

绿色算力技术创新研究报告

(2024 年)

中国信息通信研究院云计算与大数据研究所

2024年03月

版权声明

本报告版权属于中国信息通信研究院，并受法律保护。
转载、摘编或利用其它方式使用本报告文字或者观点的，应
注明“来源：中国信息通信研究院”。违反上述声明者，本院
将追究其相关法律责任。

前 言

算力是数字经济时代集信息计算力、网络运载力、数据存储力于一体的关键生产力，已成为推动各领域数字化、智能化转型的重要基石。自我国正式提出“双碳”战略目标以来，绿色低碳、节能环保成为各产业布局的底层逻辑，经济社会对生产、应用和消费绿色算力提出明确诉求。

技术创新是绿色算力发展的根本。当前，算力产业正以绿色为主攻方向，推进算力基础设施全生命周期绿色设计，围绕计算、存储、网络等核心环节加强技术攻关，构建算力资源供给与应用相融合的“数字化”+“绿色化”服务体系，以实际行动积极践行“双碳”战略，实现数字经济高质量发展。伴随产业链上下游各方的共同协作与融合发展，绿色算力在新技术、新产品、新解决方案的落地应用上取得了积极进展。

为梳理绿色算力技术创新发展态势，推动我国算力产业绿色高质量发展，中国信息通信研究院云计算与大数据研究所编制《绿色算力技术创新研究报告（2024年）》。报告从算力供需视角出发，聚焦算力设施、算力设备、算力平台、算力赋能“四位一体”绿色发展，剖析绿色算力的发展背景、概念内涵、发展演进路径和属性特征，深入探讨绿色算力在设施层、设备层和平台层的技术创新情况，总结绿色算力技术赋能的典型应用场景，最后展望我国绿色算力发展，并提出对应建议。

绿色算力相关技术、产业正处于高速发展阶段，本报告在编写过

程中，查阅和参考了行业相关材料，并对众多算力产业链核心企业展开调查研究，感谢业界对本报告的支持！如对本报告有建议或意见，请联系云计算与大数据研究所数据中心团队 dceco@caict.ac.cn。



目 录

一、 绿色算力发展态势	1
(一) 绿色算力发展背景	1
(二) 绿色算力概念内涵	3
(三) 绿色算力发展演进	6
(四) 绿色算力属性特征	9
二、 绿色算力设施层技术创新	11
(一) 建设绿色能源替代，降低供电传输损耗	11
(二) 创新储能部署方式，探索环境友好介质	13
(三) 应用自然冷却技术，突破液冷技术难点	15
三、 绿色算力设备层技术创新	17
(一) 优化服务器硬件节能，发展动态能耗管理	18
(二) 开发存储材料及工艺，促进数据技术高效	20
(三) 应用无损网络介质，推动弹性网络部署	23
四、 绿色算力平台层技术创新	24
(一) 迭代资源管理技术，调度绿色算力资源	25
(二) 简化产品开发模式，优化模型算法效率	27
(三) 完善环境监测体系，实现智能动态调优	29
五、 绿色算力技术赋能经济社会发展	31
(一) 加速科研技术革命	32
(二) 赋能产业节能低碳	33
(三) 助力社会治理格局构建	34
六、 绿色算力发展的展望与建议	36
(一) 加强绿色算力政策保障	36
(二) 夯实绿色能源底座支撑	37
(三) 巩固绿色设备技术创新	38
(四) 聚焦绿色平台能力建设	38
(五) 深入挖掘应用场景价值	39

图 目 录

图 1 绿色算力框架图	3
图 2 绿色算力发展演进阶段	6
图 3 绿色算力属性特征	9

一、绿色算力发展态势

数字经济时代，算力正在成为一种新的生产力，广泛融合到社会生产生活的方方面面，为千行百业的数字化、智能化转型提供基础动力。以数据中心、智算中心为代表的算力基础设施作为算力的重要载体，承载着支撑数字经济发展的重任。近年来，算力总规模不断扩大、用能需求不断增长，其绿色发展进程受到广泛关注，推动算力向绿色算力发展势在必行。

（一）绿色算力发展背景

政策上看，我国算力绿色发展的侧重点，正逐步从聚焦数据中心规划、选址、布局绿色转向到实现产业用能、生产和应用的全链条绿色。早在 2013 年 1 月，工信部等部门印发了《关于数据中心建设布局的指导意见》，提出要充分考虑资源环境条件，推进绿色数据中心建设。随着我国算力产业总体规模快速增长，以数据中心为代表的算力基础设施整体能耗和碳排放问题越发突出，政策开始关注用能环节绿色。2020 年 12 月，发改委出台《关于加快构建全国一体化大数据中心协同创新体系的指导意见》提出要探索电力网和数据网联动建设、协同运行机制。2021 年 7 月，工信部出台《新型数据中心发展三年行动计划（2021-2023 年）》鼓励企业探索建设分布式光伏发电、燃气分布式供能等配套系统，引导新型数据中心向新能源发电侧建设，就地消纳新能源。

“双碳”战略提出后，国家强调要坚持把节约资源贯穿于经济社会发展全过程、各领域，算力全产业链绿色发展逐渐具备政策基础。2021

年9月，国务院发布《关于完整准确全面贯彻新发展理念做好碳达峰碳中和工作的意见》提出要加快构建清洁低碳安全高效能源体系，提升数据中心等信息化基础设施能效水平。2021年12月，国家发改委等四部门发布《贯彻落实碳达峰碳中和目标要求推动数据中心和5G等新型基础设施绿色高质量发展实施方案》提出要发挥市场主体作用，强化标准引领，引入竞争机制、激励机制和成本倒逼机制，促进全产业链绿色低碳发展。2023年10月，工信部等六部门联合出台《算力基础设施高质量发展行动计划》提出要推进算力应用全产业链节能减排。

产业上看，算力市场规模稳步增长，算力市场主体日益丰富，坚持绿色发展是实现高质量发展的必然要求。随着算力应用场景不断拓展，算力需求愈发旺盛，截至2023年6月底，全国在用数据中心机架总规模超过760万标准机架，我国算力总规模达到197EFLOPS(每秒1.97万万亿亿次浮点运算)，存储总规模超过1080EB(1.08万亿吉字节)，算力总规模近五年年均增速近30%。算力产业生态主体日益丰富，以中国电信、中国移动等为代表的基础电信运营商，和以世纪互联、秦淮数据等为代表的第三方数据中心运营商，依靠建设起步早的先发优势已在本土持有大规模数据中心资源；我国传统工业企业，如国家电网、南方电网、中石油、中石化等也开始积极推动算力基础设施建设，为企业数字化转型提供支撑；阿里巴巴、腾讯、百度等互联网厂商借助平台优势，全面应用自研核心技术自建云计算、智算数据中心；随着芯片、服务器在人工智能、云计算和数据中心等领域发

挥着越来越重要的作用，国内算力设备头部厂家，如浪潮、华为、新华三、联想等在硬件算力方面不断进行技术创新与突破。

总的来看，算力市场具有产业链条长、参与主体多、应用范围广泛等特点，各企业在面对扩大经济效益、降低成本损耗等现实需求时，纷纷转向绿色技术创新和绿色管理创新，力图探索节约资源和提高能效的先进手段。以阿里、腾讯、百度、万国、中兴等企业为代表的算力产业龙头，积极践行绿色发展理念，按年度发布可持续发展报告或ESG报告，阐述其绿色发展举措，积极践行双碳目标，承诺不晚于2030年实现“碳中和”目标。

（二）绿色算力概念内涵



来源：中国信息通信研究院

图 1 绿色算力框架图

绿色算力，是算力基础设施实现绿色、低碳、可持续发展的一种

算力形式，是把自然资源、环境资源作为算力发展要考虑的关键，把经济效益和环境保护、社会责任有机融合，可实现算力的低碳清洁、高效利用与科学配置，达成算力设施、算力设备、算力平台和算力赋能的“四位一体”绿色发展。对国家而言，绿色算力是战略之所需也是发展必然，对推动我国数字经济高质量发展具有重要意义；对企业而言，绿色算力是降本增效关键之举也是提质增收实现之基，是助力产业可持续发展的新引擎。如今，经济社会对绿色算力重要性的认知不断加深，绿色算力技术也在不断创新并日益广泛地应用于各产业的数字化发展与实践。绿色算力技术是保障绿色算力发展的技术体系，主要包括材料创新、产品创新、工艺创新和手段创新等，涉及底层硬件、平台软件、核心算法等多方面，并需要考虑碳排放、水资源、废弃物、土地、生态系统多样资源。我国绿色算力技术体系正不断发展与完善，并逐步形成规模效益。

对企业而言，利用绿色算力技术驱动降本增效、提高盈利是其生产经营的根本诉求。国外领先科技企业正致力于以全方位的绿色技术驱动经营管理。谷歌¹自建100%电气化建筑Bay View利用太阳能、风力和地热产生电力，研发张量处理单元(TPU)v4超级计算机进行机器学习算法训练提高能源利用效率，利用机器学习技术防止制冷剂泄露，用海水进行冷却提高用水效率。微软²在数据中心配置锂电子电池组减少依赖煤炭或天然气的情况下保障用电稳定性，开发氢聚合物电解质膜(PEM)燃料电池技术，采用雨水收集和再利用技术重复使

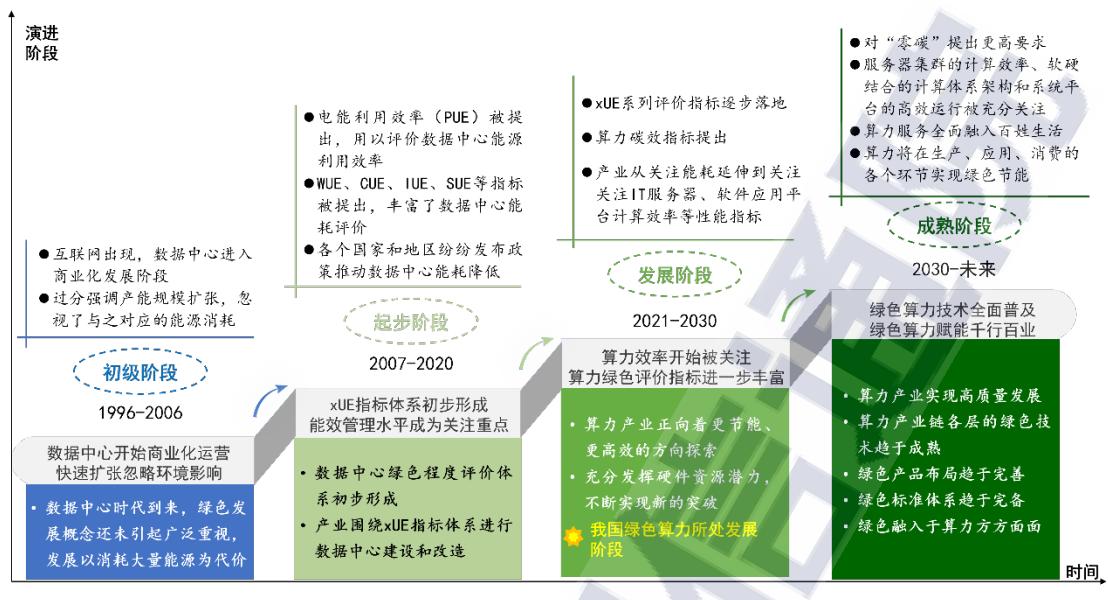
¹ 《谷歌2023环境报告》

² 微软《2022年环境可持续发展报告》

用水资源，采用直接风冷和直接蒸发冷技术冷却服务器，运行双相液体浸没式冷却技术，探索微流体技术更高效冷却芯片，探索绿色软件技术，如开发碳感知软件开发工具包（SDK）给出最低碳能源的工作负载建议，开发架构良好的框架（Well-Architected Framework, WAF）优化 Azure 云计算平台工作负载，使用资源超额订阅和收集技术提高云资源利用率等等。

“双碳”战略为我国绿色算力技术创新带来了空前重要的历史机遇，我国算力产业应充分吸收和借鉴国外成功经验和做法，一方面是要更加注重绿色可持续发展，用能上从煤炭、石油等传统化石能源向太阳能、风能、地热能、氢能等绿色可再生能源转变，减少水资源利用，探索污水、海水、雨水、中水等的开发应用场景，材料使用上向集约节约、质量效益好的新材料转变，同时促进可降解材料使用和废物回收利用；另一方面是要加强技术创新，算力涉及诸多学科和领域，包括计算机科学、数学、物理学、工程学等，坚持产学研用深度融合，激发技术在产业要素配置上的重要作用，将技术创新作为算力绿色发展的突破口，以绿色算力为基石助力经济高质量发展。

（三）绿色算力发展演进



来源：中国信息通信研究院

图 2 绿色算力发展演进阶段

1. 初级阶段（1996-2006年）

数据中心开始商业化运营，快速扩张忽略环境影响。随着互联网的出现，分散的数据资源被有效整合并通过互联网进行分发传输，数据中心进入商业化发展阶段，为了承接处理激增的互联网和金融交易等数据，大中小型数据中心均加速建设，数据中心规模和数量均显著增加。然而，在这种传统增长模式下，过分强调产能规模扩张，而忽视了与之对应的能源消耗，环境污染问题日益突出。这一阶段，数据中心时代到来，绿色发展概念还未引起广泛重视，产业发展不得不以消耗大量能源为代价。

2. 起步阶段（2007-2020年）

xUE 指标体系初步形成，数据中心能效管理水平成为关注重点。

2007 年，电能利用效率（Power Usage Effectiveness，PUE）指标被提出用以评价数据中心能源利用效率。PUE 为数据中心消耗的所有能源与 IT 负载消耗的能源的比值，PUE 值越低，说明数据中心用于 IT 设备以外的能耗越低、约节能。随后，水资源利用效率（Water Usage Effectiveness, WUE）、碳利用效率（Carbon Usage Effectiveness, CUE）、基础设施利用效率（Infrastructure Usage Effectiveness, IUE），面积使用效率（Space Usage Effectiveness, SUE）等指标又陆续被提出，有效丰富了数据中心能耗评价指标。

美国通过数据中心优化倡议（DCOI）、美国联邦数据中心整合计划（FDCCI）、联邦政府信息技术采购改革法案（FITARA）等一系列举措整合和关闭数据中心，数据中心平均 PUE 从 2.0 降低到近一半大型数据中心达到 1.5 甚至 1.4 以下。《欧盟数据中心能源效率行为准则》的最佳实践指南和《欧洲数据中心能源效率现状白皮书》规范了 PUE、SUE、DCIE 等指标。我国发布《关于数据中心建设布局的指导意见》《关于加强绿色数据中心建设的指导意见》等政策文件，引导数据中心追求先进 PUE，依据 CUE 概念制定了数据中心碳排放核算和管理体系建设的详细标准。这一阶段，衡量数据中心能耗综合管理能力及绿色程度的评价体系已经初步形成，产业围绕 xUE 能耗指标体系进行数据中心建设和改造。

3.发展阶段（2021~2030 年）

算力效率开始被关注，算力绿色评价指标进一步丰富。随着气候变化和碳排放的威胁变得前所未有地严峻，全球范围内的各个国家纷

纷提出“双碳”战略做出前瞻性回应。中国、新加坡、墨西哥等国家提出了2030年前实现碳达峰的目标，还有一些发达国家在更早的时间点实现了碳达峰，在向碳中和迈进，全球各国都在努力减少温室气体排放，以应对气候变化带来的挑战。在xUE系列评价指标逐步落地应用的同时，产业开始围绕IT设备进一步研究如何提升效率水平。

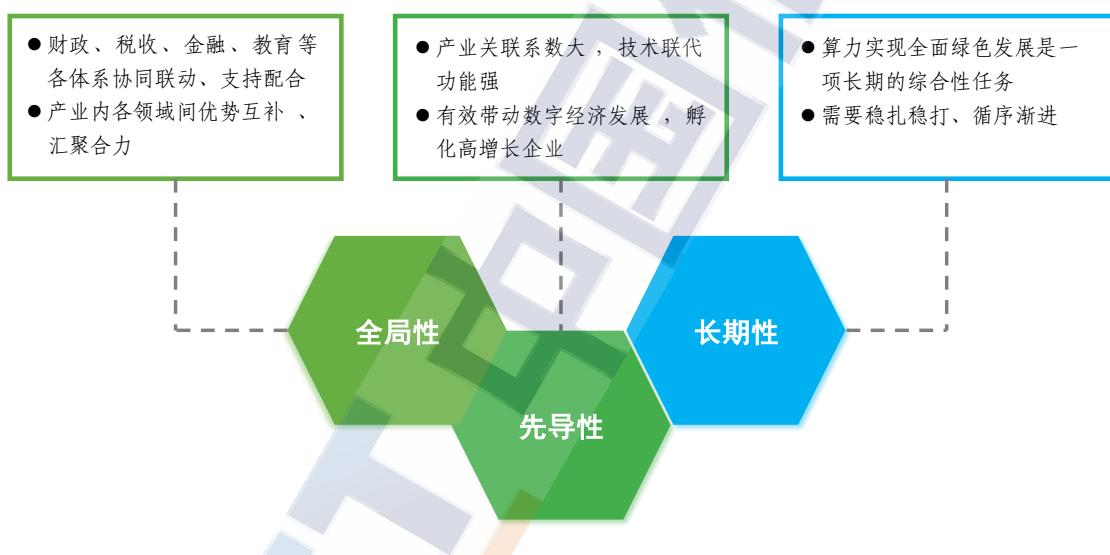
2022年，算力碳效（Carbon Emission Per Server, CEPS）指标由开放数据中心委员会提出，为服务器使用周期内产生的碳排放与所提供的算力性能的比值，可理解为服务器单位算力性能的碳排放量。上述指标的提出反映了产业从单纯关注能耗指标延伸到关注IT服务器、软件应用平台计算效率等性能指标。这一阶段，算力产业正向着更节能、更高效的方向探索，充分发挥软硬件资源潜力，不断实现新的突破。当前，我国及部分先进国家绿色算力发展处于该阶段，多元化性能评价指标被陆续提出，企业将绿色意识与业务发展深度结合作为提升竞争力的关键核心，贯穿生产经营始终。

4. 成熟阶段（2030年~未来）

绿色算力技术全面普及，绿色算力赋能千行百业。面向未来，算力产业对清洁用能实现“零碳”提出了更高要求，如应用百分百可再生能源、长时储能技术、余热回收节能技术等，用更少的能源消耗、更低的碳排放和水足迹生产更多的算力。伴随分布式计算技术的应用与普及，服务器集群的计算效率、软硬结合的计算体系架构和系统平台的高效运行被充分关注，通过面向应用环节的计算绿色化技术进一步提升算力利用效率，将服务器算力尽可能转化成可供平台端调度应

用的软件算力以节能减碳，正成为数字科技企业探索和实践的绿色技术创新前沿。随着算力赋能持续深化，包括工业、金融、政务、医疗等在内的各行各业算力需求被充分满足，算力服务全面融入百姓生活的方方面面，算力绿色标识被建立，绿色算力消费蔚然成风。算力将在生产、应用、消费的各个环节实现绿色节能。这一阶段，算力产业实现高质量发展，算力产业链各层的绿色技术趋于成熟，绿色产品布局趋于完善，绿色标准体系趋于完备，绿色融入算力的方方面面。

（四）绿色算力属性特征



来源：中国信息通信研究院

图 3 绿色算力属性特征

1. 全局性

算力与绿色发展的深度融合，既需要财政、税收、金融、教育等各体系协同联动、支持配合，又需要产业内各领域间优势互补、汇聚合力。算力覆盖产业广泛多元，进一步深化发展绿色技术，需要更多的资金支持，要靠国家财税政策引导和绿色金融相关产品推动，多管

齐下促使投资向绿色算力产业倾斜。绿色产业发展动力强劲，方兴未艾，绿色算力作为其中的重要一环，对创新型、技能型和应用型人才需求迫切，需要科研机构、高校、行业协会等各方协助，共同建设绿色算力人才队伍。引导数据中心骨干企业将项目建设在风、光资源密集地区，实现新能源发电、储能、数据中心融合发展。发挥软硬件企业各自优势，引导芯片与服务器能力相匹配，芯片产业发展与模型算法规模相适应，减少磨合和不兼容等问题，实现算力高效输出。

2. 先导性

绿色算力产业关联系数大，技术联带功能强，可有效带动数字经济发展，孵化高增长企业。绿色算力产业影响力强，是有力支撑各行业早日实现碳达峰与碳中和目标，实现绿色发展的重要引擎。具体来看，绿色算力与能源电力、半导体、集成电路、消费软件等产业关联密切，可有效带动其他产业向着绿色发展。从本质上讲，绿色算力具备算力产业和绿色产业二者的双重特征。在赋能下游应用上，绿色算力一方面可支撑如人工智能、大数据、5G、元宇宙等新兴技术发展，以更少的能耗代价进行更大规模的训练运算；另一方面可以赋能高耗能产业进行技术改造和降本增效，如带动钢铁、建筑、石油化工等重点领域构建工业互联网，实现工业制造再升级、带动智慧城市建设，提升城市“脑力”，让城市更“智慧”，生活更便利。

3. 长期性

算力实现全面绿色发展是一项长期的综合性任务，需要稳扎稳打、循序渐进。绿色算力涉及范围庞杂，因此需要从不同层面、不同视角

统筹考虑，促进其健康有序发展，推动经济社会高质量可持续发展。宏观上，为保障绿色算力发展行稳致远，制定政策规划时，要处理好绿色算力和产业需求目标之间的关系，达到二者有效平衡；推行绿色算力政策时，要注意节奏，预留适当的缓冲期，用动态发展的眼光不断丰富和完善工作思路，做到有序且量力而行。另一方面，要充分意识到，绿色算力的创新发展是一项长周期活动，涵盖产品、工艺、原料、组织管理和市场等方方面面的创新，算力从浅绿、经历中绿、深绿和完全的绿色需要时间的沉淀从而厚积薄发。未来，通过绿色算力技术的进步，绿色算力经济将成为助推和引领经济发展的新增长点。

二、绿色算力设施层技术创新

算力基础设施是集信息计算力、网络运载力、数据存储力于一体的新型信息基础设施，本文所说的算力设施，是前者的组成部分，主要是指数据中心基础设施，提供“风火水电”等保证算力正常生产，为 IT 及网络设备提供基础的动力来源、能源配送和可靠性保障。PUE 是衡量数据中心基础设施能效的主要指标，截至 2022 年底，全国在用超大型数据中心平均 PUE 为 1.36，大型数据中心平均 PUE 为 1.42，最优水平达到 1.08。全国规划在建数据中心平均设计 PUE 为 1.32 左右，超大型、大型数据中心的平均设计 PUE 分别为 1.28、1.29。算力设施层面的绿色算力技术将从供配电系统、储能系统、制冷散热系统三个部分展开。

（一）建设绿色能源替代，降低供电传输损耗

数据中心作为传统的耗能大户，电算转化的载体，运行主要依靠

电力驱动，电量消耗和随之产生的碳排放不容忽视。我国已提出了到 2025 年底，国家枢纽节点新建数据中心绿电占比超过 80%的建设目标。稳步提升太阳能、风能、水能等可再生能源电力在数据中心的应用，是算力绿色发展的关键之一。

自建可再生能源扩大消纳规模，探索“源网荷储”模式实现零碳。

数据中心自建绿电主要有两种技术手段，一是自建分布式光伏设备，提高绿色能源消纳比例。在车棚、屋顶或园区空地安装光伏组件，利用太阳能发电为数据中心供电。国内如中国电信、万国、腾讯等均有数据中心自建分布式光伏。除了利用太阳能之外，也有数据中心探索分布式风力发电。二是建设源网荷储（微电网）能源综合利用系统，实现零碳能源完全覆盖。以数据中心为负荷的“源网荷储”项目，通过建设大规模的风电或光伏设备，实现绿电自发自用和用电需求百分百覆盖。当前，合盈数据、中国广电等公司开始规划数据中心源网荷储一体化项目建设。

创新供电模式减少传输损耗，缩短电力传输路径节能增效。

供配电系统占数据中心总能耗的约 10%。近年来，高压直流技术逐渐被数据中心领域所接受，新建的互联网数据中心广泛采用了“一路市电+一路高压直流”相结合的模式，供电效率可普遍提升到 94%-95%。高压直流主要是指 240V/336V 直流供电技术，并可分为在线和离线两种应用模式。在线，是指交流电能始终经高压直流整流后为 IT 设备供电，通常会存在一定损耗；离线，是指正常情况下市电直供 IT 设备，高压直流电源仅为蓄电池提供浮充，市电中断后，转由蓄电池供电，

在这种架构下，正常情况为 IT 设备供电的电能不经过高压直流转换，此部分损耗几乎可以忽略，因此节能效果显著。若采用高压直流离线架构，供电效率可以提升到 97%以上。百度云计算（阳泉）数据中心采用“市电直供+高压直流离线”架构及自研分布式锂电系统，供电效率提升至 99.5%。

如前介绍的数据中心分布式发电、“源网荷储”新方案，除了可有效提升绿电利用外，还可以凭借其在负荷侧就近建立电源供电的特点，有效缩短“源”“荷”距离，减少传输损耗。位于国家枢纽节点集群起步区的吴江算力调度中心项目采用高效供电系统，变压器优先选择一级能效干式变压器；优化供电半径，合理选择变压器的容量、提高负荷率，使设备处于经济运行状态；采用 240V 直流系统、10kV 交流输入的 240V 直流系统等损耗率低的架构和设备；结合数据中心智慧管理系统设置能耗计量系统；选用节能型光源和灯具，结合智慧园区进行智能控制；在可利用屋面（例如指挥调度中心）设置光伏发电，用于建筑用电。

（二）创新储能部署方式，探索环境友好介质

对于数据中心而言，储能不可或缺。储能技术是推动算力在设施层实现绿色转型的重要因素，储能与供配电系统协调配合，输出稳定电力，共同保障数据中心电力系统平稳运行。

不同的储能部署方式可满足数据中心多样储能需求，促进整体用电稳定。按照部署的位置，储能可分为数据中心级储能和数据中心内储能，两种方式均可减少煤炭或天然气电力依赖，降低温室气体排放。

数据中心级储能往往将储能独立部署在数据中心楼旁，储能一方面可作为数据中心的备用电源，另一方面参与市场电力市场调峰、辅助调频等辅助服务，在提高数据中心供电可靠性的同时，也从电力市场服务中获取收益补偿，是应对极端事件、保障能源安全、支撑数据中心电力系统稳定运行的重要技术。数据中心内储能主要采用分布式结构，包括机柜级分布式供电系统(DPS)和服务器级备用锂电池组(BBU)等。DPS 分布式电源采用高效率供电模块以及全锂电池备电系统，配合能源管理系统可以实现机柜级的能源管理并进行智能调度。

在绿色低碳产业趋势下，储能介质或储能材料需具备安全可靠、性价比高、环境友好、循环寿命长等优点。储能有多种形式，对于数据中心领域，传统的铅酸电池仍是提供短暂供电的主要后备电源。然而，传统的铅酸电池存在循环寿命短、能量密度低，生命周期结束造成环境污染等问题。在双碳目标下，我国正构建以新能源为主体的新型电力系统，这给新型储能技术发展提供了支撑。数据中心主要探索的方向为锂电池储能。磷酸铁锂电池是指用磷酸铁锂作为正极材料的锂离子电池，该电池不含任何重金属与稀有金属，无毒，无污染，为绿色型环保电池。当前，谷歌、微软、AWS、特斯拉、思科、英特尔等国际科技公司以及阿里、腾讯、百度等国内互联网公司均有数据中心已经在使用锂电池供配电系统。此外，业界针对氢能燃料电池也展开多项实验与测试，证实了其作为数据中心电源的可行性。微软在 2022 年 7 月试点推出零排放的 3 兆瓦氢聚合物电解质膜(PEM)燃料电池备用电源发电机，用于数据中心多兆瓦发电需求。但同时，安

全问题仍是制约储能的关键，数据中心在部署储能时也需注重系统整体的安全性，保障不间断可持续运行。

相信未来，随着供配电系统、储能系统等多种模式融合，可以实现绿电生产、平抑峰谷、降低排放等多重目标，形成完整的能源供应体系，实现可持续发展。

（三）应用自然冷却技术，突破液冷技术难点

数据中心IT设备在运行时会持续产生热量，制冷系统的主要作用是散热制冷，维持IT设备正常运行的环境条件，属于数据中心必不可少的基础设施。对于数据中心而言，制冷系统是除IT系统能耗外最大的耗能单元，不少老、旧、小数据中心的制冷设备耗能占比高达30%-50%，制冷是影响数据中心高效节能的重要因素。

使用自然冷却技术，减少和避免使用机械制冷成为数据中心减少能耗的关键。自然冷却主要分为风侧自然冷却和水侧自然冷却两大类。在选择制冷整体设计方案时应充分综合考虑数据中心所在地区的气候特点、水资源供应情况、数据中心规模以及IT功率密度，充分利用自然冷源制冷。

直接新风自然冷却技术是散热效率最高的空气冷却方式，是将经过过滤等处理的空气直接供应到数据中心，已被微软、谷歌和雅虎等公司所采用。直接水侧自然冷却技术是采用河水、湖水、海水等自然冷源水进行冷却的技术，在我国应用取得了较大进展。如海兰信海底数据中心示范项目以海洋作为自然冷源，将服务器安放在海底的压力

容器中，通过海水的流动进行自然冷却³。采用湖水自然冷却的阿里巴巴千岛湖数据中心也已投入运行。

研发液冷技术，实现服务器高效散热大幅减少数据中心制冷用电需求。液冷技术是数据中心新兴制冷技术，随着生成式人工智能、大模型计算的需求带动 GPU、FPGA 等发热量大的元件应用发展，高功率机柜逐渐增多，液冷技术开始被提出用于解决高功率密度机柜散热需求。据 IDC 数据显示，2023 上半年中国液冷服务器市场规模达到 6.6 亿美元，同比增长 283.3%，预计 2023 年全年将达到 15.1 亿美元。当前，常见的液冷技术主要包括冷板式液冷技术和浸没式液冷技术。冷板式液冷是通过与装有液体的冷板直接接触来散热，或者由导热部件将热量传导到冷板上，然后通过冷板内部液体循环带走热量。随着冷板技术的逐渐成熟，提高冷板效率、降低冷板成本的需求也越来越强烈。传统的铜冷板存在一些限制，例如：整体冷板的重量限制，铲齿 Fin 厚度的制造限制，流阻的限制等等。联想、宁畅等创新冷板材料，如采用新型铝材料冷板、铜铝结合冷板作为核心换热部件，提升水冷板热性能提升；新华三采用全焊接+高耐压 EPDM 管设计，提升冷板耐压能力和可靠性。此外，中国移动设计院研发解耦型冷板式液冷机柜产品，推动液冷基础设施侧与 IT 设备侧解耦，推进机柜与服务器接口标准化，液冷机柜可用于通算、智算、超算、异构算力等多业务场景，采用解耦盲插、风液融合的设计理念，可与多种类型服务器适配。

³ 华彩杯算力应用创新大赛优秀案例

浸没式液冷是将发热的电子元件浸没在冷媒（冷却液）中，依靠液体流动循环带走热量。相比而言，浸没式液冷的平均 PUE 更低，散热效率更高。浪潮、宁畅、联想、新华三等主流 IT 设备厂商纷纷发力液冷产品研发，覆盖风液冷混合、冷板液冷、浸没式液冷等主流液冷技术，布局全栈数据中心液冷产品，提供从服务器、机柜、机房到数据中心的全系统液冷解决方案，实现全应用场景下差异化节能降耗。

余热回收有着广泛的应用前景，可有效帮助减少碳排。随着更大规模、更多数量的数据中心建设，服务器在运行的过程中，产生大量的余热，这些余热有着易提取、热源充足等特点，这些余热如果能被加以利用，可以加速“碳中和”进程。余热回收再利用技术是利用热泵将余热回收，用于建筑供热、生活热水等，可减少碳排放，提高能源利用效率。联想研发的温水水冷技术可实现温水液冷和数据中心热能回收，提升算力降低能耗，实现热能的回收再利用。新华三把液冷服务器中高温回水和热回收技术相结合，实现数据中心热能高效回收和利用，可实现园区的生活供热。

三、绿色算力设备层技术创新

算力设备是算力产生的源头，综合了计算、存储等 IT 和网络设备，以服务器、芯片为核心部件，实现对数据的处理与输出。现阶段，针对算力设备的绿色技术发展，主要围绕服务器计算高效、先进存储、网络传输无损等方面进行研究。

（一）优化服务器硬件节能，发展动态能耗管理

在过去很长一段时间，为了满足不断增长的用户数据处理需求，企业主要是通过扩大机架和服务器规模来提供更多算力，但是这也会导致运营成本的增加和场地空间的浪费。发达地区日益紧张的土地资源使得以扩大服务器规模来提升算力水平的数据中心建设模式难以开展。

依据指令集架构、产品形态等维度对服务器进行工艺和产品创新，实现高效节能。服务器承载处理数据和实现结果输出的功能，是算力供给的核心装备，也是数据中心中最主要的业务耗能设备，约占数据中心 IT 设备能耗的 90%。伴随社会对信息计算力需求呈指数级增长，服务器生命周期的高效节能对算力绿色发展至关重要。

从指令集架构上看，基于精简指令集（RISC）架构的处理器以其低功耗、高效能、成本低、高可靠的优势正在成为让信息计算更绿色的主力。从逻辑上来看，CPU 的指令集可以分为两种主要类型，即复杂指令集（CISC）和精简指令集（RISC）。复杂指令集的代表是 X86，由英特尔和 AMD 主导；精简指令集包括 ARM 架构、RISC-V 架构、MIPS 架构等。随着产业智能算力需求的不断增长，以及生成式 AI 对于异构算力的庞大需求，ARM 架构服务器在云游戏、数字人等新兴应用市场展现出独特优势。据投资银行 Bernstein 数据显示，我国数据中心正在加快部署 ARM 服务器，截至 2023 年一季度，我国已拥有占全球约 40% 的 ARM 服务器。据不完全统计，在中国服务器市场，当前 Arm 的份额可能已经超过了 10%。与此同时，RISC-V 近年来发

发展趋势日益迅猛，有望与 ARM 分庭抗礼。RISC-V 指令集是基于精简指令集计算（RISC）原理建立的开放指令集架构（ISA）。RISC-V 指令集完全开源，设计简单，易于移植 Unix 系统，模块化设计，完整工具链，同时有大量的开源实现和流片案例，国内外企业均有关注 RISC-V 或以 RISC-V 指令集进行开发，积极布局 RISC-V 赛道。

从产品形态上看，建设高密度服务器成为提升计算效率的重要举措。高密度服务器内，电源和风扇以共享方式进行设计，位于同一机箱内的多台服务器节点可以共享电源和风扇，一方面降低了机体的重量和空间占用，提升单位面积算力，另一方面能够提升电源和散热系统的使用效率，降低运营成本，能够进一步增加数据中心功率密度和数据中心“每平方米”的计算能力。刀片服务器是高密度服务器的一种，主要应用在商业智能分析及数据挖掘等大规模计算场景，具有耗电量低、可靠性高等优点。整机柜服务器是提升计算密度的另一种形式，依据模块化设计思路优化服务器内部架构，采用工厂预制的设计大幅缩短工期。

通过管理工具加强对服务器电源、风扇等硬件的控制力度来进行能源节约。服务器整机节能技术有功耗封顶节能技术、动态调频调压、启动低功耗、智能休眠技术、智能能耗管理、节能风扇调速等技术。

功耗封顶技术通过服务器内置的功耗测量模块来实时获取服务器的运行功耗，然后通过限制处理器或者其他部件的性能来将服务器的总功耗控制在设定的上限功耗以下，防止功耗超标的同时提高服务器利用率。

动态调频调压技术是在服务器低负载时降低 CPU 频率和电压，降低 CPU 的漏电损耗和开关损耗，大幅度降低处理器功耗。某些服务器会将预先调优好的功率模型内置于 BIOS 中，在不影响性能的情况下为工作负载实现插槽节能。启动低功耗技术包括硬盘错峰上电、刀片错峰上电、CPU 关核技术，确保服务器上电过程的功耗低于系统下的运行功耗，消除上电瞬间电流过大带来的供电风险，同时为用户节省电费开支。

智能能耗管理包括在服务器内置遥测工具，这些工具提供关键的数据和 AI 功能，帮助企业智能地监测和管理 CPU 资源、建立模型来帮助预测数据中心或网络的峰值负载，调整 CPU 频率以便在需求下降时降低用电量，还能够在有可再生能源供应时选择性地增加工作负载，从而有机会降低数据中心的碳排放量。

（二）开发存储材料及工艺，促进数据技术高效

据 IDC 数据显示，全球数据年产生量将在 2025 年达到 175ZB。为保存不断增长的海量数据，对存储容量的需求亦成几何式增长，存储耗电量激增，高性能、低能耗存储需求攀升。

数据分级、冷热数据分治、优化存储设计成为降低单位容量数据成本和功耗、提升数据存储效率和密度的有效方式。随着数据量的爆炸式增长以及数据类型的日益细化，让数据分级和冷温热数据分治可更好的降低存储能耗。冷热数据分治的原理就是将访问次数多的热数据存储到读写快的存储介质中，访问次数少的冷数据存储在不经常访问，生命周期长的存储介质中，实现数据分治降低存储能耗。固态硬

盘存储热数据是数据分治的降低能耗的关键点之一，固态硬盘（SSD）读写速度快、功耗低且发热量少，更适合热存储。固态硬盘存储介质有 FLASH 闪存芯片和 DRAM 存储芯片，闪存介质存储有高密度、高可靠、低延迟、低能耗等优势。目前，SSD 单盘容量已经达到 30TB 以上，在同等应用场景下，SSD 数据访问比磁盘快 100 倍左右，吞吐量大 100 倍，数据存储能力大幅提升，进一步降低了存储能耗。冷数据的存储介质需要满足生命周期长，碳排放低等特点，这样才能够实现低碳节能的目的，常见的冷存储介质包括机械硬盘（HDD）、磁盘、光盘等。机械硬盘读写速度相对较慢、使用寿命长，适合冷存储场景。此外，磁带、蓝光光盘等载体也常用于温、冷数据的存储，磁带生命周期内二氧化碳排放显著低于硬盘并且磁带可以大大减少电子垃圾的数量。光存储用电量仅为磁盘阵列存储的十分之一，能够有效降低数据中心 PUE 值，是未来绿色数据中心建设行之有效的存储介质。

合理的存储设计是节能降耗的关键。存算分离架构、高密度设计以及风液冷融合设计都可以大幅降低能耗。数据中心新型存算分离架构具备资源解耦以及细粒度的处理分工两个关键特征。资源解耦可以使 CPU、GPU、内存、存储等组建为相互独立的资源池，细粒度分工可以使数据处理等 CPU 不擅长的任务被 DPU 等专用数据处理器替代，可以实现能效比最优组合。存算分离架构将会大大提升资源利用率，减少数据迁移。高密化设计上，专门设计的高密存储型节点，密度达到传统存储服务器的 2-2.6 倍，结合存算分离架构，同等容量下带来能耗节约 10%-30%。风液冷融合设计上，CPU、GPU 等大功率

器件采用液冷，其他器件如存储、内存、电源采用风冷冷却，降低内存、HDD 等存储关键部件的工作温度，降低风扇转速，维护设备运转性能，达到节能降耗的目的。

数据融合技术、软件编码技术、数据重删技术和探索系统协同节能是业界研究的绿色技术热点。数据融合技术，允许多云多业务共享存储系统的文件资源，减少数据搬迁和重复存储，提升超过 30%的数据处理效率，降低约 20%能耗。软件编码技术，在大数据分析场景，采用存算分离架构后，利用数据纠删码（Erasure Code, EC）技术替代传统三副本数据冗余，可以把磁盘利用率从 30%上下提升到约 90%左右，降低能耗约 40%。数据重删技术，利用定长重删、变长重删、相似重删算法把相同数据删除，通过数据压缩、压紧算法把定长的数据块优化数据存储布局，节约存储空间。借助闪存介质带来的 100 倍性能提升，目前业界已经能够在数据库、桌面云、虚拟机等业务场景实现 2-3.6 倍的数据缩减率（重删压缩前数据总量/重删压缩后数据总量），耗能节约 50%以上。探索存储系统协同节能方式，可通过感知存储中不同控制器中 CPU 的业务压力，动态实现 CPU 降频，甚至可以根据大量业务运行数据对存储不同时期的全局负载进行建模，实现预测式的精准降频、动态节能。

算力时代，运营商存储数据量庞大，需要对海量数据应用先进存储技术进行节能优化。中国移动、中国联通等利用存算分离技术创新，大幅地降低了计算和存储服务器消耗。中国电信天翼云自主研发的存储资源盘活系统 HBlock，可以把通用服务器及其管理的存储资源转

换成高性能的虚拟存储阵列，可被视为本地盘阵的替代品，用于本地数据存储，帮助用户提高资源利用率，优化资源成本。

（三）应用无损网络介质，推动弹性网络部署

算力的传输和供给需以可靠的网络连接为基础。一方面，庞大用户规模的加速增长离不开网络传输设备的提速换代；另一方面，配合云计算、大数据、物联网、人工智能、虚拟现实/增强现实等数字技术和产业的融合应用发展，打造无损高效的网络运载服务能力，是在数字经济绿色发展道路上需要迈出的关键一步。

部署超低损耗光纤、实现全光接入网络覆盖是降低损耗、提升网络传输效率的关键。超低损耗光纤常用于长距离和高速光信号传输，具有更低的损耗、更强的抗拉强度和更小的弯曲半径，可以支持更远的传输距离和更高的传输速率。现网中光缆 G.652 光纤占比较高，超低损耗大有效面积的 G.654 光纤正在加快落地应用当中，这种光纤在 $1.55\mu\text{m}$ 处的损耗极小，仅有 0.18dB/km 。G.654E 光纤的应用，降低了衰减和非线性等因素的影响，可以减少光放站和光再生中继站的设置，有效降低网络运营成本。全光网是指光信息流在传输及交换时始终以光的形式存在，而不需要经过光/电、电/光转换。相对于无线传输和电域交换技术，全光网技术在单位比特公里传输能耗和单位比特交换能耗方面，节能效果十分显著。根据中国电信实际网络统计，与链状密集波分复用（DWDM）系统相比，ROADM 全光交换网可降低大约 50% 的能耗和机房空间、节约大约 30% 的成本。

网络动态适配技术和网络弹性智能技术可通过优化网络资源利

用减少能源消耗。网络动态适配技术在不降低网络设备处理能力的情况下，通过动态调整路由协议可以达到节能的目的。网络设备可以根据网络负载变化动态调整网络接口速率，也可以通过计算网络设备负载动态调整处理器的电压与频率，提高处理器能效。网络设备协议节能技术根据网络协议的特点，动态感知网络流量的实时变化来动态调整路由，使得在网络连通性和网络性能等约束条件下使用的网络资源最少。网络弹性智能技术可以根据需求动态调整连接和带宽，提升网络整体效率，实现整体节能的效果，例如 SRv6、FlexE 技术。SRv6 是一种基于 IPv6 平面的源路由技术，SRv6 继承了 SR（段路由）给网络带来的所有好处，如简化协议、简化网络（设计、部署、维护）、高可靠性、源路由技术、SDN 理念等，可以实现灵活的网络互连。FlexE（Flexible Ethernet，灵活以太）技术是基于高速 Ethernet 接口，为满足高速传送、带宽配置灵活等需求而发展的技术，可以提供灵活、高效、可靠的通信通道，支持多种业务处理和网络切片，提升网络性能和资源利用率，有助于网络的高效管理和运营。

四、绿色算力平台层技术创新

算力平台是以数据中心算力设备的硬件算力为基础，集成了系统、应用程序、算法等多样化工具和软件，给服务对象提供算力的平台。随着算力产业链条脉络逐渐明晰，利用平台能力汇聚算力资源，以场景化、智能化实现算力供需精准匹配，逐渐成为绿色化发展的关键。要发挥全局资源调度优化、产品开发简化和绿色监测评估智能化等功能，实现算力利用效率提高与平台层面的绿色低碳。

（一）迭代资源管理技术，调度绿色算力资源

算力资源调度是根据不同的任务和计算需求，在算力平台合理分配和利用算力资源的过程，是提高算力利用效率，减少资源浪费，保证任务高效执行的重要环节。当前，我国算力产业已经开始探索跨区域及跨集群的算力资源调度平台建设，根据规模大小及覆盖范围，可分为全国性算力调度平台和企业级调度平台。

虚拟化、池化技术是提高服务器利用率，减少能源消耗和碳排放的关键技术要素。算力由芯片、服务器等IT设备生产出来后，在用户实际使用算力之前，还需通过算力平台发挥衔接算力供需双方的桥梁作用，如何高效地将硬件算力转化为软件算力，汇集和统一硬件算力是首要和关键，相关技术突破经历了算力虚拟化和算力池化两个阶段。

算力虚拟化，是在物理服务器硬件或主机操作系统上插入精简的软件层，该软件层包含一个以动态和透明方式分配硬件资源的虚拟机监视器。多个操作系统可以同时在单台物理服务器上运行，彼此之间共享硬件资源。该技术能够将服务器CPU的平均占用率提高10%~30%，实现服务器硬件资源整合，有效节能20%~50%。

算力池化技术从虚拟化发展而来。当前，算力逐渐由串行走向并行、单核走向双核、同构走向异构，异构计算已经成为计算架构的主流，资源管纳的范围从单个节点扩展到由算力网络互联起来的全部算力节点，将多个计算资源汇聚到一个池中，通过调度算法，按需灵活调用任意程序，从而提高计算资源的利用率，降低能耗。算力池化的

一个重要特性是弹性伸缩。根据用户需求的变化，算力平台能够自动调整资源池的规模，增加或减少计算资源的数量，以满足用户的实时需求。这种能力可以提高资源利用率，并降低用户的成本。

此外，容器化、容器编排、微服务化等云原生技术使得应用程序更加轻量级、可移植和高弹性，从而实现了应用程序的快速部署、扩展和缩减，进一步减少了能源消耗和碳排放。谷歌开发的开源 Kubernetes 平台，通过有效管理容器和根据需求扩展资源来优化资源利用，可以防止过度配置，并允许更好地利用基础设施资源，从而实现节省成本的作用。

算力资源调度技术不仅适用于企业自用算力平台帮助提高资源管理效率，也适用于全国性或地区性算力平台实现更大范围的绿色。企业自用算力平台可以利用先进的资源调度技术更好地管理和使用计算资源，提高计算任务的效率和可靠性。算力产业链中上游企业依据各自的算力资源优势纷纷布局建设企业算力平台生态。广东联通天穹算力调度平台整合基础、智能、超算算力，汇聚 AI 算法仓，实现一点供应，构建算网边安一体化调度体系，实现算力资源统一管理，推动精准匹配和按需获取⁴。阿里云推出智能计算解决方案“飞天智算平台”，适配多种芯片架构，支持 X86、ARM、GPU、NPU 等多种处理器混合部署和统一调度，可以适配多种国内自研芯片，并进行应用优化。

对于覆盖用户范围更广的全国性或地区性算力平台，资源调度能

⁴ 华彩杯算力应用创新大赛优秀案例

力可以帮助用户以较低成本灵活地获取和管理低碳绿色计算资源，降低资源的浪费。这对于需要大量计算资源的应用场景，如人工智能、大数据分析等，具有重要意义。2023年，中国信通院搭建的中国算力平台·算力资源匹配平台正式上线。该平台汇聚算力资源、存力资源、运力资源等多元综合算力，集算力供给与算力需求资源于一体，实现算力供需双方在平台内进行算力交易，灵活调度算力资源，为促进算力资源的精准配置和按需获取奠定了坚实的基础。

（二）简化产品开发模式，优化模型算法效率

算力在经过虚拟化和池化后，根据负载调用任意大小的算力成为可能。计算架构提供了一种可扩展和适应不断变化的硬件和软件环境的方案。计算架构可以屏蔽异构硬件差异，减少用户跨架构编程的重编译和迁移代价，提供统一异构硬件开发工具，并提供便捷化业务开发和部署方式，实现算力的优化配置和高效利用。

无服务器与人工智能工具可帮助简化产品开发，实现算力资源高效利用。无服务器（Serverless）是一种新型的计算架构，它是一种基于事件驱动的计算模型，可以自动扩展和缩减计算资源，而无需用户管理服务器。无服务架构广泛服务于智能科学模拟、数字化政府治理、平台型算力共享等场景，提供安全可信的服务保障。它抽象了底层基础架构的细节，如服务器配置、网络和操作系统管理，允许开发人员只需专注于编写和部署他们的功能即可。例如，AWS Lambda 和 Microsoft Azure 可以允许企业在不提供或管理服务器的情况下运行代码及使用计算资源。这种方法支持根据需求进行自动伸缩，通过仅在

需要时动态分配资源来确保高效的 CPU 利用率。通过提高 CPU 利用率来达到高效利用算力资源，降低能耗的目的。

在人工智能应用开发中，人工智能应用框架可以作为平台架构的一部分，提供机器学习模型的开发、训练和部署服务。PyTorch 和 TensorFlow 是目前最流行的深度学习框架，通过计算效率优化、资源优化和压缩、硬件优化和兼容性以及社区合作和共享等方面的努力，为绿色低碳化发展做出了积极贡献。PyTorch 和 TensorFlow 等框架提供了自动微分功能，可以优化和训练模型，减少手动计算导数的复杂性，提高计算效率。支持异步计算，同时进行计算和数据传输，提高计算效率，降低能源消耗。支持将大型模型拆分为多个部分，并在多个设备上并行计算，以加速模型训练和推理，降低整体能耗。提供了模型压缩和量化技术，通过减少模型参数的数量、使用低精度数据表示和优化存储格式等方法，降低模型的能耗和计算复杂度。通过计算图优化和分析技术，可以对计算图中的操作进行重排、合并和优化，以提高计算效率和降低能耗。此外，通过内存管理和数据重用，框架提供内存管理机制和数据重用策略，避免不必要的内存分配和数据复制，减少能源消耗。除 PyTorch、TensorFlow 等人工智能框架外，还有些平台软件也针对大型 AI 模型的开发、训练进行了优化，例如开源软件 DLRover。当前，业界正在探索无服务架构和人工智能的协同效应，以期进一步提高开发效率与算力利用效率，实现技术创新。

在应用产品设计和优化上，还需要考虑发展绿色算法，优化策略和参数设置，降低代码运行能耗。编写良好的代码可以消除冗余计算，

减少不必要的循环。使用高效的算法，利用提供更快访问和处理的数据结构，可以提高应用程序的整体性能，减少不必要的计算。高效的内存管理和缓存技术对于优化也至关重要。内存池、缓存和最小化内存碎片等技术可以帮助减少内存开销并提高应用程序性能。以 OceanBase 为例，基于 LSM Tree（Log-structured Merge Tree，日志结构化合并树）的高级压缩技术可以大幅降低存储成本，支付宝某核心业务从 Oracle 迁移到 OceanBase 后，数据压缩了三分之二，由 100TB 压缩到 33TB。此外，通过自研的高性能分布式事务处理引擎、SQL 引擎以及负载均衡等机制，提升数据库的运行效率，在同等算力水平上可以承载更多的业务负载，提升资源使用效率。

（三）完善环境监测体系，实现智能动态调优

算力平台的能效监测和评估功能用于监测和管理能源消耗和碳排放，结合数据分析和人工智能技术实现工作负载的优化，旨在确保数据中心及算力设备的能源效率，同时降低碳排放，实现绿色、低碳、可持续的运营效果。

资源和碳排放监测是实现低碳运营的基础。资源监测是指对数据中心和算力设备的资源消耗进行实时监测和记录，包括电力系统、制冷系统、环境温湿度、IT 资源、IT 负载、PUE 等方面的运行概况，通过实时监控系统的关键指标了解数据中心的资源消耗情况，发现任何异常或潜在的能源浪费，并及时采取措施应对。随着业务发展中的低能耗要求提升，针对硬件算力资源，常见的监测指标从 CPU 利用率、内存使用情况、网络流量、磁盘 I/O 等延伸至面向业务的综合算

力效用指标。实时监控利用如剖析器和跟踪工具，可以帮助确定 CPU 高使用率的区域，并优化相应的软件组件。其中，CPU 利用率是国际通用的重要能效监测指标，国际标准化组织和国际电工委员会联合制定的 ISO/IEC 30134 系列标准对数据中心 CPU 利用率部分指标及测定方法进行了规定，以微软、谷歌、IBM、阿里巴巴、腾讯等为代表的一众互联网企业在日常经营活动中也会通过监控 CPU 利用率各项指标来优化服务器任务完成情况及能效表现。监控 CPU 利用率可以了解服务器的工作负载以及算力利用效率情况，采用针对性技术对算力资源进行合理分配。碳排放监测是对数据中心和算力设备的碳排放进行监测和记录。通过了解碳排放情况，可以采取措施减少碳排放，实现低碳运营。供电、制冷、服务器、网络、存储等产业链各环节亟需梳理核算碳足迹，推进算力全产业链绿色节能。

能效分析和人工智能调优是实现低碳运营的关键。能效分析是对数据中心的能源效率和性能进行综合分析。通过分析历史数据和当前数据，进行资源分配决策、容量规划和确定性能瓶颈。针对硬件算力资源，通过收集系统日志以及关键的性能指标，如响应时间、吞吐量、并发连接数等，可以深入了解系统的运行情况和潜在问题。日志分析可以帮助定位故障、发现异常行为，并为性能优化提供有价值的线索。对数据进行分析可以得出资源的利用率等情况，包括 CPU、内存、存储和网络等方面，确定系统是否存在资源瓶颈，并优化资源分配以提高整体性能。

人工智能调优是指在数据中心环境中，结合大数据处理、机器学

习、AI 算法等人工智能手段分析系统及各个节点的负载情况，然后根据这些信息自动调整任务的分配，使负载达到平衡。这使得系统能够及时响应工作负载变化和需求波动，帮助组织提前做出资源分配和调度的决策，从而减少资源浪费和缩短响应时间，更有效地管理和优化其 IT 资源，提高性能并确保系统的稳定性，提供更好的用户体验。

英国的 Deep Mind 公司用数据中心的历史数据训练神经网络系统，预测未来数小时的温度和工作负荷，实时推送出节能方案，削减了谷歌数据中心制冷能源费用达 40%。

近年来，国内数据中心开始引入 DCIM (Data Center Infrastructure management, 数据中心基础设施管理) 系统，对楼宇设施、基础设施 IT 设备进行能耗实时监控，保障数据中心的高效运行。部分公司逐渐将人工智能技术融合能耗监测系统中实现算力需求精准预测。蚂蚁集团结合人工智能和大数据资源画像能力，对上层业务应用未来 24 小时资源的使用需求进行智能预测，并综合异构算力归一、工作负载 Serverless 化、实践平台工程和配置代码化等技术，实现业务所需计算资源的快速扩缩容和统一调度，目前的预测准确度超过 90%，并结合性能持续优化（Continuous Profiling）的能力动态监测分析和实时优化，已节约数十万核资源。

五、绿色算力技术赋能经济社会发展

随着经济社会的快速发展，高精尖科研和重点行业如工业、建筑、金融、能源等对算力的需求与日俱增，需要处理大量数据进行分析和决策。人们在日常生活中对智慧城市和人工智能生成内容需求也在快

速增长，消费者对于便捷生活和趣味化娱乐有着越来越高的期望。但能耗和环境压力也随之而来，亟需绿色算力技术创新降低成本、提高效率，实现经济社会高质量发展和绿色转型同步进行。

（一）加速科研技术革命

支撑科学的研究的海量、高性能算力需求。高校和科研院所是基础研究、原始技术创新的策源地。高性能计算、理论研究和实验成为当今高校开展科学的研究的三大支柱，特别是高性能计算已成为衡量学校科研实力的重要指标。随着学校师生对高性能计算的算力需求不断增加，算力成为科研重要的计算资源和工具，助力基础科学的研究。在科研型教学模式的推动下，建立一个高效可靠的高性能计算中心成为高校信息化发展的重中之重。与此同时，人工智能正以前所未有的速度向前发展，高校和科研院所作为我国发展人工智能技术的重要高地，肩负着该领域人才培养、科学的研究和科技创新的责任。加强对人工智能相关技术的创新和探索，已成为当前各高校和科研院所备受关注的方向之一。具有人工智能的算力平台将会高效助力科研人员在科学的研究上的突破，以及在多学科交叉领域的创新。

高校和科研院所科研的数据中心需要长时间运行，且计算资源密集度具有时间差异。因此，算力资源的绿色和节能成为重要考量因素。算力平台可应用算力资源的管理、调度和均衡等技术，实现算力资源的动态化，达到绿色节能，节省科研成本。一些重要领域的计算、仿真、模拟也可以应用绿色算力，缩短研究周期，加速科研技术的创新和应用。上海交通大学计算中心“思源一号”可提供 6 千万亿次/秒的

强大算力，为上海交大基础科研和重大项目提供算力支持。该算力中心采用动静分离设计，并通过温水散热技术，系统产生的余热可为其他基础设施供热，比如实验楼球形大厅冬天供暖、食堂厨房热水加热、地下室除湿，以及拓扑量子和时间投影两个实验室的空调温控。

（二）赋能产业节能低碳

赋能工业智能低碳发展。工业领域是“双碳”战略的重点领域，其低碳发展水平直接影响全国整体“双碳”水平。我国工业领域正在向新型工业化道路转型，其中离不开应用数字化技术创新引领的产业转型升级。尤其在工业互联网领域，生产数据体量日益增大，对数据高效便捷处理的诉求愈发强烈。作为工业互联网领域新的动力引擎，绿色算力技术创新和工业互联网融合协同，一方面，“算力”可以实现工业制造生产环节数据的高效流动和制造流程质效提升；另一方面，“绿色”融入算力血脉中，随算力流动，企业应用算力时，产生的碳排放更低，为工业企业降低碳足迹提供了解决思路。位于长沙的大唐先一新能源集控中心，接入了中国大唐集团分布在湖南偏远地区的光伏、风电和水电数据，通过服务器组成强大后台，对分布在湖南全省的海量发电设备进行“少人值守、无人值班、区域管理”运行监管，进一步实现了降本增效。

赋能建筑精细化设计运维。现阶段，建筑能耗已占我国城市总能耗的30%，且该比例正在增大。随着“双碳”目标的进一步落地，建筑行业长期以来存在的发展模式粗放、生产效率低、工业化程度低、资源浪费大、建筑成本高等问题亟待解决。绿色算力技术赋能建筑行业

低碳节能，通过数字化技术应用的推广，推动建筑设计、施工和运维等各个环节的智能化和高效化，为人脸识别系统、智慧停车系统、楼宇智能运维系统、建筑能源数字化系统提供计算支撑作用，实现资源高效汇集、数据畅通流转、提高管理效能。美国的 Verdigris 公司利用人工智能技术实现非侵入式负载监控，实时自动检测建筑用能设备，包括暖通空调冷却塔、电动汽车充电桩、手机充电器、笔记本电脑等，可将客户的能源消耗降低 20%-50%。

赋能金融规模效益质量均衡发展。金融行业数字化、智能化改造一直在各行业中处于较为领先水平。数字化与业务深度融合，大数据分析和人工智能在产品创新、实时风控、精准营销、智能客服、智慧网点、量化交易等场景实现广泛应用，全面支撑金融创新业务场景、降低金融风险、节省运营成本、提升客户体验等。随着算力应用规模的持续增长，如何实现资源利用的降本增效、绿色减碳，成为金融机构共同关心的问题。蚂蚁集团自研的绿色计算体系目前已经覆盖了蚂蚁集团内部包括支付宝、网商银行在内的服务器集群。以支付宝为例，通过落地云原生服务网格将基础设施和业务解耦，极大提升了运维效率和稳定性；通过在离线等不同业务形态混合部署模式，大幅提升资源利用效率，节省上万台服务器资源。

（三）助力社会治理格局构建

搭建智慧城市信息底座。智慧城市建设成为多地的重点工作之一，城市内无时无刻不在产生海量的政务、社会数据，增加了城市治理的复杂度，智慧交通、智慧水务、智慧税务、智慧应急、智慧消防、智

慧医疗、智慧教育、智慧文旅等智慧城市应用场景的开放离不开绿色算力技术的支持，辅助智慧城市建设，提升城市“脑力”，让城市更智慧，绿色融入城市基因，打造城市生态特色，让城市更低碳，也让生活更美好。中国移动（铁岭）清河数据中心利用电厂老旧厂房，通过部署双层双联微模方，采用“工程服务+产品”的总承包模式，敏捷高效的满足了政务云业务需求，成为当地工业厂房“变废为宝”标杆案例。利用厂区闲置空地建设分布式光伏，配套储能设施，减少了市电引入和油机布装。同时，减少旧建筑拆除、新建厂房和材料运输的碳排放，具有建设周期短、节能环保、高效安全等特点。为当地带来良好的经济效益、社会效益和环保效益。在 2022 年北京冬奥会、冬残奥会期间，北京气象局搭建了 800 台高性能计算系统，满足 7×24 小时无休的天气、气候及环境气象业务应用需求，兼顾试验和业务研发的应用需求，相较于此前冬季体育盛会每隔 15 分钟更新一次气象数据，实现了冬奥历史上首次精确到分钟级和百米级的气象预报。在 2023 年杭州亚运会期间，浙江省气象局打造了包含物理及虚拟计算资源池、存储资源池及气象大数据资源池在内的综合气象算力资源平台，实现三维、分钟级、多要素气象立体综合监测网构建的“赛场精准预报模式”。

蕴育大模型时代万业革新。互联网领域，以 AI 算力为基础的智能语音助手、智能算法推荐、AI 美颜、AI 抠图早已成为各类互联网应用的标配，AI 绘画、AI 短视频制作、AI 数字人直播等应用及技术产品遍地开花，泛娱乐领域创作生态全面革新。国内外互联网厂商纷

纷布局人工智能大模型领域，大模型的发展带来了生成式AI(AIGC)应用的爆发，为用户提供突破性创新机会，甚至带来了人类生产力工具的颠覆式革新。与此同时，高能耗和高算力的需求也随着而来。大模型的数据处理量巨大，带动算力产业耗电量快速增加，GPT-3大模型每次训练，碳排放量高达552吨。大模型进一步加剧了算力焦渴，建立兼顾性能和能耗的绿色算法度量标准，构建绿色集约的大模型成为互联网领域绿色高质量发展的必然要求。腾讯发布全新的AI智能创作助手“腾讯智影”，智影数字人能实现“形象克隆”和“声音克隆”，创作者通过上传少量图片、视频和音频素材，就能得到自己的数字人分身和定制音色，进而快速生成自己的数字人播报视频，推动内容生产降本增效。

六、绿色算力发展的展望与建议

展望未来，在寻求可持续发展、实现碳达峰碳中和战略目标的时代背景下，绿色算力已成为算力发展的必选项。为实现绿色算力的高效发展，政策引导与技术创新缺一不可，共同构成了不可或缺的两大支撑。前瞻性政策的制定为绿色算力发展明确方向并提供有力保障，技术的不断创新为绿色算力的实现提供强大动力。

（一）加强绿色算力政策保障

算力基础设施绿色化、低碳化、智能化发展是我国“十四五”期间发展的主旋律，“绿色数据中心”“绿色智能的数据与算力设施”“新型智能算力生态体系”成为了各相关政策文件关注的焦点。但当前以绿色算力为施策方向，针对软件平台高效集约、算法应用绿色低碳进

行统筹规划的政策文件较少，缺乏从用能开始，经过生产调度，再到落地应用的算力全流程一体化规划设计。在政策规划上，要构建高效协同的绿色算力技术产业生态，积极引导绿色算力创新体系建设，培育全栈绿色化算力创新联合体，搭建具备节能环保绿色低碳属性的算力基础设施。支持算力企业提升自主创新水平，加大研发投入，提升核心技术竞争力，不断拓展绿色算力服务边界。在数据中心领域，围绕碳利用效率、碳中和评估、IT设备能效已经存在了相应标准和规范，但绿色算力标准体系还存在空白，对于术语、定义等基础性标准尚不完善，计算评测、分类分级方法亟需制定。未来，要建立和完善绿色算力标准体系，推进标准在算力基础设施等相关重点企业和重点领域中的应用，对算力全生命周期绿色化提供指导和建议。

（二）夯实绿色能源底座支撑

能源是算力产业发展必不可少的驱动力，算力流转中的每一个过程都与能源利用密切相关，无论是服务器的运转、制冷散热设备的驱动，还是大模型平台资源调度都离不开持续稳定的能源支持。因此，能源使用的绿色化在促进绿色算力发展中扮演着举足轻重的角色，但当前仍面临着诸多挑战，包括提高清洁能源在电力结构占比的同时保证供电安全稳定、建设数据中心微电网或源网荷储一体化项目并确保电力供需平衡、研发适合数据中心用能的储能技术、解决数据中心园区分布式能源并网问题、减少数据中心冷却和通风系统的能源消耗等问题。算力的全方位建设为算力设施领域带来了新的发展机遇，同时也提出了更高的要求，其发展将直接影响到算力绿色化的质量和效果。

未来，要全面夯实绿色算力底层基础，筑牢产业创新发展底座，引领算力能源供给侧关键核心技术创新，支持能源主体加快提升新能源电力供给比例，针对维持电力稳定不间断问题加强微电网系统研究与构建，开展储能材料探索和技术研发，针对制冷低功耗需求着力推动液冷相关技术创新、探索余热回收等节能方向。

（三）巩固绿色设备技术创新

算力设备是决定算力供给能力的关键，算力生产的多少主要取决于计算芯片的运算能力，另外还需要存储和网络等相关设备的协同配合，算力才能够更好的发挥生产力作用。我国算力设备规模正持续增长，截至 2023 年 6 月，我国在用数据中心机架总规模已超 760 万标准机架，近五年年均复合增速超 30%。实现算力设备绿色低碳运行是实现算力产业链整体节能高效的重要一步。当前，算力设备绿色发展仍面临应用成本和技术水平方面的双重制约，包括在算力设备规模不变的情况下降低总用电量、提高服务器算力碳效水平并控制相应经济成本、采用低功耗的处理器架构突破技术制约、研发先进存储提升存储设备能效、聚焦运载设备创新降低算力传输损耗。未来，要继续巩固和发展算力基础设施设备绿色化，从设备选型、系统部署、配套设施的全流程进行绿色技术研发和创新应用，支持深度应用浸没式冷板式液冷服务器，推动研发高性能的 AI 芯片，使用先进闪存介质降低能源消耗，探索存算分离架构结合高密存储带来的能耗节约。

（四）聚焦绿色平台能力建设

不同于算力设施或算力设备以实体或硬件形态存在，算力平台是

基于互联网技术发展而构建出的软件环境，使用算力平台，向上可以统筹管理算力设备输入的硬件资源，向下赋能行业应用满足工业、金融等各个行业的算力消费需求，是帮助产业实现算力灵活调配和高效应用的重要支持。算力平台服务商、云服务商等是算力平台的主要建设者。产业界对于优化算力平台能耗的关注度不足，而该环节恰好存在巨大的提升空间。当前，先进的平台正致力于采用先进软件技术复用有限的算力资源、深入业务场景通过“软硬结合”手段提高服务效率。未来，产业将聚焦软件技术使其充分激发硬件的性能潜力，包括根据服务框架、计算框架、存储引擎的属性进行硬件性能动态调节优化，运用AI和大数据能力进行历史数据分析和未来趋势预测实现工作负载绿色，进一步深化平台算法能力优化策略和参数设置，降低代码运行能耗。

（五）深入挖掘应用场景价值

算力作为数字经济时代核心生产力，在经济社会各领域都得到了广泛应用，且由于算力本身所具备的高效性和高灵活性等特点，算力在赋能应用的进程中，不断催生新技术、新模式、新业态，助力各行各业加快数字化、智能化转型，绿色化智能化赋能效果显著。随着算力在各领域应用的走深向实，算力应用正逐渐由互联网行业向交通、工业、金融、政务、医疗、教育等传统行业加速渗透，应用场景也从通用场景拓展到行业特定场景，如智能座舱、智慧城市大脑、AI数字人等。但不同应用场景对算力的需求不一，如何处理好多样化需求与低能耗低排放的关系仍是算力产业亟待解决的问题。未来，在赋能应

用层面，要继续拓展算力在数字经济、数字社会上的赋能范围，广泛建立算力绿色低碳化标识，形成绿色算力消费的碳减排方法论，引导消费者在日常生活中选择消费绿色算力，促使产业上游供给绿色算力以满足消费需求，倒逼生产方式实现算力产业整体绿色低碳发展。

中国信息通信研究院 云计算与大数据研究所

地址：北京市海淀区花园北路 52 号

邮编：100191

电话：010-62300095

传真：010-62300095

网址：www.caict.ac.cn

