

基于物联网的大数据中心能耗管理系统研究

刘建, 朱小栋

(上海理工大学管理学院, 上海 200093)
✉liujian568@pingan.com.cn; zhuxd@usst.edu.cn



摘要: “双碳”背景下, 大数据中心如何降低能耗, 发展绿色计算已成为一个迫切需要解决的问题。大数据中心的能耗管理是一项系统而复杂的工程, 在数字化浪潮中, 依托物联网技术建立一个大数据中心的能耗管理平台, 为后续的能耗优化提供一个可量化的系统则不失为一个很好的切入点。本文通过分析大数据中心涉及能耗优化的各个方面, 包括建筑物、机房设施、IT系统等, 利用物联网、云计算技术建立一个统一的能耗管理平台, 重点展示分析大数据中心能耗数据, 为各种节能技术和能耗优化方法提供一个坚实的底座。

关键词: 能耗管理; 物联网; 大数据中心; 绿色计算

中图分类号: TP399 **文献标识码:** A

Research on Energy Consumption Management System of Big Data Center based on Internet of Things

LIU Jian, ZHU Xiaodong

(Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)
✉liujian568@pingan.com.cn; zhuxd@usst.edu.cn

Abstract: Under the background of "double carbon", how to reduce energy consumption and develop green computing in big data center has become an urgent problem. Energy consumption management of big data center is a systematic and complex project. In the wave of digitalization, it is a good idea to build a large data center energy consumption management platform based on the Internet of Things (IoT) technology to provide a quantifiable system for subsequent energy consumption optimization. This paper proposes to analyze all aspects of energy consumption optimization of big data center, including buildings, computer room facilities, IT systems, etc., and use Internet of Things and cloud computing technologies to establish a unified energy consumption management platform, which focuses on the analysis of energy consumption data of big data centers and provides a solid foundation for various energy-saving technologies and energy consumption optimization methods.

Keywords: energy management; IoT; big data center; green computing

1 引言(Introduction)

新型基础设施建设(简称“新基建”)包含大数据中心、人工智能、物联网(IoT)等领域, 与传统基建相比, “新基建”内涵更加丰富, 更能体现数字经济特征。2022年是“东数西算”工程启动元年^[1], 在“新基建”“东数西算”工程引领下, 大数据中心建设正如火如荼, 与此同时绿色计算、低碳发展等理念也对数据中心能效提出了更高的要求。据《中国“新基建”发展研究报告》显示, 到2025年, 数据中心将占全球能耗的33%^[2]。在“双碳”理念逐步深化的背景下, 如何

降低数据中心能耗、发展绿色计算就越发凸显其重要性。

2 背景及相关工作(Background and related work)

关于大数据中心的能效研究, 目前业内已有不少成果, 如谷歌利用机器学习来自动管理数据中心的冷却系统, 对降低能耗产生了很好的效果。为了提高数据中心的散热效率, 李永利等人提出了一种基于机器学习的数据中心主动地板模型^[3]。国冰磊等人研究了数据库系统的能耗问题, 并以查询负载为主要建模对象构建了一个可以有效降低系统能耗的模型^[4]。

大数据中心的能耗涉及很多方面, 不仅有传统的照明、

制冷等因素，还涉及IT系统层的服务器、交换机等方面。各种因素间彼此也不是孤立的，如IT系统的负载、建筑物的材料等都会对制冷设备的能耗产生重大影响^[5]。大数据中心如何降低能耗是一个系统的问题，现有的研究都有其局限性和片面性，这主要体现在两方面：(1)现有研究大多只关注在某一层面，如制冷、供电或IT系统资源调度、性能优化等，没有综合考虑大数据中心各个层面的能效问题^[6]。(2)现有研究提出很多创新的节能手段和模型算法，但是对最基础的能耗监测和能耗数据获取没有系统的描述和展现^[7]。本文尝试从大数据中心能耗监控的角度，依托物联网技术建立一个统一的能耗管理系统，以解决目前能耗管理中面临的一些实际问题。

3 架构设计(Architecture design)

3.1 大数据中心能耗分析

根据PUE(Power Usage Efficiency)等数据中心能耗指标可以把大数据中心的能耗来源简单分为两部分：IT系统能耗和非IT系统能耗^[8]。IT系统包括服务器、网络设备、数据库系统等；非IT系统主要是指辅助设施，包括制冷设备、照明系统及其他建筑物等，具体如图1所示。区别于传统建筑物，大数据中心的能耗管理需要更多集中在IT系统层面。随着云计算时代的来临，标准化的通用服务器占据了大数据中心的半壁江山，以更细的粒度来研究计算、网络、存储等的能耗也是一个新的趋势^[9]。

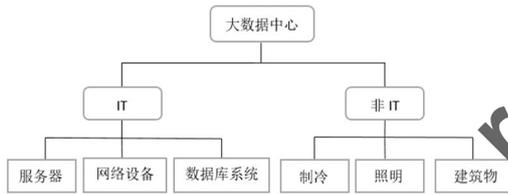


图1 大数据中心能耗分析

Fig.1 Energy consumption analysis of big data center

3.2 架构设计

系统按照逻辑架构分为采集层、传输层、应用层，具体如图2所示。



图2 系统逻辑分层

Fig.2 System logic layering

采集层：大数据中心各个设备差异较大，有传统的建筑设施如供电和供水设备等，也有服务器、交换机等IT设备，如何有效地采集并传输各种设备的能耗数据是能耗管理过程中最基本也是最核心的问题。物联网(Internet of Things, IoT)是指在互联网基础上，通过射频识别技术、传感器技术等完成任何物体与网络的交互连接，从而实现物品的信息获取

和智能应用。利用物联网技术，可以很好地将大数据中心各个设备的能耗数据统一采集到管理系统中^[10]。采集层中涉及的设备包括各类独立运行的传感设备，还有集成在空调、服务器和交换机等中的传感芯片等。

发展绿色计算除了使用绿色清洁能源外，还有一个重要方向就是降低大数据中心的电力消耗。能耗数据的采集很大一部分就是电力数据的采集，利用物联网技术来推进智能电网的发展是一个很好的方向^[11]。对于服务器的电力采集，可以通过多种方式获取更为精细的数据，如通过机柜智能PDU(Power Distribution Unit, 电源分配单元)来获取IT设备的整体用电数据，通过服务器的带外管理接口BMC(Baseboard Management Controller)或IPMI(Intelligent Platform Management Interface)来获取更详细的能耗数据^[12]。

传输层：通过采集层采集的数据经过智能网关、AI-BOX(人工智能盒子)等设备传输至应用层。鉴于很多物联网设备不具备物理网口的特性，传输网络可以采用大数据中心本地有线网络和无线网络相结合的形式。智能盒子等边缘设备对采集层的数据进行简单的压缩、去重等处理，可以极大地收敛传输层的数据量。传输层还需要考虑数据流动的安全性，必要的安全认证手段既可以规范众多分散部署的采集装置，也为应用层的设备管控提供支撑。

应用层：应用层是整个系统的中枢神经，包括设备管理、数据展现、能耗分析和监控等功能。设备管理是指对传感设备等物联网终端进行管理，包括接入、设备启停、参数配置和日志打印等操作。传输层简单清理后的数据，通过特定的模型和规则处理后用于可视化展示、大屏监控等方面。能耗分析是指对大数据中心的各个模块子系统的能耗数据进行系统性的分析评估，并提供详细的评估报告和待优化项^[12]。应用层预留公共接口和开放API,为后续其他能耗优化方案提供对接支持。应用层可以通过和其他平台的对接，如服务器的硬件管理系统和机房制冷控制系统等来测试具体能耗优化方案，并通过展示平台确认最终效果。

3.3 功能模块

管理系统采用C/S架构设计，具体模块包括仪表盘、设备管理、能耗分析、系统管理等。

(1)仪表盘：仪表盘通过各种图表系统展现大数据的能耗情况，显示维度为：大数据中心→设备类型→IT类/服务器→计算CPU。用户可以根据需求自定义仪表盘显示页面，通过勾选不同数据源组合来适应监控大屏、手机端等场景。

(2)设备管理：设备管理是整个系统的基础模块，包括网关管理、子设备信息、指令管理等功能，具体如图3所示。网关管理用于对物联网网关(IoT-Hub)进行统一管理，包括网关新增、子设备接入、安全管控等功能。在平台完成网关设备接入后，需要分别对不同网关下的子设备进行接入管理，包括IT层设备和非IT层设备，接入类型可以是物联网无线网络、有线局域网和蓝牙网络等。网关的安全管理主要用于子设备与网关间安全密钥的设置和存储。子设备信息包含设备

基础信息，如类别、名称、设备标签、运行时间等，还可以新增扩展信息，如模组信息、固件版本、模组品牌和SDK语言等。指令管理是指对发送给子设备的命令、脚本或者程序包等的管理。常规指令包含设备启停、数据包传输、参数调整等，高级指令包括如定制化的SDK包、人工智能算法，以及后续的一些能耗优化模型等。

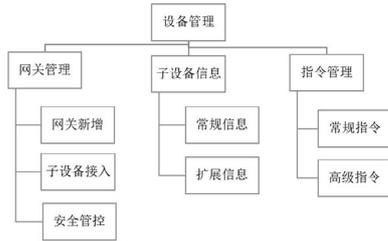


图3 设备管理

Fig.3 Device management

(3)能耗分析：能耗分析分为能耗评估和能耗监控告警两部分。能耗评估是指依托大数据平台对能耗数据进行多维度的分析和评估并最终形成报告。能耗监控告警是本系统的核心功能，包括监控项管理、模板管理、告警设置等。本系统能耗监控模块与IT系统监控如Zabbix、Prometheus等监控软件，还有第三方的设备监控平台如数据中心的动环监控平台、制冷系统监控和专业服务器监控系统等有很大的不同。主要体现在两方面：一是监控的指标更聚焦于能耗方面，采集的能耗数据也更全面且粒度更细；二是在能耗分析和评估的基础上，能耗监控能更好地利用AI告警、大数据预测等技术来提升监控告警的时效性和准确性。

(4)系统管理：系统管理主要是指平台后台配置管理功能，包括租户管理、用户管理、页面配置、区域设置等。

4 系统设计(System design)

本系统采用“自下而上”的设计方法，主要包括边缘平台数据采集传输、设备管理及指令下发和大数据分析处理等方面。

4.1 逻辑拓扑图

系统整体包含数据中心边缘平台、云端管理平台及第三方平台三个部分。边缘平台主要包括物联网子设备、物联网网关和数据清洗压缩设备。管理平台采用云环境部署，统一处理边缘平台采集的数据，并提供开放API接口与第三方服务对接。系统逻辑拓扑图如图4所示。

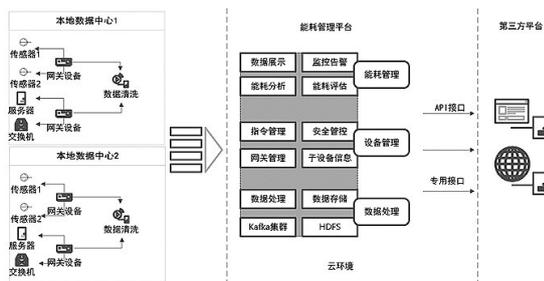


图4 逻辑拓扑图

Fig.4 Logical topology diagram

4.2 部分模块代码设计

设备管理是能耗管理系统最基础的功能模块，物联网设备的接入、信息查询及其指令管理是系统稳定运行的保障。下面以设备管理模块为例设计部分程序，图5是部分功能接口设计。

模块	功能	接口	功能描述
设备信息	查询网关列表	queryHubList	获取当前租户下的网关列表
	查询子设备列表	queryDeviceList	获取当前租户下的子设备列表
	查询网关详情	queryHubDetail	获取当前租户下的网关信息
	查询子设备详情	queryDeviceDetail	获取当前租户下的子设备信息
指令管理	Hub 扫描	sendScanMessage	下发 Hub 扫描命令
	新增指令包	addInstructionPackage	新建指令包
	删除指令包	deleteInstructionPackages	删除已经配置的指令包，可多选
	指令包发送	sendInstructionPackage	向设备发送配置好的指令包
	获取指令信息	queryInstructionPackages	获取已经配置的指令包信息
	指令包执行	executionInstructionPackages	子设备上初始化执行已经缓存的指令包
	查询执行结果	queryExecutionResults	查询指令包执行结果

图5 部分接口设计

Fig.5 Part of the interface design

服务端部分代码如下：

```

/**
 * 使用MQ向设备发送已经配置的指令包
 * @param instructionPackage
 */
public void sendInstructionPackage(
    InstructionPackage deviceId) {
    Channel===null;
    Connection conn===null;
    logger.info("加载rabbitmq配置并生成channel");
    ConnectionFactory factory===new
    ConnectionFactory();
    //设置相关参数，地址，端口，账号，密码
    String host===RabbitmqProperties.
    getProperty("host");
    String userName===RabbitmqProperties.
    getProperty("userName");
    String passWord===RabbitmqProperties.
    getProperty("passWord");
    factory.setHost(host);
    factory.setPort(5672);
    factory.setUsername(userName);
    factory.setPassword(passWord);
    PackageMessage===new PackageMessage();
    try{
    conn===factory.newConnection();
    channel===conn.createChannel();
    instructionPackage.bindDevice(deviceId);

```

```
String addPackageMessage==packageMessage.add
PackageMessage(instructionPackage,instructionPackage.
getCmd());
SendRabbitmq.sendMessage(channel,
addPackageMessage);
} catch (Exception e){
logger.info("this is addPackageMessage:"+e.
getMessage());
}finally {
if (channel !=null) {
try {
logger.info("指令包发送关闭 channel");
channel.close();
} catch (IOException e) {
e.printStackTrace();
}
}
if (conn !=null) {
try {
logger.info("指令包发送关闭 conn");
conn.close();
} catch (IOException e) {
e.printStackTrace();
}
}
}
}
```

4.3 管理系统部署拓扑图

管理系统主体使用云平台部署，业务流通过VIP01进入消息队列，数据流经VIP02进入Kafka集群最终传输至后端的Hadoop平台。管理平台的指令管理和设备管理任务通过MQ集群发布，边缘网关作为订阅者消费信息。本系统部署两套消息队列主要是为了将业务流与数据流进行隔离，确保业务系统的稳定。应用服务集群采用Java语言开发，组件服务一般部署多台，前置LVS+Nginx(LVS: Linux Virtual Server; Nginx: 高性能负载均衡软件)来提高负载和保证高可用。系统部署拓扑图如图6所示。

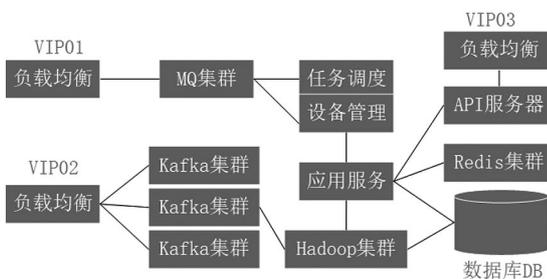


图6 部署拓扑图

Fig.6 Deployment topology diagram

5 结论(Conclusion)

本文在现有文献基础上，系统全面地分析了大数据中心各层次能耗因素，合理运用物联网、云计算和大数据等技术设计了能耗管理系统，旨在解决大数据中心能耗管理过程中的数据展示、能耗告警及能耗分析评估等问题。绿色计算是未来算力发展的新趋势，在“双碳”背景下，如何提高大数据中心的效率、降低能耗是一个持久的命题。由于时间有限，本文研究难免有不足之处，如数据采集模块的统一和规范问题、能耗分析模型的设计问题等。后续的研究除了要进一步明确本系统的设计细节，还需关注人工智能、新能源等技术在能耗方面的应用。

参考文献(References)

- [1] 叶毓睿.剖析“十四五”规划中的“软件定义”和“东数西算”[J].中国工业和信息化,2022(06):64-70.
- [2] 叶菁.云计算搭上“碳中和”国内云计算市场加速绿色变革[N].通信信息报,2021-09-08(008).
- [3] 李永利,牛弘勋,周杰,等.基于机器学习的数据中心主动地板模型研究[J].计算机仿真,2019,36(12):180-185,248.
- [4] 国冰磊,于炯,杨德先,等.面向关系数据库查询的能耗建模及计划评价[J].计算机研究与发展,2019,56(04):810-824.
- [5] 袁晓东.云计算与AI技术在数据中心供冷系统中的应用与实践[J].电信技术,2019(09):71-74.
- [6] AHMAD A, KHAN SU, KHAN HU, et al .Challenges and practices identification via a systematic literature review in the adoption of green cloud computing: Client's side approach[J]. IEEE Access, 2021, 9:81828-81840.
- [7] 潘亚刚.基于物联网的建筑物综合环境能耗监测管理系统[J]. 居舍,2021(33):172-174.
- [8] 王继业,周碧玉,张法,等.数据中心能耗模型及能效算法综述[J].计算机研究与发展,2019,56(08):1587-1603.
- [9] 吴云芸,方家琨,艾小猛,等.计及多种储能协调运行的数据中心实时能量管理[J].电力自动化设备,2021,41(10):82-89.
- [10] 刘超,孙保东.基于电力物联网的智能配用电信息采集技术研究[J].电气自动化,2022,44(03):75-78.
- [11] 高仕斌,高凤华,刘一谷,等.自感知能源互联网研究展望[J]. 电力系统自动化,2021,45(05):1-17.
- [12] 周清,张聃晟,沈子钰,等.数据中心内服务器能耗数据采集及特征分析[J].数据采集与处理,2021,36(05):986-995.

作者简介:

刘 建(1986-),男,硕士生.研究领域:云计算,物联网,IT服务管理.
 朱小栋(1981-),男,博士,副教授.研究领域:大数据管理,数据挖掘,云计算,信息安全.