

文章编号: 2095-2163(2023)09-0017-08

中图分类号: TP368.5

文献标志码: A

数据中心服务器功耗模型研究进展

岳天亮, 朱兵, 苗益川, 吕丽

(贵州大学电气工程学院, 贵阳 550025)

摘要: 数据中心为中国经济社会的蓬勃发展提供了强大的支撑力量,同时数据中心的高速发展也带来了巨大的能耗问题。数据中心服务器是能流和热流的基本单元,本文探讨了服务器的分类、不同工作负载下的机架功耗趋势;从3个方面梳理国内外研究中所使用的服务器功耗模型,并对模型的准确性以及模型应用进行了研究,以期为中心发展提供借鉴。

关键词: 数据中心; 服务器; 功耗模型

Research progress of server power consumption model in data center

YUE Tianliang, ZHU Bing, MIAO Yichuan, LV Li

(School of Electrical Engineering, Guizhou University, Guiyang 550025, China)

[Abstract] Data centers have provided a strong supporting force for the vigorous development of China's economy and society; at the same time, the high-speed development of data centers has brought about huge energy consumption problems. As the basic unit of energy and heat flow in data center servers, in this context, this paper discusses the classification of servers, the trend of rack power consumption under different workloads; compares the server power consumption models used in domestic and international studies from three aspects and investigates the accuracy of the models as well as the application of the models, in order to provide reference for data center development.

[Key words] data center; server; power consumption model

0 引言

数据中心不仅是实现多种数字化技术的重要基石,更是推动国家经济转型升级的重要动力。伴随全球数据中心产业的快速发展,中国数据中心产业发展也在不断提升,同时数据中心能源和运营损耗也日渐增加。2019年,中国云计算整体市场规模达1 334亿元,增速38.6%^[1];2019年中国大数据产业规模为5 397亿元,预计到2022年大数据产业规模将突破万亿,未来将成为世界第一数据资源大国和全球数据中心^[2];2018年数据中心总用电量为1 608.89亿千瓦时,预计到2023年数据中心总用电量将增长到2 667.92亿千瓦时,数据中心的电费占数据中心运维总成本60%以上^[3]。

数据中心由4大部分组成:电源设备、冷却设备、IT设备和其他辅助组件。IT设备和冷却设备是

两个主要组成部分,约占数据中心总能耗的90%^[4]。典型数据中心的热流和能流图如图1所示,服务器消耗电力并向热环境散热,决定了制冷系统的制冷负荷,进而影响制冷设备的用电量。而改变冷却系统运行参数会影响服务器的进气温度和气流,影响服务器的能耗。因此,服务器是数据中心热流和能流的基本单元,IT设备和冷却设备因热环境耦合。由于设备效率的不断提高,IT设备和制冷设备的节能将获得最大化,能源转换潜力体现在能源和热管理上。数据中心执行这两个管理功能的必要条件是准确的服务器功耗模型。功耗模型有助于识别优化机会和预测决策的后果,并益于更有效的管理/控制,可最大程度地节约能源。此外,服务器功耗模型在数据中心冷却系统设计和电力趋势预测中发挥着重要作用。

基金项目: 贵州省科技支撑计划项目(2017YFB0902100)。

作者简介: 岳天亮(1994-),男,硕士研究生,主要研究方向:数据中心负荷模拟;朱兵(1966-),女,硕士,高级工程师,硕士生导师,主要研究方向:供热系统规划、分析、运行与控制技术、动力工程和数据中心热管理。

通讯作者: 朱兵 Email: 452179224@qq.com

收稿日期: 2022-09-13

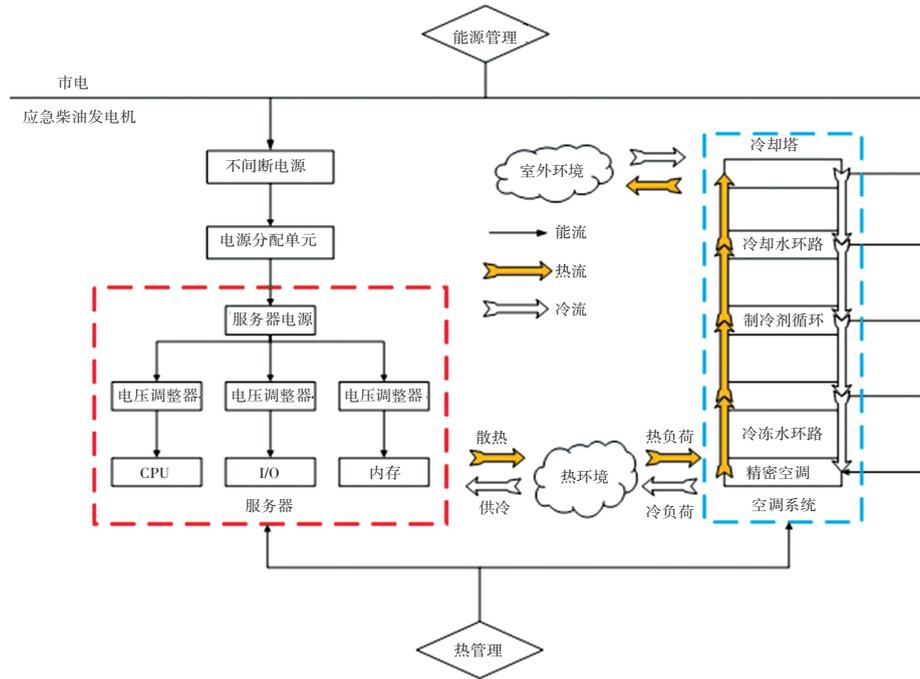


图 1 典型数据中心的能量和热流图

Fig. 1 Heat and energy flow diagram of a typical data center

1 服务器介绍

服务器是在网络环境中提供计算能力并运行软件应用程序的特定 IT 设备。服务器主要构件有：CPU、内存、芯片组、I/O 设备、存储器、外围设备、稳压器、电源和冷却系统。

1.1 服务器分类

根据服务器的配置和结构,分为机架式、刀片式和塔式。机架式服务器在数据中心应用最为广泛,安装在标准的 19 英寸机柜内,有 1U (1U = 1.75 英寸)、2U、4U 等规格;刀片服务器拥有可插入标准机架机箱的“刀片”单元,每个“刀片”均有一块带有处理器、内存和 I/O 设备的板;塔式服务器外形与结构和普通电脑主机类似。根据应用场景,服务器可以分为数据库、文件、邮件、打印、网络、游戏和应用等服务器。

1.2 服务器工作负载

基于工作负载,可将服务器分为科学研究、数据分析、业务处理、云数据、可视化和音频、通信和存储 7 大类^[5]。各类服务器的工作负载类型见表 1,各工作负载类型下的 IT 设备硬件利用率见表 2。

常见工作负载机架功耗趋势如图 2 所示,可知科学研究、数据分析工作负载机架功耗增长较快,这两种负载对服务器 CPU 的利用率高,且随着 CPU

多核、超频、超线程的发展,机架功耗增加较快。通信及存储工作负载对服务器硬盘和 I/O 利用率高,而硬盘和 I/O 的功耗增长缓慢,所以机架功耗增长缓慢。不同类型的工作负载,其功耗趋势不同,对服务器各硬件的利用程度不同,故可根据服务器工作负载的类型预测服务器的功耗。

表 1 服务器常见工作负载类型

Tab. 1 Common types of server workloads

负载类型	举例
科学研究	基于科学的应用,包括生物科学、地球科学、天气预报、工程、仿真、设计、防御、安全以及深度机器学习应用的培训
数据分析	离散数据仓库、数据分析、大数据分析和运行时深度机器学习应用程序
业务处理	管理事务、操作和客户数据库的企业级业务应用程序
云数据	维基、门户网站、社交媒体、视频分享网站、搜索引擎,在线拍卖网站
可视化与音频	维基、门户网站、社交媒体、视频分享网站、搜索引擎,在线拍卖网站
通信	有线和无线网络
存储	专用存储基础设施和服务,包括备份、分层和重复数据删除

表 2 常见工作负载类型下 IT 硬件的利用率

Tab. 2 Utilization of IT hardware for common workload types

负载类型	CPU	内存	I/O 设备	硬盘
科学研究	高	中高	低	低
数据分析	高	中	中高	中
业务处理	中高	中高	中高	中高
云数据	中	中	中	低
可视化与音频	中	中低	中低	中低
通信	中低	中低	中	低
存储	低	低	中	高

注: 中高介于高和中之间, 中低介于中和低之间。

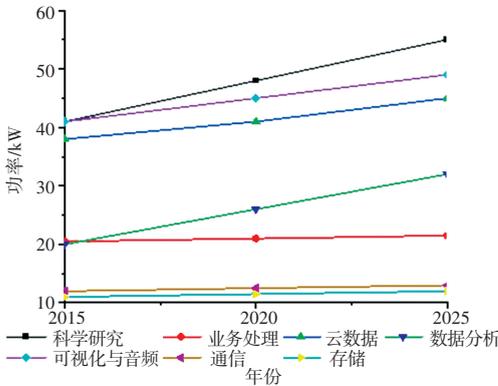


图 2 常见工作负载机架功耗趋势

Fig. 2 Common workload rack power trend

2 服务器功耗模型

建立服务器功耗模型的目的:

- (1) 估计电源管理的潜力;
- (2) 预测服务器或数据中心的能耗;
- (3) 平衡数据中心的需求: 即降低能耗和提供高可靠性。现有功耗模型涵盖了物理机和虚拟机, 可分为相加模型、动态模型和其他模型。

2.1 相加模型

服务器的功耗组成及占比如图 3 所示, 其中 CPU 在服务器功耗中的占比最大, 其次是外围设备、电源、内存、主板、硬盘/存储以及冷却风扇。可将各部分功耗直接相加来建立服务器功耗模型, 如式(1)所示。

$$P_{total} = P_{CPU} + P_{memory} + P_{I/O} + \dots + P_{disk} \quad (1)$$

其中, P_{CPU} 、 P_{memory} 、 P_{disk} 和 $P_{I/O}$ 分别代表 CPU、内存、存储和 I/O 设备功耗。

各元件都有其自身的功耗模型, 不同服务器的配置不同, 其主要耗能元件及其占比会略有不同, 如图 4 所示。

相加模型将数据中心服务器在运行过程中产生

的功耗总量表示为各子系统的能耗总和。一些学者认为 CPU、内存、硬盘、网络 and 风扇是服务器功耗模型的主要组成部分^[6-7], 其它元件的功耗与 CPU 和内存相比较小, 或与 CPU 功耗关联性较好, 即 CPU 和内存是主要组成部分; 所以服务器功耗模型是 CPU、内存和其他设备的函数^[8]

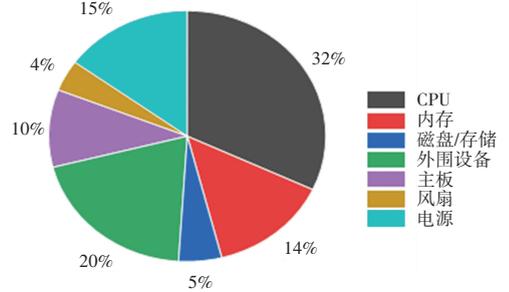


图 3 服务器功耗构成

Fig. 3 Server power consumption composition

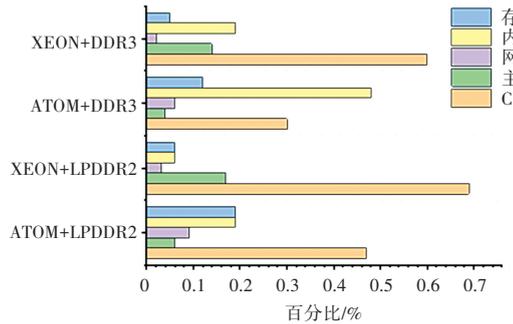


图 4 不同服务器配置下各组件的功耗占比

Fig. 4 Power consumption ratio of components in different server configurations

2.2 动态模型

服务器并不总是处于活动状态, 因此服务器功耗分为:

(1) 静态功耗 (P_{base}), 即空闲功耗, 包括 CPU、内存、I/O 等部件在空闲状态下的功耗, 取固定值;

(2) 动态功耗 (P_{active}), 即服务器功耗取决于工作负载, 服务器功耗与服务器运行状态下的负载处理量、CPU 利用率、任务处理时间和类型等变量相关。

相加模型表示为静态功耗和动态功耗之和^[9-10], 如式(2)所示。

$$P_{BA} = P_{base} + P_{active} + P_{\Delta} \quad (2)$$

其中, P_{Δ} 是修正系数(固定值或表达式)。

将 P_{base} 视为常数, 并将 P_{active} 表示为函数(线性函数、幂函数、高次多项式等)。动态模型可分为回归模型、幂函数模型和多项式模型。

从发热和散热角度出发, 服务器包括 IT 组件和散热组件, 服务器功耗模型转换为 IT 组件和散热组件总和^[11-12]。

2.2.1 回归模型

服务器功耗回归模型考虑了服务器的功能单元的功耗和性能计数器之间的相关性,捕获了固定或空闲功耗以及随着服务器功能单元活动的变化而变化的动态功耗。因此,基于回归的服务器功耗模型也被称为“幂律模型”。由于回归模型的简单和可解释,研究中大多采用回归模型,但服务器工作负载的波动频繁,这些模型不适合跟踪云计算中的服务器功耗^[13]。Fan等^[14]分析了回归模型的精度,认为回归模型可以很好地预测动态功耗,误差低于5%。然而,非线性模型的误差在1%左右,具体取决于应用对象。本文将回归模型分为:简单回归模型、多元回归模型和非线性回归模型。

2.2.1.1 简单回归模型

功耗和捕捉 CPU 活动性能计数器之间的相关性于2000年首次提出。基于此,Fan等^[14]认为服务器的整体功耗位于空闲功耗与峰值功耗区间内时,其值与CPU利用率成一元线性关系,服务器功耗 P_{server} 如式(3)所示,且该模型得到了实验验证。考虑进风温度对服务器功耗的影响,Wang^[15]等人用进风温度的二阶多项式对功耗模型进行了修正。

$$P_{server} = P_{idle} + (P_{active} - P_{idle}) \cdot u \quad (3)$$

其中, P_{idle} 为服务器的静态功耗(W); P_{active} 为服务器的峰值功耗(W); u 为服务器利用率。

Kavanagh等^[16]提出一个类似基于云系统与虚拟机的模型,针对不同应用场景使用不同的自变量,能源用户总功耗 P_x ,如式(4)所示:

$$P_x = H_{idle} + (H_{active} - H_{idle}) \cdot \frac{U_x^{uti}}{\sum_{y=1}^{U_{count}} U_y^{uti}} \quad (4)$$

其中, H_{idle} 和 H_{active} 分别为主机测量的静态功耗与动态功耗; U_x^{uti} 为指定的能源用户的CPU利用率; $\sum_{y=1}^{U_{count}}$ 与 U_y^{uti} 分别为主机上能源用户的计数和指定主机上能源用户集成员的CPU利用率。

Rezaei-Mayahi等^[17]采用简单回归模型,考虑CPU温度和进气温度的差异,研究环境温度对功耗的负面影响,假设CPU温度与服务器利用率呈线性相关,并分析了入口温度与功率增量之间的关系,并反映在公式(2)的增量中。

2.2.1.2 多元回归模型

公式(2)和公式(3)所示简单回归模型基于CPU利用率,而处理角度不同,简单回归模型能够对CPU密集型工作负载产生的能耗进行合理的预

测,但面对I/O和内存密集型应用程序引起的功耗力不从心。一些研究在服务器能耗回归模型中考虑两个及以上影响因素作为自变量,对多元回归模型进行研究。

Alan等^[18]提出CPU利用率、内存读取速度、磁盘和网络利用率函数的功耗模型,其拟合结果证明误差不超过6%,服务器的功耗 $P_{server}(W)$ 如式(5)所示:

$$P_{server} = 14.45 + 0.236 \cdot u_{cpu} - (4.47 \times 10^{-8}) \cdot u_{mem} + 0.00281 \cdot u_{disk} + (3.1 \times 10^{-8}) u_{net} \quad (5)$$

其中, u_{cpu} 、 u_{mem} 、 u_{disk} 、 u_{net} 分别为CPU、内存、磁盘以及网络的利用率。

Lent^[19]假设CPU、磁盘和I/O端口等子系统在各自的利用率方面呈现线性功耗。Witkowski等^[20]则将CPU温度添加到功耗模型中。

(3)非线性模型

Fan等^[14]提出了一个包括校准参数 r 的非线性模型,该参数使平方误差最小,如式(6)所示:

$$P_u = P_{idle} + (P_{max} - P_{idle}) \cdot (2u - u^r) \quad (6)$$

其中, r 是需要通过实验获得的平方误差最小的校准参数,且 r 取决于服务器类型; P_{max} 为服务器的峰值功耗; P_u 为服务器功耗。

在预测服务器功耗方面,式(6)中的模型比回归模型执行得更好。

2.2.2 幂函数模型

幂函数模型基于服务器动态电压频率缩放(Dynamic Voltage Frequency Scaling, DVFS)技术,考虑CPU运行频率(f)、工作电压(V)对服务器能耗影响而建立的模型。借助DVFS技术,根据处理的工作负载强度对CPU工作电压和工作频率灵活调节。对于某些型号,CPU是能耗模型的主要部分,在频率 f 下工作,动态功耗与 $V^2 \times f$ 成比例,当电压降低时,频率也降低。所有其他组件(CPU除外)的功耗是独立的,因此,服务器的功耗可以表示为与工作频率的立方关系^[21]。此外,在云计算中,考虑能源成本和性能之间的权衡,将服务器利用率和服务率视为模型的主要组成部分。然而,一些研究将服务器利用率作为表征功率模型的变量,且服务器功耗与CPU利用率之间存在幂函数关系^[22]。

2.2.3 多项式模型

针对一元线性回归模型和幂函数模型在预测非CPU密集型工作负载的服务器功耗时的不准确性问题,Zhang等^[23]使用高次多项式模型来拟合服务器功耗,发现三次多项式具有最佳拟合度。Lin等

人^[24]提出了服务器能耗与服务器利用率二阶多项式之间的数学模型。

2.3 其他模型

相加模型、动态模型是广泛使用的功耗模型,其他模型也得到了发展。BAI等^[25]使用服务器利用率和芯片温度的乘积来表示功耗。有学者还提出了考虑多核CPU、活动虚拟机数量、服务器相关变量以及CPU利用率或CPU平均温度与进风温度的功耗模型^[26-27]。

数据中心IT设备功耗与服务器处理的工作负载密切相关。负载处理量增加,服务器功耗增加,散热量上升,制冷系统出力增加,数据中心整体能耗增加。量化地研究工作负载灵活调度潜力为数据中心带来用电调节空间,有助于实现数据中心电力负荷的精准调控,关键环节建立工作负载处理过程与服务器能耗之间的映射模型,映射模型主要考虑了工作负载的到达时间、处理速率和数量等影响因素^[28-29]。

近年来,云计算技术为优化和控制数据中心的能耗提供了新的方法手段,不少学者对“云数据中心”能耗进行研究。许多学者利用计算机领域传统的负载均衡和调度、资源分配等基础或改进算法,嵌入各类能耗模型,并结合相关云计算平台调度系统对算法效果进行评估^[30-31]。此外,机器学习方法广泛用于云数据中心的功耗模型^[32]。

2.4 功耗模型对比

现有功耗模型中主要变量统计结果如图5所示。可知模型建立考虑最多的因素是CPU,其次是服务器、温度、内存和负载,网络 and 时钟频率关注较少。一些研究已经验证了模型的准确性。据统计,相加模型、动态模型的误差分别小于9%和11%;对于相加模型,大多数研究的误差小于5%,最坏情况下的错误率为10%;对于动态模型,简单回归模型、多元回归模型、幂函数和多项式等模型的误差一般分别小于4.6%、7%、6%和4%^[33]。Zhang等^[23]比较了线性、二次和三次模型,线性模型的平均误差大部分在8%以下,平均误差为2.74%;二次模型平均误差2.04%;三次多项式平均误差都在3%以下,大多数平均误差低于1.5%;发现三次模型可以很好地拟合测量数据。Lin等^[24]分析了6种功耗模型之间的误差,得出多项式模型的平均误差最小(1.615%),其次是幂函数模型(2.794%)和二次模型(2.974%);系数固定时,线性和幂函数模型的误差分别增加到5.918%和8.222%;仅对线性、幂函数和多项式三者比较发现,线性模型最大和最小误差

分别为1.41%和0.07%,幂函数模型(最大1.29%、最小0.04%),多项式模型(最大1.29%、最小0.05%),这表明3种模型都能准确的反映服务器性能特征。因此,多项式模型最适合服务器的功耗。

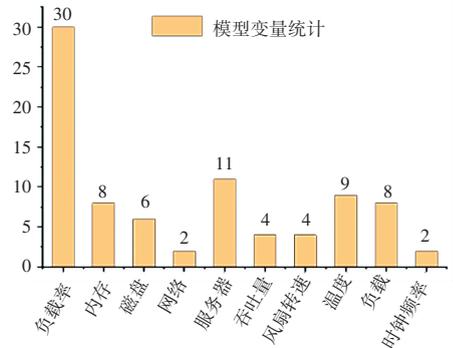


图5 服务器功耗模型主要变量统计结果

Fig. 5 Statistical results of major variables in the server power model

2.5 功耗模型应用

随着信息通信技术的发展,IT设备的功率密度和能耗不断增加,促使冷却需求和冷却系统能耗不断提高。因此,数据中心热管理和能源管理需要进一步加强,服务器作为能流和热流的基本单元,将功耗模型应用于能量预测和管理,以减少服务器和冷却系统能耗。

2.5.1 数据中心(服务器)能源管理

数据中心各部件的功耗模型用于预测各负载段功耗,优化数据中心整体能耗。利用数据中心负载部分的功耗模型,应用不同的功耗优化方法,以确保数据中心的能源效率和成本效益^[34]。为了减少服务器能耗,研究人员通过考虑最佳热环境,提出基于服务器电源模型的节能控制策略。Berezovskaya等^[35]提出了一个用于任意数据中心建模的模块化工具箱,并用该工具箱搭建SICS ICE数据中心模块1的模型,对比了模块1中所有SEE冷却器在两种策略下的能耗发现,当前策略消耗2.5 kWh,而新策略消耗0.7857 kWh,节省1.7177 kWh,12 h节能率为68.6%;Zapater等^[36]分析了计算能力、温度、泄漏和冷却功率之间的关系,并提出了一种冷却管理策略,通过设置运行期间的最佳风扇速度来最小化服务器能耗。整个集群在泄漏和风扇功率方面的节能将随着环境温度的升高而增加,在32℃时最大可节省10.3%。

服务器功耗模型可在操作期间指导服务器电源开/关、工作负载分配和虚拟化。服务器的“空闲”在实践中意味:

- (1) 服务器必须对新工作的请求立即做出反应;
- (2) 服务器允许短暂暂停;
- (3) 服务器有足够的时间唤醒^[37]。

虚拟化技术允许共享服务器、存储和网络设备,以提高资源利用率^[38];在虚拟化中,应用程序能很容易地从一个物理服务器迁移到另一个。管理人员将工作负载分配到指定服务器,其他服务器通过任务需求来关闭电源或进入睡眠模式。Noguchi 和 Nishi^[39]比较了服务器容量和传入请求总和容量,以确定是否启动服务器;Li 等^[40]为了改进传统的性能-能源模型,最大限度地减少云数据中心的能源消耗,提出了一种节能和 QoS 感知 (EEQoS) 的虚拟资源整合模型,并将 EEQoS 模型与粒子群优化算法相结合,通过设置每个 QoS 值功耗为目标函数来整合数据中心的虚拟资源,与传统模型相比,该模型平均减少了 27.2% 响应时间、31.4% 成本和 40.5% SLA 违例率,并平均提高了 13.5% 吞吐量和增加了不到 3.8% 能耗;Arshad^[41]等提出一个基于虚拟机功耗模型整合概念的能效启发式新算法,使用虚拟机整合来最小化云中高能耗,在虚拟机迁移、迁移导致的性能下降、服务水平协议违规和执行时间方面与最先进的技术相比有了显著改进。对数据中心(服务器)的电源/能量管理进行了一些尝试,但随着节能要求的不断提高,基于功耗模型的精细化控制/管理技术成为未来研究的重要方向。

2.5.2 冷却系统负荷计算

制冷供需匹配是制冷系统节能的关键。在制冷系统设计中,制冷量取决于数据中心机房的制冷需求,并通过各种参数来计算总冷却负荷,其中 IT 设备散热是主要部分,设计人员经常将 IT 设备的铭牌功率作为其散热量。根据 SPEC 公司提供的数据,服务器平均充分利用功率仅占铭牌功率的 50%,且服务器并非始终以峰值功率运行,数据中心的热负荷被显著高估,冷却设备被设计为超大尺寸,电力系统冗余过多。Cheung 等^[42]使用简单回归模型对数据中心功耗建模,并利用 2007 年至 2017 年间各个制造商提交给 SPECpower2008 数据库的 491 个数据点,通过线性回归估计服务器的最大和空闲功耗,同时将模型与一个常用的建筑模拟程序进行联合模拟,以亚热带地区一个典型的地下室设有数据中心的大型办公室为对象进行能耗模拟,结果表明该模型能够合理估算服务器、网络和配电负荷,显示了数据中心的运行状态对其运行效率的影响,包括冷却

系统的效率。

此外,服务器功耗模型还用于冷却控制或气流管理,以实现节能。Turk 等^[43]提出了冷却控制策略,通过设置每个机架的入口温度和气流,以基于动态模型以及组合冷却和工作负载管理来最小化数据中心的功率,该策略在冬季节电 1.2%,夏季节电 14.4%;Noguchi 和 Nishi^[44]提出一种主动控制器快门,防止当服务器温度低于 40 °C 时冷空气通过空闲或关闭的服务器泄漏;Ham^[11]等应用模型,考虑热特性以模拟每小时的冷却能耗,发现当送风温度高于 19°C 时,风扇能耗增加,冷却能耗也增加;Athavale^[45]等开发了基于遗传算法的静态和动态框架,用于优化数据中心消耗的冷却能量,在 IT 负载分布粒度的 3 个级别(机房、机架和行)中,机架级分布消耗的冷却功率最小,7.5 h 的动态优化测试案例表明,冷却能耗降低了 21%~50%。

2.5.3 讨论

功耗模型的应用中,准确和简单是主要要求,但两者不兼容,对功耗模型的应用有负面作用。若模型要求准确性,则需要尽可能多的考虑影响服务器功耗的因素,但只考虑服务器主要部件(如 CPU 和内存)的功耗,模型的准确度较差。例如,使用服务器的简单回归功耗模型来获得 IT 负载的功耗,用于进一步评估数据中心内部电源调节系统(IPCS)中的可靠性。Zhang 等^[23]使用服务器功耗的复杂高阶回归模型,通过在云界面中调度任务来提高服务器的电源效率,除了简单之外,还专注于模型的准确性,在使用现有模型预测或评估服务器或数据中心的功耗/能耗时,考察适用性并验证准确性。

3 结束语

服务器功耗模型在数据中心热管理和能源管理中起着至关重要的作用。本文回顾了服务器的分类和负载类型,并对现有的服务器功耗模型进行了总结和分类。此外,作者研究了这些模型的应用场景,经分析得出以下结论:

(1) 数据中心常用的是机架式服务器,不同的服务器组件配置和工作负载意味着工作重心的差异,会导致不一样的能耗变化。现有模型的应用和新模型的开发,需要考虑这些变化和服务器的工作负载。

(2) 现有的功耗模型可分为相加模型、动态模型和其他模型;而动态模型可进一步分为回归模型、幂函数模型和多项式模型。CPU 是功耗模型中考

虑最多的变量,其次是服务器、温度、内存和工作负载。但这些模型都是针对固定环境下的服务器,大部分都适用于单核系统。考虑到简单性、便利性和准确性,无论是在服务器、系统还是房间级别,线性回归和多项式模型都是最佳选择,误差分别小于7%和4%。此外,建立模型需考虑节能技术和进气温度,并应使用分段函数。

(3)建立功耗模型是为了估计能源管理的潜力,预测服务器或数据中心的功耗,平衡能源消耗和可靠性。这些模型已经在IT领域和制冷领域的一些方案中得到应用,但还需要更多的努力。服务器功耗模型可进行准确的冷却负荷计算并避免过多的冗余,但基于模型的能源和热感知管理可实现最佳的整体节能效果。

参考文献

- [1] 云计算发展白皮书[R]. 北京:中国信息通信研究院, 2020.
- [2] 中国数字经济发展白皮书[R]. 北京:中国信息通信研究院, 2020.
- [3] 张文俊,张素芳,王晓焯,等. 点亮绿色云端:中国数据中心能耗与可再生能源使用潜力研究[M]. 北京:华北电力大学,2019.
- [4] VASQUES T L, MOURA P, ALMEIDE A. A review on energy efficiency and demand response with focus on small and medium data centers[J]. *Energy Efficiency*, 2019, 12(5): 1399-1428.
- [5] ASHRAE 90454 - 2018, IT Equipment power trends (third edition)[S]. US-ASHRAE, 2018.
- [6] AHMED K M U, SUTARIA J, BOLLEN M H J, et al. Electrical energy consumption model of internal components in data centers [C]//2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe). IEEE, 2019: 1-5.
- [7] CHATZIPAPAS A, PEDIADITAKIS D, ROTSOS C, et al. Challenge; Resolving data center power bill disputes: The energy-performance trade-offs of consolidation [C]//Proceedings of the 2015 ACM Sixth International Conference on Future Energy Systems. 2015: 89-94.
- [8] DAYARATHNA M, WEN YONGGANG, FAN RUI. Data center energy consumption modeling: a survey[J]. *IEEE Communications Surveys and Tutorials*, 2016, 18(1): 732-794.
- [9] HOGADE N, PASRICHA S, SIEGEL H J, et al. Minimizing energy costs for geographically distributed heterogeneous data centers[J]. *IEEE Transactions on Sustainable Computing*, 2018, 3(4): 318-331.
- [10] KORONEN C, ÅHMAN M, NILSSON L J. Data centres in future European energy systems—energy efficiency, integration and policy [J]. *Energy Efficiency*, 2020, 13(1): 129-144.
- [11] HAM S W, KIM M H, CHOI B N, et al. Simplified server model to simulate data center cooling energy consumption[J]. *Energy and Buildings*, 2015, 86: 328-339.
- [12] GARRAGHAN P, AI-ANII Y, SUMMERS J, et al. A unified model for holistic power usage in cloud datacenter servers [C]//2016 Ieee/Acm 9th International Conference on Utility and Cloud Computing (Ucc), 2016: 11-19.
- [13] WU W, LIN W, HE L, et al. A power consumption model for cloud servers based on elman neural network [J]. *IEEE Transactions on Cloud Computing*, 2019, 9(4): 1268-1277.
- [14] FAN X, WEBER W D, BARROSO L A. Power provisioning for a warehouse - sized computer [J]. *ACM SIGARCH computer architecture news*, 2007, 35(2): 13-23.
- [15] NÖRTERSHÄUSER D, LE MASSON S, MENAUD J M. 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. 2019: 1-6.
- [16] KAVANAGH R, DJEMAME K. Rapid and accurate energy models through calibration with IPMI and RAPL [J]. *Concurrency and Computation: Practice and Experience*, 2019, 31(13): 1-21.
- [17] REZAEI-MAYAH I M, REZAZAD M, Sarbazi-AZAD H. Temperature-aware power consumption modeling in Hyperscale cloud data centers [J]. *Future Generation Computer Systems - the International Journal of Esience*, 2019, 94: 130-139.
- [18] ALAN I, ARSLAN E, KOSAR T. Energy-aware data transfer tuning [J]. 2014 14th Ieee/Acm International Symposium on Cluster, Cloud and Grid Computing (Ccgird), 2014: 626-634.
- [19] LENT R. A model for network server performance and power consumption [J]. *Sustainable Computing: Informatics and Systems*, 2013, 3(2): 80-93.
- [20] WITKOWSKI M, OLEKSIK A, PIONTEK T, et al. Practical power consumption estimation for real life HPC applications [J]. *Future Generation Computer Systems*, 2013, 29(1): 208-217.
- [21] ZHANG Mingming, WANG Songyun, YUAN Guoquan, et al. Energy-efficient real-time task allocation in a data center [C]//2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData). Chengdu: IEEE, 2016: 680-687.
- [22] WU W, LIN W, PENG Z. An intelligent power consumption model for virtual machines under CPU-intensive workload in cloud environment [J]. *Soft Computing*, 2017, 21: 5755-5764.
- [23] ZHANG X, LU J J, QIN X, et al. A high-level energy consumption model for heterogeneous data centers [J]. *Simulation Modelling Practice and Theory*, 2013, 39: 41-55.
- [24] LIN W, WANG W, WU W, et al. A heuristic task scheduling algorithm based on server power efficiency model in cloud environments [J]. *Sustainable computing: informatics and systems*, 2018, 20: 56-65.
- [25] BAI Y, GU L, 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.
- [26] 安娅楠. 基于服务器功耗模型的数据中心热环境研究 [D]. 重庆: 重庆大学, 2019.
- [27] 金超强. 基于服务器功耗模型的数据中心能耗研究 [D]. 重庆: 重庆大学, 2021.
- [28] NIU Tao, HU Bo, XIE Kaigui, et al. Spatial coordination between data centers and power system considering uncertain-ties of both source and load sides [J]. *International Journal of Electrical Power & Energy Systems*, 2021, 124: 106358.
- [29] CHANG C J, CHANG Fumin, KE J C. Optimal power consumption analysis of a load-dependent server activation policy for a data service center [J]. *Computers & Industrial Engineering*, 2019, 130: 745-756.
- [30] DONG Ziqian, ZHUANG Wenjie, ROJAS-CESSA R. Delayed best

- fit task scheduling to reduce energy consumption in cloud data centers[C]//2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData). Atlanta: IEEE, 2019:729-739.
- [31] LIANG Bin, DONG Xiaoshe, WANG Yufei, et al. Memory-aware resource management algorithm for low-energy cloud data centers [J]. Future Generation Computer Systems, 2020, 113: 329-342.
- [32] 周舟, 袁余俊明, 李方敏. 云数据中心服务器能耗建模及量化计算[J]. 湖南大学学报(自然科学版), 2021, 48(4): 36-44.
- [33] 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.
- [34] AHMED K M U, BOLLEN M H J, ALVAREZ M. A review of data centers energy consumption and reliability modeling[J]. IEEE Access, 2021, 9: 152536-152563.
- [35] BEREZOVSKAYA Y, YANG C W, MOUSAVI A, et al. Modular model of a data centre as a tool for improving its energy efficiency[J]. IEEE Access, 2020, 8: 46559-46573.
- [36] 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, 2014, 26(10): 2764-2777.
- [37] ASHRAE DATACOM SERIES BOOK 12 - 2015, Server Efficiency-metrics for computer Servers and Storage [S]. US - ASHARE, 2015.
- [38] REZAEI-MAYAHI M, Rezaad M, Sarbazi-Azad H. Tem-

- perature-aware power consumption modeling in hyperscale cloud data centers [J]. Future Generation Computer Systems the International Journal of Escience, 2019, 94: 130-139.
- [39] RASMUSSEN N. Calculating total cooling requirements for data centers[J]. White paper, 2007, 25: 1-8.
- [40] HONGJIAN LI. Energy-efficient and QoS-aware model based resource consolidation in cloud data centers [J]. Cluster Computing, 2017, 20(3): 2793-2803.
- [41] ARSHAD U, ALEEM M, SRIVASTAVA G, et al. Utilizing power consumption and SLA violations using dynamic VM consolidation in cloud data centers[J]. Renewable and Sustainable Energy Reviews, 2022, 167: 112782.
- [42] CHEUNG H, WANG S, ZHUANG C, et al. A simplified power consumption model of information technology (IT) equipment in data centers for energy system real-time dynamic simulation [J]. Applied energy, 2018, 222: 329-342.
- [43] ZAPATER M, TURK A, MOYA J M, et al. Dynamic workload and cooling management in high-efficiency data centers [C]//2015 Sixth International Green and Sustainable Computing Conference (IGSC). IEEE, 2015: 1-8.
- [44] NOGUCHI T, NISHI H. Active controlled shutter for effective cooling of servers in data center [C]//IECON 2015-41st Annual Conference of the IEEE Industrial Electronics Society. IEEE, 2015: 001668-001673.
- [45] ATHAVALA J, YODA M, JOSHI Y. Genetic algorithm based cooling energy optimization of data centers [J]. International Journal of Numerical Methods for Heat & Fluid Flow, 2021, 31(10): 3148-3168.

(上接第16页)



图7 连续12帧含干扰因素图像的检测结果

Fig. 7 Detection results of 12 consecutive frames of images with interference factors

3 结束语

本文利用气体形状和灰度分布特征提出了一种新的气体泄漏检测算法。该算法首先使用高斯滤波对输入红外视频中的各帧图像进行去噪预处理;其次,以稳定无干扰的第一帧图像作为背景帧,后续每一帧图像与背景帧做背景减除法,获取疑似气体泄漏连通区域;最后,通过形状特征和气体的灰度特征来判断连通域是否为不可见气体泄露的区域。该算法的优势在于使用了气体自身的特有规律来分析判断,即使存在外界干扰,也能更精准的排除非气体区域。

参考文献

- [1] 王建平, 李俊山, 杨亚威, 等. 基于红外成像多特征联合的乙烯气体泄漏检测[J]. 传感器与微系统, 2015, 34(6): 119-122.
- [2] 刘路民根, 张耀宗. 一种基于形状的红外图像泄漏气体检测方法[J]. 应用光学, 2019, 40(3): 468-472.
- [3] 李红娣, 袁非牛. 采用金字塔纹理和边缘特征的图像烟雾检测[J]. 中国图象图形学报, 2015, 20(6): 772-780.
- [4] 洪少壮. 基于视觉的红外气体泄漏检测算法研究[D]. 大连: 大连海事大学, 2020.
- [5] 陈长友, 杨健晟. 面向视频图像的烟雾检测算法综述[J]. 激光与光电子学进展, 2021, 58(4): 29-43.
- [6] 丁怀对. 基于运动块追踪的视频烟雾探测技术研究[D]. 合肥: 中国科学技术大学, 2015.