

Research on the Development Trends of Data Centers in the Era of Computing Power

Kaikai Yin

Beijing Branch, China Telecom Company Limited, Beijing, 100032, China

Abstract

This paper summarizes the current development status of computing power and data centers, elaborates on the challenges in data center computing power scale and performance, energy consumption and environmental impact, data center operation and security, analyzes the future development trends of data centers, explores the development of efficient, cloud based, intelligent, green and low-carbon, and sustainable data centers, aiming to provide decision-makers and researchers in the data center industry with provide valuable insights for construction and operation personnel to address future challenges and promote sustainable development.

Keywords

computing power; large model; data center; artificial intelligence computing center; computing power infrastructure

算力时代下数据中心发展趋势研究

殷凯凯

中国电信股份有限公司北京分公司, 中国·北京 100032

摘要

论文梳理了算力和数据中心发展现状, 阐述了数据中心算力规模与性能、能源消耗与环境影响、数据中心运营与安全等方面的挑战, 分析了数据中心未来发展的趋势, 探讨了高效、云化、智能化、绿色低碳化和可持续性数据中心的发展, 旨在为数据中心行业的决策者、研究人员、建设运营人员等提供有价值的见解, 以应对未来的挑战并促进可持续发展。

关键词

算力; 大模型; 数据中心; 智算中心; 算力基础设施

1 引言

2023年10月8日, 中华人民共和国工业和信息化部等六部门联合印发《算力基础设施高质量发展行动计划》明确支持算力基础设施高质量发展。数据中心是算力基础设施的核心组成部分, 是支撑数字经济、赋能千行百业数字转型的重要力量。随着互联网、大数据、人工智能等新一代信息技术的快速发展, 尤其是随着AIGC、大模型等算力新应用、新业态不断涌现, 全社会对算力的需求快速增长。

数据中心是承载并生产算力的最基础底座, 随着算力时代的到来, 数据中心行业面临多方面的挑战和机遇, 如何研判数据中心发展趋势和实现高质量发展, 是当前和今后一个时期的重要课题, 论文通过分析研究旨在为数据中心行业的决策者、研究及建设运营人员提供些许有价值的参考。

【作者简介】殷凯凯(1986-), 男, 中国山东人, 硕士, 工程师, 从事数据中心、智算中心等算力基础设施规划、建设、运营管理, 算力服务管理等研究。

2 算力及数据中心发展现状

2.1 算力发展情况

算力是集信息计算力、网络运载力、数据存储力于一体的新型生产力, 主要通过算力基础设施向社会提供服务^[1]。中国信通院将算力分为通用算力、智能算力、超算算力和边缘算力^[2]。通用算力以CPU芯片输出的计算能力为主, 超算算力主要是以超级计算机输出的计算能力为主。智能算力则以GPU、FPGA和AI芯片等输出的人工智能计算能力为主, 具备渲染、推理和模拟能力, 可面向智能驾驶、人脸识别、大模型等人工智能应用提供智算服务的一种算力服务形态^[3]。

当前中国算力规模在快速增长, 尤其是以GPU等为代表的智能算力在爆发增长。根据中国信通院数据, 截至2022年底, 中国算力总规模为180EFLOPS, 其中通用算力规模为137EFLOPS, 智能算力规模为41EFLOPS, 超算算力规模为2EFLOPS。调研数据显示预计2021—2026年期间中国智能算力规模年复合增长率达52.3%, 2026年中国智能算力规模将达到1,271.4EFLOPS(FP16)^[4]。

2.2 数据中心发展情况

数据中心是承载算力资源，对外提供高性能算力服务的关键新型基础设施。算力规模的高速发展离不开数据中心的爆发增长。根据工信部通报数据，截至2023年6月全国在用数据中心机架总规模超过760万标准机架，算力总规模达到197EFLOPS（EFLOPS是指每秒百亿亿次浮点运算次数），位居全球第二，五年年均增速超过30%。在国家“东数西算”工程的推进下，数据中心大多数集中在京津冀、长三角、粤港澳等热点区域的8个枢纽节点和10个IDC集群。目前数据中心建设呈现东西部更加均衡、规模集约化和集群化部署的特征。

根据中国数据中心产业发展白皮书（2023年）总结，数据中心正从计算中心、信息中心、云数据中心加速向算力中心演进。目前，随着AI大模型等应用快速增长，数据中心正处于从传统模式向云化、智能化、智算化、绿色低碳化转型。

3 算力时代数据中心的挑战

3.1 数据中心算力规模与性能

数据中心算力规模与性能问题是数据中心运营中的一个关键挑战。随着人工智能大模型和数字化转型的推进，数据中心必须应对日益增长的算力需求。根据网络公开资料ChatGPT训练阶段总算力消耗约为3640PF-days（即1PetaFLOP/s效率训练3640天），总训练成本约为1200万美元，AI大模型的发展需要强大的智能算力支撑。这导致了两个主要问题，即算力规模扩展和性能压力。一方面，数据中心必须不断扩展其规模，以满足不断增长的计算需求。这意味着需要更多的物理空间来容纳服务器、存储设备和网络基础设施。扩展规模的过程需要巨额的资金投入，包括建设、设备采购和维护成本。此外，这也可能会增加能源消耗，对环境产生负面影响。另一方面，性能问题是由于数据中心运行的应用程序变得日益复杂和资源密集而产生的。这些应用程序需要更多的计算能力和存储资源，以确保高效运行^[5]。然而，维持高性能并不容易，因为需要优化硬件和软件配置，以应对不断变化的工作负载。

3.2 能源消耗与环境影响

高能源消耗和环境影响问题是数据中心领域的另一个重要关注点。数据中心属于高密、高算力的信息基础设施，算力业务需求持续推升服务器和芯片性能和功率，导致整个数据中心呈现高能耗、高成本特点。在“双碳”宏观形势下，政府部门对数据中心PUE（电能利用效率）监管要求不断提高，像明确要求国家算力东、西部枢纽节点数据中心PUE分别控制在1.25和1.2以下。其中，数据中心的高能源需求导致环境问题，如碳排放和自然资源消耗的增加。解决这一问题的方法之一是采用绿色或可再生能源，以减少对传统能源的依赖，从而降低环境影响。此外，数据中心行业

还在努力提高能源效率，采取绿色节能设备、技术和最佳实践，以减少能源消耗。

3.3 数据中心运营与安全

数据中心运营与管理方面的挑战包括多方面的任务和决策。数据中心运营需要高度专业化的知识和技能，以确保硬件和软件资源的高可用性和性能，这包括监控、维护、升级和优化各种资源系统。数据中心管理也牵涉到网络与数据的安全性、合规性和隐私问题。数据中心必须应对不断增加的网络威胁和安全漏洞。黑客攻击和恶意软件威胁可能导致数据泄露、服务中断和巨大损失。因此，数据中心必须采取严密的安全措施，包括入侵检测、防火墙、安全补丁管理和访问控制，以确保数据中心的安全^[6]。保护数据与网络的安全和合法性是关键挑战之一，因此数据中心必须采取适当的安全措施，包括入侵检测、访问控制和数据加密。同时，合规性要求数据中心遵守相关法规和标准，以确保用户数据的隐私和合法性。

4 数据中心未来发展趋势

4.1 人工智能、边缘计算等新技术带来的数据中心发展趋势

人工智能和机器学习应用方面。在未来，数据中心将进一步融合人工智能（AI）和机器学习（ML）技术。这些技术的应用将推动数据中心性能和效率的提升。首先，数据中心将会完成AI大模型的训练和推理作用，这涉及大规模数据的处理和模型的训练，需要强大的计算能力和存储资源。此外，AI和ML将在数据中心中用于资源管理和故障检测。智能资源分配和自动化的运维将帮助提高数据中心的运营效率。

边缘计算应用方面，边缘计算将成为数据中心未来的一个主要趋势。边缘计算将计算资源放置在离数据生成源更近的位置，以减少延迟和提高实时性。这意味着数据中心需要在边缘设备和云数据中心之间建立更多的分布式架构。这将促使数据中心采用更灵活的硬件和软件架构，以满足边缘计算的需求。

量子计算应用方面，量子计算被认为是未来数据中心领域的一个具有巨大潜力的变革因素。它引入了一种全新的计算范式，具有破解传统加密算法、模拟复杂分子结构和解决优化问题等方面的潜在能力。因此，数据中心将需要建立量子安全的加密机制，以保护存储在其中的敏感信息，如个人数据、金融交易记录等。

区块链技术应用方面，数据中心将充当区块链网络的关键基础设施支持者，提供节点托管、交易验证和智能合约执行等服务。这些任务要求数据中心具备高度的技术专业知识和安全性，以确保区块链网络的可用性和数据的安全性。区块链技术还带来了大规模数据处理和存储需求，数据中心需要不断扩展其存储和计算资源以满足这些需求。

4.2 可持续性与绿色低碳带来的数据中心发展趋势

可再生能源的应用方面,在可持续性与绿色数据中心的发展中,可再生能源的应用是一项关键举措。数据中心行业对电力的巨大需求对环境产生了负面影响,包括碳排放和自然资源消耗。为了减轻这一影响,越来越多的数据中心开始采用可再生能源,如太阳能和风能。太阳能和风能作为可再生能源,对数据中心的可持续性和绿色运营产生了深远的影响。这些能源不仅有助于降低数据中心的碳足迹,还在经济和环境方面带来了多重好处。太阳能和风能是清洁能源来源,不产生温室气体排放,因此有助于减少数据中心的环保负担。通过将太阳能电池板和风力涡轮机纳入数据中心的能源供应链中,数据中心可以减少对传统化石燃料的依赖,减少二氧化碳和其他有害排放物的释放。这有助于满足环境法规的要求,降低环境风险,并为数据中心在社会责任方面树立良好的形象^[7]。

在运营中,数据中心需要综合考虑地理位置、气象条件和资源供应等因素,以最大化太阳能和风能的利用。此外,数据中心还需要建立可再生能源的储能系统,以便在夜间或低风能时继续供电。通过合理规划和布局,数据中心可以将可再生能源纳入其能源供应链中,实现更环保和可持续的运营。通过合理规划和布局,数据中心可以充分利用可再生能源,将其纳入能源供应链中,实现更环保和可持续的运营。

节能技术应用方面,数据中心可持续发展和环保目标的实现不仅取决于可再生能源的应用,还在于节能技术的持续发展。数据中心作为能源密集型行业,对电力的大量需求既带来了高昂的能源成本,又对环境产生了不小的压力。因此,节能技术的引入和发展成为至关重要的措施。数据中心通过采用高效的液冷服务器和存储设备,实现了显著的节能。此外数据中心运营商还在不断优化设备配置,确保它们在负载低时进入低功耗模式,进一步减少能源消耗。同时,通过使用高效的冷却设备、热通道隔离和冷热通道分离等方法,数据中心能够将冷却能源的使用降至最低。并引入智能温度控制系统,根据实际需要动态调整温度,以提高能源效率^[8]。此外,还可以引入AI智能化运维和智能能源管理系统,通过实时监测和分析数据中心的能源消耗情况,提供实时反馈和优化建议。

4.3 大算力需求带来的数据中心发展趋势

2020年以后,云计算、大数据、AI等新数字技术的加速发展,驱动数据云存储及计算、智能算力、边缘算力等需求持续增长。尤其是近一年以ChatGPT代表的生成式AI应用和大模型快速发展,带来算力需求爆发式增长。未来数据

中心算力供给结构化调整将加速,促使数据中心加速向智算中心(智能算力中心,或称智能数据中心(AIDC)、人工智能计算中心)转型,推动智算设施建设与发展进入新阶段。

目前在中国相关政策和人工智能大模型新技术的双重驱动下智算中心迎来较快发展。政策方面,从中央、国家到部分经济发达城市,近几年密集出台了一系列支持政策,内容涉及人工智能基础设施、标准体系、应用场景等多个方面,初步形成较为完整的政策体系,为加快推动算力基础设施规划建设指明方向。技术方面,受大模型、CPU、GPU异构算力、存储、网络、云平台等技术发展,各地方政府、企业掀起智算中心建设热潮。据国家信息中心与相关部门联合发布的《智能计算中心创新发展指南》显示,目前全国有超过30个城市正在建设或提出建设智算中心。

5 结语

经过近几十年的发展,随着人工智能技术的应用快速增长,数据中心正从计算中心、信息中心、云数据中心加速向算力中心演进,在此转型升级的关键阶段也面临着许多挑战和机遇。论文分析了算力时代数据中心面临的三方面挑战,提出了新技术应用、绿色低碳和大算力需求带来的数据中心发展趋势。在此基础上,希望能通过积极采取创新技术和可持续性举措,建设更为高效、云化、智能化和绿色低碳化的数据中心,为未来的数据处理和大算力需求提供安全可靠的数据中心。

参考文献

- [1] 工业和信息化部等六部门.算力基础设施高质量发展行动计划[R/OL].https://www.gov.cn/zhengce/zhengceku/202310/content_6907900.htm,2023-10-08.
- [2] 2022中国算力大会.中国算力白皮书[R/OL].<https://www.odcc.org.cn/news/p-1559872438149832705.html>,2022-08-17.
- [3] 郭亮.数据中心发展综述[J].信息通信技术与政策,2023,49(5):2-8.
- [4] 国际数据公司,浪潮信息.2022—2023中国人工智能算力发展评估报告[R].2023.
- [5] 丁巧宜,王梓耀,潘振宇,等.面向电量—调频—容量市场下的数据中心算力及电力资源规划[J/OL].电力系统自动化,2012(10):14-16.
- [6] 郭倩,杨琪媛.算力“降碳”绿色数据中心建设加速[N].经济参考报,2023-07-06(007).
- [7] 李竞元.智能化管理训练平台数据算法算力“一个中心”+“三个支撑”加速班组数字化转型[J].班组天地,2023(5):36-37.
- [8] 唐启明.智能算力时代,金融数据中心如何应对绿色节能挑战[J].中国金融电脑,2022(7):35-37.