

绿色算力技术白皮书

(2023 年)

版权声明

ODCC（开放数据中心委员会）发布的各项成果，受《著作权法》保护，编制单位共同享有著作权。

转载、摘编或利用其它方式使用 ODCC 成果中的文字或者观点的，应注明来源：“开放数据中心委员会 ODCC”。

对于未经著作权人书面同意而实施的剽窃、复制、修改、销售、改编、汇编和翻译出版等侵权行为，ODCC 及有关单位将追究其法律责任，感谢各单位的配合与支持。

www.ODCC.org.cn

编写组

项目经理：

张一星 中国信息通信研究院

杨硕 中国信息通信研究院

工作组长：

郭亮 中国信息通信研究院

贡献专家：

王月 中国信息通信研究院

张磊 中国电信集团有限公司

曹恺 蚂蚁科技集团股份有限公司

彭晋 蚂蚁科技集团股份有限公司

黄挺 蚂蚁科技集团股份有限公司

季雨洁 蚂蚁科技集团股份有限公司

孙曦 蚂蚁科技集团股份有限公司

徐明灼 蚂蚁科技集团股份有限公司

杨豫京 宁畅信息技术（杭州）有限公司

李素红 宁畅信息技术（杭州）有限公司

王海岩 宁畅信息技术（杭州）有限公司

赵雷 宁畅信息技术（杭州）有限公司

陈光 联想（北京）信息技术有限公司

崔吉顺 联想（北京）信息技术有限公司

郝京阳 联想（北京）信息技术有限公司

司晓霞 联想（北京）信息技术有限公司

周弘立 新华三技术有限公司

汪新新 新华三技术有限公司

前言

随着我国产业结构持续优化升级，算力已成为推动各领域数字化、智能化转型的重要基石。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，为推动数字中国建设提供了有力支撑。2020年9月我国正式提出“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。

技术创新是算力发展的根本。当前，算力产业正以绿色发展为主攻方向，推进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产业技术攻关，加大算力服务研发力度挖掘高效无感应用潜力。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，本《白皮书》从算力供需视角出发，提出了绿色算力的内涵定义，总结了绿色算力的属性特征和发展演进阶段。聚焦算力设施、算力设备、算力平台等层次的绿色技术发展态势，分析算力赋能行业绿色发展的典型应用场景，总结我国绿色算力发展面临的挑战，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，我们必须不断创新，持续探索，才能为发展高效、节能、低污染、科学集约的绿色算力奠定坚实的基础。期待本《白皮书》能为大家带来启示，为推动绿色算力革命贡献一份力量。

由于时间仓促，水平所限，错误和不足之处在所难免，欢迎各位读者批评指正。如有意见或建议请联系编写组。

www.ODCC.org.cn

目录

版权声明.....	I
编写组.....	II
前言.....	III
一、绿色算力发展态势.....	1
（一）绿色算力定义内涵.....	1
（二）绿色算力属性特征.....	4
（三）绿色算力发展演进.....	7
二、算力设施的绿色技术发展态势.....	10
（一）绿电消费比例稳步提升，电能传输效率有效提高.....	10
（二）储能部署方式多元可靠，储能设备环境友好效率高.....	11
（三）自然冷却技术快速发展，液冷技术进入全面普及阶段.....	13
三、算力设备的绿色技术发展态势.....	17
（一）服务器硬件节能方向广泛，动态能耗管理智能化发展.....	17
（二）先进存储大幅削减能源消耗，数据软件技术有效提升存储效率... ..	20
（三）网络传输设备向无损化发展，智能弹性网络进一步增强传输效率... ..	23
四、算力平台的绿色技术发展态势.....	25
（一）算力资源调度全局化，促进低碳算力选择应用.....	25
（二）平台产品开发智能化，实现算力资源高效利用.....	28
（三）能效监控指标多元化，深度融合人工智能技术.....	32
五、实现算力赋能绿色.....	35
（一）绿色算力加速科研技术革命.....	35
（二）绿色算力赋能产业节能低碳.....	36

(三) 绿色算力构建社会治理新格局 38

六、 绿色算力发展的挑战与建议 41

(一) 加强绿色算力政策保障 41

(二) 夯实绿色能源底座支撑 41

(三) 巩固绿色设备技术创新 42

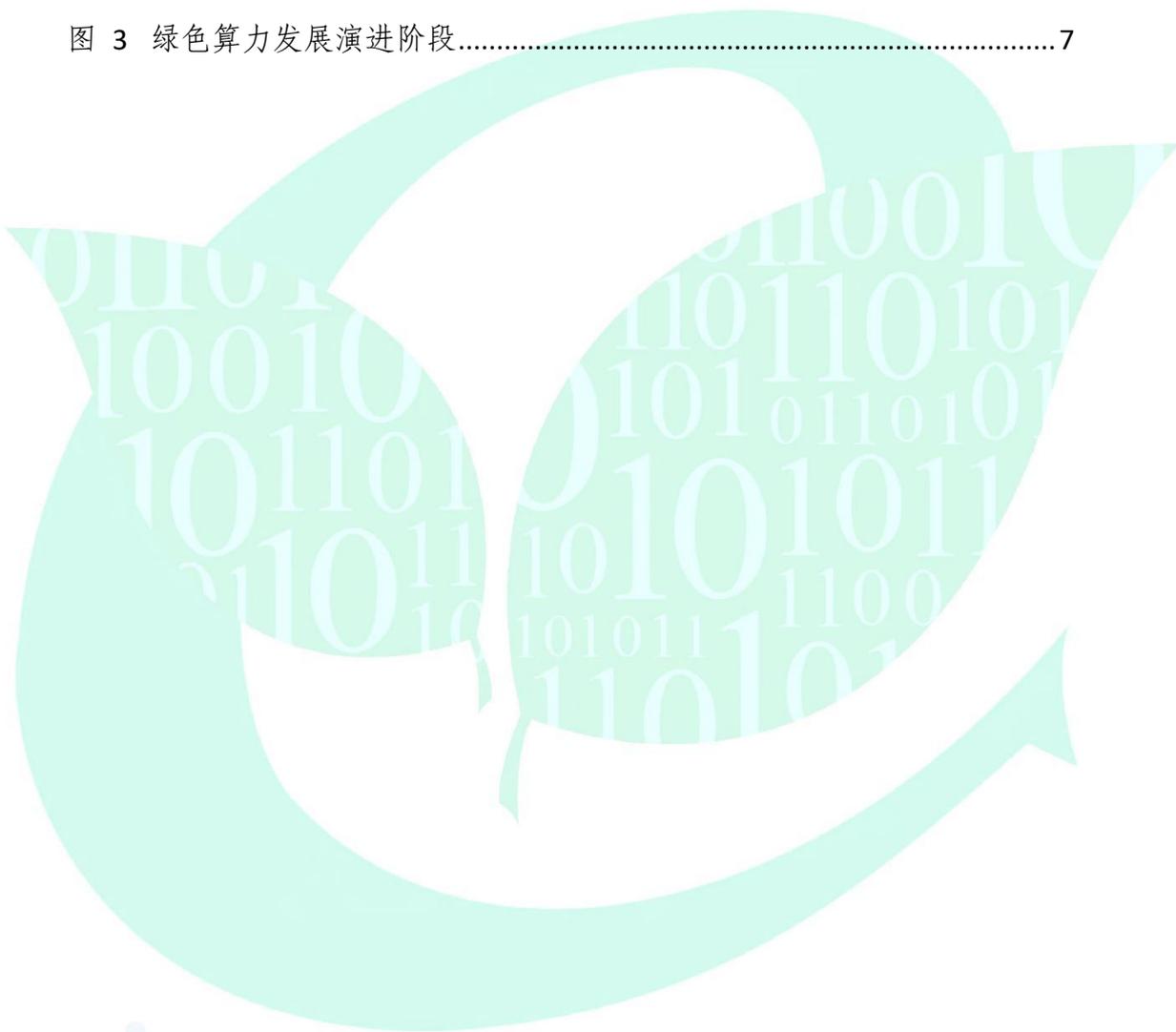
(四) 聚焦绿色平台能力建设 43

(五) 深入挖掘应用场景价值 44

www.ODCC.org.cn

图 目录

图 1 绿色算力框架图.....	3
图 2 绿色算力属性特征.....	4
图 3 绿色算力发展演进阶段.....	7



www.ODCC.org.cn

绿色算力技术白皮书

一、绿色算力发展态势

(一) 绿色算力定义内涵

随着数字经济快速发展，我国产业结构持续优化升级，传统产业增效提质，新兴产业加速布局，算力底座的重要性日益凸显，并成为推动各行业数字化升级的关键要素。在国家“碳达峰、碳中和”战略目标的引领下，高效利用资源，减少能源消耗，降低碳排放和污染物排放，推动算力绿色发展成为了时代需求和行业共识。

政策上看，算力绿色发展的关注重点正逐步从聚焦数据中心规划、选址、布局绿色，过渡到实现产业用能、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。随着我国数字经济蓬勃发展，算力产业规模快速增长，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，用能环节节能减排成为必然。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。在“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。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，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，如华为、

联想、新华三等在硬件算力方面不断进行技术创新与突破。算力产业具有产业链长、参与主体多、应用范围广泛等特点，从能源资源角度出发，算力产业各环节发展均受到资源要素和环境压力的制约；从技术发展角度出发，算力产业发展需要底层硬件、核心算法和平台软件等多方面技术的支持。“双碳”战略为我国绿色低碳创新发展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内外成功经验和做法，一方面是要更加注重可持续发展，用能上从煤炭、石油等传统化石能源向太阳能、风能、氢能等绿色可再生能源转变，材料使用上向集约节约、质量效益好的新材料转变；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，要充分发挥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。

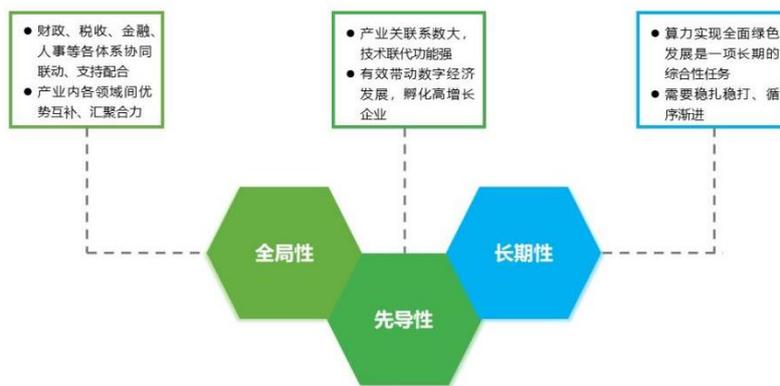


来源：中国信息通信研究院

图 1 绿色算力框架图

算力绿色低碳发展，是国家战略之所需也是行业发展之必然，未来算力将朝着高效、节能、低污染、科学集约等方向持续发力。绿色算力是把环境资源作为算力产生的关键要素，把生产过程“绿色化”作为绿色算力的主要内容，把应用结果“绿色化”作为绿色算力的更高追求，自下而上实现算力在算力设施、算力设备、算力平台和算力应用全体系绿色化的研究领域。绿色算力的目的是打通算力供需通路，完成算力产业用能转型、生产高效、应用和需求场景相匹配。随着绿色算力理念在各层面的深化和实践，绿色算力技术的不断创新和进步，绿色算力正日益广泛地应用于各产业的数字化发展与实践，我国绿色算力体系正不断发展与完善，并逐步形成规模效益。

（二）绿色算力属性特征



来源：中国信息通信研究院

图 2 绿色算力属性特征

全局性。算力与绿色低碳的深度融合发展，既需要财政、税收、金融、人事等各体系协同联动、支持配合，又需要产业内各领域间优势互补、汇聚力。算力覆盖产业广泛多元，进一步深化发展绿色低碳技术，需要巨大的资金支持，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。我国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，引导芯片能力与服务器性能相匹配，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。在绿色算力经算网调度传输，赋能工业、交通、金融、互联网等提质增效的过程中，绿色算力已深度融入了产业循环各领域的全周期和全体系。

先导性。绿色算力产业关联系数大，技术联代功能强，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低碳化发展。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、石化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。类比绿色电力和电力，绿色算力与算力的关键区别在于将节能、低碳、循环、高效等理念全面贯彻到算力生产、应用、赋能、消费等各个环节，为算力发展注入“绿色基因”，实现效率最大、环境友好与社会协调发展。

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，推动我国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，通过算力绿色低碳技术的进步，绿色算力经济将成为助推和引领中国经济发展的新增长点。

（三）绿色算力发展演进



来源：中国信息通信研究院

图 3 绿色算力发展演进阶段

初级阶段：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，绿色化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。

起步阶段：数据中心 PUE 指标被提出，绿色节能成为产业焦点

2007 年，电能利用效率（Power Usage Effectiveness, PUE）由美国绿色网格组织 TGG 提出，用以评价数据中心能源利用效率，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，PUE 的提出为衡量数据中心绿色节能情况提供了可量化的评价标准。各个国家和地区纷纷发布政策推动数据中心 PUE 值降低，PUE 成为全球通用的评价数据中心能源效率的指标。这一阶段，数据中心绿色评价体系已经初步形成，产业围绕 PUE 进行数据中心建设和改造。

发展阶段：数据中心综合能耗评价体系初具规模，算力效率逐渐引起关注

伴随产业发展，PUE 本身的局限性逐渐显露，如无法体现基础设施和 IT 设备内部各模块节能情况、不同规模数据中心的 PUE 标准难统一、缺乏对 IT 效能等要素的评价。仅凭 PUE 指标已经无法全面真实地评价数据中心绿色节能情况。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）陆续被提出，构成了衡量数据中心能耗综合管理能力及绿色程度的评价体系。2022 年，算力碳效（Carbon Emission Per Server, CEPS）由开放数据中心委员会提出，用以评价服务器计算效率，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，该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 及服务器计算效率。这一阶段，算力产业正向着更节能、更高效的方向探索，充分发挥硬件资源潜力，不断实现新的突破。当前，我国绿色算力发展处于该阶段，多元化评价指标被陆续提出，相关企业将绿色意识并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。

成熟阶段：算力生产绿色低碳，绿色算力赋能千行百业

面向未来，算力产业对清洁用能提出了更高要求，如百分百应用可再生能源、应用储能技术、采用余热回收节能技术等，用更少的能源消耗、更低的碳排放生产更多的算力。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，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。

二、算力设施的绿色技术发展态势

算力设施为数据中心提供最基础的动力来源、能源配送和可靠性保障，是保证数据中心 IT 设备运行的必要基础设施，当前，面向算力设施的绿色技术发展，主要围绕供配电系统、制冷散热系统、储能系统三个部分展开。

（一）绿电消费比例稳步提升，电能传输效率有效提高

电力是支撑算力发展的根本，数据中心作为传统的耗能大户，电算转化的载体，其绿色低碳发展面临严峻挑战。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，已是大势所趋。现阶段我国大部分算力需求集中在数字经济发展较发达的东部沿海地区，东部地区缺乏能源，电力以火电为主，使用石油、煤炭、天然气等不可再生能源发电，造成的大气和环境污染问题已引起产业界的广泛关注，提升新能源使用比例、降低碳排放成为产业发展的必选项。除了建立集中式绿电发电站供电外，鼓励分布式发电引导就地消纳也是有效提高新能源供电量的手段。近几年，数据中心在建筑屋顶及外立面铺设分布式光伏，在数据中心产业园区建立分散式风电，通过大量的建设，大幅提升新能源的供电量，促进能源的可持续发展、改善环境，树立绿色企业形象、提升自身的竞争力。当前，包括京津冀、成渝、宁夏、内蒙古等在内的四个国家算力枢纽节点均对可再生能源使用率提出了明确目标。

为确保“业务零中断”，数据中心 IT 负荷往往需要“365×24”小时不间断运行。通过缩短电力传输路径、减少转换层级、提升电力转换效率，实现传输线路的高效供电，是数据中心供配电系统的技术热点。当前，新建的互联网数据中心广泛采用了“HVDC+市电直供”相结合的模式，供电效率可提升到94%-95%。规划以数据中心为主要负荷的微电网、“源网荷储”新方案，实现在负荷侧就近建立电源供电，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，降低电能损耗；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。通过以上措施，节电量达到158万kWh/年；可再生能源发电量约为10.12万kWh/年。

（二）储能部署方式多元可靠，储能设备环境友好效率高

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侧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。在算力领域，储能可分为数据中心级储能和数据中心内储能，为高效存储和利用能

源、减少温室气体排放提供保障。数据中心级储能指的是储能作为数据中心的备用电源，同时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，供电系统效率高达96%，整体寿命高达10年，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。分布式供电模式把属于传统大型UPS的耗电划到了机柜内，极大地降低了PUE值。

储能有多种形式，包括电池储能、氢储能、压缩空气储能等。绿色低碳产业趋势下的储能设备将具备安全可靠、生命周期内性价比高、环境友好、循环寿命长等优点。氢能燃料电池是将氢气和氧气的化学能直接转换成电能的发电装置，其主要优势是对环境无污染、运行安静无噪声、发电效率高，当前业界已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。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。数据中心在应用磷酸铁锂电池时，可重点利用其能量密度较高和常温倍率性能良好的特点。未来，为实现绿色算力，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，形成完整的能源发电体系，实现可持续发展。

（三）自然冷却技术快速发展，液冷技术进入全面普及阶段

对于数据中心而言，制冷系统是影响节能表现的重要因素。在选择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。随着大型数据中心规模持续增加，尽可能使用自然冷却技术，减少和避免使用机械制冷成为数据中心减少能耗的关键。自然冷却技术可分为直接新风自然冷却、间接空气自然冷却、自然冷源水自然冷却、蒸发式冷却水自然冷却等。相比而言，直接新风自然冷却技术是路径最短的自然冷却方式，也是散热效率最高的空气冷却方式。空气质量较差、温湿度不合适的数据中心主要采用水系统自然冷却技术。针对风冷冷却，宁畅研发蝶翼散热系统，CPU 散热器采用 8 热管直触 X 形高流量散热器，散热面积增加 42%，最大支持 400W 风冷散热，并支持液冷 CPU 风冷散热。联想开发八爪鱼式散热器，采用 8*D8 热管，前 4 后 4 的交叉排布的设计，大大提高了热管的传热效率，可以支持 500W-550W 的处理器风冷散热。

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。实现高功率机柜的有效散热成为数据中心关注的重点问题。根据机柜的热密度等级，依次可选择的冷却方式有房间空调、列间空调、机架内部冷却、液冷冷却等。相比而言，液冷技术具有制冷能效高、节约机房空间的优势，可分为直接冷却和间接冷却两种。目前，间接冷却以冷板式

液冷技术为主，直接冷却以浸没式液冷技术为主。液冷技术是使用液体作为热量传输的媒介，可直接导向热源带走热量，其突破了气候条件和地域的限制，可以在全国任何地区进行应用，即使是华南、华东等温度高的地区也可以应用液冷技术进行冷却，最终促使数据中心的制冷电力消耗大幅度降低。

具体而言，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。传统的铜冷板也受到了一些限制，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。联想开发了新型铝材料冷板作为核心换热部件，利用 3DVC 技术，增强了铝冷板的均温性，流阻值比传统铜冷板减少了 30%，实现水冷板热性能提升 10%，成本优化 20% 以上，突破了冷板换热极限。联想全覆盖海神水冷技术能够实现单机柜高密度水冷节点支持 72 个，单机柜功率 100kW 以上，实现高效的算力冷却。宁畅采用铜铝结合冷板，实现铜高效导热、铝减轻重量、无电化学反应风险。聚焦漏液安全，在冷板接头采用宝塔倒刺+螺纹，两级防漏液机制，漏液风险降低 80%。采用四级漏液检测机制，0.05ml 漏液量即可精准监控，1 秒内即可完成漏液报警。新华三采用全焊接+高耐压 EPDM 管设计，冷板耐压能力提升 50% 以上，冷板可靠性大幅提升。此外，新华三设计的 GPU 液冷冷板模组，冷头水嘴采用极简等流三通设计，不增加管路复杂度的基础上，实现流量均分的并联设计，散热能力提升 20%，

流阻降低 30%。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，采用着解耦盲插、风液融合的设计理念，可与多种类型服务器适配，赋能算力基础设施绿色低碳化转型。**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。宁畅 B7000 液冷 TANK 在浸没液冷技术之上，独创开孔及液体流道设计，进一步提高散热效率，PUE 可低至 1.05。新华三 H3C Uniserver Cube 机架箱式浸没液冷系统，将单服务器作为一个密封浸没液冷系统，并采用融合架构，实现了冷板式和浸没式两种路线的优势融合，解决了机房空间利用率低，液冷散热占比小，以及冷却液用量大成本高等问题。相比而言，浸没式液冷的平均 PUE 更低，散热效率更高，是未来数据中心算力设施建设的发展方向。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%。宁畅、联想、新华三、浪潮等主流 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。

此外，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，利用热泵技术将余热回收利用，有着广泛的应用前景，尤其是在我国北方地区，可有效帮助用户降低用热成本。随着更大规模、更多数量的数据中心建设，将有更多的设

备，产生更多的余热，这些余热如果能被加以利用，可以加速“碳中和”进程。联想在关注散热效率和能源效率提升的同时，聚焦能源再利用效率。通过综合能源效率 PUE 和能源再利用效率 ERE 平衡设计，采用热回收技术，对废热回收利用。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，既可以提升 5%的算力，还可以降低 42%的能耗，并创造高达 90%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。

www.ODCC.org.cn

三、算力设备的绿色技术发展态势

算力设备是算力产生的源头，综合了计算、存储和网络等 IT 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）服务器硬件节能方向广泛，动态能耗管理智能化发展

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。

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、处理器个数、产品形态等维度，服务器被分为了多种类型，产业循着上述方向对服务器的高效节能做出优化。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，以 ARM 架构为主流代表。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。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。宁畅 B8000 整机柜服务器，采用全液冷整机柜设计，设备散热效率提升 50%，功率高达 100KW/柜，PUE 最低可至 1.09。并在服务器内存池化未来形态上进行深度探索，采用水、电、网三路盲插设计，比传统部署周期提升 30 倍，可实现高密算力、轻松部署、绿色节能、快速交付、智能运维。

除了上述从计算需求角度提高服务器能效外，还可以通过管理工具加强对电源、风扇等其他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、智能能耗管理等技术。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。

联想从芯片、整机、系统等不同层次进行技术创新，在芯片级别做到了服务器的功率监控；适用业界最高运行效率的电源，并通过智能休眠技术，自适应动态调整工作电源状态；在服务器内部使用高效散热设计，均匀化散热器热流及流阻，采用航空涡轮设计技术，通过双马达冗余热插拔风扇，使风扇在较低转速下即可处于最高效率工作点，在增加散热的同时，节省能源。在系统层次上通过负载均衡技术、监控数据分析预测技术等实现自适应节能。新华三

液冷动态智能控温管理平台，支持服务器级别的智能调温。当机柜内服务器业务负载变化时，智能控温管理平台可监测服务器温度，通过调整泵速和电动调节阀开度，动态调整二次侧回路的温度和流速，实现随着业务的不同维持温度动态平衡。当服务器正常运行时，动态降低泵转速，降低能耗，在部分服务器温度升高时，动态提高泵转速，保证业务进行。动态能耗管理智能化，可有效降低能耗15%左右，做到了IT监控和基础设施调控的融合创新。

（二）先进存储大幅削减能源消耗，数据软件技术有效提升存储效率

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。热存储是指将数据存储在日常访问的存储介质上，具有访问速度快的优点，适合频繁访问的数据。固态硬盘（SSD）读写速度快、功耗低且发热量少，更适合热存储，即热数据的存储场景。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。冷存储通常使用离线介质将数据保存在不需要实时访问的位置，通常用于存储备份、归档或低访问频率数，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几十分之一，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，数据中心存储采用高密设计和风液冷设计来降低能耗。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%。风液冷设计上，采用风液冷技术与数据存储结合已经成为趋势。CPU、GPU 等大功率器件采用液冷，其他器件可采用风冷/液冷，显著降低了内存、HDD 等存储关键部件的工作温度，可降低风扇转速 50% 左右，维护设备运转性能。

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 35% 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 33% 提升到 91%，降低能耗

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。

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，通过标准 iSCSI 协议提供分布式块存储服务，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。以天翼云云计算业务调动系统为例，存储资源盘活技术通过盘活天翼云公司内部存量的服务器资源，实现存储资源的再利用，将原有的存储资源利用率从 17%提升至 65%，盘活的存储空间组成分布式存储资源池，节约存储空间 17.8P，当年节约碳排放 231.4 吨，在一年内节约了 187.4 万元支出。在天翼云科技有限公司为长城紫晶提供的解决方案中，该技术盘活了内部约 100TB 的存储空间。在天翼云内蒙古和林格尔云存储 OOS 机房，通过部署存储资源盘活技术，每年为电信内蒙古园区降低了 12 万元的预算，当年减少 81 吨标准煤消耗，减排 330 吨二氧化碳。

（三）网络传输设备向无损化发展，智能弹性网络进一步增强传输效率

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前提下，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。

www.ODCC.org.cn

四、算力平台的绿色技术发展态势

算力平台是以数据中心算力设备的硬件算力为基础，集成了云原生操作系统、AI（人工智能）操作系统、智能监控系统等多样化工具和框架，给服务对象提供通用算力、智能算力和高性能算力的算力服务平台。算力平台的关键在于实现了算力资源与计算要求有效对接、算力与用户需求有效适配、算力在节点间灵活调度和智能管理，根据实际业务场景，提供针对性的服务，依据业务需求进行软件和算法优化，实现算力高效流转、无感调度、随取随用。麦肯锡早期曾做过一个调研，商用和企业数据中心的服务器利用率很少超过 6%，而高达 30% 的服务器处于“昏睡”状态，一直在耗费电力但没有提供有用的信息服务。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，逐渐成为产业关注的重点和绿色化发展的关键。具体而言，即聚焦软件技术、软硬结合等手段，激发拓展平台能力，发挥全局资源调度优化、智能产品开发技术支撑和绿色监测评估优化等功能，实现算力利用效率提高与平台层面的绿色低碳。

（一）算力资源调度全局化，促进低碳算力选择应用

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、云原生技术通过软件能力优化平台性能，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%。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，**算力池化技术**从虚拟化发展而来，是云原生技术的一种，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。算力平台会集中管理多个计算资源，并维护一个资源池，这些资源可以是物理服务器、虚拟机实例、容器等。资源管理系统负责对资源进行监控、分配、回收和维护，确保资源的可靠性和可用性。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。同时算力平台可以做到有效地按需来针对用户的计算任务进行调度，将任务分配给合适的计算资源。调度算法通常考虑任务的优先级、资源的可用性、任务的依赖关系等因素，以实现公平、高效的资源分配。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。

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。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6月，中国信通院联合中国电信共同发布我国首个实现多元异构算力调度的全国性平台——“全国一体化算力算网调度平台（1.0版）”。该平台汇聚通用算力、智能算力、高性能算力、边缘算力等多元算力资源，实现不同厂商的异构资源池的算力动态感知与作业智能分发调度。特别在 AI 训练作业调度流程中，可实现跨资源池/跨架构/跨厂商的异构算力资源调度。云厂商以及

算力设备商等依据各自的算力资源优势也纷纷布局建设企业算力平台生态。新华三发布了支持 AIGC 大算力调度的“傲飞算力平台”，可与业务深度结合提升工作效率，持续优化算法和积累公域数据，为用户提供最新知识能力。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。

（二）平台产品开发智能化，实现算力资源高效利用

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。

平台工程和无服务是帮助简化产品开发，实现算力资源高效利用的架构模式。平台工程（Platform Engineering）主要指应用与系统软件协同设计、透明研发、部署解耦，系统软件以模块化方式交付，从而提升产研效率，加快业务迭代速度。平台工程是一门设计和构建工具链与工作流的学科，可以为云原生时代的软件工程组织提供自助服务能力，平台工程的目的正是让应用研发者尽可能简单、无痛地参与到规模化的 DevOps 工作中，是让开发人员自助服务从而达到降本增效的目的。无服务（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。使开发人员能够专注于编写和部署代码，而无需管理底层基础设施。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。无服务对开发人员透明地处理资源分配和管理。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。国内使用无服务架构的企业有腾讯的云函数（Serverless Cloud Function, SCF）。云函数是腾讯云为企业和开发者们提供的无服务器执行环境，帮助在无需购买和管理服务器的情况下运行代码。只需使用平台支持的语言编写核心代码并设置代码运行的条件，即可在腾讯云基础设施上弹性、安全地运行代码。

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。平台工程可以提供统一的开发平台和工具链，包括人工智能应用框架的集成和支持，以简化人工智能应用的开发过程，提高开发效率和质量。而无服务架构可以提供计算资源的管理和扩展服务，为人工智能应用框架和平台工程提供支持。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。

计算效率优化上，PyTorch 和 TensorFlow 等框架提供了自动微分功能，能够自动计算导数。这样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。同时，框架支持异步计算，能够同时进行计算和数据传输，最大程度地减少计算等待时间，提高计算效率，降低能源消耗。此外，模型并行和数据并行，框架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。

资源优化和压缩技术上，通过模型压缩和量化，框架提供了模型压缩和量化的技术，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。同时通过图优化和计算图分析，框架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。

硬件优化和兼容性上，通过硬件加速器支持，这些框架与硬件供应商合作，支持各种硬件加速器，如图形处理器（GPU）和神经网络处理器（NPU），充分利用硬件加速器的优势，提高计算效率，降低功耗。同时，通过模型库和算法优化，框架提供了丰富的模型库和算法优化，针对不同硬件平台进行优化，使得模型在特定硬件上运行更高效，提高能源利用率。此外，丰富跨平台兼容性，框架在不同操作系统和硬件平台上都具备较好的兼容性，使得开发者可以在各种设备上灵活使用，从而降

低硬件升级的需求，减少资源消耗。社区合作和共享上，PyTorch 和 TensorFlow 等框架拥有庞大的开源社区，开发者可以共享经验、优化代码和模型，减少重复劳动，提高能效。同时，框架提供了丰富的教程和文档，帮助用户理解和使用高效能的模型开发和训练方法，提高用户的技术水平，进而在实践中促进绿色低碳化发展。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 DLRouter 可以帮助 AI 模型开发人员更专注模型架构设计，而无需关心工程实现方面的问题如硬件加速、分布式运行等，可以让大型 AI 模型的分布式训练变得简单、稳定、快速且绿色。

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。在算力设备上运行的应用程序，其架构和设计可以影响服务器的利用率。没有设计成利用并行处理或分布式计算能力的应用程序可能无法有效利用算力资源。高效的软件设计和优化在最小化 CPU 利用率方面起着至关重要的作用。编写良好的代码可以消除冗余计算，减少不必要的循环。同时使用高效的算法，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。同时，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。在内存管理方面，当前使用了各种技术来优化应用程序的性能，如内存映射文件、虚拟内存和物理内存等技术。在缓存技术方面，使用了 CPU 缓存、磁盘

缓存和网络缓存等。以 OceanBase 为例，基于 LSM Tree 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）能效监控指标多元化，深度融合人工智能技术

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。**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，可以采取减少碳排放的措施，实现低碳运营。

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。然后，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。随后通过监测资源利用率，可以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 demand 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%。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。

www.ODCC.org.cn

五、实现算力赋能绿色

算力赋能是算力产业链的最终环节，是指将算力作为驱动生产方式变革和数字化转型的关键要素，应用于金融、交通、能源等行业，赋能各行各业实现能源效率提升、数字化转型推进及供应链优化。绿色算力是算力在节能环保方向的进一步细化，通过采用更环保的算力设施、更高效的硬件设备、更优化的系统设计等方式使算力具备绿色属性。绿色算力的更高追求是实现赋能应用的绿色化，具体而言，用户在学习和应用时既能利用绿色算力的数字化优势提高生产效率，又能利用绿色算力的绿色化优势推动经济社会节能降碳，充分发挥全局性、先导性和长期性特征，实现经济社会高质量发展和绿色转型同步进行。

（一）绿色算力加速科研技术革命

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”可在极为有限的空间内提供 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，实现超高计算密度与功耗密度。应用了联想第五代温水水冷技术，系统 PUE 可以达到 1.1。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。如此余热回收再利用，替代原来需要消耗的电力、燃气能源，可实现每年 950 吨、约 10%比例的额外碳排补偿。2022 年，“思源一号”在中国高校 HPC 算力排名中位列第一。

（二）绿色算力赋能产业节能低碳

绿色算力助力工业智能低碳发展。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，通过联想的高性能高可靠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。

绿色算力助力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%。

绿色算力助力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，硬件资源成本节约可达 40%，节省上万台服务器资源。

（三）绿色算力构建社会治理新格局

绿色算力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算力技术的支持，算力辅助智慧城市建设，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，仅用 6 个月即完成项目交付，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。利用厂区内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，项目年均减排二氧化碳 6037 吨。同时，减少旧建筑拆除、新建厂房和材料运输碳排放 600 吨，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。新华三在济南北方医疗大数据项目中，大量采用预制化技术，减少数据中心建设过程中的碳消耗，提高了效率，为绿色算力的信息化基础做出创新。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，联想助力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”。

绿色算力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，带动数据中心耗电量快速增加，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，以 176B 参数规模的大模型为例，FP16 精度下，训练过程内存峰值会达到 2800G+，需要 35 张以上 80G 显存的 GPU 才能容纳。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”，智影数字人能够实现“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，腾讯智影还接入了数字人直播，可实现 7×24 小时不间断开播。

www.ODCC.org.cn

六、绿色算力发展的挑战与建议

（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”期间的发展的主旋律，“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。未来，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情况的储能技术、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，已建设 130 条干线光缆，存储规模不断扩大。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。从产业链整体看，由算力平台服务商、云服务商等构成了产业链中游，当前产业对于优化算力平台的能耗的关注不足，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。



ODCC公众号



ODCC订阅号