



# 上海市算力基础设施 发展报告 (2024年)

上海市通信学会“算力浦江”专业委员会  
2024年11月



# 「水木人工智能学堂」

水木AI知识荟 & 交流群 📣

📖 每日分享行业报告、行业资讯等！

🔗 链接海量AI行业精英！

🎉 不定时进行名校名企行活动！

🚀 足不出户，尽在水木AI知识荟！

🔥 扫码添加小编微信，免费进水木AI交流群

交流  
社群



去噪  
星球



去噪星球 每日仅需0.5元

公众号：水木人工智能学堂

## 编写委员会

主编：

陈皆重

副主编（按姓名首字母排序）：

姜欣      李果      梁艳      廖运发      刘义

刘杰      石忆      邬骁斌      严钧      肖晴

郑荣良

编委会专家（按姓名首字母排序）：

蔡玉鑫      常永波      陈俊琰      邓志辉      房思哲

龚沛东      贺夕彧      何艳      纪容      连达

刘晶晶      陆闻婕      黎建伟      苗晓春      彭莉

沈飞      王静      汪君      吴家焯      余乐丰

郑为展      朱纪华

此外，特别鸣谢为本报告做出贡献的国家（上海）新型互联网交换中心、华为上海代表处、燧原科技、浪潮信息、上海邮电设计咨询研究院、上海对外经贸大学等单位相关专家。

## 前 言

当前，全球人工智能技术加速发展，生成式 AI 大模型、高算力智能芯片、大规模智算中心等新质生产力不断突破，国家“东数西算”工程全面实施，算力基础设施建设提速增效。在此背景下，上海加快落实人工智能、算力基础设施建设国家战略，制定发布了系列算力基础设施布局及产业提质发展的政策文件，加快构建算力产业生态集群。

本报告旨在系统梳理上海算力基础设施发展的最新情况，全面展现全市数据中心的算力规模、存力规模、算网运力、空间布局、绿色算力、产业生态等方面的发展现状。从多维度数据分析出发，以期着力体现上海算力产业发展的内涵和优势。分析发现，截至 2024 年 6 月，全市已建数据中心标准机架数 57.4 万个，建成通用算力规模达 7.7EFLOPS（FP32）；同时上海智算中心建设表现亮眼，全市已建大型智算中心 12 个，智能算力总规模超过 54EFLOPS（FP16），处于全国领先地位。

本报告还围绕算力设备、算力中心、算力平台、算力应用四个方面，遴选呈现上海算力产业创新发展的最新代表成果，详细展示“上海市十大算力产业先锋案例”主要内容，为行业提供借鉴参考。报告最后还分析了上海算力基础设施发展面临的挑战，以及进一步发展的意见建议。

本报告中存在不足之处，敬请大家批评指正。

## 目 录

一、国内外算力进入高速发展快车道 .....	1
(一) 技术突破开启算力产业新篇章 .....	1
(二) 智能算力引领全球算力新浪潮 .....	3
(三) 我国算力产业持续引领新发展 .....	6
二、上海市算力产业发展呈现新格局 .....	13
(一) 算力规模能级显著提升 .....	13
(二) 算网运力水平持续提高 .....	16
(三) 绿色算力体系加快构建 .....	19
(四) 算力空间布局愈发完善 .....	21
(五) 算力产业构建更加丰富 .....	24
三、上海市算力创新发展先锋成果 .....	28
(一) 算力设备，引领前沿创新突破 .....	28
(二) 算力中心，重点项目群星闪耀 .....	39
(三) 算力平台，打造最强解决方案 .....	64
(四) 算力应用，赋能行业百花齐放 .....	79
四、上海算力产业发展挑战与建议 .....	93
(一) 面临挑战 .....	93
(二) 发展建议 .....	94

## 一、国内外算力进入高速发展快车道

### (一) 技术突破开启算力产业新篇章

**AI 大模型技术催生强劲算力需求。**当前，OpenAI o1、Gemini、Sora 等大模型相继发布，在人机交互、内容生成等领域展现出了卓越的性能，但其庞大的参数规模和复杂的训练过程对算力提出了极高的要求，促使各行业对算力的需求急剧增加。以 GPT-3 大模型为例，其模型参数量约为 1746 亿，训练一次需要的总算力约为 3640PF-days，即以每秒一千万亿次计算，需要运行 3640 天。2023 年推出的 GPT-4 参数量进一步扩大到 1.8 万亿，是 GPT-3 的 10 倍，训练算力需求上升到 GPT-3 的 68 倍，在 2.5 万个 A100 上需要训练 90-100 天。根据大语言模型综述论文《A Survey of Large Language Models》显示，不同模型的训练数据规模差异显著。例如，LLaMA 使用了约 14000 亿个 token，而 PaLM 则使用了约 7800 亿个 token，如此巨量数据训练推动算力需求急剧增加。在整个大模型行业的持续推动下，算力产业将加快技术创新步伐，以满足日益增长的人工智能及其他领域的需求。

**高算力芯片推动算力产业升级。**在人工智能大模型技术发展驱动下，AI 芯片计算能力正极速提升。云侧方面，2023 年 12 月 7 日，AMD 发布面向 AI 推理和训练芯片 MI300A；2024 年 3 月英伟达发布 B200 芯片，其由两个 Blackwell GPU 和一个 Grace CPU 芯片组合而成，AI 性能为每秒 20 千万亿次浮点运算，是此前 H100 GPU 性能的 5 倍；端侧方面，英特尔 Lunar Lake 芯片计



划于 2024 年末用于高端笔记本电脑，将使用 Lion Cove 性能核心和 Skymont 效率核心，集成新的 Battlemage GPU 架构和强大的 NPU，具备 45 TOPS（TOPS：处理器每秒钟进行一万亿次操作）以上的 AI 处理能力，总体平台 AI 性能可达到 120 TOPS。这些芯片采用了先进的制程工艺和架构设计，不仅提高了算力的供给能力，也推动了计算架构的创新和发展，为深度学习等领域提供了更高效的计算解决方案。同时，高算力芯片的发展也促进了芯片设计、制造等相关产业的进步，为算力产业的可持续发展奠定了坚实的基础。

**多技术协同升级推动先进计算持续发展。**一方面异构计算已成为智能计算领域高算力的主流架构形式。AI 大模型、区块链等多元技术创新应用蓬勃发展，计算领域正经历一场深刻的变革，步入智能计算的新纪元。这一转型不仅加速了计算技术的迭代速度，还深刻影响着计算产业的格局，促使其向更加智能化、高效化的方向重构。在此过程中，搭载各类异构计算加速芯片的 AI 服务器或先进计算平台，其作为算力供给的核心载体，正日益成为支撑智能应用广泛部署的基础设施。另一方面，先进计算领域展现出活跃的体系化创新态势。技术创新持续覆盖基础工艺、硬件、软件、整机不同层次，如 7nm 及 5nm 工艺升级、互联持续高速化、跨平台化演进，软硬耦合加速智能计算进入 E 级时代等。长期看，随着量子计算、光计算、类脑计算等前沿计算技术创新步伐的不断加快，2035 年后先进计算将逐步开启非经典计算规

模化落地应用的发展阶段。

## (二) 智能算力引领全球算力新浪潮

全球算力规模保持高速增长。随着大数据的深度渗透、物联网的广泛铺设，以及人工智能技术的突破性进展，全球算力正展现出强劲且稳定的增长态势。IDC 数据显示，到 2027 年，全球非结构化数据将占到数据总量的 86.8%，达到 246.9 ZB，全球数据总量从 103.67 ZB 增长至 284.30 ZB，复合年均增长率为 22.4%，保持稳定增长态势。根据《中国算力发展报告（2024 年）》报告数据显示，截至 2023 年底，全球计算设备算力规模已达到 910 EFLOPS（FP32），同比增长 40%。其中，通用算力规模为 551 EFLOPS（FP32），智能算力规模为 335 EFLOPS（FP32），超算算力规模为 24 EFLOPS（FP32）。近年来，全球算力结构经历了深刻变革，智能算力作为新业态迅速崛起，截至 2023 年底，全球智算规模同比增长 136%，增速显著超过算力总规模增速。

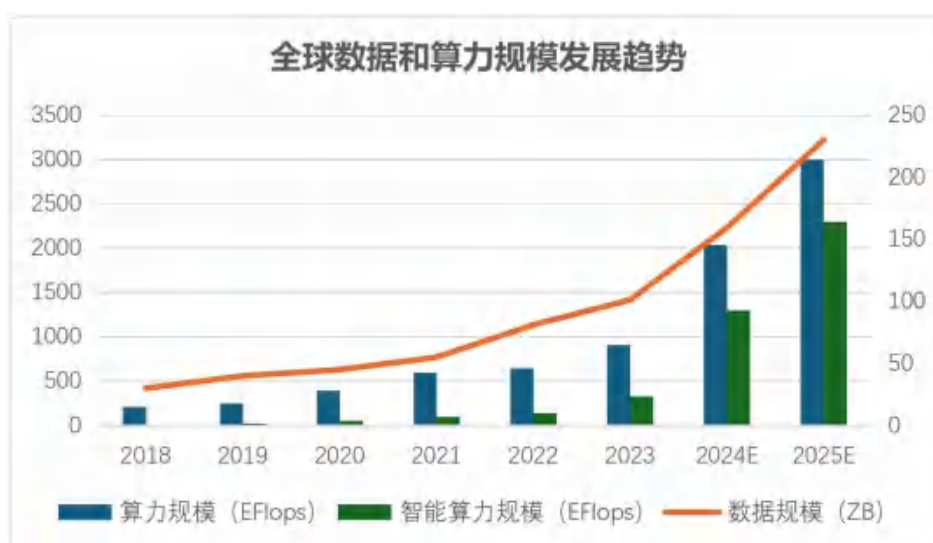


图 1-1： 全球数据（右轴）、算力规模和智能算力规模（左轴）发展趋势  
来源：国家网信办、中国信通院、前瞻产业研究院、IDC、Gartner 数据



**国际巨头引领智能算力发展新态势。**在当下智能算力引领全球算力新浪潮的大势之中，英伟达、微软、英特尔等国际巨头展现出强劲的发展态势与创新活力，成为推动这一浪潮的重要力量。英伟达凭借其强大的 GPU 产品线，在 AI 与数据中心市场中确立了领先地位，硬件创新方面，英伟达持续推出基于 Ampere、Hopper 等先进架构的 GPU 及 Grace CPU，针对 AI 与 HPC 任务进行了深度优化，展现了其在硬件设计上的领先地位；平台与软件生态方面，通过 CUDA 平台，英伟达为开发者提供了高效的编程模型与工具集，促进了 GPU 并行计算能力的广泛利用，同时，cuDNN、TensorRT 等软件库与框架的推出，进一步简化了深度学习与机器学习应用的开发与部署流程。作为英伟达的主要竞争对手，AMD 推出 Ryzen 与 EPYC 系列处理器，强调多核心性能优势，满足 AI 与机器学习任务的多样化需求，通过收购赛灵思，加强了 FPGA 与自适应计算领域的布局；通过 ROCm 等平台提供优化软件工具包与开发者资源，增强了对主流机器学习框架的支持，促进了硬件与软件生态的深度融合。微软在智算芯片领域的布局主要体现在其 Azure 云服务上，Microsoft Azure Maia 100 加速器与 Cobalt CPU，采用 5nm 制程技术，同时提升 Azure 平台对 OpenAI、Copilot 等服务的支持能力，进一步优化用户体验与服务性能。算力中心投资方面，据中国信通院数据，2024 年 1 至 9 月，美国亚马逊、微软和谷歌云服务三巨头披露的 AI 数据中心建设计划总规模超过 1300 亿美元。地区分布主要

集中在欧洲（580 亿美元，占比 44%）、美国本土（306 亿美元，占比 23%）和东南亚（260 亿美元，占比 20%），此外，在中东、日本、南美洲、非洲也有布局。2024 年 9 月，马斯克旗下 xAI 在美国孟菲斯州建设的数据中心 Colossus 上线，其规模达到 10 万张 H100 GPU 加速卡，计划未来扩展至 20 万张。

**全球主要国家和地区持续加码推进智能算力发展。**在全球战略竞争中，智能算力已成为各国争夺的新高地，全球主要国家和地区纷纷加快智能算力战略布局进程。2024 年 9 月，美国白宫宣布启动一系列举措，加强 AI 数据中心基础设施建设，以智算中心打造全球科技竞争的新引擎，全力推进在美国本土建设数据中心。这一系列战略行动不仅彰显了美国政府对智算中心建设的高度重视，更表明了智算中心作为未来科技创新与经济发展的核心驱动力的重要地位。此外，白宫成立智能数据中心基础设施工作组，扩大对联邦、州和地方当局数据处理中心的许可。欧盟于 2024 年 7 月 9 日正式实施了《欧洲高性能计算联合企业条例》（《EuroHPC 修正案》），进一步扩展了欧盟高性能计算联合企业的职能范围，明确将 AI 工厂的开发与运营纳入其中。该修正案的核心内容包括：开发和运营 AI 工厂、增强超级计算能力和支持 AI 技术创新。通过该修正案，欧盟高性能计算联合企业将有权购买、升级和运营 AI 工厂，显著提升欧洲超级计算机对科研和工业用户的可及性。日本政府近年为 5 家日本企业提供总额超过 700 亿日元的补贴，用于打造人工智能超级计算机，以减少

对美国技术的依赖。2024年，加拿大政府斥资20亿美元以增强人工智能领域的的数据计算与处理能力；启动面向公众的人工智能算力设施咨询服务，该服务专注于推广人工智能算力设施，提高其在研究机构、企业及个人用户中的知名度与可及性。

### **(三) 我国算力产业持续引领新发展**

算力是数字经济时代的新质生产力，算力网是支撑数字经济高质量发展的关键基础设施。我国对算力产业的发展予以高度关注，通过精准的政策布局持续引领着算力产业的发展航向。《中国算力发展报告（2024年）》显示，2023年，中国数据生产总量达到32.85 ZB，同比增长22.44%，呈现稳步增长态势。截至2023年底，我国提供算力服务的在用机架数达810万标准机架，同比增长24.2%。截至2024年6月，算力规模达到246 EFLOPS（FP32），其中智算规模为76 EFLOPS（FP32），同比增速超过65%，持续引领产业变革；存力规模达到1200 EB，其中先进存储容量占比超过25%，彰显了中国在全球算力版图中的强劲增长势头。



图 1-2：我国在用算力中心标准机架规模

来源：中国信息通信研究院

**算力网络运载能力优化显著。**据中国信通院公开信息，截至 2023 年底，国家级骨干直联点已增加至 26 个，骨干网互联带宽扩容至 40T。在接入网络方面，58.3% 的在用算力中心已接入骨干网；从接入带宽来看，算力中心的平均出口带宽达到 821Gbit/s，在用机架的平均带宽约为 885Mbit/s。根据对时延达标率、高速光传输网络端口数、光层容量、IP 层容量和综合接入节点 OTN 覆盖率五个指标的评估，广东、江苏、湖北、四川、浙江、山东、河北、上海、贵州和云南位居前十。各枢纽间和从枢纽到各地市的时延表现良好，大部分省市基本形成 20ms 时延圈和 5ms 时延圈，部分省市的光层和 IP 层时延性能仍需进一步提升。针对智算业务的需求，高速光传输网络端口数量持续增加；省际干线的 400G 端口数量已接近 9000 个，占比达到 4%。光层容量方面，各枢纽之间的总带宽配置为 923.6Tbps，远高于实际使用带宽。IP 层的带宽配置为 300.75Tbps，基本满足行业应用需求。综合接

入节点的 OTN 覆盖率达到 73.31%，逐步增强对用户全光品质的支撑能力。



图 1-3: 运载力综合水平得分

来源：中国信息通信研究院

**智算产业上升为国家战略竞争点。**在近年来国家信息基础设施建设的宏伟蓝图中，智算中心的崛起成为了至关重要的战略支点。2020年4月20日，国家发展改革委正式界定了新型基础设施的范围，将智能计算中心作为算力基础设施的重要代表，纳入信息基础设施的广泛范畴。随后，2021年5月，国家发改委联合多部门颁布了《全国一体化大数据中心协同创新体系算力枢纽实施方案》，旨在构建全国性的算力网络枢纽节点布局，以推动算力资源的优化配置。进入2022年，国家发改委等四部门联合发出通知，批准在京津冀、长三角、粤港澳大湾区、成渝地区双城经济圈以及内蒙古、贵州、甘肃、宁夏等八个关键区域启动国家算力枢纽节点的建设，并规划了十个国家数据中心集群，标志着“东数西算”工程的全面铺开。2024年7月，国家发展改革



委联合有关部门印发《数据中心绿色低碳发展专项行动计划》提出强化人工智能重大生产力布局，综合考虑数据中心新增用能用水需求和资源保障能力，引导智算中心规范化建设，推动形成集聚效应。

表 1： 算力基础设施相关国家政策（2023-2024）

时间	出台单位	发布政策	政策要点
2024年8月	工业和信息化部等十一部门	《关于推动新型信息基础设施协调发展有关事项的通知》	优化布局算力基础设施。支持数据中心集群与新能源基地协同建设，推动算力基础设施与能源、水资源协调发展。加强本地数据中心规划，合理布局区域性枢纽节点，逐步提升智能算力占比。
2024年7月	国家发展改革委、工业和信息化部、国家能源局、国家数据局	《数据中心绿色低碳发展专项行动计划》	鼓励各地区面向本地业务需求，采取高端替换、增减挂钩、重组整合等方式，促进小规模、低效率存量分散数据中心向集约化、高效率转变。到2025年底，国家枢纽节点地区各类新增算力占全国新增算力60%以上，国家枢纽节点算力资源使用率显著超过全国平均水平。
2023年12月	国家发展改革委等五部门	《深入实施“东数西算”工程加快构建全国一体化算力网的实施意见》	统筹通用算力、智能算力、超级算力协同计算。到2025年底，综合算力基础设施体系初步成型。国家枢纽节点地区各类新增算力占全国新增算力的60%以上，国家枢纽节点算力资源使用率显著超过全国平均水平。
2023年10月	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》	到2025年，计算力方面，算力规模超过300EFLOPS，智能算力占比达到35%。
2023年2月	中共中央、国务院	《数字中国建设整体布局规划》	优化算力布局，促进东西部算力联动，引导通用数据中心、超算中心、智能计算中心、边缘数据中心等合理梯次布局。

来源:政府各部门官网整理



国内企业积极布局智算芯片研发。华为、百度、阿里巴巴等互联网巨头加快 AI 芯片自研步伐，并在各自云平台率先试用。一些初创公司从芯片底层架构设计入手，试图通过“弯道超车”抓住机遇。华为打造全栈自主 AI 基础软硬件。基于昇腾系列处理器构建全栈 AI 计算基础设施及应用，通过丰富的产品形态打造面向“端、边、云”的全场景 AI 基础设施方案。在基础硬件方面，推出 Atlas 系列产品；在软件方面，构建全栈全生命周期的模型开发工具链。推出异构计算架构 CANN，构建自主开源全场景 AI 框架 MindSpore 以及一系列应用使能工具。百度昆仑芯 2 代搭载自研第二代 XPU 架构。XPU-R 在性能上实现了 2 至 3 倍的提升，作为国内首款采用 GDDR6 显存的通用 AI 芯片，其内存带宽高达 512GB/s。平头哥发布首款 Arm 服务器芯片。倚天 710 由平头哥自主设计研发，采用先进架构，具备高能效、高带宽等特点，兼容 Armv9 架构，支持 8 通道 DDR5，最高速率 4400MT/s，峰值总带宽 281GB/s。海光 DCU 依托开放生态加速推广。海光信息推出的 DCU 产品采用 GPGPU 路线，性能强悍、生态完善，已在人工智能、大数据处理、商业计算等领域规模化应用。其提供自主开放的完整软件栈，兼容“CUDA”“ROCm”生态，支持主流深度学习框架与应用软件。

上海市以其独特的科技创新氛围和资源优势，孕育出了一批杰出的智算芯片企业。沐曦构建全栈 GPU 芯片产品线。曦思®N 系列，专为 AI 推理设计；曦云®C 系列，适用于千亿参数级别

的 AI 大模型训练及通用计算任务；曦彩®G 系列，专注图形渲染领域。燧原科技于 2024 年 7 月推出面向数据中心大规模部署的新一代人工智能加速卡燧原®S60，满足大语言模型、搜广推及传统模型的需求，具有模型覆盖面广、易用性强、易迁移易部署等特点。天数智芯研发的天垓 150 基于通用 GPU 架构，被用于构建大模型训练异构算力集群，并取得了国际 SOTA 水平。在智源研究院的 Aquila2-70B-Expr 大模型混合训练项目中，使用天垓 150 构建多个节点集群并完成了模型的预训练，训练结果达到了国际领先水平，展现了其在异构计算中的强大混合训练能力。壁仞科技于 2024 年 9 月发布了其自主研发的异构 GPU 协同训练方案——HGCT，成为业界首个实现三种及以上异构 GPU 混合训练同一大模型的技术先驱，异构 GPU 间的通信效率上达到 98% 以上，端到端训练效率维持在 90% 至 95% 之间。益思芯科技 DPU 智能网卡产品实现 P4 可编程的网络加速、NVMe-oF 存储加速、RDMA 等核心功能，并和国产 GPU 厂商完成了智算中心国产算力底座的构建；FTTR 全光网络芯片已经率先完成流片和商用供货，处于国内领先水平。

表 2：上海市智算芯片企业及产品介绍

企业名称	产品名称	性能特点	应用场景
沐曦	曦思、曦云、曦彩	MXN 系列 GPU（曦思）用于 AI 推理，为人工智能应用提供强大的算力支持。MXC 系列 GPU（曦云），算力卓越，如，MXC500 GPU 的算力高达 15TFLOPS，具备 CUDA 兼容性。MXG 系列 GPU（曦彩），专注于图形渲染领域，支持各种复杂的图形处理需求，为用户带来卓越的图形表现效果。	可广泛应用于人工智能、智慧城市、数据中心、云计算、自动驾驶、科学计算、数字孪生、元宇宙等前沿领域。
燧原	邃思	邃思 2.0 芯片基于人工智能领域专用处理器架构设计，基于独家 GCU-LARE®技术支持专属 P2P 通道，提供互联高带宽低延时，高效应对多卡间高并发高流量的数据通信需求，提升分布式训练性能。邃思 2.5 人工智能推理芯片基于第二代 GCU-CARA®架构，作为云燧®i20 推理卡的算力核心，提供强劲的全精度 AI 算力、先进的存储方案。	广泛应用于票据识别、智能客服、精准营销、智能风控、绿色智算中心、泛互联网、智慧城市、金融等多个关键行业和场景。
天数智芯	天垓 150	基于通用 GPU 架构，兼容国际主流的 GPU 计算模型，确保与各类主流系统和应用的无缝对接。支持国内外的主流 AI 生态系统。并兼容如 TensorFlow、PyTorch 等深度学习框架，用户可以在多个平台上自由选择所需工具，享受灵活性。性能显著提升，具备处理大规模训练和复杂计算的强大能力，高效应对复杂的 AI 计算任务。	为智慧城市的关键应用提供高性能计算支持，如交通流量分析、环境监测。作为数据中心的核心理算单元，大幅提升数据处理的效率，支持云计算和大数据处理等高密度计算任务的高效运行。
壁仞	BR100	峰值算力超过了英伟达目前在售的旗舰计算产品 A100 GPU 的三倍，创出全球算力纪录，16 位浮点算力达到 1000T 以上、8 位定点算力达到 2000T 以上，其单芯片峰值算力更是跃升至 PFLOPS 量级，标志着中国的 GPU 芯片迈入“每秒千万亿次计算”新纪元。通过 chiplet 封装技术，在提升性能的同时降低功耗，能效比出色。	专为 AI 训练、推理及通用计算设计，具有高算力和高能效比，适合大规模 AI 处理。此外，还应用于生物科学、工业设计、生产制造、农业、航天航海、地质勘探与宇宙探索等多个领域。

来源:相关企业官网整理

## 二、上海市算力产业发展呈现新格局

### (一) 算力规模能级显著提升

近年来，上海持续加快算力基础设施建设，已初步形成以智能算力为主导，智能算力、通用算力、超算算力协同发展的算力产业体系。根据中国信息通信研究院《中国综合算力指数 2024 年》显示，上海算力综合指数全国排名第三，连续三年排名居前，在综合算力领域具有持续竞争力。

上海算力规模保持稳步增长，浦东新区保持领先。截至 2024 年 6 月，全市已建数据中心标准机架数为 57.4 万个（以功率 2.5KW 为一个标准机架）。全市数据中心在用标准机架数 46.4 万个，上架率约 69.3%，整体 PUE 在 1.5 左右。全市已建数据中心通用算力规模 7.7 EFLOPS（FP32），在全国范围内处于前列。从区域分布来看，已建数据中心标准机架数排名前三的依次为浦东新区 28.0 万个、宝山区 7.5 万个、青浦区 5.7 万个。在用标准机架数排名前三的依次为浦东新区 22.3 万个、上架率 65.2%；宝山区 6.3 万个、上架率 90.6%；青浦区 2.7 万个，上架率 60.4%。



图 2-1: 上海各区域数据中心标准机架数及上架率

来源: 上海市通信管理局

从数据中心规模类型看, 上海在用数据中心 (指单数据中心物理标准机架数超过 100 个) 数量为 127 个, 其中超大型数据中心 13 个、大型数据中心 46 个、中小型数据中心 68 个, 主要分布于浦东新区、宝山区。另据统计, 目前全市数据中心 IT 机房建筑面积已超 500 万平方米, 且以商业用电为主, 总配电量超 415 万 KVA, 相较于全国均处于前列。

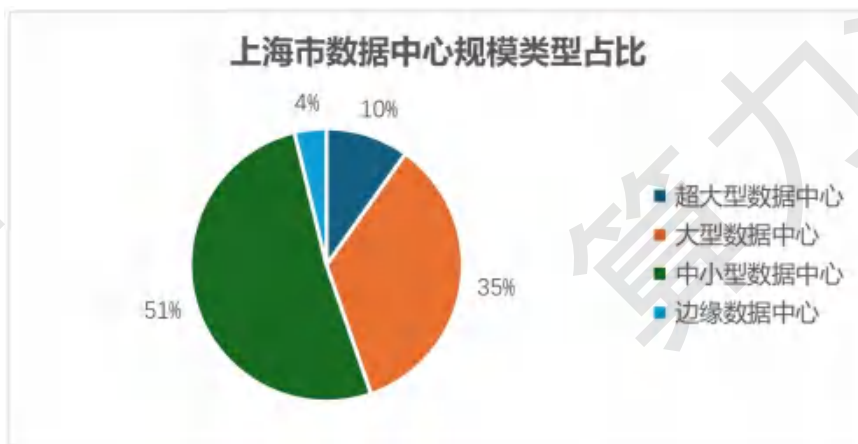


图 2-2: 上海市数据中心规模类型

来源: 上海市通信管理局



**上海智能算力建设规模初显。**上海智能算力部署提档升级，截至 2024 年 6 月，全市已建大型智算中心 12 个，智能算力总规模超过 54 EFLOPS（FP16），其中，浦东新区智算规模占全市比重 38.8%，松江区智算规模占全市比重 47.7%，即全市超八成的智能算力集中在浦东新区、松江区。单体智算中心层面，临港新片区商汤智算中心、松江区科技网大数据计算中心、松江区腾讯智算中心智算规模排名前三。此外，中国电信、中国移动、中国联通三大运营商总计已在临港新片区投产智算近 7 EFLOPS（FP16）。预计到 2025 年底，上海智能算力总规模将超过 70 EFLOPS（FP16）。

**上海超级计算布局稳步推进。**上海超级算力主要依托上海超算中心，以及复旦大学、上海交通大学、上海科技大学等高校院所建设落地，目前已形成一定规模超算供给能力，加快助力上海科研计算、科技研发、AI4S 攻关突破。未来，上海市将进一步支持上海超算中心高性能计算资源升级扩容，构建自主核心软硬件深度应用、高性能计算与智能计算多元融合的先进算力平台。

**上海存力规模持续位居全国前列。**根据中国信通院发布的《中国存力发展报告（2024 年）》显示，我国存力规模稳步发展，2023 年增速约 20%，存力总规模达到 1200 EB，东部各省份存力发展水平较高，全国占比近 61%。其中，上海以存力总规模近 116 EB 的量级在全国各省级行政区规模排名第四。从存力分指数来看，根据《中国综合算力指数（2024 年）》显示，上海



在全国省级行政区存力分指数排名第三，且与去年持平，仅次于广东、江苏。近年来，上海引领全国数据存储技术突破，先进存储、国产存储占比不断走高，SSD 硬盘使用规模、分布式存储占比进一步提升，基础电信运营商存储核心数据、重要数据灾备覆盖率达到 100%。预计 2025 年上海存力容量总规模将达到 175 EB，先进存储容量比例将达到 30%。

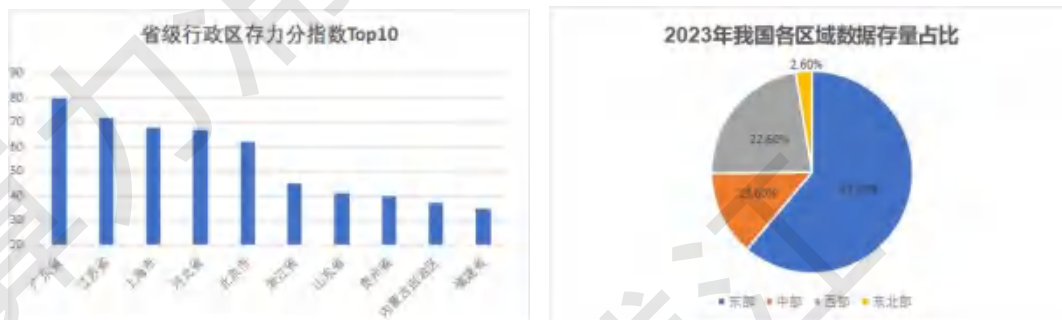


图 2-3: 全国各省级行政区存力指数及各区域数据存量  
来源：中国信息通信研究院

## (二) 算网运力水平持续提高

上海高度重视算网运力布局发展，加快构建泛在互联的高水平网络基础设施，不断提升 5G-A 网络、万兆光网的覆盖广度和应用深度。

上海大部分数据中心已接入省级以上骨干网络。据统计，全市数据中心接入国家级骨干网 16 个、省级骨干网 75 个，占比分别为 13%、59%，即七成以上数据中心均已接入省级以上骨干网络，超全国平均水平十六个百分点。同时，接入城域网的数据中心有 19 个，占比 15%，未来上海将积极推进此类型数据中心纳入省级以上骨干网，实现跨城、跨区域网络连接互联互通。此外，全市支持 IPv6 协议的数据中心已达 54 个，占比 43%，显著高于

全国平均水平。面向未来网络技术持续演进，上海算力中心网络正向超大规模、超高带宽，超低时延、超高可靠等方向发展。

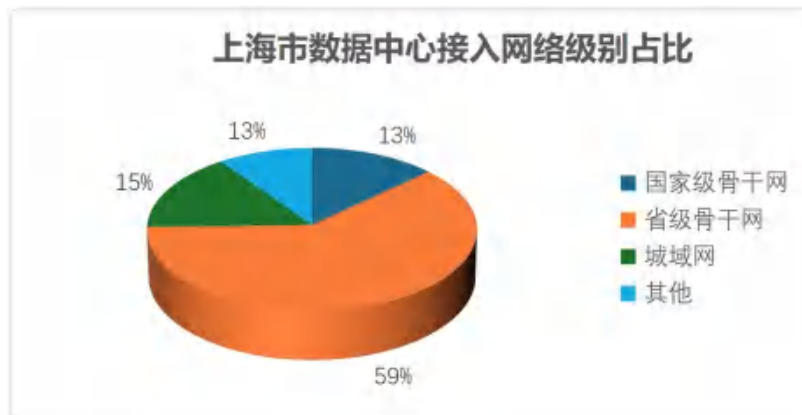


图 2-4: 全市数据中心接入网络级别占比

来源：上海市通信管理局

上海算力中心网络通信质量不断优化。时延圈层面，上海已打造链接全国的算力网络“5-10-20”三层时延圈，其中 5ms 时延圈覆盖长三角省市，5ms~10ms 时延圈覆盖长三角周边省份（如江西、山东、湖北、河南、福建、湖南），10ms~20ms 时延圈覆盖京津冀枢纽、成渝枢纽、粤港澳大湾区枢纽、贵州枢纽、甘肃枢纽、宁夏枢纽、内蒙古枢纽以及其它省份（如广西，辽宁，海南，陕西，山西，吉林，云南）。



图 2-5: 上海节点 5-10-20 三层时延圈

来源: 中国信息通信研究院

时延数据层面，京津冀、长三角、粤港澳、成渝、贵州等全国其它枢纽的重点城市访问上海算力枢纽节点的网内平均时延为 12.66ms，网间平均时延为 13.42ms，全国各省市访问上海算力枢纽节点的网内平均时延为 13.95ms，网间访问平均时延为 15.29ms，这些时延数据较 2023 年均有一定程度的优化。上海市域内，上海各数据中心节点之间的网络质量较好，平均访问时延仅为 1.92ms；上海各数据中心节点至全国其它枢纽和 31 省的性能接近，平均时延均不超过 16ms。

算力调度方面，上海构建了一张城市级高速全光算力环网。上海以基础电信运营商、上海新型互联网交换中心为主体构建了“3+1+N”的算力网络调度体系，其中三大电信运营商负责跨区域的算力网络调度，新型互联网交换中心负责本市算力中心之间的算力网络调度。截至 2024 年 10 月，新型互联网交换中心已完

成 19 个网络节点建设，覆盖全市 24 家重点数据中心企业，其基于裸金属和容器级别的算力调度交易平台 2.0 版本已正式商用，目前和 11 家云算平台互联互通，累计归拢通算资源 6334P、超算资源 101P、智算规模 1816P。长三角算力枢纽建设方面，上海启动长三角（上海）算力互联互通平台建设工作，平台建成后将在长三角地区形成“跨地域、跨主体、跨架构”的算力资源标准化互联互通，实现算力资源高效供需匹配。基础电信企业已启动超过 3500 公里本地和长途光缆工程建设，打通长三角一体化示范区算力枢纽节点和芜湖枢纽节点的算力网络。

### **（三）绿色算力体系加快构建**

上海正加快打造智能算力基础设施，对新建智算中心 PUE 值、存量改造智算中心 PUE 值、智算中心内绿色能源使用占比、液冷机柜数量占比等绿色算力关键指标提出了全新的要求。

**全市数据中心 PUE 能效结构优化显著。**近年全市数据中心进行了 PUE 能效测试，其中 88.5%的数据中心依托自身技术团队进行 PUE 测试，11.5%数据中心则邀请独立第三方进行测试。根据有效填报 PUE 测试结果显示，有 19 家数据中心的实测 PUE 在 1-1.3 之间，占全部数据中心的 14.9%；PUE 在小于 1.5 的数据中心数量为 56 个，占比 44.1%，显示了全市数据中心 PUE 能效结构正有效改善。大部分实测 PUE 数值稍高的数据中心正在开展节能改造或提出节能改造计划，其中已完成节能改造的数据中心数量达 23 个。基础电信运营商则通过“老、旧、散、小”

数据中心升级逐步将 IT 设备平均在网年限缩减到 5 年内，以降低能耗实现绿色算力。

**全市数据中心逐步转向新型制冷方式。**据统计，全市数据中心采用风冷制冷方式有 59 家，占比 46%，采用水冷、混合制冷方式的数据中心数量分别为 45 家、17 家，占比分别为 35%、13%。随着数据中心高性能服务器、高功率机柜的使用量越来越大，先进计算中心等新型技术设施建设越来越多，数据中心的制冷方式结构也发生了显著改善，水冷、液冷、混合制冷等新型制冷方式需求也越发强劲，其实际部署占比正加速走高，将助力数据中心能源利用更加高效。

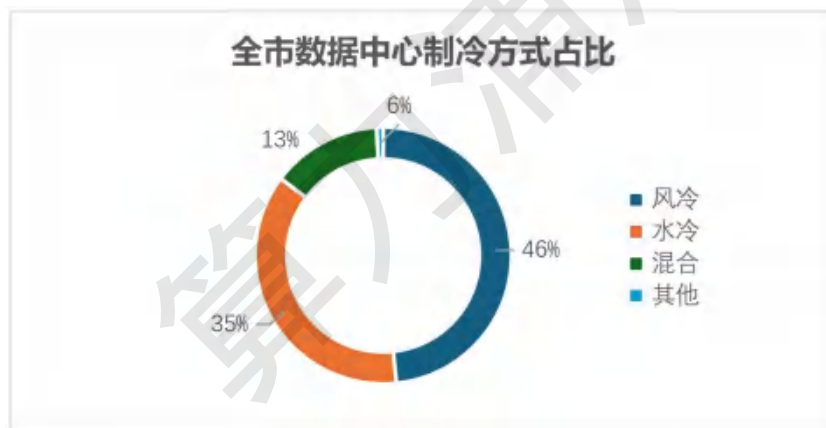


图 2-6: 全市数据中心制冷方式占比

来源：上海市通信管理局

**绿色清洁能源使用正成为新风向。**据不完全统计，全市数据中心使用太阳能、水能等形式清洁能源的数量为 37 个，占比 30.3%，也就是说平均每三家就有一家数据中心使用了绿色能源，其中有 4 个数据中心实现了 100%清洁能源使用率，分布于浦东新区、嘉定区和杨浦区。上海数据中心清洁能源来源方式主要有自建清洁能源设施、购买绿证、绿电市场化交易、源网荷储一体



化项目等，其中自建清洁能源设施占比最高达 54%。此外，上海积极推动氢燃料电池、生物燃料等绿电在数据中心的试点，全面普及智能光伏在数据中心的应用，光伏供电占数据中心整体电量的 5%左右，推动万国外高桥等多个数据中心开展零碳数据中心试点，实现用能的全面绿色化。

#### **(四) 算力空间布局愈发完善**

依据上海算力基础设施布局统一要求，上海数据中心落地选址更趋优化，尤其是大规模智算中心正形成向临港新片区、松江区、长三角一体化示范区等地聚集的态势。

从区域分布来看，上海数据中心在全市各区基本均有分布，其中浦东新区以 59 个、全市占比 46.5% 的优势排名第一，紧随其后的分别是宝山区 16 个，青浦区 9 个，静安区 8 个。虹口区、杨浦区、奉贤区等区域数据中心数量较少，长宁区、崇明区则暂未有数据中心体现。同时，上海大力推动基础电信企业启动边缘数据中心建设规划，截至 2024 年 6 月底，各基础电信企业依托分布在上海的接入机房初步形成了 46 个边缘数据中心节点。





图 2-7：全市数据中心数量区域分布

来源：上海市通信管理局

从全市空间布局来看，上海已建数据中心在市域范围内呈现出分区域聚集态势。其中浦东金桥、外高桥、闵行浦江镇及周边区域，以及黄浦区、静安区市中心区域范围聚集态势显著，该地区以中小型传统 IDC 布局为主。另外，青浦区、松江区、临港新片区等区域，数据中心数量逐步增多，该地区主要以落地高性能计算中心、超大规模智算中心为主。总体上，全市数据中心分布和上海大数据、云计算、软件信息、人工智能、智能网联汽车、智能制造等产业发展空间布局高度吻合。

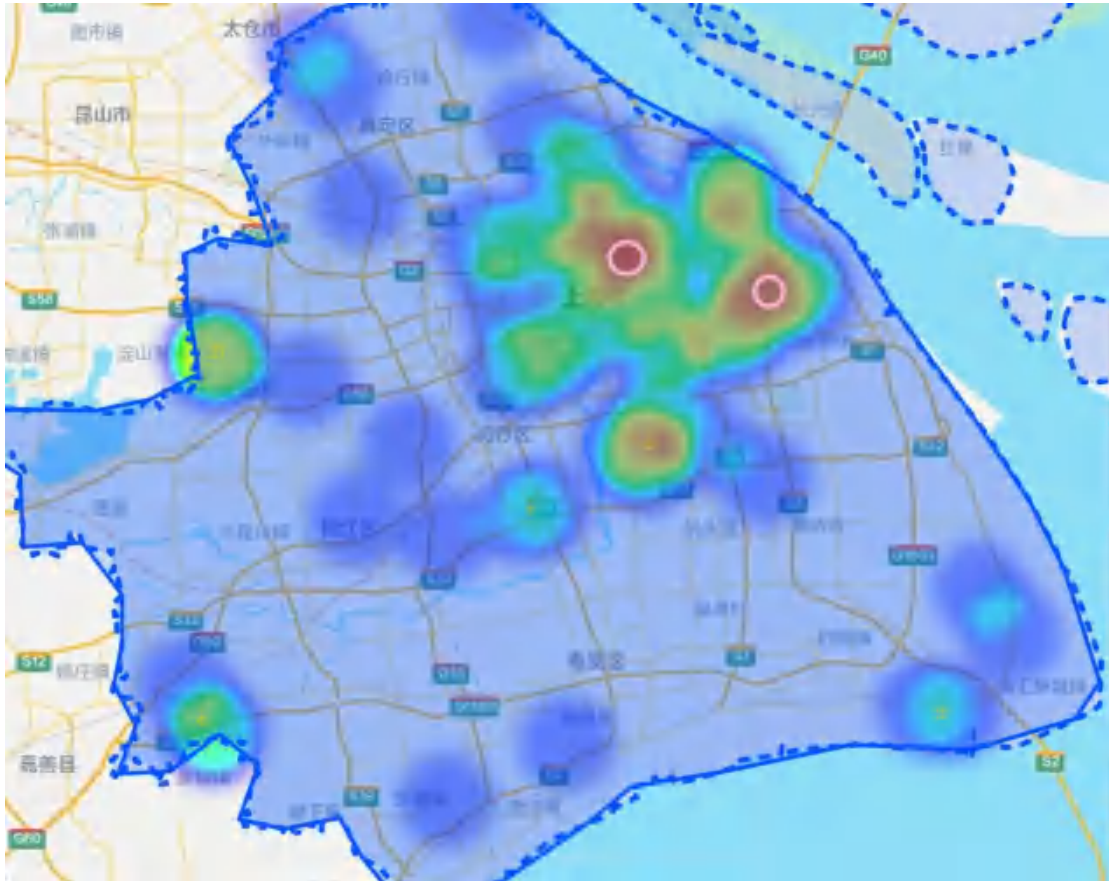


图 2-8: 全市数据中心物理地址空间分布情况

来源: 上海市通信管理局

智算中心空间布局方面，上海已落地 12 个大型智能计算中心，智算总规模已超 54 EFLOPS (FP16)，在数量规模、算力供给等方面均处于全国前列。上海智能计算中心形成了“三平台多中心”的空间布局态势，“三平台”即上海算力交易平台（新型互联网交换中心）、上海市智能算力公共服务平台（仪电集团）、上海市人工智能公共算力服务平台（上海超级计算中心）。“多中心”分别是临港、松江、浦东、嘉定和青浦等地区，落地了具有代表性的智能算力中心。尤其是临港新片区，已建设包含全国首个 5A 级智算中心的商汤临港 AIDC，以及三大运营商承建的大规模智算集群等，临港新片区已成为上海乃至全国智算基础设

施最密集、规模最大的地区之一。

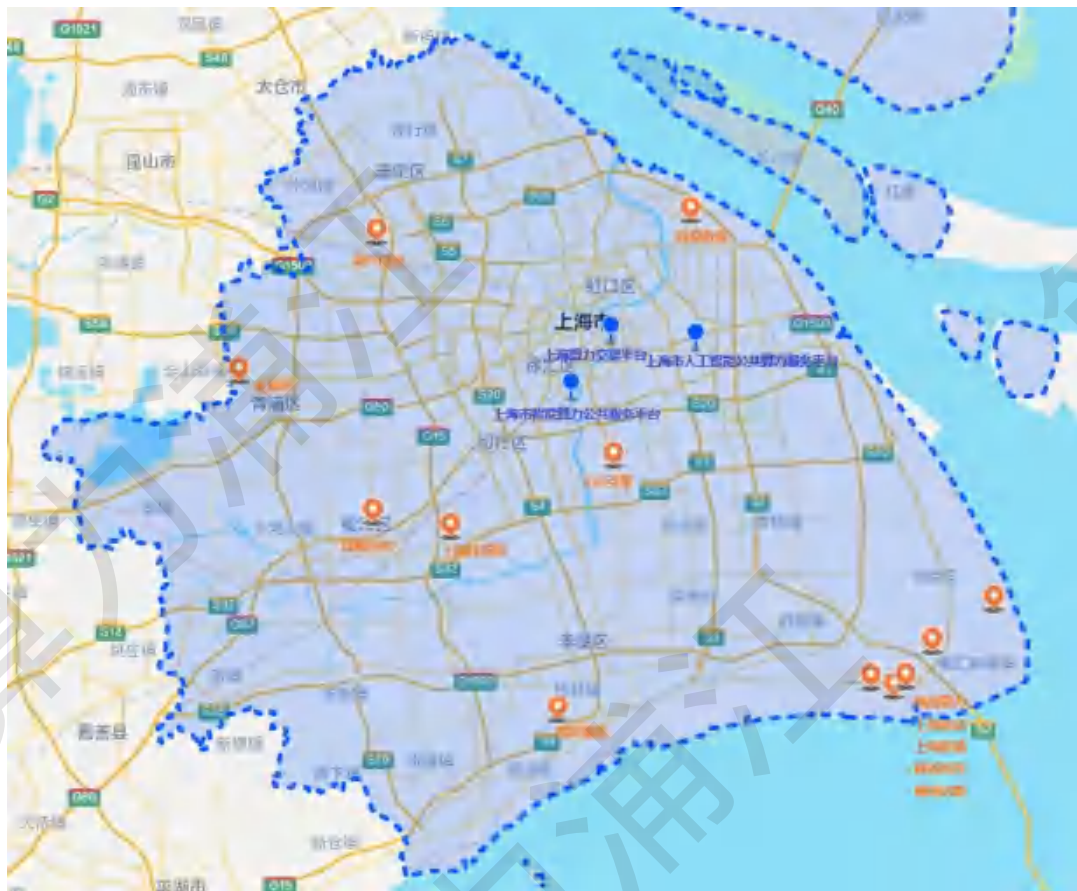


图 2-9: 全市智算中心空间布局图

来源：上海市通信管理局

### (五) 算力产业构建更加丰富

近年来，上海始终将落实网络强国、数字中国、质量强国战略放在首要位置，聚焦推动高质量发展，制定了一系列算力基础设施布局及产业发展的相关政策规划，依托行业组织加快构建算力产业生态圈，打造具有影响力的算力产业高地。

制定系列算力高质量发展政策，优化产业顶层设计。2020年，上海市人民政府印发《上海市推进新型基础设施建设行动方案（2020-2022年）》提出要“打造亚太一流的超大规模人工智能计算与赋能平台”，有力支撑上海人工智能产业发展。

2021-2022 年期间，上海连续发布《上海市全面推进城市数字化转型“十四五”规划》《上海市数字经济发展“十四五”规划》等文件，均高度重视算力高质量发展，提出推动建设绿色数据中心，加快打造全国一体化大数据中心体系的上海枢纽节点。2022 年 6 月上海市通信管理局正式印发《新型数据中心“算力浦江”行动计划（2022-2024 年）》，作为国内首个聚焦算力主题进行体系化布局的省市级政策文件，将上海算力基础设施建设推向又一个新阶段。2023 年 4 月，上海市经济信息化委印发《上海市推进算力资源统一调度指导意见》，提出深化推进本市算力资源高效、开放、有序使用；同年 11 月连续印发《上海市推动人工智能大模型创新发展若干措施（2023-2025 年）》，提出打造市级智能算力统筹调度平台，构建规模化先进算力调度和供给能力。2024 年 3 月，上海市通信管理局联合 11 个部门联合印发《上海市智能算力基础设施高质量发展“算力浦江”智算行动实施方案（2024-2025 年）》，提出推动上海算力基础设施向智能算力基础设施演进，对绿色算力关键指标提出了新的要求。



表 4：近年上海市算力基础设施建设主要政策

时间	出台单位	发布政策	政策概要
2024年 3月	上海市通信管理局等十一部门	《上海市智能算力基础设施高质量发展“算力浦江”智算行动实施方案（2024-2025年）》	加快推进临港、青浦、松江、金山、张江等重点区域大型智算中心建设。根据上海市各特色产业园区发展需要，建设以算力多样化、网络全光化、应用容器化为主要特征的边缘智算中心，提升多元化异构算力算效水平。
2023年 10月	上海市经济和信息化委员会等六部门	《上海市推动人工智能大模型创新发展若干措施（2023-2025年）》	打造市级智能算力统筹调度平台，构建规模化先进算力调度和供给能力。对符合条件的智算中心，在能耗指标等方面予以绿色通道优先支持；加快临港、金山、松江等重点智算产业集聚区建设。
2023年 10月	上海市经济和信息化委员会、上海市商务委员会	《推动区块链、大模型技术赋能生产性互联网服务平台发展实施方案》	实施智能算力加速计划，加快大模型算力基础设施建设。建设大模型测试评估与协同创新中心，推动大模型产业健康发展。
2023年 9月	上海市人民政府	《上海市进一步推进新型基础设施建设行动方案（2023-2026年）》	打造超大规模自主可控智能算力基础设施。支持有关创新平台牵头建设自主可控智能算力重大科技基础设施，打造基于自主可控通用人工智能芯片、自主可控光电混合计算芯片、自主可控训练框架、自主可控全光交换网络的超大规模智能算力集群。
2023年 4月	上海市经济和信息化委员会	《上海市推进算力资源统一调度指导意见》	开展全市算力基础设施及算力资源输出能力排摸，形成算力清单。基于算力资源底数，推动头部企业接入市人工智能公共算力服务平台，构建一体化算力调度服务体系和平台基础框架，实现算力资源统一编排。
2022年 6月	上海市通信管理局	《新型数据中心“算力浦江”行动计划（2022-2024年）》	探索围绕算力布局、算力网络、算力赋能、算力生态、绿色算力、算力安全等六大重点任务推动上海算力基础设施高质量发展。
2022年 6月	上海市人民政府办公厅	《上海市数字经济发展“十四五”规划》	推动建设绿色数据中心，强化算力统筹和智能调度，提升数据中心跨网络、跨地域数据交互能力，实现数据中心自动化能效调优，提升数据中心能效密度。
2021年 10月	上海市人民政府办公厅	《上海市全面推进城市数字化转型“十四五”规划》	推动数据中心存算一体集约化布局，加快打造全国一体化大数据中心体系的上海枢纽节点。

来源：上海各政府机构官网



**发挥行业组织优势，共建算力产业良好生态。**上海积极依托上海市通信学会、上海市算力网络协会、上海市集成电路行业协会、上海市数商协会等行业组织，汇聚了成百上千家算力产业链上下游重点企业、专业服务机构，共同打造算力产业生态共荣圈。其中，上海市通信学会“算力浦江”专委会已构建由基础电信、数据中心、云服务、服务器、算力芯片，算力存储、绿色节能等各领域 100 余家企业组成的算力基础设施产业生态圈，尤其是依托上海算力芯片产业的领先优势，专委会中算力芯片企业占比已超过 17%以上；同时，依托生态圈建立了算力基础设施领域标准体系，启动立项了 10 个算力网络团体标准，其中《零碳数据中心设计与建设技术规范》《数据中心浸没式液冷系统单相冷却液技术指标和测试方法》《智算中心能力验收规范》《智算中心算力性能评估测试方法》等 4 个标准已正式发布。融合应用方面，专委会组织算力应用场景创新征集活动，探索算力+医疗、算力+交通等创新场景，遴选了 100 个算力应用新场景，并于 2023 年底发布算赋百业百景图。

**全面加强算力中心运行安全管理。**上海已构建起了从算力服务业务准入和合规监管、算力中心建设申请和验收、算力中心安全运行、算力网络性能监测、网络和数据安全管理等方面一整套全生命周期的监管体系。围绕数据中心网络安全和数据安全，开展“磐石行动”网络安全攻防演练和“浦江护航”数据安全专项行动。安全生产方面通过通信建设工程质量监督管理流程开展实

地核查，全方位保障算力中心稳定运行。同时，在工业和信息化部统筹指导下，上海开展算力网络性能监测和算力安全运行监测工作，每季度发布上海市算力网络监测报告。

### 三、上海市算力创新发展先锋成果

#### （一）算力设备，引领前沿创新突破

近年来，上海在智能算力领域发展势头强劲，特别是大算力 AI 芯片、数据存储、服务器，以及 5G-A、万兆光网等高速网络设备取得显著成就，不仅提升了上海的科技创新能力，更为全球数字经济的发展注入了强劲动力。

在大算力 AI 芯片方面，上海依托雄厚的科研实力和完善的产业链，已成功量产多款高性能智能芯片，广泛应用于自动驾驶、智能制造等领域。这些芯片不仅提升了相关领域的计算效率，更为 AI 技术的普及和深化应用提供了有力支撑。

数据存储与服务器方面，上海拥有众多全球领先的科技企业，他们构建了高效、稳定的数据存储和服务器系统。这些系统不仅可以支持海量数据的存储和处理，还具备高度的安全性和可扩展性，为上海乃至全球的数字化转型提供了坚实保障。

在高速网络设备领域，上海已经实现了 5G 网络的深度覆盖，为智慧城市、自动驾驶等前沿应用提供了强大的网络支持。同时，上海还在网络设备方面取得重要突破，研发出多款高性能、低延时的网络设备，为 5G 技术的普及和深化应用奠定了坚实基础。预计到“十四五”期末，上海将成为 5G 网络部署和融合应用的

全球标杆城市。

## 典型先锋案例一：

### 上海忆芯实业有限公司

#### 高性能企业级 SSD 存算一体主控芯片及方案的研发与应用

##### 企业简介

忆芯科技，国内领先的高端 PCIe SSD 主控芯片和成品盘供应商，业务方向覆盖消费级、工业级和企业级，为各行业的信息化发展提供高质量芯片级底层保障。忆芯科技是国家专精特新“小巨人”企业、国家知识产权优势企业，拥有从底层算法到芯片设计，再到解决方案等经验丰富的技术团队。

##### 案例概述

该项目首创性的将高性能存储控制器芯片架构与 AI 加速架构、数据库分析加速架构、国密加解密安全模块进行异构融合，通过领域定制化设计，实现了高性能存算协同加速的一体化芯片。搭载该存算一体芯片的企业级 SSD 解决方案，在电力、水利、石化等关键基础设施行业成功推广，实现了产业化示范应用。



## 解决方案

### （一）主要技术创新点：

1、自主研发存算一体技术，首次在存储芯片中集成神经网络处理单元，提升 AI 大数据应用算力能效比达 12TOPS/W，填补国内部分技术空白，属于国内首创技术。

2、提出存储器高可靠和高安全技术，解决了闪存颗粒擦写次数少和数据泄密风险等国际难题，提高了固态硬盘的使用寿命，并加强了敏感数据的安全性，实现了高可靠国产替代，技术达到了国际领先水平。

3、创造性地提出高性能软件定义存储控制器架构，自研高性能、灵活的企业级 SSD 控制技术，解决了缺少中国自主可控的高性能企业级 SSD 方案的难题，顺序写性能超过国外方案 40%，随机读性能超过国外方案 20%，随机写性能超过同类产品 84%，并且写延迟降低到 7  $\mu$ s，比国外同类产品低 50%。

### （二）关键成效：

1、吞吐带宽和延迟：多核同构 CPU 同时控制存内计算的 AI 加速 NPU，矩阵加速计算单元，数值统计加速计算单元，高速 NAND 存储控制单元；片上一致性总线保障数据吞吐带宽从存储侧到数据计算侧延迟小于 1.5  $\mu$ s，总吞吐带宽大于 14GB/s，远超目前 PCIe 固态硬盘到主机的最大带宽和读写延迟能力。

2、数据分析能力：存算一体架构，硬件设计 8TOPS，规模电路算力规格实际端到端推理应用效果超过国外 32TOPS 平台，AI



存算核心算效比超 12TOPs/W。计算索引加速能力，键值数据库并发效率提升 10-15 倍，主机内存占用减小至传统数据库服务 1/6。统计加速能力，最高提升时序类数据库写入速度 6 倍，读取速度 30 倍，保障时序数据库聚合计算能力综合提升 0.5-3.7 倍。向量数据库加速支持，512 维向量相似度计算吞吐达到 10 万次/秒。

3、数据安全可信能力：自研硬件加解密功能模块结合可计算存储 CPU 服务调用管理，可在整体主机 CPU 启动操作系统前完成所有操作指令的记录和态势感知，并通过端到端的验签及加解密流程完成高安全的存算数据交互。

### （三）应用价值：

该项目实现了高性能 PCIe4.0 接口主控芯片和存储方案，解决了国家缺少有竞争力的自主研发存储主控芯片和兼容性强，安全可靠的国产存储方案的难题。此外，该项目突破传统计算体系结构，实现了更高效的存算一体架构。

### 成果效益

该项目解决方案主要面向金融、能源、通信、AI 及互联网等领域，主要服务于工业级市场和企业级市场，客户包括世界/中国 500 强、央国企、党政机关等，为数据中心、个人电脑及 OEM 等多种应用场景提供低能耗、高算力、高可靠、高安全的 AI 计算存储一体化设备。该项目产品及解决方案在电力、能源、国产信创行业完成了产业化示范应用落地。项目自推出后累计销售收入额超 10 亿。

该项目通过自主研发的存储控制技术和方案，有效解决了存储

设备在产品寿命、稳定性、数据保护等方面的技术瓶颈，实现了存储技术的国产化替代，保障了国家信息安全，推动了产业技术进步。存算一体技术的创新为 AI 大数据应用提供了强大的算力支持，对科技进步和社会发展产生了积极影响。

## 典型先锋案例二：

### 优刻得（上海）数据科技有限公司

#### “丰收一号” 国产 GPU 千卡集群

#### 企业简介

UCloud 优刻得是中立、安全的云计算服务平台。自主研发 IaaS、PaaS、大数据流通平台、AI 服务平台，推出公有云、私有云、混合云、专有云等全线云产品，为政府、AI 大模型、工业互联网、运营商、教育、医疗、零售、金融、互联网等各行业用户，提供全面的数字化转型升级服务。

2020 年 1 月，UCloud 优刻得正式登陆科创板，成为中国第一家公有云科创板上市公司，同时成为了中国 A 股市场首家“同股不同权”的上市企业，开创了中国 A 股资本市场及公司治理的先河。

UCloud 优刻得在全球设有 31 个可用区，遍及国内、东南亚、欧洲、北美、南美、非洲等 25 个地域，结合内蒙古乌兰察布、上海青浦两大自建数据中心，构建全球化数字信息基础设施。目前已为全球超过 5 万家企业级用户提供云服务。

#### 案例概述

该项目旨在响应国家新一代人工智能发展战略，聚焦国产 GPU 千卡集群搭建，进一步加强国产化高性能算力基础设施建设，优刻得科技股份有限公司联合上海优算丰信息技术有限公司、沐曦集成电路（上海）有限公司，基于上海沐曦第一代旗舰产品曦云 C500 系列 GPU 芯片的 GPU 服务器，于 2023 年 12 月开始投资建设了首

个国产 GPU 千卡集群“丰收一号”。“丰收一号”集群搭载了 2560 颗上海沐曦第一代旗舰产品曦云 C500 系列 GPU 芯片的 GPU 服务器。该项目建设地点为优刻得（上海）数据科技有限公司位于上海市青浦区南厝江路 130 号的数据中心。

“丰收一号”集群设计总算力规模约 640P，计划部署合计 320 台 GPU 服务器以及配套存储服务器、管理服务器和网络设备等。目前“丰收一号”集群一期 160 台设备于 2024 年 4 月验收上线，二期 160 台于 2024 年 10 月验收上线。目前“丰收一号”集群已完成上海商汤智能科技有限公司、中国信息通信研究院、北京智源人工智能研究院 FastMOE 等标杆模型的训练和测试工作、香港科技大学等机构的若干模型的训练测试（包括但不限于 LLama-7B/70B、Bert-110M、StableDiffusion 1.5、InternLM-7B 等模型）。该项目建成后将持续为推动上海、长三角乃至全国算力基础设施迭代升级提供引擎动力，加速培育新质生产力。

## 解决方案

优刻得（上海）数据科技有限公司联合上海优算丰信息技术有限公司和沐曦集成电路（上海）有限公司，以沐曦集成电路（上海）有限公司的曦云 GPU 芯片为核心，构建了一个 2560 张卡的智算集群，该集群位于优刻得青浦算力中心。

### 1. 技术体系

1) GPU 服务器：项目使用了 2560 颗上海沐曦第一代旗舰产品曦云 C500 系列 GPU 芯片，构建了 320 台高性能 GPU 服务器。基

于强大的计算能力和高效的能耗比，能够支持大规模模型训练和推理任务。

2) 软件栈技术：沐曦芯片采用全自研的 MACA 软件栈，从底层驱动到上层开源框架均采用全自研技术，兼容主流 GPU 软件生态，包括英伟达的 CUDA 生态，实现从英伟达生态到国产 GPU 生态的零成本迁移。

3) 存储系统：项目配备了高性能的存储系统，支持高速数据读写和大容量数据存储。存储系统采用了分布式存储架构，确保数据的高可用性和可靠性。

4) 网络架构：项目采用了多种组网方式，包括 InfiniBand (IB)、RDMA over Converged Ethernet (RoCE) 和 TCP/IP 协议，确保数据传输的低延迟和高吞吐量。网络架构的设计考虑了安全性和扩展性，能够支持大规模集群的高效运行。

5) 管理平台：项目提供了一套完整的管理平台，支持集群统一调度和管理。管理平台包括资源调度模块、任务管理模块、监控模块和安全管理模块，能够实现资源的动态分配和任务高效执行。

6) 应用适配：该项目在上层的大模型/AI 应用层面进行了完善的应用适配，包括 LLaMa 系列、GPT、GLM 等主流模型的适配验证，赋能政务、金融等行业的大模型应用落地。

## 2. 实施方案

### 1) 项目规划与建设

2023 年 12 月，项目正式启动，2024 年 3 月完成所有设备的安



装和调试。2024年3月至4月，项目团队进行了系统的集成和测试，确保各个组件之间的兼容性和稳定性。2024年4月，一期160台设备顺利验收上线。2024年10月，二期160台设备顺利验收上线，整个项目正式投入运行。

## 2) 运维与管理

项目成立了专门的运维团队，负责集群的日常运维和管理。运维团队成员具备丰富的云计算和高性能计算经验，能够及时处理各种技术问题。在监控与维护方面，项目采用了先进的监控系统，实时监测集群的运行状态，包括CPU、GPU、内存、网络和存储等关键指标。运维团队通过监控系统及时发现和处理异常情况，确保集群的稳定运行。

## 3. 关键成效

### 1) 生态兼容优势:

项目采用沐曦自研的GPU软件栈MXMACA®，全面兼容CUDA，支持数千个开源应用，确保用户可以无缝迁移现有的CUDA应用。项目原生支持PyTorch框架，覆盖了全部2500多个算子，确保用户可以高效进行模型训练和推理。支持所有主流开发框架，包括PyTorch、TensorFlow、MXNet、JAX、Caffe等国际主流人工智能开发框架，以及PaddlePaddle、OneFlow等国内自研人工智能开发框架，全面赋能各领域的AI应用。

### 2) 集群扩展与稳定性优势:

项目支持8卡全互联，互联带宽高达896GB/s，确保数据传输

的低延迟和高吞吐量。实现了商业千卡集群的稳定无故障运行超过14天，确保了集群的高可靠性。项目支持从128卡到1024卡的线性扩展，扩展线性度超过94%，确保了集群的高效扩展能力。

### 3) 多样化组网方式接入:

项目具备强大的网络接入能力，支持多种组网方式，能够满足不同业务场景的需求。无论是需要低延迟和高吞吐量的数据传输，还是对网络安全性的严格要求，项目都能提供相应的网络支持。

### 4) 高性能存储系统:

项目接入了优刻得自研高性能存储产品UPFS，显著提升了存储吞吐性能。实际测试结果显示，存储系统的性能达到了预期目标，Checkpoint速度较传统存储提升了近10倍。

## 4. 应用成效

针对高校及科研机构，项目已经成功支持了上海商汤智能科技有限公司、中国信息通信研究院、北京智源人工智能研究院FastMOE等机构的模型训练和测试工作。这些机构通过“丰收一号”集群，大幅提高了科研效率，加速了科研成果的产出。项目为多所高校提供了高性能计算资源，支持了人工智能、机器学习等前沿学科的教学和科研活动。高校师生通过“丰收一号”集群，能够进行大规模的实验和研究，提升了学术水平和创新能力。

针对政府方，项目为政府部门提供了高性能计算资源，支持了城市管理、公共安全、环境保护等领域的数据处理和分析。政府通过“丰收一号”集群，能够更好地制定政策和规划，提升公共服务

水平

针对大模型研发创新型企业，项目为企业提供了强大的计算资源，支持了产品开发、数据分析和业务优化。企业能够高效率、低成本的开发，快速响应市场变化，提升竞争力。

### 成果效益

通过项目为上海、长三角乃至全国人工智能产业的转型升级提供引擎动力，加速培育新质生产力。项目的实施将帮助模型厂商克服算力瓶颈，加速大模型技术的创新应用落地。针对传统算力配置与资源管理痛点，项目旨在革新算力资源管理，提升行业算力效率。推动智能算力基础设施云化，解决资源碎片化，提高运维效率，优化算力资源使用，降低成本。项目满足市场需求，促进产业链上下游协同，推动算力产业高质量发展。

该项目建成后对行业发展具有重要意义。一是推动国产 AI 芯片的应用，加速国产技术的成熟与迭代。二是平台将降低企业使用 AI 技术的门槛，促进中小企业的数字化转型。三是通过提供高性能的 AI 算力服务，项目将支持更多创新 AI 应用的孵化和成长。

## (二) 算力中心，重点项目群星闪耀

上海在超大规模智算中心、边缘智算节点等方面均取得了显著进展，不仅推动了上海数字经济的高质量发展，也为全国乃至全球的智算中心建设提供了宝贵经验。

上海在超大规模智算中心建设方面走在全国前列，如商汤临

港人工智能计算中心获得全国首个 5A 级智算中心算力性能认证，成为智算中心建设的“样板间”，远超规划建设任务，可支撑多个千亿参数超大模型同时训练。

此外，上海还积极推动边缘智算节点建设，以满足日益增长的实时数据处理需求。边缘智算节点的部署有助于降低数据传输延迟，提高数据处理效率。这些节点通常部署在靠近数据源或用户终端的位置，以实现更快速的数据处理和分析。如以漕河泾开发区为代表的领先园区开始探索园区级的定向算力服务供给，目前漕河泾开发区已具备 50 PFlops（FP16）规模的园区智算算力专属服务，未来规划扩展至 200 PFlops（FP16）规模，为园区自身及园区企业提供普惠的 AI 算力供给服务，加速园区/企业的数智化转型升级。

### 典型先锋案例三：

#### 上海科技网络通信有限公司

#### 城市智算底座 - 仪电智算中心（松江）

##### 企业简介

上海科技网是上海市属国资下唯一的中立智算中心运营商。短期内实现从 IDC 运营商转型成为优质智算中心服务商的跨越，为客户提供“高可用智算中心，高性能算力服务”。公司已具有超过 18 年 IDC 运营管理、13 年云服务及 23 年城域网经验。建设和运营的仪电智算中心（松江）目前已获得国标 A 级机房认证和等保三级安全备案，符合金融业信息系统、满足 NFTC 金融数据中心认证要求，并斩获长三角绿色低碳示范数据中心、中国 IDC 产业金融信创数据中心、长三角高质量算力解决方案提供商及 2023 年度中国 IDC 产业智算中心优质服务商大奖等多项荣誉。

##### 案例概述



图 1：仪电智算中心（松江）园区效果图

仪电智算中心（松江）一期，在 2019 年底取得上海市经信委首批数据中心用能指标后，于 2020 年正式启动项目建设并于 2022



年投入运营，同时，除了基础设施建设，仪电智算中心（松江）项目已具备健全完善的智算中心运维能力，全面通过 CQC8302 运维体系审核。目前已有包括人工智能大模型公司、智算科技企业及金融行业头部企业等多家大型客户入驻，成为其稳定、高可用算力和云计算主生产环境。智算中心斩获长三角绿色低碳示范数据中心、中国 IDC 产业金融信创数据中心、长三角高质量算力解决方案提供商及 2023 年度中国 IDC 产业智算中心优质服务商大奖等多项荣誉。



图 2（左）：2023 年度中国 IDC 产业智算中心优质服务商大奖

图 3（右）：长三角高质量算力解决方案提供商

伴随智算需求的爆炸式增长，上海科技网积极筹建仪电智算中心（松江）二期，为加快形成全国一体化算力体系、培育算力产业生态贡献力量。2023 年 7 月，该项目作为上海市 2023 年重大功能性事项导入松江新城。2024 年 2 月，入选上海市 2024 年重大工程清单。项目投产后，将成为全国算力规模领先的智算中心，进一步推动智算塑造新质生产力，支撑“两个强国”战略发展要求，推动数字经济做强做优做大。项目进一步体现绿色设计，应用液冷、储能等新技术，持续优化算力设施电能利用效率、水资源利用效率、碳利用效率，提升算力碳效水平。

### 案例详情

“十三五”以来，随着我国“互联网+”、大数据、数字经济等国家战略的深入推进以及5G网络、人工智能、工业互联网等新型基础设施的加速建设，上海科技网提前规划和投资建设仪电智算中心（松江）一期，在2019年底以综合评分第一的排名取得了上海市经信委首批数据中心用能指标，并于2020年启动项目建设，于2022年成功建成。随着仪电智算中心（松江）投入运营，上海科技网成为提供“高可用智算中心、高性能算力服务”的优质智算中心服务商。目前已有包括人工智能大模型公司、智算科技企业及金融行业头部企业等多家大型客户入驻，成为其稳定、高可用算力和云计算主生产环境。



图4：仪电智算中心（松江）一期实景

2022年12月，ChatGPT横空出世，人工智能再次进化。模型所使用的数据量和参数呈指数级增长，带来智能算力需求爆炸式增加。同时，2024年的《政府工作报告》也提出，要适度超前建设数字基础设施，加快形成全国一体化算力体系，培育算力产业生态。在此背景下，上海科技网积极筹建仪电智算中心（二期）项目。2023

年7月，该项目作为上海市2023重大功能性事项导入松江新城。2024年2月，该项目入选上海市2024年重大工程清单。二期项目全面投产后，将建设成为全国算力规模领先的智算中心，将进一步推动智算塑造新质生产力，支撑“两个强国”战略发展要求，推动数字经济做强做优做大。

仪电智算中心（松江）一期和二期项目是驱动数字经济、赋能科技创新进步的新型智算中心，实现了“智”的飞跃。

相较于传统数据中心，智算中心的主要不同有：一是架构不同，智算中心采用虚拟化技术和云服务提供计算和存储资源，意味着智算中心可以更加灵活高效响应客户需求，并且能够做到快速部署新的应用和服务；二是性能更优，智算中心使用无损高速网络和分布式存储技术，可以更好满足客户对高性能计算和大数据处理的需求；三是单机柜功率，智算中心搭载了GPU，单机柜功率显著提升，能耗也会随之大幅增加，对智算中心的运营带来比较大的成本挑战。

仪电智算中心（松江）项目从设计之初便以智能敏捷、安全可靠、绿色低碳等为理念，持续提升算力综合供给能力，着力强化运力高效承载，不断完善存力灵活保障。通过领先的体系架构设计，通过多种“黑科技”加持，真正实现绿色低碳发展。部署了华东地区最大的间接蒸发冷系统、ECC智能控制等先进技术，降低机房制冷功耗，实现全年自然冷却时间达到40%，单机房模块PUE达到1.1，整体PUE小于1.29。能够更好地实现可持续、绿色低碳运营。

同时，仪电智算中心（二期）项目将进一步体现绿色设计，加快高能效、低碳排的算网存设备部署，推动软硬件协同联动节能。根据规划，二期项目将应用液冷、储能等新技术，进一步优化算力设施电能利用效率、水资源利用效率、碳利用效率，提升算力碳效水平，整体 PUE 将小于 1.2。

### 成果效益

仪电智算中心（松江）是作为上海市国资体系超大规模智算中心，为上海夯实数字新基建、培育算力产业集群、加快智慧城市建设提供重要支撑！

除安全可靠的基础设施建设及稳定高效的运维能力外，其节能降耗措施和效果对同类型智算中心也有一定示范效应，在全行业中处于较好水平。在上海市通信学会、上海市通信学会新型数据中心专家委员会开展的《“算力浦江”-2022 新型数据中心创新发展案例集》征集评选中，松江大数据计算中心（暨仪电智算中心）入选“上海市新型数据中心创新发展案例 -- 高能效”名单，取得了较好的社会效应。

作为优质智算中心服务商，上海科技网在未来将进一步深耕高等级智算中心集聚园区，同时推进人工智能、数据中心、算力生态等先进产业集群一体化发展和协同创新，并携手合作伙伴共建算力生态圈，为上海城市数字化转型发挥更大作用。

## 典型先锋案例四：

### 上海有孚网络股份有限公司

#### 临港有孚绿色智算中心

#### 企业简介

上海有孚网络股份有限公司是领先的新型云计算数据中心服务提供商。公司创立于 2001 年，总部位于上海，在北京、广州、深圳等地均设立有分支机构。我们致力于通过提供绿色多元的算力支撑，让企业数字化更简单！

公司以新型算力中心服务、异构算力服务及专有云服务为主营业务，在北上广深超一线城市拥有多个高标准云计算数据中心及算力集群，并建立了以智算平台为载体的跨区域云网协同、算网融合的算力服务能力，可为用户提供含算力资源服务、基础云服务、智算等新一代计算平台及异构算力资源。

公司已完成高新技术企业、上海市专精特新企业、上海市企业技术中心、上海市科技小巨人企业等认定，并获得由工信部颁发的“国家新型数据中心典型案例”、由工信部和中国信通院颁发的“数据中心低碳等级”认证、中国信通院评审的“中国算力中心服务商十强”等权威认证。

#### 案例概述

随着国家“双碳”战略的深入实施与《算力基础设施高质量发展行动计划》的推进落实，算力中心作为数字经济的关键支柱，正逐步向绿色低碳转型。在此背景下，临港有孚绿色智算中心积极响



应国家政策，引领算力行业绿色发展趋势。

依托规模化投入和公司绿色理念的坚定践行，临港有孚绿色智算中心主动融入光伏、储能等一系列绿色能源解决方案和技术革新，并广泛应用智能照明、能效优化平台等多种节能技术装备优化能源使用效率，为智算中心的绿色运行奠定了坚实基础。

在技术创新方面，临港有孚绿色智算中心取得了突破性成就。有孚率先采用特斯拉 Megapack 储能系统，打造上海首个规模化智算中心场景储能示范项目，并成为中国首个商业化示范应用项目；同时，依托基于人工智能和大数据分析的能效优化平台，实时监测和分析算力中心的能耗情况，通过 AI 节能算法，提供优化建议，实现 PUE 值的持续降低，推动算力中心的绿色可持续发展。

临港有孚绿色智算中心的建设，不仅提升了智算中心的能效，降低了运营成本，还为社会带来了广泛的生态效益。此外，其成功经验和技術路径也为绿色产业链的发展提供了有力支持，进一步赋能了数字经济的新质生产力，推动了算力基础设施的迭代升级，为实现国家碳达峰和碳中和目标贡献了积极力量。

## 解决方案

### （一）技术体系与实施方案：

临港有孚绿色智算中心技术体系的核心在于“绿色+算力”的深度融合，通过采用先进的绿色技术与能源管理系统，打造高效、低碳、可持续的算力基础设施。具体技术体系包括以下几个方面：

#### 1、储能技术前沿应用：临港有孚绿色智算中心率先引入了特

斯拉 Megapack 储能系统，通过先进的智能算法优化充放电策略，减轻电网负荷，实现削峰填谷，显著提升能源利用效率。初期部署规模达 32MWh，预计能有效降低智算中心对电网的依赖，并在用电高峰期减轻电网负荷，展现了其在能源管理方面的创新优势。

**2、光伏发电普及实践：**临港有孚绿色智算中心充分利用临港地区丰富的太阳能资源，结合园区实际情况，一期实际铺设面积约 8000 平方米，规模达 88 万 kWh。光伏的部署不仅减少了对传统能源的依赖，还显著降低了碳排放，体现了智算中心在绿色能源应用上的积极探索。

**3、智能照明系统革新：**临港有孚绿色智算中心采用了 LED 光源与智能控制系统相结合的智能照明方案。通过实时感知环境人员变化进行开关灯设置，结合智能面板等手动控制设备，可实现自动控制、现场控制和值班室远程控制相结合，有效避免了不必要的电能浪费，照明系统节能率可达 40% 以上，展示了智算中心在节能降耗方面多模式探索。

**4、雨水回收系统构建：**为进一步提升水资源利用效率，智算中心内建立了雨水回收再利用系统，这一系统不仅降低了对市政供水的依赖，还有助于维护园区绿色生态环境，展现了智算中心在环保方面的综合努力。

**5、智算中心能效优化平台：**基于人工智能和大数据分析的能效优化平台，对智算中心能耗信息进行实时监测和分析，通过数据接入、分析、挖掘，平台能够精准识别能耗瓶颈并提供优化建议，

助力 PUE 值的持续降低。

## （二）关键成效

**1、能效提升：**依托储能系统、光伏发电及能效优化平台等绿色技术结合，临港有孚绿色智算中心的 PUE 值显著降低，相比传统数据中心提升能效显著。

**2、水资源利用率优化：**通过雨水回收系统等节水措施，中心有效降低了水资源消耗，WUE 表现优异，展现了绿色智算中心在节水环保方面的贡献。

**3、低碳运营，减排显著：**光伏发电与储能系统的结合，不仅大幅提高了绿电比例，还显著减少了二氧化碳排放，为碳达峰碳中和目标贡献了重要力量。

**4、项目智能化管理：**通过智能管理平台提供全面的数据分析和决策支持，帮助运营管理团队优化资源配置和运维流程，运用大数据分析技术解决能耗瓶颈，挖掘更多优化机会方案，为智算中心的长远发展注入源源不断的绿色动力。

**5、运营成本有效降低：**可量化指标中，绿色设备与系统的接入，年节约电费近千万元，显著降低项目的运营成本，实现了绿色运营和经济运营。

## （三）应用价值

通过上述绿色低碳项目的实施，临港有孚绿色智算中心在能效提升和环境保护方面取得了显著成效。项目不仅降低了 PUE 和 WUE 值，还显著减少了碳排放和水资源消耗，实现了经济效益和

环境效益的双赢。

临港有孚绿色智算中心的绿色低碳项目积极响应了国家节能减排和绿色算力政策要求，为智算行业的绿色可持续发展树立了标杆。同时，通过引入先进节能技术和智能管理系统，智算中心的能效显著提升，运营成本得到有效控制，为客户提供了更经济、更环保的算力服务。此外，随着市场对绿色算力的需求日益增长，临港有孚绿色智算中心的绿色低碳升级项目满足了客户对环保、高效的算力服务需求，增强了市场竞争力。项目的成功经验也为智算行业的绿色转型提供了可借鉴的模式和路径，有助于推动整个行业向低碳、环保、可持续的方向发展。

### 成果效益

**1、从市场经济效益来看，**有孚智算中心的绿色升级项目，通过提升能效和充分利用可再生能源，显著降低了电力消耗，为企业用户提供了成本效益更高的算力服务。此举不仅强化了企业的品牌形象，符合了环境、社会和治理（ESG）标准，而且吸引了对绿色算力有高度认同的用户群体。

**2、从绿色产业效益来看，**项目的实施也进一步推动了绿色全产业链的发展，加速了技术的创新和迭代，增加了光伏制造、储能开发和智控系统等相关领域的就业机会。同时，项目孵化了与绿色智算中心相关的市场新机会，为绿色经济的增长提供了新动力，为经济社会的绿色转型和高质量发展做出了积极贡献。

**3、从环境效益来看，**临港有孚绿色智算中心通过光伏、储能

技术减少对传统能源的依赖，有效降低碳足迹；同时推动能源结构优化，显著提升绿电比例，为环保事业和能源可持续发展贡献显著力量。



## 典型先锋案例五：

### 中国移动通信集团上海有限公司

中国移动智算中心（上海）--绿色智算助力发展新质生产力

#### 企业简介

中国移动通信集团上海有限公司是中国移动（香港）有限公司下属全资子公司，经营上海地区的数字蜂窝移动电话、IP 电话、互联网接入业务及相关的信息服务、技术开发、技术服务等业务。主要从事移动通信、IP 电话和互联网等网络的设计、投资和建设；移动通信、IP 电话和互联网等设施的安装、工程施工和维修，经营与移动通信、IP 电话和互联网业务相关的系统集成、漫游结算清算、技术开发、技术服务、广告业务、设备销售等；出售、出租移动电话终端设备、IP 电话设备、互联网设备及其零部件，并提供售后服务。

#### 案例概述

中国移动智算中心（上海），位于上海移动临港数据中心园区，作为华东地区最优质的互联网数据中心资源之一，也是中国移动七大一类 IDC 中心之一。立足“一个定位”（立足华东、服务全国、辐射海外），把握“两大主题”（创新与服务），贯彻“三大理念”（高科技、信息化、绿色环保），完善“三大功能”（通信生产基地、信息服务基地、海缆登陆站），凭借强大的平台处理能力、存储能力、云计算能力和移动通信服务能力，有效集聚、整合上下游产业链，助力上海五大中心建设再上台阶。

中国移动智算中心（上海），积极响应国家“碳中和、碳达峰”政策，始终遵循“安全、高效、绿色、智能”的原则高标准建设，秉承最大能源效率，最小环境影响的宗旨，围绕极简、低碳的核心理念，贯穿项目设计、施工、运维各个阶段。采用优化供电架构、AI智慧调节、绿色智能运营等创新方式，有效实现了年均 PUE1.28。

中国移动智算中心（上海）不仅是数字经济高速发展的“助推器”，也是传统产业数字化转型的“新引擎”。中国移动正全力推进全国一体化算力网络之一，长三角国家枢纽节点建设，加快推动数字经济和实体经济深度融合，以网络化之正拓数字化、智能化之新，助力“数字中国”现代化建设，为中国经济的高质量可持续发展提供坚实基础。

### 解决方案

目前，上海移动 IDC 机架能力超 6 万架，已形成超 100pflops 的算力资源能力。其中作为两大核心节点之一，中国移动智算中心（上海）致力打造全新的绿色数据中心、推动算力和网络能力一体化拉通，引领算网融合创新，激发澎湃数智能量。建设阶段创新提出“平台化”设计理念，采用模块化设计，为大数据、AI、云计算等各个模块的发展预留空间，并将“数字孪生”技术赋能传统基建工程管理，达到数据中心智能建造 5A 级水准，使其可以兼顾新型数据中心近期和远期的需求，以绿色先进的数据中心资源满足城市算力需求。

#### 1、绿色新型基础设施：

**PUE 智能调优技术：**利用大数据和 AI，在推动大型数据中心的制冷走向“智冷”，降低 PUE 的同时，可实现数据中心能耗的进一步优化。算法充分挖掘 L1 系统内部匹配最优运行以及 L1 和 L2 的联动调节节能空间，实现 PUE 实时智能调优，随运行生命周期持续优化。系统级 AI 能效寻优方案配合间接蒸发冷却技术可实现 PUE 降低 3%-5%，配合水冷冷冻水方案可实现 PUE 降低 8%-15%。

**高效电力模块技术：**采用一体化集成方案，具备一体化集成、安全可靠、节省机房占地面积和能源、安装省时、省力、省心、架构兼容、部署快速灵活和一体化集中监控等特点，电力模块方案，内置高效模块化 UPS，采用铜排预制缩短供电链路，将原有的数据中心供配电系统链路供电效率从 94.5% 提升到 97.5%，方案显著降低能源消耗。工厂预制可以提前在工厂进行测试，保证产品质量，同时整体运输到现场，可以大幅缩短交付时间。

**采用磁悬浮冷水主机：**通过磁悬浮冷水机组的合理配置，使其平均负荷维持在 70% 左右，发挥其部分负荷性能极高的特性，保证冷水机组的高效运行。采用变频直驱结构，在装机初期或者机楼低负荷期，大大提高冷源系统的运行效率；磁悬浮轴承，运行过程中无摩擦，减少了机械损失；压缩机免维护，提高运行效率；避免油膜覆盖，充分发挥高效换热铜管的性能，长时间运行效率无衰减。

**间接蒸发冷却机组：**利用外界冷源进行间接换热及水喷淋蒸发冷却技术，有效延长自然冷却使用时间，显著降低数据中心制冷系

统能耗。间接蒸发冷却智能温控解决方案采用一体式架构，内置 DX 补冷，各功能部件在工厂预制集成，并进一步实现冷电融合，暖通融合及 AI 能效优化，提供了新一代极简、绿色、智能、安全的创新温控节能解决方案。

**5G 赋能智慧建造与运维：**数据中心建设从传统方式向数字化、智能化建造演进过程中，需要实现数据资产的贯穿，建造管理的全生命周期可以从统一的数据模型中结合业务特征，构建各阶段的数字化能力。

## 2、智算中心布局：

智算中心布局融入公司算力网络布局体系，按照“集中训练、分布推理，统一管控、弹性调度，自主可控、绿色低碳”的原则，以“N+X”为目标体系构建技术领先、绿色节能、服务全局的智算中心。

智能计算全栈平台致力于通过算力集群一体化交付模式，协调 AI 应用、平台、基础设施等产品一同推进云上技术产品体系化发展，可快速满足未来 TaaS、MaaS 和第三方平台的接入和未来演进需求。

智算中心将构建标准统一的智算中心软硬件全栈技术体系，初期提供 IaaS（基础设施即服务）和 PaaS（平台即服务）服务为主、逐步提供 TaaS（任务即服务）和 MaaS（模型即服务）服务为主的多层次智算服务能力。其中，TaaS 是用户只需关注业务本身，无需关心应用、平台、基础设施的组织形式和状态；MaaS 是基于大模型的新型商业模式，为下游应用提供安全、高效、低成本的模型使

用。

该项目智算服务器国产化芯片服务器占比 100%。风冷普通 GPU 服务器占比 100%。统筹考虑项目采用芯片类型异构、网络组网成熟度差异、承载业务属性不同等方面，按照集群方式进行组网建设，每个集群独立组网，相互隔离。集群由扣卡模组服务器组成，采用框盒直连组网架构。

智算资源管理架构包括智算中心硬件如服务器、分布式存储、组网设备等、智算资源管理平台如容器基础设施管理器 CIM、容器基础设施引擎 CIE、智算 AI 管理。

智算中心硬件：由服务器、分布式存储、组网设备等组成。服务器包括通算服务器、智算服务器，CPU 包括 Intel、AMD 等 x86\_64 架构以及华为等 aarch64 架构，GPU 包括 NVIDIA、华为等。分布式存储包括块存储、文件存储和对象存储。组网设备包括常规以太网组网以及 IB、RoCE 等高速无损网络。

智算资源管理平台：容器基础设施管理器 CIM 是资源池内节点、存储、网络等基础设施的管理系统，负责将资源池内节点、存储根据组网要求构建资源池，缓存容器镜像/包，并根据管理命令，调用 CIE 对容器业务进行编排，参与容器的生命周期管理过程。容器基础设施服务引擎 CIE 是容器的直接供应者，它负责根据 CIM 的要求，分配和组织资源、构建和交付容器、执行容器管理命令。

智算 AI 管理平台：智算 AI 管理平台负责智算业务的镜像/包仓库，负责对接智算资源管理平台进行任务生命周期管理，负责对



接分布式存储进行存储管理等。智算 AI 管理平台包含其依赖的所有基础软件，例如操作系统及虚拟化/容器化系统等。

智算中心拟建设新一代智能计算全栈平台，打通 AI 平台、容器软件、弹性计算实例、存储和参数网络交付流程，提供高性能 AI 训练和推理集群服务，支撑智算服务体系产品化、标准化演进。

### 成果效益

中国移动智算中心（上海）PUE 小于 1.3，同时具备“部署灵活、定制交付”等特点，可高效满足行业需求。

项目建设基于安全管控智能化、质量管理精细化、进度管理可控化、造价管控网络化和施工绿色化的“五化”目标，在设计、施工数字化的基础上，引入 5G、AI、云平台、大数据和边缘计算等技术，依托 CPMS 系统能力升级，完善中国移动智能建造方案，在集团引领下项目先行探索 5G+智慧建造，作为未来数据中心建设总结经验。同时，在设计及建造阶段数字孪生基础上，形成一整套基础数据库，作为运维阶段“数字孪生”的基础，实现运维仿真、主动响应式服务等典型应用场景。结合基础 BIM 模型、GIS、大数据、AI 技术等的集成，形成运维人员与物理信息系统 HCPS 的有效互动，全面提升“数智化”运维能力，实现故障及时发现并预警，并通过 AI 专家算法提供网络故障参考解决方案。

中国移动智算中心（上海），坐落上海移动临港数据中心园区内，园区先后获得了上海市科学技术委员会互联网数据中心超低能耗技术应用示范；工信部、中国信息通信研究院、开放数据中心委

员会（ODCC）颁发全国首个 DC-Tech 数据中心智能建造最高等级 5A 级认证；CQC A 级认证；绿色网格 TGGC&开发数据中心委员会 ODCC 颁布的数据中心绿色等级 5A 级认证；白玉兰奖；申安杯奖；第二届“新绿杯”信息通信行业赋能碳达峰碳中和创新大赛全国二等奖、第二届“华彩杯”算力应用创新大赛全国二等奖、2024 年度智算中心典型案例、2024 年度中国互联网协会助力经济社会数字化转型特别推荐案例等。

## 典型先锋案例六：

### 城地香江（上海）云计算有限公司

#### 上海城地临港智能科技创新产业园项目

#### 企业简介

城地香江（上海）云计算有限公司是上市公司城地香江投资设立的全资子公司。为用户提供数据中心投资、设计、运营、运维一体化服务的第三方企业。长三角地区最大的第三方数据中心服务运营商之一，6年安全可靠的数据中心托管及管理服务经验专注于在上海、北京、深圳核心位置提供精品高标永不停歇的数据中心，上海区域重点布局（拥有超过22000台机柜，总功率超160MW）。

#### 案例概述

上海城地临港智能科技创新产业园项目位于上海浦东新区倚天路843号，园区内建设两栋高层丙类厂房（1#厂房、2#厂房）、3#110变电站、4#办公楼、5#门卫；一期规划2#厂房为数据机房楼约3000个（单机柜8kW）。二期规划1#数据机房楼IT机柜约3000个（单机柜8kW）。园区总体布置图见图1。本数据中心是按照美国TIA-942及国家GB50174的A级数据中心标准建造机房，配备高端网络设备和完善的机房物理设施，符合Tier3+标准建设要求，符合绿色IDC标准。



图 1: 园区鸟瞰效果图

本 IDC 为实现绿色低碳，尽可能降低运行 PUE，综合考虑利用水侧自然冷却+风侧自然冷却（间接蒸发冷）、水源多联机热回收、光伏发电、预制混凝土保温夹心板等绿色低碳材料、错峰蓄冷系统等创新技术。在屋面设置间接蒸发冷却空调机组，进行风侧自然冷却。根据室外气候条件（气温低、空气干等），间接蒸发冷却空调机组分别采用“干模式”“湿模式”“混合模式（机械补冷）”等。室外空气通过“换热器”直接与 IT 机房内空气进行换热（关闭冷源），充分利用自然冷源。风侧自然冷却，全年完全自然冷却时间在 5 个月以上。通过上述技术的应用，目前 IDC 综合运行 PUE 均控制在 1.28 以下，较好的实现了节能降碳目标。

### 解决方案

上海城地临港智能科技创新产业园数据中心主要应用的创新技术有：水侧自然冷却+风侧自然冷却（间接蒸发冷）、水源多联机热回收、光伏发电、预制混凝土保温夹心板等绿色低碳材料、错峰蓄冷系统等。该项目 IDC 制冷系统采用冷冻水系统，系统采用水冷冷水机组+自然冷却形式，每台冷机配置 1 台板式热交换器。机

房空调末端采用下送风冷冻水精密空调，冷却塔采用开式冷却塔，置于屋面。空调系统运行模式包括冷机冷源模式、预冷混用模式、自然冷源模式。空调冷冻水采用一次泵变流量系统，环网形式布置，同时设置了满足系统满载运行情况下供应冷量 15 min 的蓄冷装置。

为延长自然冷却时长，项目提高冷冻水供回水温度至 18/24℃。冷却系统、冷源控制系统分为三态冷源模式自动转换：冷源系统依据室外湿球温度及冷塔出水温度与设定值的关系判断模式。三态模式为冷机冷源模式、预冷混用模式、自然冷源模式。

各模式控制信息及激活条件如下：

#### 1) 冷机冷源模式

(1) 冷机冷源模式激活，系统完全由冷水机组提供冷源。

(2) 激活条件：室外湿球温度  $\geq$  冷机冷源模式室外湿球温度设定（默认 17℃，可更改）。

#### 2) 预冷混用模式

(1) 该模式多用于春秋季节，为冷机冷源模式的初期及末期。冷冻机冷却水进水温度低于开启要求，此时通过阀门的切换使冷却水先经过板式换热器一次侧，冷冻水则经过板式换热器二次侧。板换一次侧的冷却水与二次侧的冷冻水进行能量交换，使一次侧冷却水出水温度升高至冷冻机开机要求，充分利用自然冷源中的能量使冷冻机达到开机要求。

(2) 激活条件：自然冷源模式湿球温度设定（默认 11℃，可更改） $<$ 室外湿球温度 $\leq$ 预冷混用模式湿球温度设定（默认 17℃，



可更改)。

### 3) 自然冷源模式

(1) 此时通过阀门的切换使冷却水经过板式换热器一次侧，冷冻水则经过板式换热器二次侧。一次侧的冷却水与二次侧的冷冻水进行能量交换，使板换二次侧的冷冻水出水温度保持在 18℃，完全利用自然冷源中的能量为系统提供冷源。

(2) 激活条件：室外湿球温度  $\leq$  自然冷源模式湿球温度设定（默认 11℃，可更改）。为最大限度地利用室外自然冷源，减少过渡季自然冷却方式与机械制冷方式频繁切换，项目配置了板式热交换器分别与冷水机组串接，室外低温冷却水先进入板换进行换热，然后再进入冷水机组换热。当室外湿球温度达到 11℃ 时，板式换热器能提供全部制冷，此时冷水机组可停机。

为紧跟“双碳”政策，更好地实现节能减排，更加充分利用自然冷源，本数据中心使用了间接蒸发冷却技术，该技术利用湿球温度低于干球温度的原理，通过非直接接触式换热器将室外空气或者经加湿预冷处理室外空气冷量传递给数据中心内部较高温度回风，实现风冷和蒸发冷却相结合。间接蒸发冷却技术利用湿球温度远低于干球温度的原理，对室外空气进行喷淋降温，增加间接自然冷却的时间。在一年中大部分时间内，采用风侧高效换热器冷却和蒸发冷却相结合的技术，空调机组运行在自然冷却模式下，可以节约 60% 以上的空调系统能耗。

该项目于一、二期（1#，2#楼）屋面分别规划设置 40 台间接

蒸发冷却空调机组。间接蒸发冷却空调机组按 N+1 冗余设置，根据机房分别配置。其运行示意如图 2 所示。

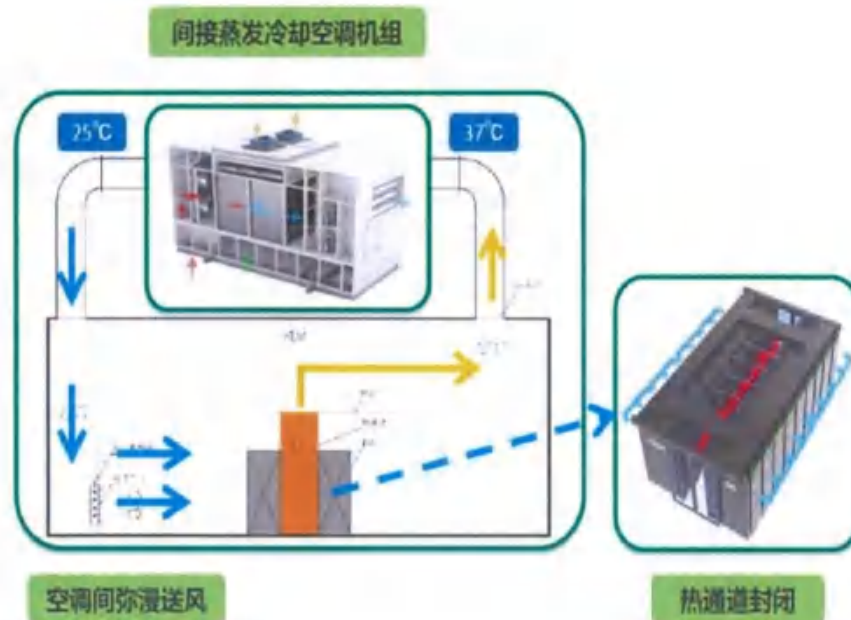


图 2： 间接蒸发冷却空调机组运行示意图

该项目使用的间接蒸发冷却技术共有 3 种运行模式：①室外温度相对较低的时段，通过高效间接换热单元，进行自然冷却。②室外温度相对偏高的时段，通过高效间接换热单元和蒸发冷却单元，进行自然冷却。③高温时段，通过高效间接换热器、蒸发冷却单元进行自然冷却，并开启补充制冷。

通过应用多维度节能技术，本数据中心实现了高质量绿色数据中心目标，投运后平均运行 PUE 值低于 1.28。本数据中心运行 PUE 达到长三角先进水平，较好地实现了节能降碳目标。

### 成果效益

本项目目前通过采用多维度节能技术数据中心 CLF 可低至 0.2，数据中心整体 PUE 为 1.28，相比基于传统数据中心 (PUE1.5)，

总节能可达 2463.12kW，换算到全年共节约标准煤  $2463.12 \times 8760 / 10000 \times 1.229 = 2651.8$  吨 (tce/a)，按照 2021 年省级电力平均二氧化碳排放因子 (上海)  $0.5834 \text{kgCO}_2/\text{KWH}$  计算二氧化碳减排能力  $12587.9 \text{tCO}_2/\text{a}$ 。按照未来一、二期全部投产可形成节能能力为节约标准煤 11368.9 (tce/年)，二氧化碳减排能力  $53967.7 \text{tCO}_2/\text{a}$ 。

城地香江 (上海) 云计算有限公司始终重视环境保护，积极响应国家建设高算力、低能耗数据中心的号召，担当社会责任。在上海城地临港智能科技创新产业园数据中心项目的设计及建设上，通过对新材料、新技术、新系统、新能源的综合利用，首度尝试将“清洁能源”“高效制冷”“能源回收”等技术组合应用于数据中心中，使得数据中心的能源利用率可以得到明显提升并取得了超出预期的实际效果，这在临港片区乃至上海地区数据中心行业中属于领先地位。为绿色数据中心的设计建设提供了非常宝贵的借鉴意义，十分具有全行业推广意义。

### (三) 算力平台，打造最强解决方案

随着人工智能技术的不断迭代升级和政策持续推动，上海持续加强新型算力基础设施建设，优化算力资源配置，打造算力调度、算力+大模型平台等典型案例，推动算力产业高质量发展，为数字化转型和智能化升级提供了有力支撑。

在算力调度方面，上海大力开展算力基础设施及算力资源输

出能力摸排，形成市级层面算力清单，并构建一体化算力调度服务体系 and 平台基础框架，实现算力资源统一编排。国家（上海）新型互联网交换中心推出了全国首个基于交换中心的算力调度交易平台，能够为用户提供算力调度、算力托管业务服务，缓解算力结构不均衡、供需不匹配等问题。漕河泾公共算力服务平台后续将持续为园区企业提供专业的算力和大模型应用综合服务，涵盖国产算力适配、异构算力资源调度、概念验证（Proof of Concept）专业算力服务等，致力于为园区企业提供稳定、高效的国产算力环境，精准分配算力资源、实现资源利用的最大化，以及通过算力服务帮助企业 and 科研机构提供可行性、创新性实验验证等。

在**算力+大模型平台**方面，上海推动算力赋能应用水平提升，加强底层芯片、技术创新、应用赋能和产业链协同，构建新型数据中心算力生态体系。比如，随着中国电信在上海临港智算园区的首个国产单池万卡液冷算力集群正式交付并投入商业运营，实现了单一集群内万卡高速互联，满足万亿级参数大模型训练所需的多机多卡并行、高吞吐无损通信等需求，满足了天翼视联星辰海纳行业大模型基于天翼云一站式智算服务“慧聚平台”进行大模型基座能力预训练和场景微调的需要，利用平台提供的训推一体化算力服务，提高行业大模型研发效率。

## 典型先锋案例七：

### 上海市新型互联网交换中心有限责任公司

基于新型交换中心的跨主体、多元异构算力调度交易平台

#### 企业简介

国家（上海）新型互联网交换中心（SHIXP）是由工业和信息化部批复，在上海市委、市政府、上海市通信管理局、上海临港新片区管委会的牵头组织下，于2021年12月20日正式揭牌并启动运营。SHIXP是实现网络之间流量集中交换的新型信息基础设施，具有吸纳成员多、业务种类丰富、融合创新潜力大等特点。当前SHIXP已形成了以端口服务为核心，“一点接入、多点连通”为特色，面向企业提供企业互联、企业上云、云网融合、算力交易、国际服务、安全服务等六大场景的服务能力。先后获得了第五届“绽放杯”5G应用征集大赛上海区域赛金奖、首届“金灵光杯”中国互联网创新大赛新型工业化赛道优秀奖，SHIXP正努力成为创新业务发展活跃的交换中心。

#### 案例概述

SHIXP围绕全国一体化算力网络国家枢纽节点对“网络互连互通”的业务需求，面向落实“算力浦江”智算行动计划重点任务，结合自身“一点接入、多点连通”的业务特点，充分发挥“中立、开放”的试点定位优势，构建了“1+3+4+N”的总体建设方案：

(1) “1”表示“一中心”：即算力调度中心，包含算力管理平台，交易服务平台。



(2) “3”表示“三底座”支撑：即基于算力基础设施、通信网络基础设施和国家（上海）新型互联网交换中心，充分聚合计算、存储、网络三大资源。

(3) “4”表示“四大保障体系”：即算力评估标准体系、算网服务标准体系、算力交易标准体系、算力安全防护标准体系。

(4) “N”表示“多元服务场景”，推动算力资源与行业场景的深度融合。



## 解决方案

该项目方案依照“三层架构、五大平台”体系搭建：

(1) IXP 算网平台。采用“同城双备、南北双环”架构方案搭建扁平化算力专网，主要由核心层 P 设备和接入层 PE 设备组成，客户可由接入层 PE 设备或 SW 设备接入 SHIXP 网络，它可将算力供需双方在网络层面高效互联，通过算力资源与网络资源状态的协

同调度，将算力数据通过短路由就近转发，调度到匹配的计算节点，实现优化区域算力网络架构的目的。同时，通过融合运用 SDN、Flex E、NFV、SRv6、Smart OTN 等新型算力承载技术，实现了算网“按时、按需”灵活调整的业务需求。

(2) 多元异构算力互联互通调度平台。通过与公有云、私有云、行业云、超算中心、智算中心、科研高校等不同主体算力中心的云网平台打通，实现多方算力的互联互通编排调度。具体能力包括：多方算力接入、异构算力纳管、算力一体化度量、算力需求智能解构、多方算力动态编排、算力自动化调度开通等能力等。

(3) 算力调度交易门户。实现算力的交易过程支撑、交易参与方的管理、交易过程。提供具体能力包括：算力意图订购，算力方案推荐，算力可信交易，用户管理、合作伙伴管理、多态算力商品封装及管理、交易管理、订单管理、计费结算、服务管理、商城管理、运营管理、算力统一控制台等功能。主要分为以下功能层：

算力接入层：算力作为核心生产力，在各行各业中承担着至关重要的角色。但不同算力平台其技术架构、管理概念、管理模式、服务能力、资源规格、API 开放程度与协议等都存在较大的差异，极大增加了算力资源的使用复杂度。需要通过算力交易平台拉齐各类异构算力平台的管理概念、服务模型并适配各平台 API 从而实现不同主体的算力网络和平台之间的算力交互、协同和共享，为算力需求方提供更高效、便捷的算力服务。算力接入层便是用于对接多方异构算力，实现多方接入、异构适配。包括算力标识管理、算力

接入适配、算力接入鉴权、算力接入控制等核心能力。

**算网资源层：**算力资源具备多样化的内核和分布式泛在特性。算力资源管理是针对多样化的算力资源进行调度、分配和协同的管理过程。通过全量资源的可视化管理，实现资源的可靠使用和可信管控。资源管理对象应覆盖计算、存储、网络、应用、安全等资源。算网资源层便是用于聚合管理多方算网资源，包括计算资源管理、存储资源管理、网络资源管理、资源信息核验和资源状态跟踪等核心能力。

**编排管理层：**编排管理层负责全网算力资源的统一管控和智能编排，是算力编排、调度、开通的关键模块和智能中枢，作为算力需求方和算力供应方的智能算力撮合调度器，通过对多方算力资源如运营商算力、云服务商算力等的度量建模和产品设计，完成对异构多样化算力的能力画像，并通过对算力需求的智能解构匹配，实现对算力需求的智能方案推荐、编排调度，完成算力资源的调度部署和开通。编排管理层的核心亮点是智能解构、编排调度，包括算力度量、算力感知、智能解构、编排调度、算力视图、监控运维等核心能力。

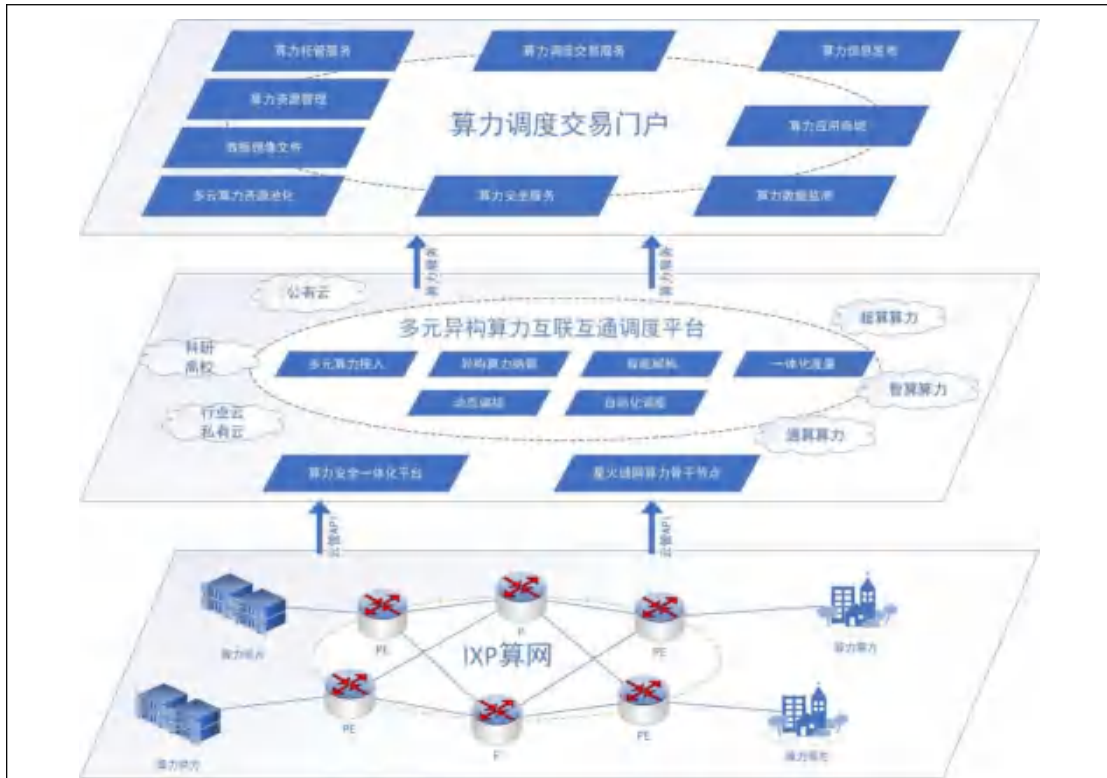
**服务运营层：**算力网络基础设施的发展，算网大脑能力的成熟演进，也促进算力交易运营的发展。算力交易作为数字经济领域的一个热门话题，对于推动数字经济的发展和创新具有重要意义。建设高效、安全、稳定的算力交易平台，是实现算力交易生态健康发展的关键之一。服务运营层便是用于建设算力交易的全链条能力，

核心目的是实现生态合作、公平运营，包括用户管理、合作伙伴管理、产品管理、交易管理、订单管理、计费结算、服务管理、商城管理、运营管理、系统管理等关键能力。

算力应用层：算力应用层主要是基于服务运营层提供的丰富能力，构建面向多方角色的交互门户。包括算力交易统一门户、园区算力交易门户、行业算力交易门户等。算力交易统一门户为主门户，园区算力交易门户、行业算力交易门户为不同类子门户，各地园区可定制自己的特色园区门户，各行各业也可定制自己的特色行业门户，促进特色门户发展、万千应用市场建设、各类专区运营，促进算力应用百花齐放，赋能千行百业。

(4) 算力安全一体化平台。基于人工智能、大数据分析等最新技术，为平台内供方、需方、第三方服务商等企业提供网络安全、数据安全、信息安全一体化的安全解决方案，支持本地化部署和 SaaS 化部署。

(5) 星火链网算力骨干节点。通过建设全国首个星火链网算力骨干节点基于国家级区块链的可信算力服务系统，从交易订单上链存证、计算过程上链存证、结算订单上链存证、算力券登记与核验四个方面入手为算力交易平台提供可信保障，为算力网络、算力交易市场提供安全可信保障服务。



该项目当前已建成一张高速全光网算力调度网络，总长约 1000 公里，包括 2 个核心节点、5 个扩展节点和 12 个延伸节点共 19 个算网节点的建设部署，可覆盖全市约 70% 数据中心，重点解决了数据中心间互联成本高、转发效率低和安全合规等问题，基本形成了本市“3+1+N”算力网络调度体系；与联通云、阿里云、华为云、腾讯云、优刻得、火山引擎、有孚云、AWS 等多家国内外头部云服务提供商，商汤科技、万国数据、有孚网络、世纪互联等头部数据中心已实现预接入，可为全市重点产业园和企业用户提供更为便利的网络服务和云算服务；全国首个算力交易集中平台——上海算力交易平台于 2024 年 4 月 1 日在临港新片区正式发布商用，可为算力供需双方提供一个“中立、开放、公平、可信”的算力交易环境，绕过中间环节一线拉通算力供需双方，通过降低企业用算成本，



激发了上海算力市场活跃度。

### 成果效益

该项目作为全国首个算力调度平台，面向供需侧全产业链，加速构建上层芯片设计、软件服务规模发展，对数据中心的运力进行统一管理和交易，推动了算力市场的规范化和透明化，从而提升了整个行业的效率和竞争力。同时，作为产业链中的关键环节，也改变了传统的价格机制和服务方式，支持定价策略设置，实现了更加公平和公正的交易环境。此外，项目的成功实施，也将进一步推动行业技术的创新和发展，提升整个产业的技术水平。该项目商业模式主要是通过为算力供方、需方、服务商等企业提供算网端口接入服务和算力交易服务。以打造“中立、开放、公平、可信”第三方算力服务平台创新服务模式，实现为供方企业提升算力资源利用率，为需方企业解决算力短缺和弹性需求的难题。算网端口以日/周/月租的形式收取端口费，算力交易以合同额收取一定比例的服务费，并形成收入来源。该项目已实现与联通云、有孚网络、万国数据、优刻得、中科曙光等 8 家企业达成战略合作，与 11 家云算平台实现互联互通，归拢“通算、超算、智算”各类算力超 8000P，吸引入驻企业超 200 家，初步形成了本市算力资源合理流通态势。

## 典型先锋案例八：

### 中国电信股份有限公司上海分公司

搭建大模型基准测试工具和适配优化平台，

标准引领推进国产算力生态

#### 企业简介

中国电信股份有限公司上海分公司是中国电信的省级分公司，业务主要包括：

基础电信业务：固定网本地电话业务、固定网国内长途电话业务、固定网国际长途电话业务、IP 电话业务、卫星国际专线业务、因特网数据传送业务、国际数据通信业务、公众电报和用户电报业务、26GHz 无线接入业务、国内通信设施服务业务等。

增值电信业务：在线数据处理与交易处理业务、国内因特网虚拟专用网业务、因特网数据中心业务等第一类增值电信业务，以及存储转发类业务、呼叫中心业务、因特网接入服务业务和信息服务业业务等第二类增值电信业务。

全业务综合信息服务：固话语音、移动通信、宽带互联网接入、卫星通信、互联网电视、信息化应用等在内的全业务综合智能信息服务。

#### 案例概述

在人工智能时代，算力基础设施作为支撑大模型技术应用和发展的关键，其重要性日益凸显。上海市积极响应国家战略，通过《算力基础设施高质量发展行动计划》和“算力浦江”智算行动实施方

案，致力于推动算力基础设施的高质量发展。

上海电信在上海市通信管理局和上海市通信学会的指导下，联合上海人工智能实验室等科研院所、业界头部企业制定了《智算中心性能测试评估方法》《智算中心能力验收规范》两项团体标准。

大模型基准测试和适配优化平台（以下简称“翼芯平台”）基于两项团体标准针对大模型算力国产化适配难、优化难、协作难、复用难、使用难等问题开发，填补了国内空白在大模型算力的国产化适配和优化领域国内行业领先，满足通算+智算场景下信创客户打单需求和内部云信创演进需求。

翼芯平台围绕标准、评测、适配、优化、复用 5 大核心内容开展多架构自动化适配、多模型场景的自动批量测试、训练推理跨层一体优化、中央式资产库和软硬件智能推荐等技术创新，构建具有统一标准和适配优化流程、自动化的批量测试套件、智能化分析调优方法和工具、云原生集群管理、可复用的 AI 资产库和智能推荐的一站式平台。

### 解决方案

技术体系：

翼芯平台是行业内首个面向大模型算力国产化评估测试与适配优化的一站式服务平台，其主要包括：

#### （1）统一标准自动评测

算力标准与用例，提供电信自定义的统一算力测试标准和一些包括配套的测试指标，测试用例，测试方法等。

自动化测试套件：平台提供统一的自动化测试套件，基于测试结果和计算架构智能选择优化策略和优化技术组合，提高国产芯片的有效算力。

## （2）适配优化

迁移适配工具集：集成包括模型格式转换、训练代码迁移、精度角度、调试优化等工具，采用 CICD 进行适配的全链路评估跟踪，以流程自动化实现对多种计算架构的适配，缩短 50%的适配时间，可覆盖 5 种以上主流大模型与 5 种以上国产软件栈间的相互转换，可向上对接多开源社区、多模型格式、多类型的大模型，无缝转换引入丰富模型生态。

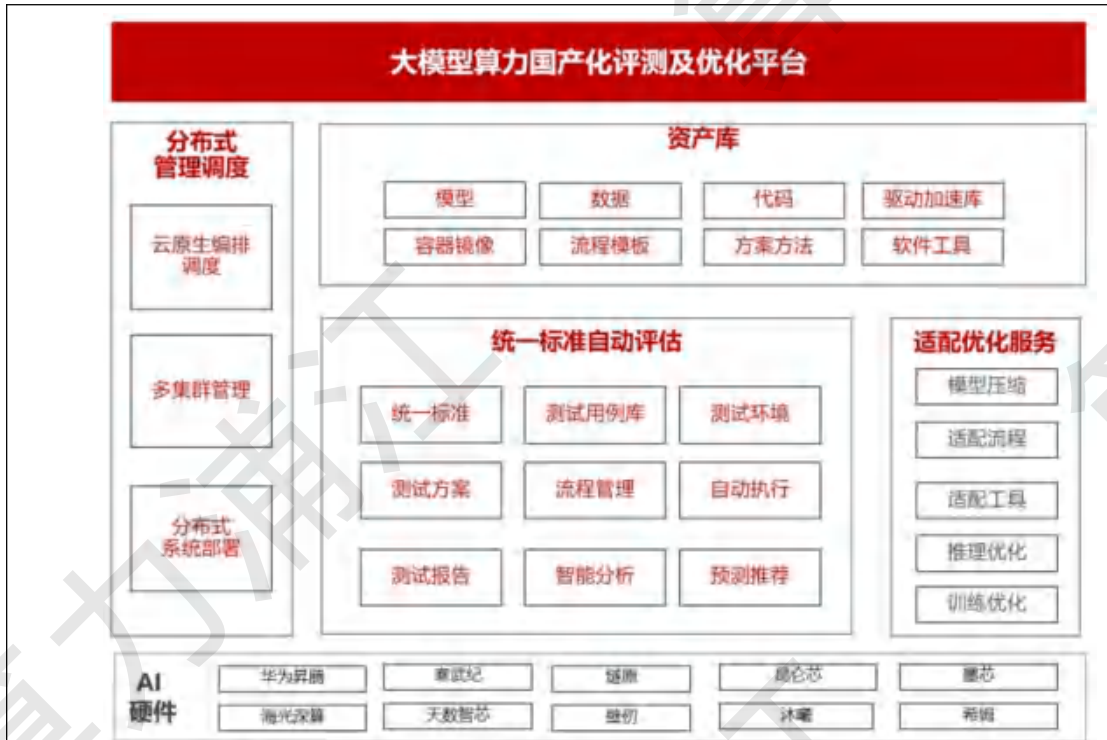
智能优化代理：平台智能分析性能卡点堵点，智能推荐不同大模型的软硬件组合，通过智能优化代理调整训练超参和推理框架优化参数，平均提高 20%以上的大模型推理性能。

## （3）统一资产库

提供统一 AI 资产库，可快速复用适配流程、驱动和加速库、适配后的运行代码、测试数据集、大模型和配套容器镜像等，资产库可减少 80%的重复开发和重复调试适配。

## （4）分布式管理调度

主要包括云原生编排、多集群管理、分布式系统部署，可调用本地环境或远程集群运行测试用例和适配优化。



图：各模块业务逻辑：

实施方案：

(1) 组网方式：组网架构采用数据中心传统的 spine-leaf 网络拓扑架构，翼芯平台常接入智算中心的管理/业务网络，并支持多个翼芯平台的分布式组网互联和对远程的智算集群的管理调度测试验证。



网络共设计 5 个区域，分别为出口区、带内管理区，带外管理区域、测试适配优化平台区以及被测国产 GPU 服务器区。

(2) 系统架构：



统一算力性能度量评估：面向多种大模型测试，自动化批量化测试，测试过程采集全量数据用于后续分析。自动化评估性能精度分析和性能瓶颈分析，进行根因分析追踪，给出解决卡点和堵点的建议。

迁移适配，加速落地：多种模型格式转化，多种 AI 框架无缝对接，代码级的算子迁移，数据格式转换，标准化适配流程全过程跟踪，加速大模型国产化落地。

模型压缩，优化提速：基于计算图优化和通信并行的训练优化，模型压缩量化和推理加速，基于计算架构智能选择优化策略和优化技术组合。

资产复用，智能推荐：适配后的运行代码、测试数据集、国产硬件驱动和编程框架、AI 框架、大模型容器镜像等，从资产库中推荐用所需的硬件、模型或技术栈等。

关键成效：

基于统一标准，通过统一的模型、数据集、测试脚本（脚本已适配多家国产芯片）提供适配方法和统一的流程，可有效形成规模推广效益。

形成统一测试验收标准，包含功能、计算、存储、网络、大模型能力、管理平台能力、安全性、稳定性、生态兼容性等 9 大维度 30+用例 100+指标，以全面客观的测试标准指引算力中心测试验收，促进算力生态的培育。

应用价值：

对于中小企业而言，翼芯平台可以有效降低大模型测试的环境搭建成本，并通过复用平台的能力，实现大模型的快速优化，在更广泛的社会层面上带来深远而积极的影响。

### 成果效益

**经济社会效益：**翼芯平台将算力中心评估标准应用于算力基础设施自动化评测，支持远程评测，可视化结果展示，可大幅的节省验收的人力和资源成本，提升评测效率。同时基于统一标准，通过统一的模型、数据集和适配多家国产芯片的测试脚本，提供了完整的适配方法和统一的流程，可有效形成规模推广效益。

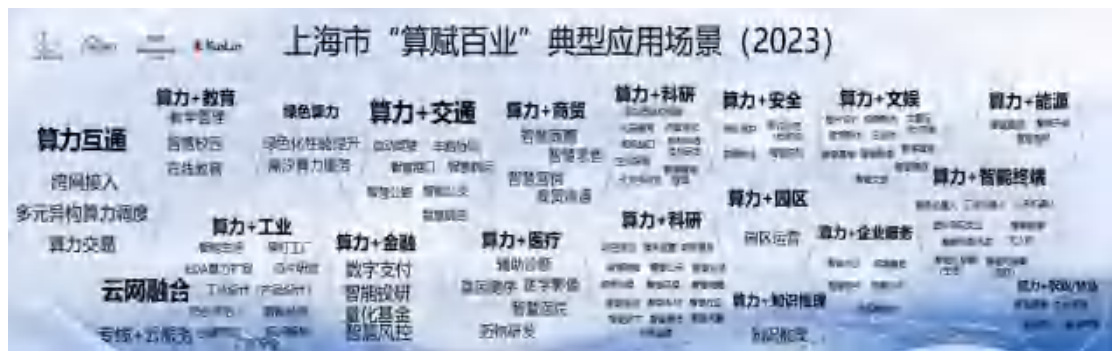
本平台将有效降低大模型测试的环境搭建成本，并通过复用平台的能力，实现大模型的快速优化，在更广泛的社会层面上带来深远而积极的影响。

**商业模式：**大模型基准测试和适配优化尚未有成熟的产品和服务，是一个潜在的新兴市场，预计能为各类中小型的大模型服务提供商提供完整的大模型基准测试和适配优化服务。翼芯平台投产后预期可以为公司带来营收，成为公司的业务增长点。

**应用推广现状：**中国电信上海公司在临港信息园建设了万卡国产算力资源池，翼芯平台依托该物理资源，搭建了国产 GPU 测试环境和大模型评测环境。在万卡资源池的硬件集成、软件集成、验收测试等环节中，翼芯平台研发团队积累了对千卡昇腾集群的性能优化经验。目前翼芯平台已经对主流的国产大模型和国产 GPU 卡进行了内部评测和适配优化，取得了良好的效果。

#### (四) 算力应用，赋能行业百花齐放

作为数字经济时代的核心引擎，算力正持续赋能各行各业，应用覆盖制造、政务、金融、文旅、生物医药等领域，驱动智能化转型与服务创新。上海市作为我国的经济枢纽与科技创新的前沿阵地，高度重视算力这一关键生产要素，加速算力技术在各行业领域内的应用探索与实践融合，逐步构建算力赋能的行业生态体系，带动行业数字化转型和智慧城市建设。



**算力互通优化资源配置。**跨网间算力接入方面，通过构建算力网络全栈产品体系，上海市实现了对多元算力资源的灵活配置与高效协同，为产品研发、生产制造、销售服务等数字化场景提供了强大的算力支撑。该体系不仅实现了数据共享、平台支撑和协同处置等服务，还极大地提升了数字化场景的运行效率和智能化水平。天翼云、华为云、阿里云等国内知名云计算企业已接入国家级多元异构算力调度，通过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，实现了算力的优化配置。

**算网融合助力资源高效整合。**上海电信应用专线+云服务场景，提供云聚 SaaS 综合性业务支撑系统。该系统基于运营商城市基础设施投资与建设业务，开通专线接入云厂商算力平台，通

过采购云厂商算力资源，显著提升了云业务的运行速度、便捷性与连续性，同时降低企业入云成本，有效改善了用户体验。多云纳管与多网协同方面，上海联通通过搭建云擎平台，将公有/私有云、网络设施、终端设备、视频图像等资源统一接入、多云纳管与统筹整合；基于对资源核心性能指标、状态指标等信息的对接收敛，实现了对资源的实时监测、自动预警与闭环管理。这提升了资源可靠性与安全性，为企业提供了更加稳定、高效的算力服务。AI+赋能算网融合发展，中企通信依托中企云时代®SmartCLOUD 算法平台，构建云网融合架构，以及智能算网大脑，充分整合智赋“云网安”一站式 ICT 服务资源，满足用户对算力服务的需求。

**算力节能助力行业绿色发展。**算力节能领域正引领上海市迈向更为绿色高效的发展道路。在数字化绿色低碳方面，上海市企业积极运用先进技术和创新模式，推动数据中心、工业等领域节能减排，助力城市实现碳中和目标。其中，万国数据作为领先的数字基础设施服务商，万国“上海三号”数据中心通过智能化工具动态多维指标与自动数据分析，实现对能效的精益化管理，为长期实现 IDC 极致效能创造了可能。同时，增强水、电、风、热等子系统末端数字化程度，让节能管理迈入了智能时代。商汤科技建设的人工智能计算中心基于微模块机房搭配液冷机柜，微模块机空调末端采用行间空调，贴近服务器机柜，实现空调就近制冷，热源转移，有效提升了空调运行效率，实现能源保护并节



约成本。优刻得上海青浦云计算中心通过屋面铺设光伏发电系统，实现了清洁能源的利用，不会产生任何污染和温室气体排放。同时，结合高压直流与智能化运营等手段，减排效果显著，为行业的可持续发展提供了有力保障。

算力深度融入上海市行业发展，为工业领域注入新活力。在智能制造方面，上海市企业积极整合技术、资金、人员、产业链上下游等资源，综合应用物联网与智能计算，打造智慧园区、智能工厂和智能车间。例如，在智慧园区领域，天翼视联研发的星辰海纳行业大模型具备强大的视觉分析与理解能力，可以提供工业园区 AI 监控场景智能监控解决方案，并凭借其大模型泛化能力强、识别精度高、支持面向细分场景的小样本训练等优势，在以智慧园区为代表的场景应用中取得了良好的规模发展势头。在智能工厂领域，宝武钢铁则依托宝信工业互联网平台优势，结合人工智能、数字孪生、大数据等 IT 技术，打造超融合架构集群赋能“黑灯工厂”，提供“全流程管控+数字化赋能”双驱动智能工厂整体解决方案，实现了无人化、集控化、一键化为一体，极致、高效、安全的智能工厂车间。在智能车间领域，华友钴业开发的华友新能源智能制造大数据中心通过建设智能车间和 SAP、ERP、MES、WMS 等信息系统，实现了系统间的集成应用，打造智能制造生态体系。

此外，上海市还在标识解析、数据共享等方面积极探索算力应用。通过条码、电子标签等载体采集数据，并将标识与信息、



地址关联，实现了工业大数据的按需共享和合理流转，为工业互联网平台开展资源定位、监测认证和信息共享提供了有力支撑。在算力的加持下，工业数据的潜能得以加速释放，智能引擎可以更有效地调配生产资源、重构生产流程，提高制造业生产力。随着算力的提升，“车路协同”“车网互联”的智能网联汽车步入快车道，“安全、畅通、低碳、高效”的交通网络正在加速构建。

## 典型先锋案例九：

### 支付宝（中国）网络技术有限公司

#### 绿色计算技术在蚂蚁消费金融业务中的实践

##### 企业简介

支付宝自 2004 年起步，始终致力于数字支付开放平台的建设和发展，于 2011 年 5 月获人民银行首批颁发《支付业务许可证》，主营互联网支付、移动电话支付、银行卡收单、预付卡的发行与受理（仅限于线上实名支付账户充值）四项支付业务，业务范围覆盖全国。支付宝现已研发担保交易、快捷支付、条码支付、刷脸支付等创新支付技术，服务于商业经营、便民缴费、交通出行等不同场景下的数字支付需求，为超 10 亿用户、8000 万商家提供支付服务保障，助力实体经济蓬勃发展。

##### 场景需求

如何用绿色计算技术推动消费金融的双碳实践

1.突破传统计算设备能源浪费和污染问题：传统的计算设备使用方式往往会带来大量能源的浪费和环境的污染，所以急需进行绿色计算技术的研究和应用，以实现双碳战略的目标。

2.实现数据中心产业中上游环节节能减排：目前行业内，数据中心碳中和的路径可以概括为“少用电、用绿电、碳抵消”。通过探索可再生能源的使用，基础设施、IT 设备节能，智能运维管理等技术手段，分别实现数据中心产业中上游对应环节的节能减排。

3.助力产业下游挖掘使用效能潜力：通过技术手段合理分配计

算资源，提升单位算力所能支撑完成的业务量，将资源利用率提升到最大幅度，能够有效降低单位业务应用所需能耗，从而减轻对于能源的需求，降低对环境的压力。

### 解决方案

#### (1) 通过池化手段实现算力归一，提升稳定性

通过算力归一化，把不同硬件架构的 CPU 的实际计算能力进行对齐，对上可屏蔽不同 CPU 的架构差异，体现出统一的计算能力，在控制业务应用算力使用复杂度的前提下，充分享受硬件升级带来的高计算能效红利。此外，运行在不同的 CPU 上的业务会被分配不同的 CPU 核数，让运行在不同 CPU 上的业务性能都保持一致，充分使用了每款 CPU 的资源，有效提升了混合集群资源的效率。

#### (2) 统一调度，提升系统性能

基于开源的 Kubernetes 打造了容器调度平台，其采用的弹性调度架构将数据中心的服务器标准化成统一云化资源池，并通过统一容器 API 为蚂蚁消金业务应用提供在智能弹性容量、分时智能调度以及在离线混部等云原生容器编排能力，帮助上层各类业务实现了资源统一管理、应用容量高可用保障、异构硬件无感知屏蔽、基础资源最大化分时复用等能力，达成了机器资源使用率大幅提升的效果。

#### (3) Serverless 无服务器化，实现工作负载的弹性伸缩与调度

结合 Serverless 的理念以及内部应用发布模式的现状，把在线

应用分为三种发布模式：应用级发布模式、模块级发布模式、函数级发布模式，并同时推进三种发布模式的 Serverless 化演进。

#### （4）资源弹性伸缩，提高运维效率

在业内率先提出了全新的基于 AI 算法的端到端弹性伸缩预测决策一体化系统，整体性能达到国际领先水平，实现了机器资源按需伸缩，提高了运维效率和业务峰值活动支撑能力。

#### （5）建立绿色观测体系，持续优化工作负载资源消耗

建设了持续性能分析（Continuous Profiling）技术，可以实时地、低开销、可持续地观测和分析工作负载的资源消耗，识别性能瓶颈，优化系统的性能表现，提升工作负载的处理速度和降低处理成本。

通过 GreenOps 绿色减碳平台，为蚂蚁消金业务提供了一站式的碳排放观测、分析、优化的能力。包括：持续性观测、碳排放/CPU 水位情况、多维度的账单、碳排放分析、优化建议、一键优化执行等功能。

### 成果效益

通过上述绿色计算技术在蚂蚁消金的成功应用，实现了蚂蚁消金累计减少储存 6.8PB，累计缩容 20360Core，累计减少 357 吨碳排放，实现单笔交易 SCI 同比下降 58%。项目不仅有力帮助了蚂蚁消金积极响应国家双碳战略，同时在行业内也具有广泛的应用价值。

## 典型先锋案例十：

### 中国联合网络通信有限公司上海市分公司

#### 高性能算力赋能红色文化元宇宙

##### 企业简介

中国联合网络通信有限公司上海市分公司是中国联通在上海的重要分支机构，坚定不移贯彻落实网络强国、数字中国、科技创新、国企改革等重大决策部署，坚持联网通信业务和算网数智业务协调发展，充分发挥“网、云、数、用、安”数字技术新优势，深度参与和服务上海现代化产业体系，在善政、兴业、惠民层面做了大量的实践，努力成为千行百业首选的“数字伙伴”。

在安全数智的联通云大体系架构之上，上海联通全面推进“云算焕新”工作。一是构筑服务不同行业、云边协同的通算、智算、超算资源，提供按需而选多样算力服务。二是打造全光底座、泛在覆盖、极速互联、弹性调整的入云联接、云间组网，提供网随算动的算力网络服务。三是自主研发多云运营平台和算网大脑，实现资源统一度量、算法智能编排，提供算网一体的智能调度服务。四是服务千行百业客户的个性化需求，提供从咨询规划、资源建设，到上云迁移、专业运营的云网数安一体化集成运营服务。通过构筑“联接+感知+计算+智能+安全”一体化算网服务，倾力打造为中国式现代化强劲算力引擎的国家“首选云”。

##### 解决方案

通过该项目建设，致力于形成“五新”新格局，即打造中共一大纪念馆数字化转型新动能、红色文博场馆服务延伸新模式、红色



文化教育新思路、算力赋能红色文化产业升级新标杆、云网算助力元宇宙场景推广新实践。同时，结合中共一大纪念馆、联通等多方能力，打造基于算力网络支撑、结合元宇宙技术运用、实现应用领域创新的标杆实践。通过项目复制，能够促成更多依托于算力的应用项目落地，持续拉动算力、云网等发展和应用场景推广。

围绕问题及需求，“数字一大”项目形成了如图 1 所示的云计算技术架构。基于上海联通目前整体算力网络架构，依托丰富的算力布局、强大的运力网络以及广泛的覆盖，通过算网大脑 2.0 的整体资源纳管及调度能力，为“数字一大”定制化提供实时云渲染应用，最终打造了元宇宙穿越探索、山海游览、三会一课三大云渲染体验。

图 1：项目云计算技术架构



**算网层。**通过联通机房中搭建定制化的数字一大 GPU 服务器集群专区，包括 21 张高性能显卡 V100S 以及基础 NVIDIA 4090 显卡资源池，融合高速网络及带宽，打造云渲染算力资源池。结合运营商优势，提供高保障算网资源。

**调度层。**算网大脑 2.0 可实时监控和收集“数字一大”APP、小程序及 PC 端等不同终端用户的宽带使用情况、延迟及丢包率等，并将这些内容传递到编排器，生成智能调度策略，进行统一调度资

源。联通算网大脑平台如图 2 所示。



图 2: 联通算网大脑平台

能力层。实时云渲染平台接受用户发起的渲染交互指令，进行渲染工作，并将渲染后的元宇宙穿越探索场景、元宇宙山海游览场景、元宇宙三会一课场景画面编码生成音视频流，通过网络实时传回客户端，实现云端渲染效果的实时展现。联通云渲染平台如图 3 所示。

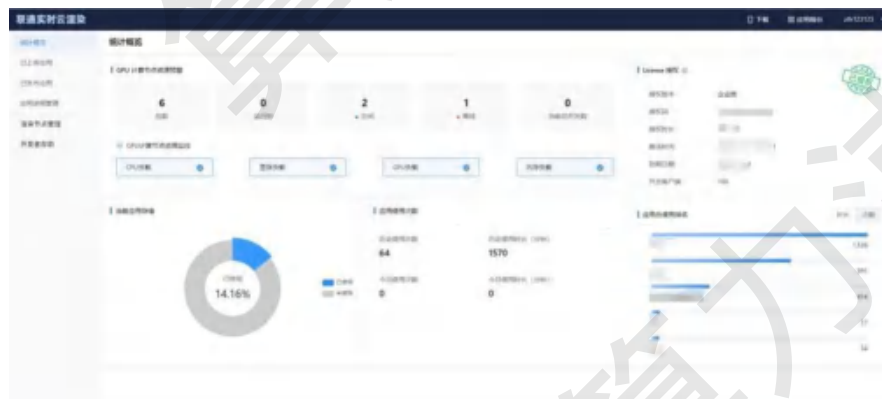


图 3: 联通云渲染平台

应用层。最终实现三大场景的云渲染支撑，其中元宇宙穿越探索场景中可以实现多人组队同游；元宇宙山海游览场景可以实现与

大量珍贵文物及宝贵红色艺术作品的高清晰度翻阅、近距离互动；元宇宙三会一课场景中可实现千人同屏、共学共建。

在算网大脑的赋能下，“数字一大”项目的云渲染技术路径如图 4 所示。具体流程：①用户在 APP、小程序、PC 端等终端上进行操作，并发出渲染请求；②上海联通数字一大联通云服务专区对用户请求进行处理，其中云渲染平台对于渲染需求进行初步解析与验证，并提交至联通算网大脑 2.0 进行调度申请；③算网大脑 2.0 收到调度申请后，分析请求的细节，如渲染任务的规模、复杂度、所需资源类型（CPU、GPU）、存储需求、网络要求等。根据分析结果，算网大脑选择最佳的基础资源池进行任务调度，同时考虑资源的可用性、成本和地理位置等因素。目前数字一大项目核心使用的资源池为中国联通浦江云数据中心资源，此云数据中心建筑面积超过 5 万平方米，可提供丰富的算力资源；④资源调度之后，由云渲染平台执行渲染任务。将需求分散至多个不同的渲染节点，进行并行渲染后，将渲染结果统一返回云渲染平台；⑤云渲染平台将结果返回至用户端。

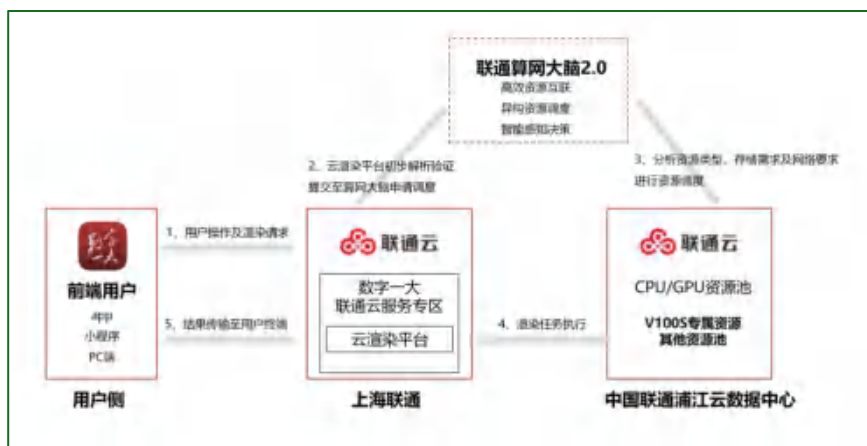


图 4：项目云渲染流程

依托高性能 GPU 算力集群的强大支撑，“数字一大”可承载 20000 人同时在线，共游 100000 平方米的元宇宙场景，并为用户提供高清的画面显示、流畅的交互操作，提升用户的沉浸感。同时，“数字一大”能够支持会议模式下的 1000 路高并发，为基层党建工作注入了创新活力。

基于整体的云算架构设计，最终打造形成了以“数字一大”小程序为一点入口、以“数字一大”APP 为主要载体、以“初心之旅”大空间 VR 等多种服务为触达方式的线上线下融合元宇宙服务矩阵。



图 5：“数字一大”服务矩阵

“数字一大”项目在世界人工智能大会、长三角国际文化产业博览会、深圳国际文化产业博览交易会等多项展会亮相，吸引了大量观众前来体验，并获得了央视新闻、人民日报、新华社等央媒以及文汇报、东方网等市级媒体的宣传报道。



图 6：“数字一大”在 2024 世界人工智能大会上亮相

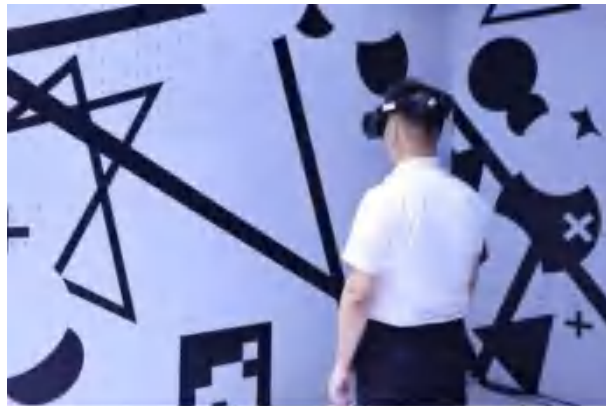


图 7：“数字一大”大空间 VR 吸引众多游客前来体验



图 8: 人民日报报道“数字一大”

## 成果效益

“数字一大”项目的成功落地，将算力、元宇宙等技术与文旅



行业的实际应用需求紧密结合，创新云渲染和算力调度方式，实现了中共一大会址等核心建筑以及《共产党宣言》等珍贵文物的 1:1 孪生复刻建模的高清显示，并保障了用户与建筑、文物等物品的流畅交互，在提升用户体验的同时创新伟大建党故事讲述方式，促进红色文化传播与弘扬。同时，基于高性能 GPU 算力集群支撑，项目实现了游览模式与会议模式的高并发，实现了天南海北的同志共游一大会址、共上一堂党课，以数字技术赋能基层党建的创新。作为标杆项目，“数字一大”展示了算力技术在文旅领域的巨大潜力和广阔前景。

“数字一大”项目具有非盈利性的特点，但其线上线下融合、服务外延扩大的模式，有效提升了中共一大纪念馆的品牌影响力，间接促进了红色文创产业的发展。同时，该项目作为开放性数字平台，为党组织、党员及社会大众提供高效便捷的红色服务，串联多个红色文化资源，创新了红色文化的传播方式，以更为丰富、多元的方式展现红色足迹和伟大建党精神。总的来看，“数字一大”通过线上 APP 和小程序及线下 VR、AR 体验等服务载体，创新展现了红色文化的内涵，其已实现多次国家级展会亮相，累计曝光 9000 万次，吸引了 17.7 万人实际体验，为其他算力技术融入文化产业的实践提供了宝贵的参考价值。

## 四、上海算力产业发展挑战与建议

### (一) 面临挑战

#### 1、AI 大模型发展推动算力结构升级，上海面临高质量算力供给和国产高端芯片短缺挑战

随着 ChatGPT 等人工智能大模型的兴起，全球算力需求正以前所未有的速度增长。一方面，大模型训练所需的算力资源极为庞大，对计算性能和存储能力提出了极高要求。上海虽已建成多个高性能计算中心和智算中心，但在面对多个大规模、高复杂度的 AI 大模型训练任务时，仍将存在算力供给压力。另一方面，算力资源的分配和调度也面临挑战，表现在找算力难，调算力难，用算力难等方面，尤其是具备高可用性、高性价比、高灵活性的高质量算力供给更为缺乏。然而，国产算力在性能、稳定性、兼容性上还有差距，尚未形成全栈服务能力，高端 AI 芯片短缺，且缺乏丰富的工具集、数据集、参数集，进一步导致市场上算力供需矛盾突出。

#### 2、绿色计算需求迫切，上海面临算力中心节能减排和可持续发展的挑战

算力中心的运行消耗大量能源，同时产生大量热量和二氧化碳等废弃物，对环境和生态系统构成威胁。在全球气候变化和绿色低碳发展背景下，算力中心的节能减排和可持续发展问题愈发凸显。上海作为国际大都市，对环境保护和可持续发展有着更高的要求。算力中心需要采用更先进的节能技术和设备，如液冷散

热、太阳能光伏板等，以降低能耗和减少碳排放。同时，算力中心还需进一步优化布局和运营策略，如采用分布式部署、动态调整算力资源等，以提高能源利用效率和资源利用率。

### 3、全球科技创新竞合并存，上海面临算力产业人才与技术短缺的挑战

算力产业的发展离不开人才和技术的支撑。然而，当前算力产业面临着人才短缺和技术瓶颈的双重挑战。一方面，算力产业需要跨学科、跨领域的复合型人才，如算法工程师、数据科学家、硬件工程师等。这些人才不仅需要具备扎实的专业知识，还需具备创新思维和解决问题的能力。然而，当前市场上这类人才供不应求，成为制约算力产业发展的关键因素之一。另一方面，全球算力产业的技术更新迭代速度极快，对技术创新和研发能力提出了极高要求。上海虽已拥有一批具有国际竞争力的算力企业和科研机构，但在面对全球算力技术的快速发展时，仍需加强技术创新和研发投入，推动算力技术的突破和应用。

#### (二) 发展建议

针对上海算力产业面临的算力需求与供给不平衡、算力中心节能减排和可持续发展、算力产业人才与技术短缺三大挑战，以下提出三点针对性的发展建议，旨在推动上海算力产业的高质量发展：

##### 1、构建多元化算力供给体系，优化算力资源配置

为解决算力需求与供给不平衡的问题，上海应构建多元化的

算力供给体系，优化算力资源配置。

一是统筹政府相关专项资金支持算力基础设施重大项目建设，在土地、网络、电力、能耗指标及市政配套建设等方面给予相应支持，加快建设高性能计算中心、国产化智算中心、园区级边缘智算设施和算力服务平台等，提高算力供给能力。支持上海异构国产芯片在单节点内以及跨节点的互联互通验证，同时吸引更多的国产芯片企业和算力调度、算力度量等研究机构联合开展验证，形成更大范围的共识和产品化的能力。

二是建议加快国家算力互联互通体系建设，推动各个建设的算力调度平台统一通过算力互联互通体系开展互联，制定互联互通平台和各个智算资源之间，以及与调度平台之间的统一标准，便于形成一张全国性的算力调度的互联互通体系。加快算力标识、算力度量等技术问题的研究，解决技术堵点，持续推动算力向电力一样按需获取的目标演进。

三是建议尽快开展算力登记、标注工作，逐步完成算能聚合，算能供需匹配，适度冗余的机制。针对长三角区域发展，依托市级第三方算力交易中心中立、合规、市场化特点，尽快构建长三角算能登记认证平台，统筹算力建设规模及形态，完善供需对接；同时，推动算力资源的跨区域、跨行业共享，打破信息孤岛，实现算力资源的优化配置。

## **2、推动绿色算力发展，实现节能减排和可持续发展**

针对算力中心节能减排和可持续发展的问题，上海应积极推

动绿色算力发展，实现节能减排和可持续发展。

一是加强算力中心的节能技术研发和应用，如采用先进的液冷散热技术、太阳能光伏技术、新型储能技术等，降低算力中心的能耗和碳排放。推动算力中心的智能化运营和管理，通过智能调度、动态调整算力资源等手段，提高算力中心的能源利用效率和资源利用率。

二是构建国内国际算力大通道，推动上海成为国际算力数据贸易港，大力推进国内国际算电协同、算数协同，发展绿色算力。建议以上海为起点，连通中西部省份，构建国内算电协同大通道；同时，以上海为起点，覆盖东亚地区，构建国际安全互信的算力-数据大通道，实现双向算力和数据的互通，推动算力走出去，数智产业引进来。

三是建议用能管理逐渐从能耗“双控”向碳排放“双控”转变，对于在算力中心所用的绿色能源，不再纳入能耗总量和强度控制，对于绿色能源就近建设的重点智算中心，对电力资源进行直供试点，不额外增加上网电价。

### **3、加强人才培养和技术创新，推动算力产业高质量发展**

一是加强高校和科研机构与企业的合作，推动产学研用深度融合，培养跨学科、跨领域的复合型人才。通过设立奖学金、实习实训基地、职称认定等方式，吸引更多优秀人才投身算力产业。

二是加强技术创新和研发投入，推动算力技术的突破和应用。鼓励企业加大研发投入，加强与国际先进企业的合作与交流，引



进和消化吸收国际先进技术，提高自主创新能力。

三是加强算力产业的国际合作与交流，推动算力技术的国际化和标准化发展。通过参与国际标准制定、举办国际学术会议等方式，提高上海算力产业的国际影响力和竞争力。

上海市通信学会“算力浦江”专业委员会  
2024年11月

