AI 兴起，智能算力浪潮来袭

核心观点

- **ChatGPT 等通用大模型的成功极大拉动对智能算力的需求，目前我国人均智能算力仍处于全球中等水平，发展智能算力对我国拥有重大战略意义。**随着 ChatGPT 带来的新一代 AI 浪潮，国内外 Bert、GPT4、文心一言等通用大模型相继发布，基于大模型的多场景应用也不断拓展，万亿级别参数的大模型以及各种垂直行业的应用极大地驱动了对智能算力的需求。人工智能一直是中美两国科技竞争的重要领域，中国的数据优势较为明显，然而算法和智能算力却明显落后于美国。近年来，我国智能算力快速增长，然而我国的人均算力仍处于中等水平，落后于美、英、德等国家。人均算力的水平与一国的智能化水平高度相关，我国积极发展智能算力、打造智算中心是打造国际竞争力、发展综合国力的关键。
- **国家推动多地智算中心建设，由东向西逐步扩展。科技部发文推动人工智能公共算力平台建设，要求使用一定比例国产算力以及国产开发框架。**据国家信息中心，未来 80%的场景都将基于人工智能。而普惠大众的智能算力就是 AI 发展的基础资源，将要像水、电一样驱动科技发展、社会进步，智算中心正式实现这些科技创新的源泉。近年来，国家制定一系列政策，在全国范围内推动信息基础设施建设，智算中心的建设是其中的关键环节。当前我国超过 30 个城市正在建设或提出建设智算中心，一般智算中心的的起步算力目标是 100P，整体布局以东部地区为主，并逐渐向中西部地区拓展。另外，科技部出台政策推动人工智能公共算力平台建设，要求在混合部署的公共算力平台中，自主研发芯片所提供的算力标称值占比不低于 60%，并优先使用国产开发框架，使用率不低于 60%，此举将推动我国 AI 芯片等国产化进程。
- **AI 发展将拉动 AI 芯片、服务器、AI 算力云服务等需求。** AI 需要多元异构算力提供支持，将极大拉动 GPGPU、TPU、NPU 等 AI 芯片的需求。目前，英伟达占据中国 AI 芯片市场 80%的份额，海光信息、寒武纪等芯片厂商崛起在即，产品性能提升明显，有望逐步实现国产替代。浪潮信息、中科曙光等服务器厂商也将持续受益于 AI 浪潮，浪潮信息的 AI 服务器在世界市场和中国市场均蝉联第一位，是 AI 服务器行业的顶尖巨头。而中科曙光是高性能计算的龙头，响应国家号召建设曙光 5A 级智算中心，覆盖全算力精度，赋能人工智能应用场景落地。拓维信息和四川长虹则是华为的亲密合作伙伴，依托“昇腾+鲲鹏”打造 AI 服务器。另外，用云服务提供 AI 算力的方式可以减少部署和管理本地计算基础设施的复杂性，优刻得、深桑达旗下中国电子云等第三方中立厂商有望参与 AI 云服务，持续受益于 AI 算力需求的提升。

投资建议与投资标的

随着智能计算资源需求的大幅增加， AI 芯片、AI 服务器及云计算算力需求将持续提升。

- **AI 芯片需求快速增长，国产化替代在即。**建议关注澜起科技(688008，买入)、海光信息(688041，买入)、寒武纪-U(688256，未评级)。
- **AI 计算需求提升有望持续拉动 AI 服务器需求，**建议关注浪潮信息(000977，未评级)、工业富联(601138，买入)、联想集团(00992，未评级)、中科曙光(603019，买入)、拓维信息(002261，未评级)、四川长虹(600839，未评级)。
- **着 AI 算法的计算需求不断增加，将有越来越多的企业使用云计算平台来满足其计算需求，中立云计算厂商有望持续受益。**建议关注深桑达 A(000032，未评级)、优刻得-W(688158，未评级)。

风险提示

AI 技术发展不及预期风险；芯片供应不足风险；国产化进度不及预期风险；通用大模型被禁用风险。

行业评级	看好（维持）
国家/地区	中国
行业	计算机行业
报告发布日期	2023 年 04 月 08 日

证券分析师

浦俊懿 021-63325888*6106
pujunyi@orientsec.com.cn
执业证书编号：S0860514050004

蔺剑 021-63325888*8514
kuaijian@orientsec.com.cn
执业证书编号：S0860514050005
香港证监会牌照：BPT856

陈超 021-63325888*3144
chenchao3@orientsec.com.cn
执业证书编号：S0860521050002

谢忱 xiechen@orientsec.com.cn
执业证书编号：S0860522090004

联系人

杜云飞 duyunfei@orientsec.com.cn
覃俊宁 qinjunning@orientsec.com.cn

相关报告

大模型应用百花齐放，AI 发展进入新时代	2023-03-27
OpenAI 发布插件功能影响 C 端生态，B 端应重视 OA 与办公软件入口潜力	2023-03-26
英伟达 GTC 召开，AI 应用广泛落地	2023-03-23

目录

图表目录

一、AI 带动智能算力需求，中国人均智能算力处于中等水平

1.1 ChatGPT 等通用大模型的发展拉动对智能算力的需求

机器学习进入大模型时代，ChatGPT 等通用大模型的训练迭代极大拉动对智能算力的需求。模型成功部署后，推理也将需要大量智能算力做支撑。从模型训练角度来说，据 J. Sevilla 等发布的文章《Compute Trends Across Three Eras of Machine Learning," 2022 International Joint Conference on Neural Networks (IJCNN)》，机器学习的训练计算大概可以分为三个时期。第一个时期为 2012 年之前，训练算力大致遵循摩尔定律，约每 20 个月翻一番。而进入深度学习时代，算力翻倍的速度加速至 5-6 个月。2015-2016 年左右开启了大模型时代，在这个时期，计算量增长变慢，翻倍时间约为 10 个月。但整体的训练计算量比深度学习时代的系统大 2 到 3 个数量级(OOM)。从 2022 年底，随着 ChatGPT 成功带来的新一代 AI 浪潮，国内外 Bert、GPT4、文心一言等通用大模型相继发布。这些大模型的训练需要千亿、甚至万亿级参数，以及上千 GB 的高质量数据，大模型的训练迭代将极大地拉动了智能算力的需求。另外，日后随着模型的成熟落地和推广，模型推理所需的智能算力也将逐渐增加，占比不断提高。

图 1：机器学习进入大模型时代，训练算力量级大幅提升，每十个月翻一倍

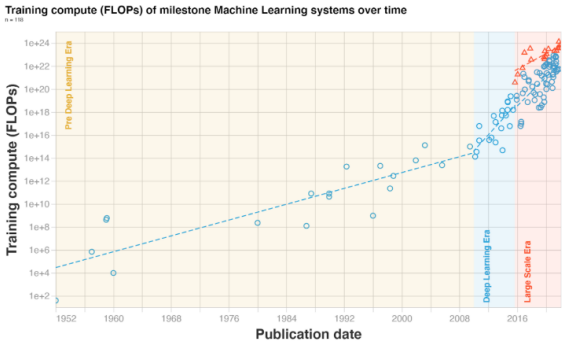
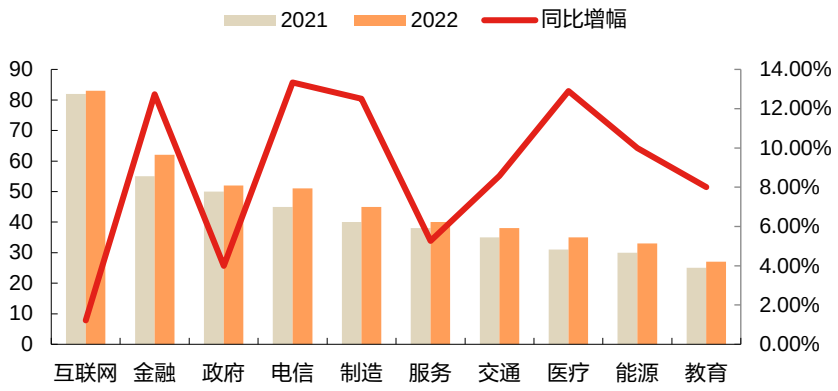


Figure 1: Trends in n=118 milestone Machine Learning systems between 1950 and 2022. We distinguish three eras. Note the change of slope circa 2010, matching the advent of Deep Learning; and the emergence of a new large scale trend in late 2015.

数据来源：J. Sevilla, L. Heim, A. Ho, T. Besiroglu, M. Hobbhahn and P. Villalobos, "Compute Trends Across Three Eras of Machine Learning," *2022 International Joint Conference on Neural Networks (IJCNN)*, Padua, Italy, 2022, pp. 1-8, doi: 10.1109/IJCNN55064.2022.9891914、东方证券研究所

除了通用大模型的训练，垂直行业大模型的训练、基于通用大模型的微调的行业应用也需要大量的智能算力做支撑。垂直行业的大模型训练也需要大量的智能算力，另外，基于大模型的多场景应用也不断拓展。AI 渗透千行百业，拉动智能算力规模高速增长。2022 年，各行各业的 AI 应用渗透度都呈不断加深的态势，尤其是在金融、电信、制造以及医疗领域，为实现业务增长、保持强大竞争力、从而占据更大的市场份额，企业纷纷入局 AI 领域，通过新技术提升传统业务用户体验，人工智能应用增长迅速。据 IDC 和浪潮信息联合发布的《2022-2023 中国人工智能计算力发展评估报告》，预计到 2023 年年底，中国将有 50%的制造业供应链环节采用人工智能技术实现业务体验提升。在未来，随着 AI 技术对传统行业赋能作用日益凸显，催生出更大智算需求成为必然。

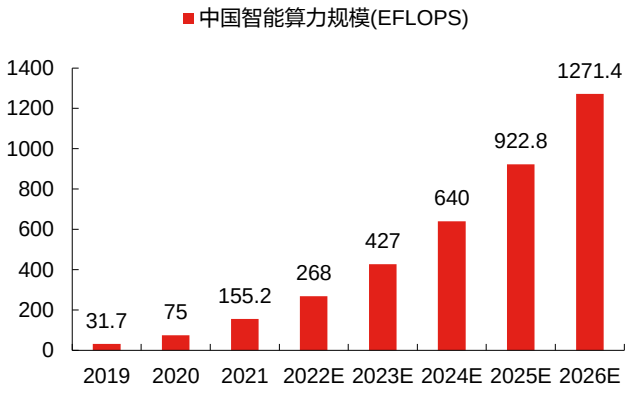
图 2：中国人工智能行业应用渗透度（%）及同比增长



数据来源：IDC、浪潮信息、东方证券研究所

智能算力规模高速增长，中国智能算力预计 2026 年突破 ZFLOPS 量级。在大模型取得突破、应用场景的广泛开拓与深入发展的背景下，智能算力需求将在未来几年迎来井喷。随着数据量高速增长、数据要素化进程推进、同时算力模型复杂度日益提升，智能算力作为释放数据价值的必要工具，其需求与规模快速增长。OpenAI 分析显示，从 2012 年以来，最大规模的 AI 模型训练中所需要的计算量，每 3.5 个月便翻倍一次。相比于摩尔定律 2 年的倍增期，算力需求具有远高于芯片承载计算量的增长速度。同时据 IDC 数据与预测，2021 年中国智能算力规模达到 155.2EFLOPS，并在之后的几年始终保持稳健增长态势，预计到 2026 年将突破进入每秒十万亿亿次浮点计算级别，智能算力实现 1,271.4EFLOPS 的庞大规模，2021-2026 年期间，预计年复合增长率达到 52.3%。

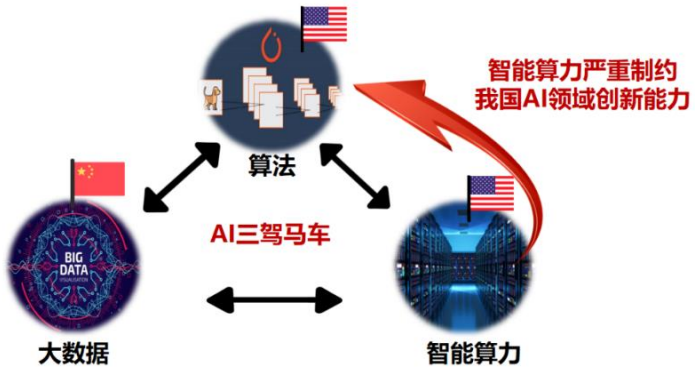
图 3：我国智能算力规模（EFLOPS）



数据来源：IDC、东方证券研究所

发展智能算力对我国拥有重大战略意义。人工智能一直是中美两国科技竞争的重要领域，而中美在 AI 的三大技术数据、算法、算力中各有优势。中国的数据优势较为明显，然而算法和智能算力却明显落后于美国。智能算力对于我国发展 AI、数据要素、以及社会全方位的发展都有着重大的战略意义。

图 4：AI 的三驾马车——数据、算法、智能算力



数据来源：郑纬民院士学术报告《人工智能算力基础设施的设计、评测与优化》、东方证券研究所

智能计算的算力精度主要为单精度、半精度或整形，推理相比训练所需算力精度更低。算力可分为通用算力、智能算力以及超算算力，对应着三种计算模式：基础计算、智能计算以及超级计算。不同的场景所需的算力种类不同，其对应的计算精度不尽相同。如一些产业数字化的场景对精度要求不高，通用算力（基础算力）即可满足需求。而例如天体物理、气象研究、航空航天等高精尖科研领域需要能够支持复杂运算、性能高的双精度算力，即超算算力。而对于人工智能的模型训练及推理来说，处理文字、语音、图片或视频等需求较大，单精度、半精度、甚至整型的计算即可满足应用需要。一般来说，相比于模型训练，模型推理所需的算力精度较低，很多场景 Int8 即可满足需要。

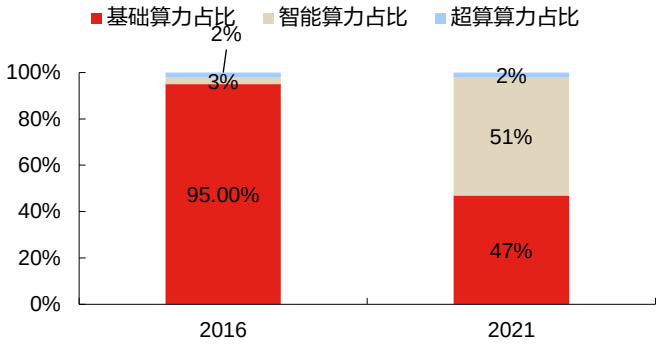
图 5：AI 计算场景愈加丰富，对智能算力的需求增加

数据来源：华为、东方证券研究所

1.2 我国智能算力逐年提高，但人均智算仍处于中等水平

我国智能算力占比已经超过通用基础算力达到 51%，预计未来持续高速增长，增长速度远超通用基础算力。我国智能算力占算力的比重也增长迅速。据中国信通院，2016 年，智能算力在我国算力中的占比仅为 3%，而 2021 年，我国智能算力占比已经超过基础算力，达到 51%，成为算力快速增长的驱动力。IDC 计算得出 2021 年到 2026 年期间中国智能算力规模年复合增长率为 52.3%，与同期预测得出的 18.5%通用算力规模年复合增长率相对比，可以展现出智能算力的井喷式增长态势。

图 6：2016-2021 年中国算力结构变化



数据来源：中国信通院、东方证券研究所

从智能算力总额来看，美、中处于领先地位。从人均智能算力的角度，中国仍处于全球中等水平。据《中国算力指数发展白皮书（2022）》，美、中的智能算力处于全球领先地位，分别占全球比重的 45%和 28%。然而从人均算力的高低来衡量，美国、英国、德国等国家的人均算力普遍高于 1000GFlops，而我国的人均算力处于中等水平。据 IMB 研究发现，人均算力的水平与一国的智能化水平高度相关，我国积极发展智能算力、打造智算中心是打造国际竞争力、发展综合国力的关键。

二、积极布局智算中心建设，孵化大模型助力 AI 发展

2.1 我国加快智算中心布局，三十城积极响应

2.1.1 智算中心是“东数西算”的关键，为社会提供智算资源

据国家信息中心，未来 80%的场景都将基于人工智能，所占据的算力资源将主要由智算中心提供。AI 大模型已经成为国家、企业和科研院所积极发展、重点投入的大方向。而普惠大众的智能算力就是 AI 发展的基础资源，将要像水、电一样驱动科技发展、社会进步。智算中心正式实现这些科技创新的源泉。

近年来，国家制定一系列政策，在全国范围内推动信息基础设施建设，智算中心的建设是其中的关键环节。2020 年 4 月 20 日，国家发展改革委首次明确新型基础设施范围，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范畴。2021 年 5 月，国家发改委等四部门联合发布了《全国一体化大数据中心协同创新体系算力枢纽实施方案》，提出布局全国算力网络枢纽节点。2022 年 2 月 17 日，国家发改委、中央网信办、工业和信息化部、国家能源局联合印发通知，同意在京津冀、长三角、粤港澳大湾区、成渝、内蒙古、贵州、甘肃、宁夏等 8 地启动建设国家算力枢纽节点，并规划了 10 个国家数据中心集群。“东数西算”工程正式全面启动。而智算中心承载的以模型训练为代表的非实时性计算尤为适合“东数西算”。

我国智算中心加快布局，孵化多行业大模型，推动 AI 应用落地。随着全国一体化算力网络和“东数西算”工程的部署，我国智能计算中心也加快布局。根据国家信息中心与相关部门联合发布的《智能计算中心创新发展指南》，当前我国超过 30 个城市正在建设或提出建设智算中心，一般智算中心的的起步算力目标是 100P，整体布局以东部地区为主，并逐渐向中西部地区拓展。比如天津的智能计算中心项目一期工程覆盖 850 余家企业及科研院所；成都智算中心去年上线，聚焦智慧医疗、智慧办公等应用场景。根据 ICPA 智算联盟统计，截至 2022 年 3 月，全国已投运的人工

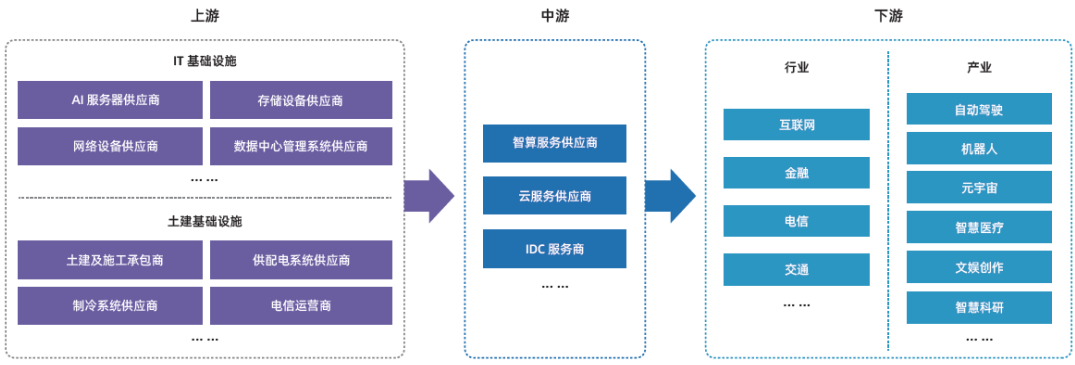
智能计算中心有近 20 个，在建设的人工智能计算中心超 20 个。各地方也结合本地产业特色，加快人工智能应用创新，聚合人工智能产业生态，例如武汉人工智能计算中心陆续孵化出紫东.太初、武汉.LuoJia 等大模型，加速推动 AI 在多模态交互、遥感等领域的落地应用。

图 7：多地积极推动智算中心建设

数据来源：郑纬民院士学术报告《人工智能算力基础设施的设计、评测与优化》、东方证券研究所

智算中心产业链涉及多关键环节，可提供数据服务、算力服务、算法服务、生态服务等全面服务。智算中心产业链上游主要由AI服务器供应商、存储设备供应商、网络设备供应商以及数据中心管理系统提供商等IT基础设施提供商以及土建施工承包商、供配电系统供应商、制冷系统提供商土建基础设施商。其中，AI芯片以及AI芯片间的互联是决定智算性能的关键。目前，AI加速计算主要是采用CPU系统搭载GPU、FPGA、ASIC等异构加速芯片，我国AI芯片在性能和软件生态方面还有待进一步提高。而NVLink和OAM两种高速互联架构可以保障多加速器间进行高速互联通信，提升模型训练效率，满足各领域场景和复杂的AI模型的计算需求。另外，大模型分布式训练对计算、存储都有高性能、易扩展的要求，同时低延迟、高带宽的网络有助于保障AI集群训练的高效。智算产业链中游为智算服务提供商、云服务供应商、IDC服务商等。行业下游则为互联网、金融等行业及各种先进产业的落地应用，为企业和科研单位提供数据服务、算力服务、算法服务、生态服务等多元化服务。

图 8：智算中心产业链



数据来源：智能计算中心创新发展指南、东方证券研究所

2.1.2 国家推动公共算力平台建设，引导使用国产算力及开发框架

科技部出台政策推动公共算力平台建设。依照《新一代人工智能发展规划》，科技部启动“人工智能驱动的科学研究”（AI for Science）专项部署工作，同时发布《科技部办公厅关于开展国家

新一代人工智能公共算力开放创新平台申报工作的通知》，提出要推进 AI 领域的模型与算法创新工作，加快推动国家新一代人工智能公共算力开放创新平台建设，支持高性能计算中心与智算中心异构融合发展。

公共算力平台对自主研发芯片以及国产开发框架的使用提出要求，推动 AI 基础设施国产化浪潮。公共算力平台的建设指引（下称《指引》）中要求在混合部署的公共算力平台中，自主研发芯片所提供的算力标称值占比不低于 60%，并优先使用国产开发框架，使用率不低于 60%。此外，在算力方面，对 AI 训练和推理的常用规格进行要求，16 位浮点 (FP16)性能应达到 400PFLOPS，32 位浮点 (FP32)性能应达到 200PFLOPS，16 位整型 (INT16)性能应达到 400POPS。同时，《指引》针对如今大模型训练需求的井喷需求，同时对环境承载能力提出要求，提出平台应配置成熟易用的人工智能全栈运行环境，能够运行千亿级参数的预训练模型，且同时实现至少三种典型人工智能业务场景的解决方案。

2.2 各地打造大算力智算中心，为社会提供澎湃算力

2.2.1 北京昇腾人工智能计算中心成立，长期实现 1000P 的算力规模

2023 年 2 月，北京昇腾人工智能计算中心正式点亮，首次采用市场化运作模式，智算中心打造产业发展新模式。2023 年 2 月 17 日，2023 中关村论坛首场系列活动——北京人工智能产业创新发展大会成功举办。在会上，北京昇腾人工智能计算中心正式点亮。该计算中心由北京市门头沟区政府联合中关村发展集团、华为公司携手打造，可为企业和科研单位等提供昇腾 AI 澎湃算力服务。北京昇腾人工智能计算中心实现了两个“首发”：首先，该中心采用市场化运作模式，积极探索人工智能“根技术+产业资本+产业政策”建设新模式。二是以自主创新根技术为依托，通过搭建“公共算力服务平台”、“应用创新孵化平台”、“产业聚合发展平台”、“科研创新与人才培养平台”四个平台，形成“普惠算力+创业服务+创新平台+优质人才”的产业发展新模式，打造具有特色产业的“京西智谷”。

图 9：2023 年 2 月，北京昇腾人工智能计算中心正式点亮及生态签约仪式

数据来源：华为计算、东方证券研究所

北京昇腾人工智能计算中心一期的算力规模实现 100P，未来将达到 1000P。北京昇腾人工智能计算中心一期算力规模达到 100P，而其首批签约的 47 家企业以及科研范围的算力使用规模预计为 248P。未来，该智算中心将持续扩容：短期将实现 500P 的算力规模，长期将为企事业单位、科研院所提供 1000P 的普惠算力，助力人工智能产业蓬勃发展。

北京昇腾人工智能计算中心赋能金融、医疗、出行等多行业的大模型。在会上，该智算中心与多模态量化金融大模型及其应用——MTGFinTech 正式签约，MTGFinTech 支持金融计量、金融随机分析模型，应用于风险管理、产品定价、资产管理等场景。与智算中心合作的还有全球首个视觉为中心的自动驾驶大模型——神行，该模型能够打通感知到预测全流程，推动自动驾驶规模化落地，应用于驾驶辅助、共享出行、城区配送等方面。

图 10：北京昇腾人工智能计算机中心大模型合作启动

数据来源：华为计算、东方证券研究所

2.2.2 商汤承建临港 AIDC 智算中心，算力规模远超同业水平

商汤承建的上海临港 AIDC 人工智能计算中心算力规模超群，可满足同时训练 20 个千亿参数量大模型。商汤 AIDC 计算中心占地超过 13 万平方米，一期机柜数量 5000 个，现已正式投入使用。中心通过部署 2.7 万块 GPU 实现超过 4910 Petaflops 的总算力供给，实现大幅超过业界绝大多数在 200-300 Petaflops 区间浮动的智算中心算力规模，通过完成行业十倍的规模提升实现了跨越式提高的用户体验。在庞大规模的算力支撑下，商汤 AIDC 可以同时满足训练 20 个千亿参数量大模型的性能要求，算力可支持的单个最大模型的参数量超过万亿，能够充分满足各大厂商积极研发大模型的基础设施需求，支撑中国 AI 智算产业高速发展。商汤 AIDC 的优点不仅仅局限于规模一项，还同时达到了低碳节能的绿色环保发展要求。商汤通过采取各种能源优化措施、实施离心系统并部署工业冷却制冷剂，通过提高每摄氏度 3%到 5%的冷却效率，有效实现 80%的降低幅度，年均 PUE 优化至 1.28，与国内其他数据中心相比能耗平均水平低约 10%。AIDC 是支撑商汤大模型的有力基础计算平台，同时也是向社会输送普惠算力的坚实底座。

图 11：上海临港 AIDC 人工智能计算中心

数据来源：商汤官网，东方证券研究所

三、芯片、服务器、云计算厂商有望持续受益于 AI 算力需求提升

随着智能计算资源需求的大幅增加，我们认为 AI 芯片、AI 服务器、以及云计算算力需求将持续提升。

- **AI 芯片需求快速增长，国产化替代在即**：AI 计算需要多元异构算力提供支持，将极大拉动 GPGPU、AISC 等 AI 芯片的需求。中国 AI 芯片市场规模有望快速增长，据艾瑞咨询发布的《2022 年中国人工智能产业研究报告(Ⅴ)》，预计 2027 年达到 2164 亿元。目前，英伟达凭借其 AI 芯片的超高性能，占中国加速卡市场的 80%以上。海光信息、寒武纪等巨头坚持迭代升级，其产品性能日益提升，有望获得更多市场份额，实现国产替代。另外，随着大模型的成熟部署，对性能要求稍低的推理芯片的占比将日益提升， 也有益于国产 AI 芯片占比提升。
- **AI 计算需求提升有望持续拉动 AI 服务器需求，浪潮信息、中科曙光等 AI 服务器龙头有望持续受益，拓维信息、四川长虹基于华为“鲲鹏+昇腾”的生产 AI 服务器，也有望受益。**相比于 AI 芯片，AI 服务器的技术壁垒较低，浪潮信息、中科曙光等中国服务器厂商占领了 AI 服务器绝大部分市场，有望持续受益于智能算力需求提升。其中，浪潮信息的 AI 服务器在世界市场和中国市场均蝉联第一位，是 AI 服务器行业的顶尖巨头。而中科曙光是高性能计算的龙头，响应国家号召建设曙光 5A 级智算中心，覆盖全算力精度，赋能人工智能应用场景落地。拓维信息和四川长虹则是华为的亲密合作伙伴，依托“昇腾+鲲鹏”打造 AI 服务器。
- **随着 AI 算法的计算需求不断增加，将有越来越多的企业使用云计算平台来满足其计算需求，中立云计算厂商有望持续受益。**云计算厂商可以为企业提供灵活的计算资源，帮助企业更好地管理其计算需求，提高其计算效率和灵活性，减少部署和管理本地计算基础设施的复杂性。除了华为云等国内外云计算巨头，优刻得、深桑达旗下中国电子云等第三方中立厂商也在积极参与 AI 云服务，有望持续受益于 AI 算力需求的提升。一方面，这些厂商可以作为客户除云大厂外的第二选择。另外，这些厂商具有中立方的身份优势，不会和下游客户产生竞争，更容易获得客户信任。

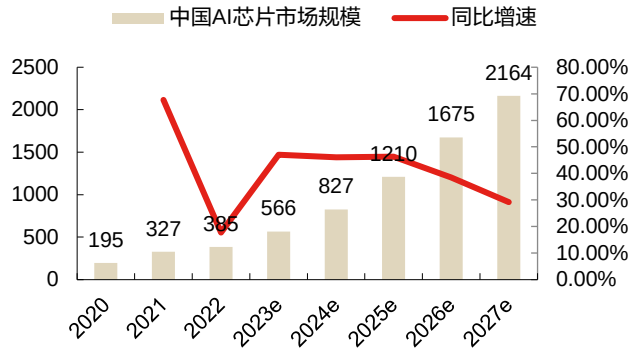
3.1 AI 芯片需求上涨，国产替代在即

AI 需要多元异构算力提供支持，拉动 AI 芯片需求。人工智能算法需要从海量的图像、语音、视频等非结构化数据中挖掘信息。从大模型的训练、场景化的微调以及推理应用场景，都需要算力支撑。而以 CPU 为主的通用计算能力已经无法满足多场景的 AI 需求。以 CPU+AI 芯片（GPU、FPGA、ASIC）提供的异构算力，并行计算能力优越、具有高互联带宽，可以支持 AI 计算效力实现最大化，成为智能计算的主流解决方案。服务器中的 CPU 和 AI 卡的数量并不固定，会根据客户应用需求调整，对于 AI 服务器来讲，较为常见的是配备 2 个 CPU，以及八个 AI 卡。而相比于 AI 服务器，传统的通用服务器则以 CPU 为主。因此，AI 的发展将极大拉动 GPGPU、TPU、NPU 等 AI 芯片的需求。

中国 AI 芯片市场将保持高速增长，AI 推理芯片份额有望持续提升，国产化 AI 芯片占比有望提升。2022 年，中国的 AI 芯片市场规模约 385 亿元。随着 AI 发展以及智算中心建设浪潮，该市场预计将保持高增长趋势。据艾瑞咨询测算，到 2027 年，中国的 AI 芯片市场规模预计将达到 2164 亿

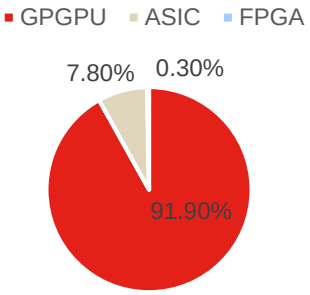
元。另外，在我国高性能及 AI 服务器中， GPGPU 凭借其优秀的性能和通用能力占比 92%, 剩下份额由 AISC 和 FPGA 分享。随着 AI 模型的优化落地，AI 推理芯片的占比将日益提升。据艾瑞咨询，2022 年，中国 AI 训练芯片以及 AI 推理芯片的占比分别为 47.2%和 52.8%。预计 2027 年，中国 AI 训练芯片与推理芯片的比例将分别达到 23.7%与 76.3%。相比于 AI 训练芯片，推理芯片的性能要求以及精度要求较低，部分国产 AI 芯片凭借其良好性能以及性价比能够满足推理端的需求，我国 AI 芯片国产化占比有望提升。

图 12：中国 AI 芯片市场规模（亿元）



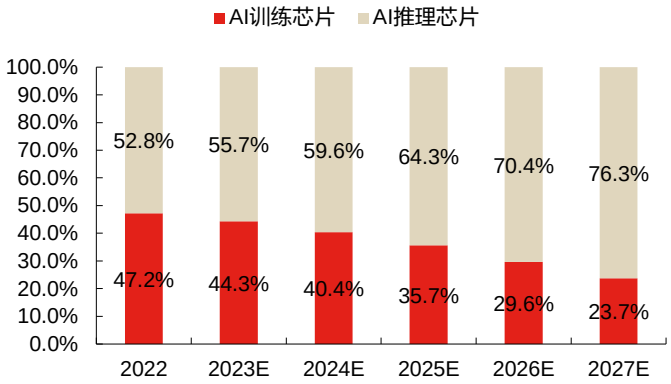
数据来源：艾瑞咨询、东方证券研究所

图 13：2021 年，中国高性能及 AI 服务器中各类 AI 芯片占比



数据来源：IDC、东方证券研究所

图 14：2022-2027 年中国 AI 芯片训练、推理比例

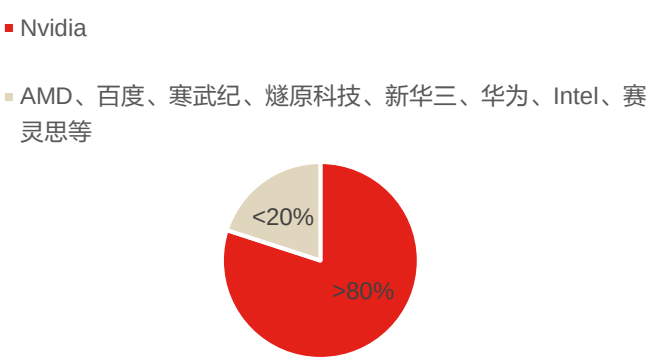


数据来源：艾瑞咨询研究院、东方证券研究所

AI 芯片领域的三类玩家。大模型的训练需要大规模的训练数据以及强大的计算资源，需要多卡多机协同完成。这对 AI 芯片本身的性能，以及多卡多机的互联提出了很高的要求。目前，在 AI 芯片领域，有三类玩家。一种是以 Nvidia、AMD 为代表的实力强劲的老牌芯片巨头，这些企业积累了丰富的经验，产品性能突出。另一种是以 Google、百度、华为为代表的云计算巨头，这些企业纷纷布局通用大模型，并自己开发了 AI 芯片、深度学习平台等支持大模型发展。如 google 的 TensorFlow 以及 TPU，华为的鲲鹏昇腾、CANN 及 Mindspore。最后是一些小而美的 AI 芯片独角兽，如寒武纪、壁仞等。

英伟达占据 80%以上中国加速卡市场份额，国产 AI 芯片亟待发展。根据 IDC 的数据显示，2021 年中国加速卡的出货数量已经超过 80 万片，其中 Nvidia 占据了超过 80%的市场份额。剩下的份额有 AMD、百度、寒武纪、燧原科技、新华三、华为、Intel 和赛灵思等品牌。

图 15：2021 年，中国 AI 加速卡市场份额占比

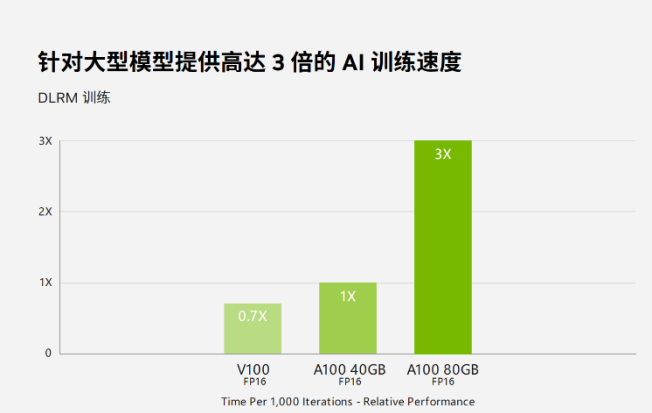


*以出货量口径
数据来源：IDC、东方证券研究所

3.1.1 英伟达：全球 GPU 龙头

英伟达占据芯片市场绝对优势。长期以来，英伟达在高端 GPU 市场占据绝对主导地位，现如今已量产的主流 A100 芯片相比前代产品 V100，性能得到显著提高，代表当今高端芯片水平。最新一代 H100 芯片也已经亮相，即将量产。天数智芯数据显示，2021 年英伟达在中国云端 AI 训练芯片市场的份额达到 90%。据 IDC，在 2021 年中国出货的 80 多万张加速卡中，英伟达占据超过 80%份额。芯片的研发周期较长，英伟达具有绝对先行优势，虽然目前国内企业突破英伟达垄断仍然任重道远，但寒武纪、华为 AI 芯片快速发展，有望逐步进行国产替代。

图 16：A100 的性能表现是 V100 的三倍



数据来源：NVIDIA 官网、东方证券研究所

受制裁影响，英伟达对部分产品性能进行“阉割”，推出“中国版芯片”A800、H800。2022 年 10 月，美国发布了针对中国的先进计算与半导体产品的出口管制，限制美国企业向中国出口先进高端芯片设备。在新管制的限制下，英伟达的 A100、H100 被禁止售卖给中国，而采用 12nm 工艺、性能较低的 V100 GPU 芯片不在管控之列。针对此次制裁，英伟达对 A100 的部分性能进行“阉割”，推出 A800。相比于 A100，A800 在单卡计算性能上没有差别，但是互联带宽从 600GB/s 下降到了 400GB/s，在一定程度上影响了如大模型训练等多卡互联场景的性能。目前，A800 已实现量产，并在中国规模化落地应用。英伟达还推出了旗舰芯片 H100 的替代版 H800，目前还未量产。

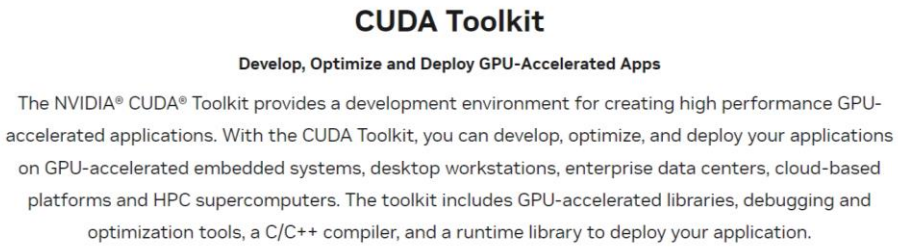
表 1：英伟达不同型号通用 GPU 参数规格对比

产品规格/型号	H100	A100	A800	V100
FP32	67 teraFLOPS	19.5 teraFLOPS	19.5 teraFLOPS	8.2 teraFLOPS
FP16 Tensor Core	1979 teraFLOPS*	624 teraFLOPS	624 teraFLOPS	16.4 teraFLOPS
INT8 Tensor Core	3958 TOPS	1248 TOPS	1248 TOPS	
GPU 显存	80GB	80GB	80GB	32GB
GPU 显存带宽	3.35TB/s	2039 GB/s	2039 GB/s	1134 GB/s
互连技术	NVLink：900GB/s PCIe 5.0：128GB/s	NVLi：600 GB/s PCIe 4.0：64 GB/s	NVLi：400 GB/s PCIe 4.0：64 GB/s	NVLi：300 GB/s PCIe 4.0：32 GB/s
最大热设计功率 (TDP)	700W	400W	400W	300W
发布时间	2022.03	2020.03	2022.11	2017.5

数据来源：NVIDIA 官网，东方证券研究所整理

CUDA（ Compute Unified Device Architecture ）是英伟达在 GPU 领域形成垄断的利刃。CUDA 是 NVIDIA 开发的一种并行计算平台和编程模型，它允许开发人员使用 C/C++语言在 NVIDIA GPU 上进行高性能计算。CUDA 的出现使得 GPU 不再只是用于图形处理，而是为英伟达的 GPU 提供了强大的计算能力，高效的数据处理能力和良好的并行性能。另外，CUDA 使英伟达能够在各个领域占据重要地位。例如，CUDA 可用于加速深度学习、人工智能、医疗、气象预测、金融建模等众多领域的计算任务，其开发平台应用广泛。

图 17：CUDA 使英伟达 GPU 在人工智能领域优势突出

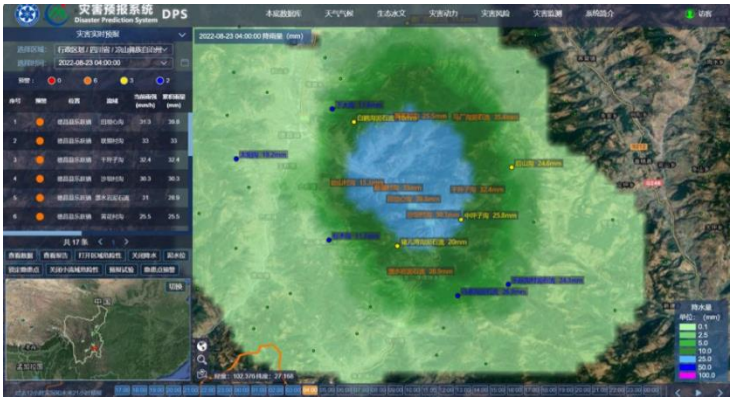


数据来源：Nvidia 官网、东方证券研究所

3.1.2 海光信息：国产高性能 CPU 和 GPGPU 领军企业

海光信息专注于研发、设计和销售高端处理器（CPU 以及 GPGPU），持续技术创新、产品迭代。海光信息的主要产品为应用于服务器和工作站等设备中的通用处理器（CPU）和协处理器（DCU，即 GPGPU）。海光处理器性能出众，同时软硬件生态丰富、工具链完整、应用迁移成本低。另外，海光 CPU 与 DCU 虽脱胎于 AMD，但经过多年独立自主研发迭代，已经实现自主可控、安全可靠，是国产芯片之光。 目前，苏州昆山、成都等多地超算中心已经搭载海光 CPU 与 DCU，为社会提供优质算力。

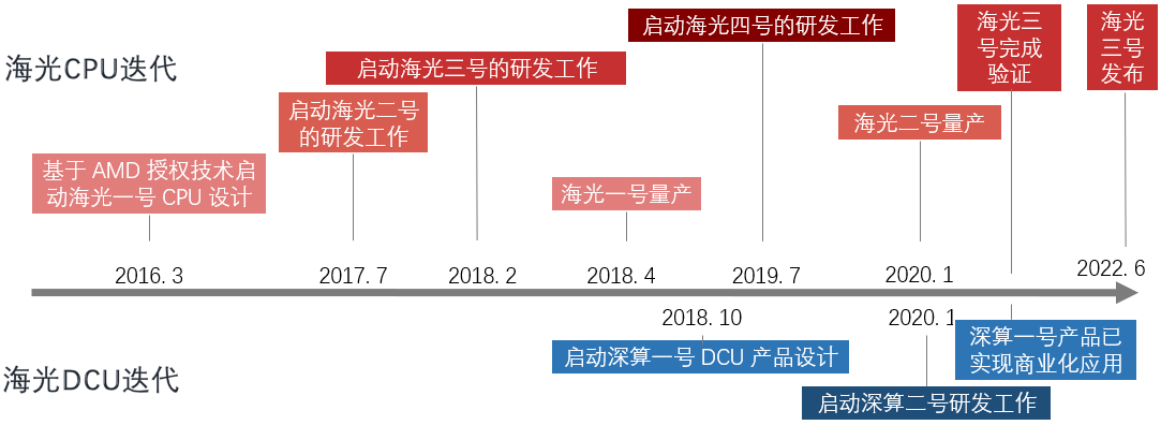
图 18：搭载海光 DCU 的成都超算中心为山地灾害风险模拟与险情预报系统提供算力支持



数据来源：澎湃新闻、东方证券研究所

海光 CPU 一、二代均已商业化，三代初亮相，四代有序研发中。海光 DCU 一代已商业化应用，二代研发中。公司持续技术创新和演进，坚持走“销售一代，验证一代，研发一代”的产品开发策略。公司建立了完善的高端处理器的研发环境和流程，持续开发多代产品，产品性能不断提高，同时功能不断完善丰富。海光 CPU 的四代产品中，海光一号和海光二号均实现了商业化应用，海光三号已亮相发布会，海光四号处于研发阶段。海光 DCU 于 2018 年启动 DCU 第一代产品深算一号的产品研发，于 2020 年 1 月启动了深算二号的研发，截至 2022 年 6 月，深算一号已实现商业化应用。

图 19：海光 CPU 与海光 DCU 演变情况



数据来源：海光信息招股书、东方证券研究所绘制

海光 DCU 某些硬件性能与英伟达的 A100、AMD 的 MI100 相近。海光 DCU 双精度计算能力突出。据北京大学高性能计算系统中标公告（HCZB-2021-ZB0364），海光信息的 DCU Z100 的通用计算核心达到 8192 个。其关键性能指标实现：FP64 10.8TFlops，显存 32GB HBM2，对比全球芯片巨头的高端 AI 芯片不遑多让。英伟达 A100 的相关指标为：FP64 9.7 TFlops、显存 40/80GB HBM2。AMD MI100 的相关指标为：FP64 11.5 TFlops、显存 32GB HBM2。

图 20：海光 Z100、NV A100、AMD MI100 性能对比

	Nvidia A100	AMD MI100	Hygon Z100
FP64(全精度算力)	9.7 TFLOPS	11.5 TFLOPS	10.8TFLOPS
GPU 显存	40/80GB HBM2	32GB HBM2	32GB HBM2

数据来源：采招网、英伟达官网、AMD 官网、东方证券研究所绘制

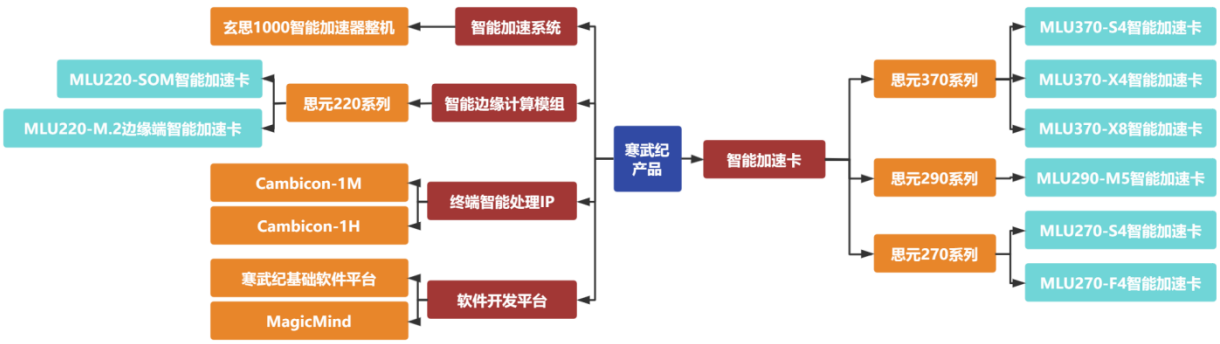
海光 DCU 生态丰富，工具链完整。海光的 DCU 脱胎于 AMD，兼容主流生态——开源 ROCm GPU 计算生态，支持 TensorFlow、Pytorch 和 PaddlePaddle 等主流深度学习框架、适配主流应用软件。ROCm 又被称为类 CUDA，现有 CUDA 上运行的应用可以低成本迁移到基于 ROCm 的海光平台上运行。

2022 年，海光发布国内首个全精度（FP64）异构计算平台，该平台搭载 CPU 海光三号和 DCU 海光深算，涵盖数值模拟、AI 训练、AI 推理所需的多样算力，实现了智能计算与数值运算的深度融合。同时，此平台可全面支持 TensorFlow、PyTorch、Caffe2 等主流 AI 深度学习框架，目前已超过 1000 种应用软件部署在该平台上。

3.1.3 寒武纪：国产 AI 芯片先行者

寒武纪始终深耕芯片研发，不断推陈出新、实现技术进步。寒武纪成立于 2016 年，专注人工智能芯片产品的研发与创新。公司成立之初便开始了对 AI 芯片领域的探索创新。并在 2016 年年底成功研发出全球首款 AI 手机芯片——寒武纪 1A 。2017 年，这款芯片被搭载于华为的高端系统级芯片麒麟 970，应用于 Mate10 手机，并获得了广泛好评。芯片可以在功耗极低的前提下，涵盖人脸识别、语音识别、图像增强等多种功能。此后，寒武纪又陆续推出了多款 AI 芯片产品，包括云端训练芯片 MLU100、边缘推理芯片 MLU270、车载推理芯片 MLU290 等 。这些产品都具有高性能、低功耗、高集成度等特点，在图像识别、语音识别、自然语言处理等领域都有着优异的表现。

图 21：寒武纪产品体系



数据来源：寒武纪官网，东方证券研究所

思元 370 是寒武纪的首款训练推理一体芯片，也是其云端产品的第三代。思元 370 采用了 7nm 制程工艺，并成为首款采用 Chiplet 技术的人工智能芯片。该芯片最大算力可达 256TOPS(INT8)，

是上一代云端推理产品思元 270 算力的两倍，同时该芯片还支持 LPDDR5 内存，内存带宽是 270 的三倍，因此可以在板卡有限的功耗范围内为人工智能芯片分配更多的能源，从而输出更高的算力。思元 370 智能芯片还采用了先进的 Chiplet 技术，支持灵活的芯粒组合，仅用单次流片便可以实现多款智能加速卡产品的商用。目前，该公司已推出三款加速卡：MLU370-S4、MLU370-X4 和 MLU370-X8，包含应用于计算密度高的数据中心、针对专注人工智能推理相关业务的互联网厂商需求和应用于对算力带宽要求高的训练任务，满足用户的多样化需求。

新一代训练芯片寒武纪 590 还未量产，据悉训练能力突出。寒武纪最新一代云端智能训练芯片思元 590 还未正式发布，据寒武纪董事长在 2022 WAIC 上介绍，思元 590 采用全新的 MLUarch05 架构，实测训练性能较在售产品有了显著提升。思元 590 可提供更大的内存容量和更高的内存带宽，其 PCIe 接口也较上代实现了升级。

图 22：寒武纪训推一体思元 370 系列

数据来源：寒武纪官网、东方证券研究所

3.1.4 百度昆仑芯：性能优越、生态蓬勃，是支持文心一言的坚实底座

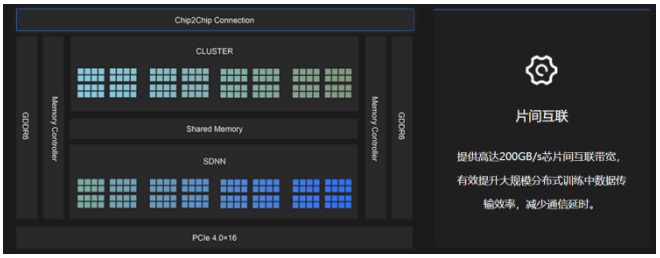
脱胎百度，昆仑芯大力投入研发设计，自主芯片进展迅速。昆仑芯前身是百度内部的智能芯片及架构部门，2011 年 6 月成立并开始探索 AI 计算与芯片相关研究，公司于 2021 年完成业务分拆，成为一家独立公司并完成融资。公司早期主要工作为依托 FPGA 芯片完成人工智能的计算加速，2017 年部署的 FPGA 芯片累计超过 12000 片。昆仑芯研发与应用落地并重，探索深耕技术与应用落地。昆仑芯 2018 年开始着眼 AI 芯片的自研工作，重磅推出昆仑芯一代，正式步入研发轨道。2020 年,第一代昆仑芯实现大规模部署，2021 年昆仑芯二代重磅推出，成为业界前列的 AI 芯片之一。2022 年，第二代昆仑芯实现了在数据中心、工业领域、自动驾驶等多领域的大规模部署与落地，攻克技术落地难关。

图 23：昆仑芯发展历程

数据来源：公司官网，东方证券研究所绘制

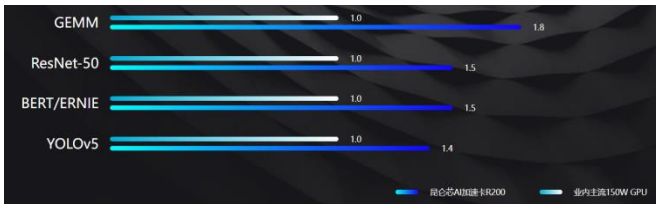
昆仑芯科技持续深耕技术领域与芯片开发，其中昆仑芯二代性能出众、领跑同业。昆仑芯二代 AI 芯片使用新一代昆仑芯 XPU-R 架构，实现通用性、易用性与高性能的显著提升，提供 256TOPS@INT8 以及 128 TFLOPS@FP16 两种类型算力，覆盖多种性能要求。不仅如此，在技术角度，相比于第一代昆仑芯的 14 纳米支撑，昆仑芯二代 AI 芯片基于 7nm 工艺打造，同时也是业界第一颗配备 GDDR6 高速显存的 AI 芯片，相比一代产品通用计算核心算力提升 2-3 倍，可为数据中心高性能计算提供强劲 AI 算力。除此之外，昆仑芯为提升产品的适配性，持续着力优化芯片架构等底层和新技术，以适配人工智能应用与其他各类算法，所支持的算法涵盖 TensorFlow、Pytorch、PaddlePaddle 等主流深度学习开发框架，同时在性能测试中展现出了比业界主流方案更加优秀的性能功耗比和性价比。到 2022 年底，昆仑芯二代已经出货数万片，是为数不多的经过互联网严苛场景淬炼的 AI 芯片。不仅昆仑芯二代性能出众，昆仑芯已着手新一代芯片的研发设计，目前第三代、第四代 AI 芯片的研发工作已经启动，三代产品预计将于 2024 年实现量产落地。

图 24：昆仑芯二代，片间互联可以达到 200GB/s



数据来源：公司官网、东方证券研究所

图 25：昆仑芯二代 AI 加速卡 R200 与业界主流方案测试性能对比



数据来源：公司官网、东方证券研究所

重视生态构建与落地可能性探索，昆仑芯为行业树立先行标杆。AI 芯片的技术研发是基本前提，但对大规模场景的需求也使得落地问成为“卡脖子”的难点堵点。为充分发挥二代 AI 芯片的性能优势，昆仑芯致力构建具备高延续性与软硬件适配性的自研架构，实现了包含框架、服务器、CPU 与操作系统在内的完整生态组成，甚至完成了非主流产品的适配工作。不仅如此，依托百度大平台，昆仑芯完成了与百度飞桨与百舸的原生适配，高度匹配基于飞桨的推理和训练模型产品，深度耦合百度的 AI 生态，同时实现了包括搜索、小度、无人驾驶在内的 AI 业务场景落地，以 AI 芯片为基点，实现完整 AI 计算生态的构建。

图 26：昆仑芯软件架构

数据来源：公司官网，东方证券研究所

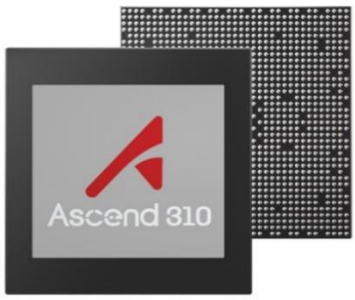
3.1.5 华为：打造完善“鲲鹏+昇腾”生态

华为提前布局，确立“一云两翼、双引擎”的产业格局。其中的一云指华为云，“双引擎”指围绕“鲲鹏”与“昇腾”两大中心打造的基础芯片族，两翼指智能计算业务以及智能数据与存储业务。早在 2004 年，华为就开始投资研发第一颗嵌入式处理芯片，维持着“量产一代、研发一代、规划一代”的芯片研发节奏，通过全栈创新构建面向世界的普惠服务，在智算领域提供面向端、边、云的“鲲鹏+昇腾+x86+GPU”的多样性算力。在“双引擎”中，华为构造基于“昇腾”系列的 AI 处理器和基础软件的 Atlas 人工智能计算解决方案，覆盖模块、板卡、小站、服务器、集群等丰富的产品形态，全方位覆盖深度学习领域的训练过程。

着眼构筑硬件算力底座，华为推出昇腾 310 和昇腾 910。昇腾 310 是华为首款全栈全场景人工智能芯片，采用自研华为达芬奇架构 NPU，通过集成丰富计算单元提高 AI 计算的完备度，可以输出 16TOPS@INT8, 8TOPS@FP16 两种类型算力，可以有效承接 AI 训练与推理使用，具有广泛的适用性。且芯片功耗仅为 8W，助力实现低碳化、绿色化算力供应，同时，芯片具备全 AI 业务流程加速，通过大幅提高性能有效降低部署成本。而随着研发深入，华为推出了业界算力最强的 AI 处理器——昇腾 910，该芯片通过借助研华为达芬奇架构 3D Cube 技术，实现了业界最佳 AI 性能与能效，半精度（FP16）算力达到 320 TFLOPS，整数精度（INT8）算力达到 640 TOPS，完全达到设计规格，更进一步满足 AI 训练与推理使用。不仅如此，华为还推出了 Altas300 系列 AI 加速卡，涵盖 Atlas 300I Duo 推理卡、Atlas 300I Pro 推理卡、Atlas 300V Pro 视频解析卡与 Atlas 300T Pro 训练卡，促进智能算力的场景化落地，满足涵盖互联网、运营商、金融等领域对 AI 训练与高性能计算的智算需求。

图 27：昇腾（HUAWEI Ascend）310

图 28：昇腾（HUAWEI Ascend）910



数据来源：全爱科技官网，东方证券研究所

数据来源：全爱科技官网，东方证券研究所

华为构建 CANN 人工智能框架，助力 AI 生态健康高速发展。华为不仅注重企业自身发展，同时也积极承担社会责任，维护健康开发者生态。作为昇腾 AI 异构计算架构，CANN 从 2018 年初露峥嵘，到目前最新发布的 6.0 版本，在使能 AI 开发效率和性能方面始终保持业界前列。经过数年的研究积累和技术优化，CANN 现已包含 1400 余个高性能算子，已完成主流 AI 框架的算子加速需求覆盖，不断深入完善支撑神经网络训练和推理加速的服务供给。作为昇腾 AI 基础软硬件平台的核心，CANN 肩挑上层深度学习框架与底层 AI 硬件，实现两者之间的互联互通。框架全面支持昇思 MindSpore、飞桨（PaddlePaddle）、PyTorch、TensorFlow 等主流 AI 框架，提供能够覆盖众多典型场景应用的 900 多种模型，兼容多种底层硬件设备，提供强大异构计算能力，从模型优化、系统分析、模型部署等多维度帮助开发者扫除重重障碍。

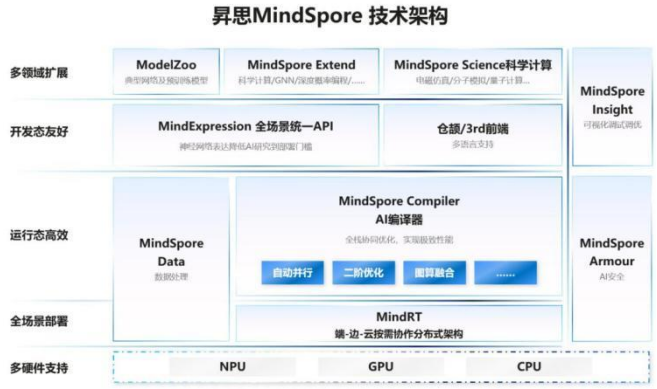
图 29：华为 CANN 人工智能框架技术架构

数据来源：中关村在线，东方证券研究所

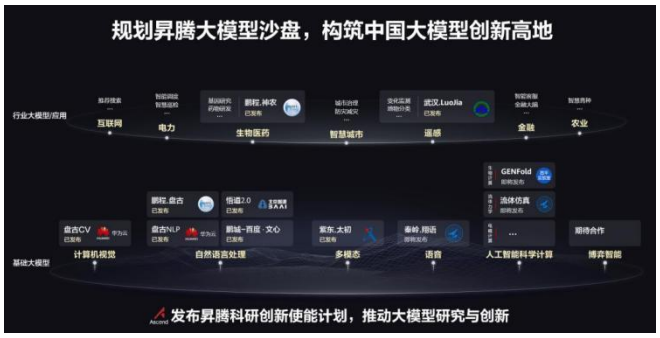
为科研机构提供研发底座，昇思 MindSpore AI 框架为大模型发展提供蓬勃活力。在 ChatGPT 爆红带领的大模型研发时代浪潮下，昇思 MindSpore AI 框架原生支持训练的特性成为其后续发展大模型的关键。不仅如此，框架的开源也促进科研机构以其为根基，打造了一系列大模型，其中包括鹏城实验室的业界首个 2000 亿参数中文预训练语言模型鹏程.盘古和面向生物医学领域的鹏程.神农大模型，中科院自动化所的业界首个图文音三模态大模型紫东.太初，以及武汉大学的全球首个智能遥感框架及数据集武汉.LuoJia。从这些大模型来看，华为已经形成了一套清晰和成熟的支持大模型发展的路径，即使科研机构在其技术底座上不断进行模型开发。在可预见的未来中，借助着昇腾 AI 提供的强大算力底座，昇思 MindSpore AI 框架将得到高速发展，并与算力底座相辅相成、互相抚弄，所形成的良性循环将进一步为大模型的研发注入活力。

图 30：华为昇思 MindSpore 技术架构

图 31：以昇思 MindSpore 为基，大模型高速发展



数据来源：机器之心，东方证券研究所

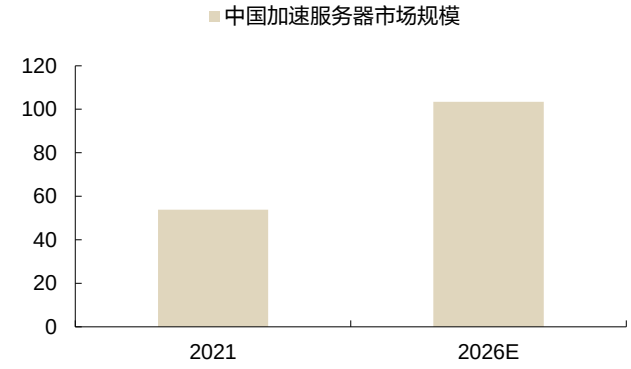


数据来源：机器之心，东方证券研究所

3.2 中国 AI 服务器市场有望快速增长

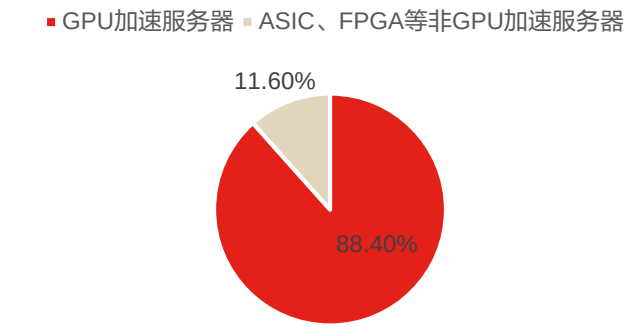
中国加速服务器市场有望快速增长，GPU 服务器仍占主导，非 GPU 服务器增长迅速。据 IDC，2021 年，中国加速服务器市场规模达到 53.9 亿美元，同比增长 68.6%。其中，GPU 服务器仍然占据主导地位，市场份额近 90%。与此同时，NPU、ASIC 和 FPGA 等非 GPU 加速服务器以 43.8%的增速实现了 11.6%的市场份额，达到 6.3 亿美元。另外，IDC 预测，到 2026 年，中国加速计算服务器市场将达到 103.4 亿美元。

图 32：中国加速服务器市场规模（亿美元）



数据来源：IDC、东方证券研究所

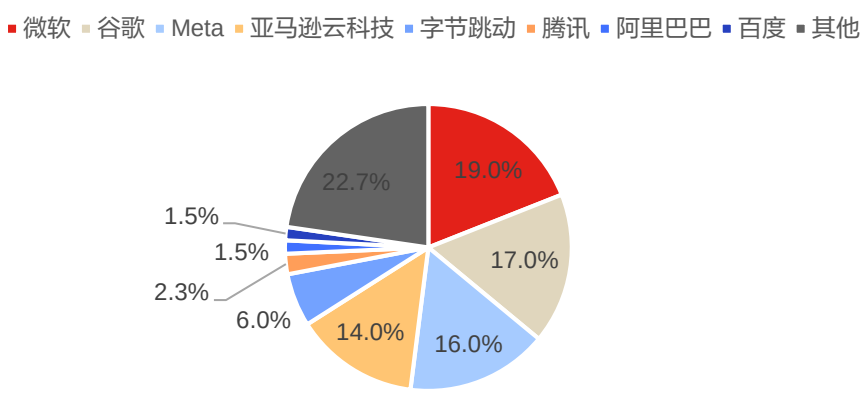
图 33：2021 年，中国 GPU 和非 GPU 加速服务器市场份额占比



数据来源：IDC、东方证券研究所

全球云计算巨头是 AI 服务器采购的主力军。 2022 年，在 AI 服务器采购方面，北美科技巨头微软、谷歌、Meta、亚马逊云科技仍然占据着最大的市场份额，合计达到 66%。随着中国对 AI 发展建设的重视，AI 基础设施建设也在加速，字节跳动、腾讯、阿里巴巴以及百度的 AI 服务器采购量分别为 6%、2.3%、1.5%以及 1.5%。

图 34：2022 年，各业者服务器采购量占比



数据来源：TrendForce、东方证券研究所

3.2.1 浪潮信息：全球 AI 服务器第一大品牌商

浪潮电子信息产业股份有限公司是浪潮集团有限公司旗下三家上市公司之一，是全球领先的 IT 基础架构产品、方案及服务提供商。公司主要从事服务器等云计算基础设施产品的研发、生产和销售，业务覆盖计算、存储、网络三大关键领域，提供包括云计算、大数据、人工智能、边缘计算等全方位数字化解决方案。公司以“智慧计算”为战略，构建开放融合的计算生态，为客户构建满足多样化场景的智慧计算平台，全面赋能传统产业的数字化、智能化转型与变革，重视算力基础设施的建设和发展，为中国数字经济发展提供充足的算力支持。

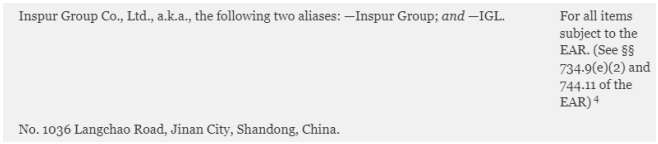
图 35：浪潮 AI 服务器为 AI 巨头长期保持深入合作

数据来源：浪潮信息官网、东方证券研究所

浪潮集团被美国列入实体清单，浪潮信息更改注册地址免于受限。2023 年 3 月 6 月，美国将浪潮信息的控股股东浪潮集团新列入商务部实体清单。该实体清单明确列出了浪潮集团的地址，即中国山东省济南市浪潮路 1036 号。浪潮信息原本的公司注册地与浪潮集团相同。为应对此次制裁，2023 年 3 月 6 日，浪潮发布公告，修改公司注册地址为“济南高新区新泺大街 1768 号齐鲁软件园大厦 B 座 B302”，希望在此次实体清单事件的受到的影响尽可能减小。

图 36：浪潮集团被列入实体清单

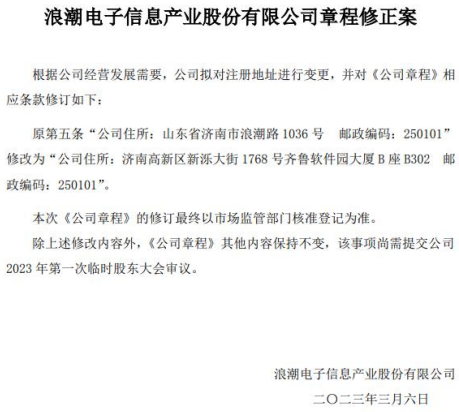
图 37：浪潮信息更改公司地址公告


Inspur Group Co., Ltd., a.k.a., the following two aliases: —Inspur Group; *and* —IGL.

For all items subject to the EAR. (See §§ 734.9(e)(2) and 744.11 of the EAR)[4]

No. 1036 Langchao Road, Jinan City, Shandong, China.

数据来源：联邦公报、东方证券研究所


浪潮电子信息产业股份有限公司章程修正案

根据公司经营发展需要，公司拟对注册地址进行变更，并对《公司章程》相应条款修订如下：

原第五条“公司住所：山东省济南市浪潮路 1036 号 邮政编码：250101”修改为“公司住所：济南高新区新泺大街 1768 号齐鲁软件园大厦 B 座 B302 邮政编码：250101”。

本次《公司章程》的修订最终以市场监管部门核准登记为准。

除上述修改内容外，《公司章程》其他内容保持不变，该事项尚需提交公司 2023 年第一次临时股东大会审议。

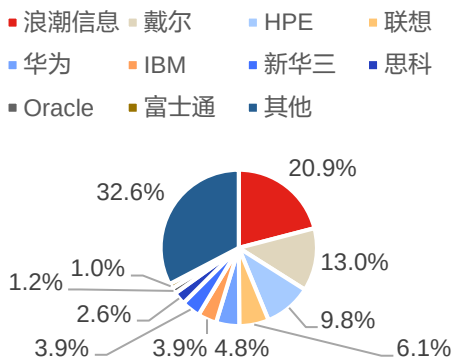
浪潮电子信息产业股份有限公司

二〇二三年三月六日

数据来源：浪潮信息、东方证券研究所

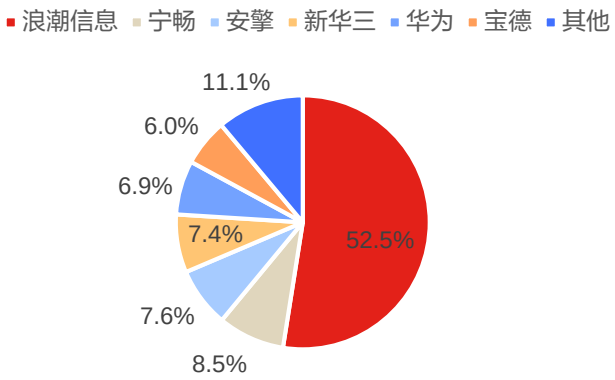
浪潮信息的 AI 服务器在世界市场和中国市场均蝉联第一位，是 AI 服务器行业的顶尖巨头。根据国际数据公司 IDC 发布的 2021H2《全球人工智能市场半年度追踪报告》，2021H1 全球 AI 服务器市场规模 156 亿美元，浪潮信息在世界 AI 服务器市场占有率达 20.9%，份额同比提升 3.6 pct，销售额同比增长 68.3%，蝉联全球第一。另外，据 IDC 发布的《2021 年下半年度（H2）中国加速计算服务器市场报告》，在中国市场，浪潮 AI 服务器占有率达 52.5%，连续 5 年保持中国 AI 服务器市场份额超过 50%。

图 38：2021H1 全球 AI 服务器市场份额比例



数据来源：IDC、东方证券研究所

图 39：2021H2 中国 AI 服务器市场份额比例



数据来源：IDC、东方证券研究所

3.2.2 工业富联：全球服务器 ODM 龙头

工业富联已形成相关技术、产品储备。2022 年上半年，公司推出模块化服务器，包括支持 X86 与 ARM 架构的运算模块、管理模块与接口模块。2022 年，公司首发两款经权威机构认证的基于 ARM 架构主流高性能多核云服务器，为全球云服务提供商及企业数字化转型提供强大助力。新一代先进冷却技术及解决方案是公司未来成长的重要支撑之一，公司持续加大数据中心节能技术的研发，推出先进冷却解决方案，通过沉浸式与机柜式液冷散热系统，实现节约成本及提升效率的目标。同时，公司积极开拓 HPC 相关业务，取得了国内外大型云服务商客户认可，有望分享到 HPC 行业快速成长红利。

图 40：工业富联子公司鸿佰科技自研的先进液体冷却解决方案

数据来源：鸿佰科技、东方证券研究所

携手英伟达提高产品性能。公司宣布采用基于 NVIDIA HGX、OVX 和 CGX 系统设计的 NVIDIA Grace CPU 和 NVIDIA Grace Hopper Superchip，以满足超级数据中心及边缘运算等更高的算力需求。NVIDIA Grace CPU 是专为现代数据中心设计的突破性中央处理器，与当今领先的处理器相较，其提供最高的性能和 2 倍内存容量及能效。它可以满足需要高效能运算、数据分析、数字孪生、云端游戏等对运算能力具有严格要求的应用，透过性能、容量、能源效率和可配置性的再度提升，为需要超大规模计算应用的相关服务，提供更佳资源。Grace Hopper Superchip 是把 NVIDIA Grace CPU 与基于 NVIDIA Hopper 架构的 GPU 配对，此集成模块可服务于高效能运算和大规模 AI 应用程序，可将大至兆字节运算的应用程序性能提高 10 倍，为科学家和研究人员提供强大的运算效能支持。公司也将推出搭载 NVIDIA Grace CPU 超级芯片的新服务器系统，该系统将在模块上使用 NVLink-C2C 和 LPDDRX 连接两个 CPU 芯片，取消了传统服务器的 DIMM 插槽，大幅提高散热效能，该系统也弹性支持额外的高性能 PCIe 卡和 DC-SCM 模块。

图 41：NVIDIA Grace CPU 以及 Grace Hopper Superchip

数据来源：挖贝网、东方证券研究所

公司 CSP 业务优势明显，生产经营全球化布局。公司持续投入自主技术研发，业务覆盖数据中心、云服务、高性能计算、边缘计算等领域，主要客户涵盖全球市场占有率较高的头部品牌商、北美前三大 CSP 服务商、国内头部 CSP 服务商及互联网应用服务企业，出货量位居全球领先地位。公司前瞻性布局全球化生产制造基地及供应链，在中国大陆、中国台湾、匈牙利、捷克、越南、墨西哥、美国等多个国家及地区建立制造基地；在中国大陆、中国台湾、中国香港、美国、新加坡、捷克、匈牙利、墨西哥、越南、印度、日本等多个国家及地区均开展经营业务，可以满足全球客户的全球交付需求。公司的全球数字化管理系统，可实现柔性调配生产与供应链资源，高效、迅速地满足客户区域生产及全球交付需求，通过对芯片、工业软件等领域的投资布局，为客户提供更低价、更强韧的一站式供应链服务，有效抵御宏观经济风险冲击，为业务持续增长保驾护航。

3.2.3 联想集团：全球第三大服务器品牌

公司已经成为全球第三大服务器提供商，位据全球 HPC 榜单 TOP500 榜首，具备交付全球客户的能力，ISG 为代表的业务集团有望成为第二增长曲线。

基于联想集团在高性能计算方面的专业与经验，韩国气象厅携手联想，为其建造最新高性能计算机——“五号”，并已在韩国气象厅下属的国家气象高性能计算中心正式投入运行。从硬件性能来说，“五号”是全世界最顶尖的高性能计算机之一，峰值性能达到 50PFLOPs，即 5 亿亿次每秒浮点运算能力。以当今全球高性能计算机顶尖性能榜单来估计，“五号”可以排进前十名。在实际运算中，“五号”被分成两台互为备份的高性能计算机“Maru”与“Guru”，它们互相配合防止天气模型预测失误。在 2021 年 6 月发布的全球高性能计算机 TOP500 榜单中，“Maru”与“Guru”分别位列全球第 23 位和 24 位。“五号”基于联想 ThinkSystem SD650-V2 服务器架构，引入了最新的处理器核心和领先的网络和存储技术，同时使用联想服务器的核心技术之一“海神”直接式温水水冷技术。

图 42：联想为韩国国家气象局提供高性能计算机

数据来源：联想官网、东方证券研究所

紫金云在甘肃省运营的大型数据中心，是助力“东数西算”工程的集群之一。因此，紫金云需要大规模扩展计算能力，增加新的高性能计算和大数据存储资源以满足项目需求。同时，紫金云数据中心作为甘肃省发展数字经济的重要基础平台，对高性能计算和大数据存储资源有极高要求。数据中心不仅要求打造一个技术领先的高性能计算平台，还需要成熟的建设运营经验和相关技术人才推进落地。紫金云选择了联想的 ThinkSystem 服务器和存储解决方案。具体来看，联想为整个紫金云搭建了高性能计算平台的系统、存储、网络、登录管理系统、集群系统软件等整个运行环境：1）安装了 100 个联想 ThinkSystem SD630 V2 高密度机架式计算节点（搭载第三代英特尔至强可扩展处理器）及 50 个联想 ThinkSystem SR670 V2 GPU 节点（搭载 4 个 NVIDIA A100 Tensor Core GPU）；2）部署了一个高度并行化的存储系统：用于 IBM Spectrum Scale 的联想分布式存储解决方案，总存储容量接近 10 PB；并通过高速的 Mellanox InfiniBand HDR 网络架构与联想服务器相连。

图 43：紫金云高性能计算平台五大特点

高性能计算节点	强大的单节点计算性能	并行存储系统架构	节点间网络高速带宽访问	核心数据备份功能
• 包含高密度机架式计算节点和GPU节点两类，支持异构加速。 • 配置50台联想GPU服务器，每台服务器采用了4块NVIDIA A100 GPU显卡，以满足科学计算、人工智能等应用场景。	• 采用主流计算节点配置，配置100台CPU计算节点，采用目前先进的英特尔®至强® Platinum 8358 32C 2.6GHz处理器，提供强大的单节点计算性能。	• 性能优越支持大规模I/O并发处理。 • 存储系统裸容量近10PB，总聚合I/O读带宽：50GB/s，写带宽：45GB/s。	• 采用高速HDR Infiniband高速网络互联 • 计算、存储网络采用目前业界先进100Gb HDR Infiniband高速网络，实现计算和存储网络融合设计，全线速无阻塞。	• 保障数据安全

数据来源：联想官网、东方证券研究所

3.2.4 中科曙光： 高性能计算龙头

中科曙光是中国信息产业与高性能计算领域领军企业。中科曙光成立于 2006 年，背靠中国科学院计算所，经过多年发展，公司于2014年在上海证交所上市。作为中国核心信息基础设施领军企业，中科曙光专注于高性能计算领域，基于公司多年积累的高端计算优势，积极布局智能计算、云计算、大数据等领域，主要业务涉及高端计算机、存储产品、云计算服务、网络安全产品等。目前，中科曙光主要有两款智能计算服务器产品：X785-G30 与 X785-G40。

图 44：中科曙光 X785-G30：HPC、深度学习训练/推理

数据来源：中科曙光官网、东方证券研究所

图 45：X785-G40：训练与推理功能的全能型 GPU 服务器

数据来源：中科曙光官网、东方证券研究所

中科曙光拥有国际领先的 5 大智能制造生产基地、7 大研发中心，在全国 50 多个城市部署了城市云计算中心。公司充分发挥高端计算优势，大力发展人工智能、云计算、大数据等先进计算业务。根据国家规划、产业发展和政企用户需求，公司打造了全新智能算力基础设施-曙光 5A 级智算中心，提供全场景人工智能计算服务，包括算力供给、算法优化、数据服务和行业应用。曙光5A级智算基础设施采用分布式异构并行体系结构，搭载多种类型的芯片，可提供多样化的算力供应，覆盖全算力精度，以满足不同的人工智能应用场景和多种用户需求为区域和行业的智能化发展提供支持。

图 46：中科曙光 5A 级智算中心

图 47：曙光智算中心布局

数据来源：中科曙光官网、东方证券研究所

数据来源：中科曙光官网、东方证券研究所

3.2.5 华为： 打造超强 AI 集群，提供 AICC 全栈解决方案

具备超强全栈能力，华为打造超强算力 AI 集群与昇腾 AICC 全栈解决方案。凝结数千颗昇腾 910 AI 处理器，华为推出 Atlas 900 AI 集群，在全球范围内立于算力产业巅峰。集群通过整合多种高速接口，借助华为集群通信库与作业调度平台，充分发挥昇腾 910 的强大性能，可提供相当于 50 万台 PC 计算能力，达到 256P～1024P FLOPS @FP16，ResNet-50@ImageNet 性能居业界前列。集群在液冷方面也有出众表现，通过部署单柜 50KW 混合液冷系统来支撑支撑＞95%液冷占比，从而有效节省机房空间 79%，增加数据中心算力密度。不仅如此，华为响应国家号召，推出昇腾 AICC 全栈解决方案，实现从底层基础到落地应用的全面覆盖，集成异构计算架构 CANN、全场景 AI 计算框架 MindSpore、全流程开发工具链 MindStudio、昇腾应用使能 MindX 四大核心软件，建设具备复杂训练与数据处理能力的人工智能计算中心，可满足不同行业的 AI 领域模型开发、训练和推理等多样化计算需求。通过提供廉价算力显著节约成本，从而有效提升研发效率，产出更大的经济效用。

图 48：华为昇腾 Atlas 900 AI 集群

数据来源：华为官网、东方证券研究所

华为与鹏城实验室共建鹏城云脑，打造超级 AI 算力平台。作为设备提供商之一，华为积极参与“东数西算”计划落地，通过布局通用计算的鲲鹏计算以及 AI 计算的昇腾计算，来实现高效普惠的计算范式，其研发的搭载鲲鹏、昇腾处理器的 Atlas 900 集群达到了全球训练最快的极高性能。鹏城实验室具有强大的研发人才资源，华为与其合作共建，围绕'鲲鹏+昇腾'双引擎全面启航计算战略，构建“鹏城云脑”工程。“鹏城云脑”是一项重要的大型科学设施，旨在支持国家重大战略、满足基础研究需求以及推动数字经济发展。目前，已经完成了开源开放的 AI 技术试验平台

“鹏城云脑Ⅰ”，正在建设兼顾研究和赋能的大规模 AI 算力平台“鹏城云脑Ⅱ”，“鹏城云脑Ⅱ”可提供不低于 1000P ops 的整机 AI 计算能力和 64PB 的高速并行可扩展存储，配备 200PB 存储和百 GB 级网络传输速率，其 AI 算力处于国际领先水平。两千亿参数中文 NLP 大模型盘古以及生物医药领域的大模型神农都是在鹏程云脑上研发。另外，智能超级算力平台“鹏城云脑 Ⅲ”的预研已经启动，

图 49：鹏城云脑机房

数据来源：鹏城实验室官网，东方证券研究所

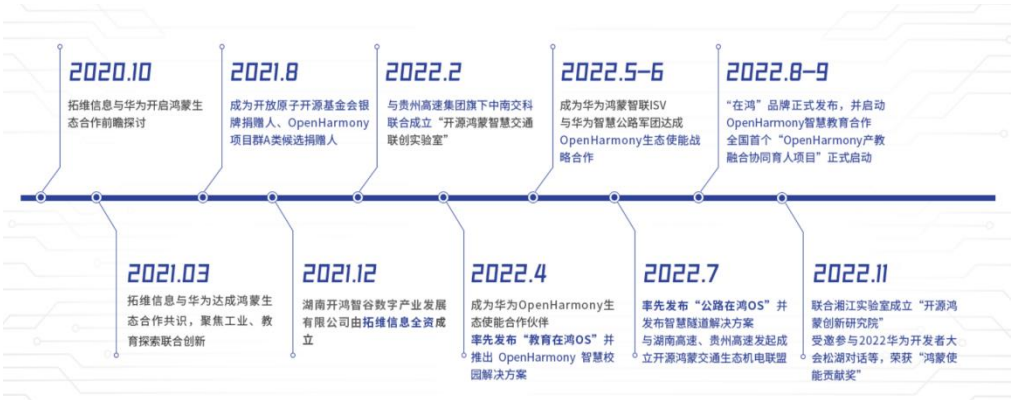
图 50：“鹏城云脑 II”连续三届获得“AIPerf500”榜单冠军

数据来源：绿色计算产业联盟，东方证券研究所

3.2.6 拓维信息：华为“鲲鹏昇腾”战略合作伙伴

拓维信息是中国领先的软硬一体化产品及解决方案提供商。公司成立于 1996 年，2008 年上市。业务覆盖政企数字化、智能计算、鸿蒙生态，布局全国 31 省、海外 10+国家，聚焦数字政府、运营商等重点领域和行业，服务超过 1500 家政企客户，提供全栈国产数字化解决方案和一站式全生命周期的综合服务。拓维信息紧紧跟随数据经济发展浪潮，已构建了智能计算、鸿蒙生态与政企数字化三大类型产品。同时，作为华为云首批“同舟共济”战略合作伙伴，拓维信息自 2017 年起开始布局华为云业务，携手华为和众多软硬件生态伙伴打造全栈国产数智化产品及解决方案，并基于华为技术底座，由子公司湘江鲲鹏构造“兆瀚”自主计算产品。

图 51：拓维信息与华为携手共创生态



数据来源：拓维信息官网，东方证券研究所

依托华为“昇腾+鲲鹏”技术架构，拓维信息积极构建 AI 智算解决方案。在 AI 智算催生大量算力需求的背景下，拓维信息与华为联合探索算力生态建设可能，其子公司湘江鲲鹏先后发布基于鲲鹏处理器和昇腾处理器的数十款兆瀚 AI 产品，包括 AI 推理、AI 训练、AI 小站、AI 集群、智能边缘等。同时，作为华为“昇腾智造”及“昇腾智行”双领域解决方案合作伙伴，拓维信息积极探

索落地场景，研发出基于昇腾的 AI 稽核、质检等解决方案，助力解决经济产业发展中的痛点堵点。不仅如此，拓维信息积极推动国产智能算力的发展，在贵州、甘肃两个算力枢纽与其他生态建设者一同探索“东数西算”国家工程的落地实践，深入参与长沙、重庆人工智能算力中心建设。其中，拓维信息在贵州与贵安产控、云上贵州共同注资设立的云上鲲鹏以“平台+生态”方略，助力多领域行业数字化转型，目前已在教育、交通等 8 大行业研发出了 33 个优秀行业应用案例，致力促进大数据与实体经济深度融合。

图 52：“兆瀚”产品体系

数据来源：湘江鲲鹏官网、东方证券研究所

3.2.7 四川长虹：参股华鲲振宇提供澎湃算力

立足华为技术底座，助力长虹转型升级，华鲲振宇参与数据设施新基建。2020 年 6 月，四川长虹为实现计算产业的战略化转型升级，联手各大资方建立华鲲振宇，并由华为提供技术底座，全面负责基于“鲲鹏+昇腾”处理器的“天宫”自主品牌系列产品的生产销售全流程服务。华鲲振宇积极参与包含江西、福建、成都、济南在内的多省市智算中心建设，公司研发的“天宫”昇腾 AI 全栈基础软硬件平台为中心提供核心基础设施。以成都为例，华鲲振宇承建的人工智能算力平台提供高达 300P 的 AI 计算能力，相当于 15 万台高性能计算机的计算能力。在“天宫”昇腾技术底座提供的坚实基础之上，成都智算中心已和近 100 家企业与高校科研机构合作共创，孵化近 200 个人工智能解决方案。

图 53：华鲲振宇参建的成都智算中心

数据来源：成都市工业经济和信息化研究院，东方证券研究所

图 54：天宫 AI 训练服务器 AT800 Model 9000

数据来源：华鲲振宇官网东方证券研究所

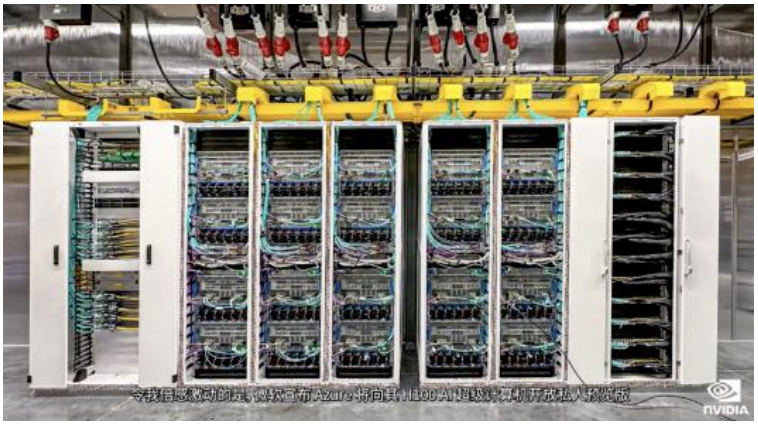
AT800 Model 9000 算力与能效兼备，有效满足数据中心绿色化要求。华鲲振宇注重研发，近期推出的基于鲲鹏+昇腾技术架构研发的 AI 训练服务器 AT800 Model 9000 展现出了优越性能，在

ChatGPT 高速发展催生巨额算力需求下，可以有效填补当前算力不足的供给缺口。服务器已实现覆盖部件到整机的全栈自主可控，可广泛应用于 AI 大模型开发、训练和推理。不仅如此，AT800 还可以在保证整体高性能和稳定性的前提下进行全栈适配调优，可以有效满足不同场景下的业务要求。服务器在算力密度、网络带宽和能效比方面表现卓越，不仅能够提供 2.56 PFLOPS FP16 超强算力，8*100G RoCE v2 高速接口有效缩短 10~70%的芯片间跨服务器互联时延，还提供业界 1.3 倍的 2.56 PFLOPS/5.6 kW 超高能效比，促进数据中心绿色化、提高能源利用率。

3.3 众多厂商积极布局“云上”AI 算力

英伟达发布 DGX 云服务，提供云上算力。近日，“军火商”英伟达发布了 DGX 云服务，提供专用的 NVIDIA DGX AI 超级计算集群，并且搭配了 AI 软件，使客户可以通过网络浏览器访问 AI 超算。每个 DGX 云实例由八个 A100、或是 H100 80G Tensor Core GPU 支持。目前，该服务已经与 Azure、Google GCP、以及 Oracle OCI 开展合作。

图 55：英伟达发布 DGX 云服务，提供云上算力



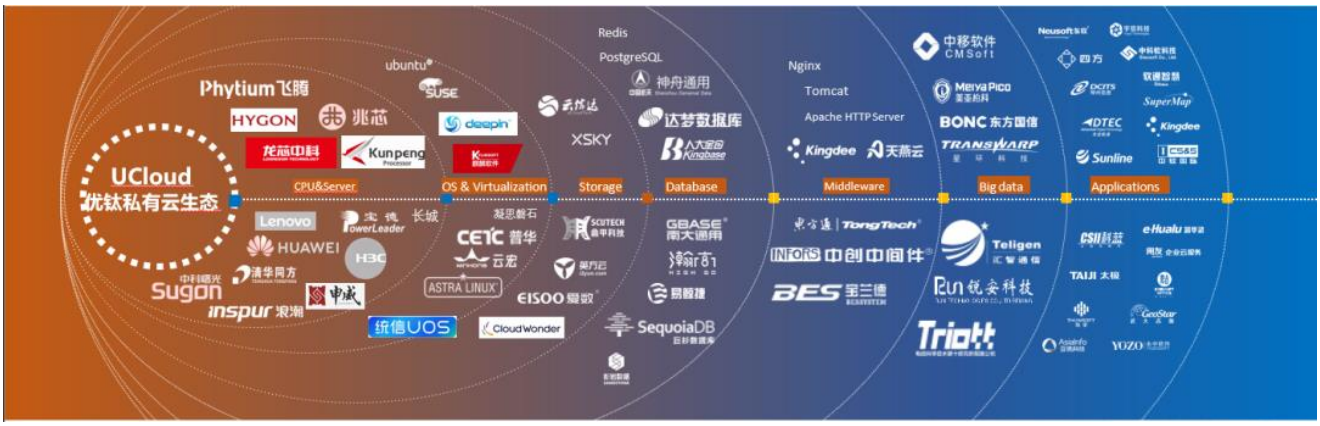
数据来源：英伟达、东方证券研究所

基于云计算的 AI 能力逐步得到验证。用云服务提供 AI 算力的方式可以减少部署和管理本地计算基础设施的复杂性。随着 AI 技术的发展，中国的 AI 公有云服务业随之增长。据 IDC，2022 年中国 AI 公有云服务市场规模将达到 74.6 亿元，较 2021 年呈上升趋势。目前，中国 AI 公有云服务厂商市场格局较为稳定。据 IDC，2022H1，百度智能云在中国 AI 公有云服务市场占比第一，阿里云位居第二，且与第一名差距逐渐减小。华为云、腾讯云紧跟其后，市场份额不断提升。除了头部云厂商，优刻得、深桑达旗下中国电子云等第三方中立厂商也有望参与 AI 云服务，持续受益于 AI 算力需求的提升。

3.3.1 优刻得提供多种云计算服务，积极适配 AI 领域智算需求

成立于 2012 年的优刻得是中国第一家公有云科创板上市公司。优刻得自主研发并提供计算、网络、存储等 IaaS 和基础 PaaS 产品，不涉足客户业务领域，致力建立中立、安全的云计算服务平台。公司依托全球部署的 31 大高效节能绿色云计算中心和国内 11 个线下服务站，提供涵盖公有云、私有云和混合云在内的多种云服务，并基于云计算推出综合性行业解决方案。同时优刻得已完成与数十家厂商的兼容性测试和互认，联合形成信创生态，可以通过一个云平台兼容管理多款信创体系 CPU 芯片。

图 56：优刻得私有云生态体系



数据来源：优刻得官网，东方证券研究所

积极适配爆发智能算力需求，优刻得为 AI 客户提供算力支持。UCloud 自 2018 年起，加快建设、应用部署在国家算力枢纽节点的上海青浦云计算中心和内蒙古乌兰察布云计算中心，通过自建数据中心提供自由部署、负载灵活的定制化服务，为互联网和传统行业的大中型客户提供更具性价比的定制化“混合云”解决方案。两大数据中心结合既有计算资源，实现“前店后厂”式的云资源规模化延展效应。在承建的乌兰察布云计算中心中，优刻得部署多种 GPU 高性能计算产品，机柜设计功率覆盖 4.4-8.8kW，可满足用户对机房等级、系统架构、单机柜功率等多样使用需求的量身定制。同时，为更好满足由大模型训练推理带来的井喷算力需求，优刻得积极研发，目前公司提供的 GPU 服务器可支持多种卡型 GPU 资源，包括 A100、V100S、MI100 多种类型显卡，适配训练、推理等用户需求，同时将智能算力与其他公有云资源打通，降低管理难度、实现便捷调度，为 AI 客户提供算力支持，促进人工智能产业发展。

图 57：内蒙古乌兰察布云计算中心

数据来源：优刻得官网，东方证券研究所

图 58：上海青浦云计算中心

数据来源：优刻得官网，东方证券研究所

3.3.2 深桑达建立中国电子云，致力建设自研数据底座

以 PKS 技术架构体系为基，深桑达成立中国电子云，提供安全算力服务。中国电子云成立于 2021 年，是深桑达以中国电子 PKS 自主安全计算体系为底座建立的数据基础设施，包含国产化自研可信的计算架构和分布式云原生操作系统，提供涵盖 IaaS、PaaS 和 SaaS 能力的专属公有云服务。中国电子云依托的 PKS 体系脱胎于中国电子多年的深厚技术积累，名称中的“P”代表飞腾 CPU，“K”指的是麒麟操作系统，“S”即安全，中国电子云的自研原生技术架构是保障其安全性的最大依托，体系中使用的飞腾 CPU 和麒麟操作系统均为中国电子自主研发，同时中国电子云交付的 60%硬件设备基于国产芯片，有效防范过度依赖进口芯片带来的断供危险。凭借着

在安全方面的出众表现，PKS 体系在信创领域具备最大的市场份额，飞腾 CPU 达到 74%的市场份额，麒麟操作系统甚至高达 87%。

图 59：PKS 体系技术架构

数据来源：中国电子官网，东方证券研究所

国资云赋予中国电子云独特信任优势，与智源研究院联手探索本土硬件适配，助力实现智能算力国产化。随着云计算领域不断发展推进，政务机构和大型国企催生了巨额上云需求，市场份额不断扩张，发展空间广阔。云计算开源产业联盟数据显示，2021 年我国政务云市场规模达 802.6 亿元，预计 2023 年将增长至 1203.9 亿元。而中国电子云脱胎于中国电子，出身天生赋予其国内政企领域的充分信任，且自研架构也从技术角度保障了安全性，有效防范后门风险、漏洞风险和断供风险，保障业务过程中的“本质安全+过程安全”。但目前，中国电子云底层硬件中大比例使用的国产芯片在实现应用迁移时面临挑战，一定程度上阻碍了原生架构的建设，难以实现完备的智能算力本土替代化。为建立更完备的自研技术体系，中国电子云与智源研究院联手，共同建立“大模型国产算力云平台开放实验室”，合作探索实现国产 CPU 的大模型适配部署，未来有望以国产算力支撑 AI 领域的算力需求。

3.3.3 中科曙光：人工智能云计算平台提供稳定高效算力

曙光人工智能云计算平台提供一站式深度学习训练与实时推理。曙光人工智能云计算平台解决方案可以提供基于云的 GPU 计算服务。该系统以主流深度学习框架为基础，支持 Caffe、TensorFlow 等多种主流深度学习框架。该平台与 Kubernetes 和 Docker 容器技术相结合，提供实验环境、离线任务和在线服务三大功能，支持业务从模型研究、批量训练到在线预测的全流程打通。该解决方案提供一站式深度学习训练/实时推理、图形图像处理以及科学计算等，大大简化了企业构建深度学习平台的难度，提高了资源使用率，降低了业务投入成本，使用户更加专注于深度学习应用本身，是目前性价比最高的整体 AI 训练与推理解决方案之一。

图 60：人工智能云计算平台解决方案

数据来源：中科曙光官网、东方证券研究所

四、投资建议及相关标的

随着智能计算资源需求的大幅增加， AI 芯片、AI 服务器及云计算算力需求将持续提升。

- **AI 芯片需求快速增长，国产化替代在即。**建议关注澜起科技(688008，买入)、海光信息(688041，买入)、寒武纪-U(688256，未评级)。
- **AI 计算需求提升有望持续拉动 AI 服务器需求**，建议关注浪潮信息(000977，未评级)、工业富联(601138，买入)、联想集团(00992，未评级)、中科曙光(603019，买入)、拓维信息(002261，未评级)、四川长虹(600839，未评级)。
- **随着 AI 算法的计算需求不断增加，将有越来越多的企业使用云计算平台来满足其计算需求，中立云计算厂商有望持续受益。**建议关注深桑达 A(000032，未评级)、优刻得-W(688158，未评级)。

五、风险提示

AI 技术发展不及预期风险：若我国人工智能技术发展不及预期，有可能影响对智能算力的需求。

芯片供应不足风险：若 CPU 及 GPU 等高性能芯片供应不及预期，有可能影响算力供给。

国产化进度不及预期风险：若 AI 芯片国产化进度不如预期，有可能影响国产 AI 芯片供给。

通用大模型被禁用风险：若 ChatGPT 等通用大模型被禁用，有可能影响对智能算力的需求。

分析师申明

每位负责撰写本研究报告全部或部分内容的研究分析师在此作以下声明：

分析师在本报告中对所提及的证券或发行人发表的任何建议和观点均准确地反映了其个人对该证券或发行人的看法和判断；分析师薪酬的任何组成部分无论是在过去、现在及将来，均与其在本研究报告中所表述的具体建议或观点无任何直接或间接的关系。

投资评级和相关定义

报告发布日后的 12 个月内的公司的涨跌幅相对同期的上证指数/深证成指的涨跌幅为基准；

公司投资评级的量化标准

买入：相对强于市场基准指数收益率 15%以上；

增持：相对强于市场基准指数收益率 5%～15%；

中性：相对于市场基准指数收益率在-5%～+5%之间波动；

减持：相对弱于市场基准指数收益率在-5%以下。

未评级 —— 由于在报告发出之时该股票不在本公司研究覆盖范围内，分析师基于当时对该股票的研究状况，未给予投资评级相关信息。

暂停评级 —— 根据监管制度及本公司相关规定，研究报告发布之时该投资对象可能与本公司存在潜在的利益冲突情形；亦或是研究报告发布当时该股票的价值和价格分析存在重大不确定性，缺乏足够的研究依据支持分析师给出明确投资评级；分析师在上述情况下暂停对该股票给予投资评级等信息，投资者需要注意在此报告发布之前曾给予该股票的投资评级、盈利预测及目标价格等信息不再有效。

行业投资评级的量化标准：

看好：相对强于市场基准指数收益率 5%以上；

中性：相对于市场基准指数收益率在-5%～+5%之间波动；

看淡：相对于市场基准指数收益率在-5%以下。

未评级：由于在报告发出之时该行业不在本公司研究覆盖范围内，分析师基于当时对该行业的研究状况，未给予投资评级等相关信息。

暂停评级：由于研究报告发布当时该行业的投资价值分析存在重大不确定性，缺乏足够的研究依据支持分析师给出明确行业投资评级；分析师在上述情况下暂停对该行业给予投资评级信息，投资者需要注意在此报告发布之前曾给予该行业的投资评级信息不再有效。

免责声明

本证券研究报告（以下简称“本报告”）由东方证券股份有限公司（以下简称“本公司”）制作及发布。

。本公司不会因接收人收到本报告而视其为本公司的当然客户。本报告的全体接收人应当采取必要措施防止本报告被转发给他人。

本报告是基于本公司认为可靠的且目前已公开的信息撰写，本公司力求但不保证该信息的准确性和完整性，客户也不应该认为该信息是准确和完整的。同时，本公司不保证文中观点或陈述不会发生任何变更，在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的证券研究报告。本公司会适时更新我们的研究，但可能会因某些规定而无法做到。除了一些定期出版的证券研究报告之外，绝大多数证券研究报告是在分析师认为适当的时候不定期地发布。

在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议，也没有考虑到个别客户特殊的投资目标、财务状况或需求。客户应考虑本报告中的任何意见或建议是否符合其特定状况，若有必要应寻求专家意见。本报告所载的资料、工具、意见及推测只提供给客户作参考之用，并非作为或被视为出售或购买证券或其他投资标的的邀请或向人作出邀请。

本报告中提及的投资价格和价值以及这些投资带来的收入可能会波动。过去的表现并不代表未来的表现，未来的回报也无法保证，投资者可能会损失本金。外汇汇率波动有可能对某些投资的价值或价格或来自这一投资的收入产生不良影响。那些涉及期货、期权及其它衍生工具的交易，因其包括重大的市场风险，因此并不适合所有投资者。

在任何情况下，本公司不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任，投资者自主作出投资决策并自行承担投资风险，任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

东方证券研究所

地址：上海市中山南路 318 号东方国际金融广场 26 楼

电话：021-63325888

传真：021-63326786

网址：www.dfzq.com.cn