

Abstract: To achieve the energy saving in data centers, it is crucial to accurately construct power consumption models for data center servers. The traditional server power consumption model based solely on CPU utilization ignores the dynamic characteristics of server power delay changes influenced by temperature variables, and cannot distinguish the phenomenon of server power layering under different CPU working states of fluctuation or step, resulting in significant model errors. A data center server power consumption model based on temperature estimation considering CPU working state (PMTC) is proposed. PMTC captures the dynamic characteristics of server power consumption delay changes which are affected by CPU core temperature delay changes considering the temperature variables during the model establishment process, eliminates model errors caused by CPU working states based server power consumption layering. During the power consumption calculation process, accurate estimation of the CPU core temperature is performed by judging the CPU working state, avoiding additional measurement of temperature variables. By comparing the PMTC with traditional server power consumption models through the built power consumption testing platform, it is verified that the proposed PMTC restores the dynamic characteristics of server power consumption delay changes without adding additional temperature measurement devices and inputting data dimensions, significantly reducing model calculation errors.

This work is supported by National Key R&D Program of China (No. 2018YFA0702200).

Key words: load modeling; data center; server; power consumption model; thermoelectric coupling