

谷歌AI芯片团队并入云部门 旨在对抗微软和OpenAI?

本报记者 李玉洋 上海报道

在由 ChatGPT 掀起的新一轮人工智能(AI)竞赛中,科技巨头谷歌也得奋力追赶。

科技媒体 *The Information* 报道,谷歌确认将其专门负责 AI 芯片开发的团队转移至谷歌云(Google Cloud)部门,以增强云计算团队的竞争力,追赶强敌 AWS 和微软 Azure。

《中国经营报》记者注意到,谷歌此前还宣布将 DeepMind 和 Google Brain 两个人工智能研究实验室合并为一个叫 Google DeepMind 的部门。

全联并购公会信用管理委员会专家安光勇对记者表示:“这表明谷歌正在采取积极措施来应对竞争对手

AI服务很大程度依托于云

谷歌把 DeepMind 和 Google Brain 合并,可见谷歌集中优势兵力、全力投入竞赛的决心。

据悉,谷歌的 AI 发展历史可追溯到 2000 年代初期,彼时谷歌开始研究机器学习算法,通过组建或收购的方式,打造起自己的 AI 研发力量。其中,外界较为熟悉的是 DeepMind 和 Google Brain。前者成立于 2010 年,2014 年被谷歌以 5 亿美元的价格收购,其后一直作为独立部门进行运营,2016 年因其研发的 AlphaGo 人工智能围棋程序击败了韩国围棋九段棋手李世石而声名大噪;后者也交付了不少备受瞩目的项目,其中包括 Transformer 大模型,它被认为是 ChatGPT 开发的基石。

如今谷歌把 DeepMind 和 Google Brain 合并,可见谷歌集中优势兵力、全力投入竞赛的决心。“在谷歌计算资源的支持下,将所有 AI 人才整合成一个专注的团队,将大大加快我们在 AI 方面的进展。”对于将 DeepMind 和 Google Brain 合并,Alphabet 首席执行官桑达尔·皮查伊(Sundar Pichai)此前发文说。

“AI 请求本身就是基于大规模数据中心集群,未来 AI 服务很大程

度,并在 AI 和云计算领域继续保持领先地位。”而国内一位从事云计算代理服务的业内人士认为,这些举措能增加谷歌云的能力,“现在看来国内云计算市场不会受到影响”。

大模型是目前科技领域的新焦点,而云计算成为大模型技术落地的最佳平台和基础设施。谷歌、微软、亚马逊、阿里巴巴、百度、腾讯、华为等已发布或正在研发相关大模型产品的公司,大部分有自己的云计算平台、云产品和生态。

据了解,当前大模型的训练、推理的算力支持主要来自 GPU(图形处理器),但这类芯片却供不应求。“因为 AIGC 和 ChatGPT 4 近期火热,市场对 GPU 需求非常大,而主流 GPU 就是英伟达的 A100 或 A800,这些产品供不应求,货期高达

2 个月。”在某国际云计算大厂工作的程前(化名)告诉记者,目前包括微软、谷歌、亚马逊以及国内 BAT 等大厂都在进行 AI 芯片开发,做 AI 推理或训练时,自研 AI 芯片能够提供比较及时的算力支持,避免对英伟达的依赖并降低运营成本。

记者还注意到,谷歌在 4 月初宣称将 4000 个第四代 TPU(张量处理器,是谷歌为机器学习定制的专用芯片)串联在一起,构建起的超级计算机的运行速度要比采用英伟达 A100 GPU 的同等机器快 1.7 倍,效率高出 1.9 倍。谷歌还表示已经部署了数十台采用第四代 TPU 打造的超级计算机,供内部使用和外部通过谷歌云使用。对于这一情况的更多细节,记者联系谷歌方面,截至发稿未获答复。



云计算成为大模型技术落地的最佳平台和基础设施。视觉中国/国度都要依托于云,而 ChatGPT 本身就是部署在云上的。”资深产业分析师黄辉锋对记者表示,ChatGPT 发布后,谷歌一直很着急。

安光勇认为,不管是把 AI 芯片转移至云计算部门,还是把两个人工智能实验室合并,目的有两个,一是应对微软和 OpenAI 的威胁,二是推进自身 AI 技术的发展。

根据 *The Information* 的报道,谷歌将负责 AI 芯片研发的工程团队转移到 Google Cloud,是为了提高 Google Cloud 出售 AI 芯片给那些租用其服务器公司的能力。“Google Cloud CEO Thomas Kurian 一直在向 Google 的高层争取更多资源,以更好地装备团队,与来自 OpenAI 的 AI 软件微软的 Azure 云业务以及即

将加入的亚马逊 AWS 开展竞争。”该报道如此提到。根据谷歌母公司 Alphabet 近期发布的 2023 年第一季度财报,Google Cloud 销售额达到 75 亿美元,同比增长 28%,这是自 2020 年开始公布业绩以来,该部门首次实现盈利,营业利润为 1.91 亿美元。目前,Google Cloud 业务合计占 Alphabet 总收入的 10%。

根据市场调研机构 Canalis 发布的 2022 年第四季度全球云基础设施数据,AWS、微软 Azure 和谷歌云分别是全球前三大云服务供应商;而 IDC 发布的 2022 年下半年中国公有云市场报告显示,前五大市场份额的云厂商分别是阿里云、华为云、天翼云、腾讯云和亚马逊云。

微软、谷歌均欲加大 AI 投资

虽然谷歌 AI 的热度还无法与 ChatGPT 相比,但其至少已经意识到自身短板,并积极补强。

作为谷歌在新一轮 AI 竞赛中的劲敌,同一天发布财报的微软拿出了一张优异成绩单。财报显示,微软截至 3 月 31 日的季度收入为 529 亿美元,同比增长 7%,好于预期。

在报告期内,微软宣布追加数十亿美元投资 OpenAI,并表示将利用后者的人工智能模型开发新版必应搜索引擎,并将 GPT-4 接入到 Office 工具,推出了 AI 版 Office“全家桶”Microsoft 365 Copilot。

“世界上最先进的人工智能模型与世界上最通用的用户界面——自然语言——相结合,开创了一个新的计算时代。”微软首席执行官萨蒂亚·纳德拉(Satya Nadella)在财报新闻稿中表示,“在整个微软云服务中,我们是帮助客户从数字支出中获得最大价值并为下一代 AI 进行创新的首选平台。”

在 ChatGPT 引领的 AI 浪潮的助推之下,微软智能云服务在报告期内实现了不俗的增长。具体来看,包括 Azure 公共

微软、谷歌均欲加大 AI 投资

虽然谷歌 AI 的热度还无法与 ChatGPT 相比,但其至少已经意识到自身短板,并积极补强。

作为谷歌在新一轮 AI 竞赛中的劲敌,同一天发布财报的微软拿出了一张优异成绩单。财报显示,微软截至 3 月 31 日的季度收入为 529 亿美元,同比增长 7%,好于预期。

在报告期内,微软宣布追加数十亿美元投资 OpenAI,并表示将利用后者的人工智能模型开发新版必应搜索引擎,并将 GPT-4 接入到 Office 工具,推出了 AI 版 Office“全家桶”Microsoft 365 Copilot。

“世界上最先进的人工智能模型与世界上最通用的用户界面——自然语言——相结合,开创了一个新的计算时代。”微软首席执行官萨蒂亚·纳德拉(Satya Nadella)在财报新闻稿中表示,“在整个微软云服务中,我们是帮助客户从数字支出中获得最大价值并为下一代 AI 进行创新的首选平台。”

在 ChatGPT 引领的 AI 浪潮的助推之下,微软智能云服务在报告期内实现了不俗的增长。

具体来看,包括 Azure 公共

自研 AI 芯片的优势

大模型竞争之外可能还有芯片选择之争。

根据英伟达公布的信息,训练一次 1750 亿参数的 GPT-3 需要 34 天、使用 1024 张 A100 GPU 芯片,同时 OpenAI 可能至少需要 3.24 万张 A100 用于日常推理。

旷视科技联合创始人、CEO 印奇则透露,要做成 GPT A100 GPU 芯片,硬件投入就需要 20 亿元,中国目前可用作大模型训练的 A100 芯片总共只有 4 万张。

据 *The Information* 报道,包括亚马逊、微软、Google 和甲骨文等主要云计算供应商都在限制客户对云服务器的使用,一些客户租用硬件的等待时间长达数月,核心问题仍是 GPU 等算力芯片供不应求。

云、企业服务、SQL Server 和 Windows Server 在内的微软智能云业务部门实现营收 220.8 亿美元,同比增长了 16%;而包括 Dynamics、LinkedIn 和 Office 在内的生产力和业务流程部门公布收入 175.2 亿美元,同比增长约 11%。纳德拉表示,Teams 通讯应用程序在报告期内每月有超过 3 亿活跃用户,高于上一季度的 2.8 亿。

在财报电话会上,微软 CFO 艾米·胡德表示,微软接下来会加大云基础设施,特别是与 AI 相关的投资,并且在消费者业务方面如 Windows 的活跃设备数量、Edge 和游戏业务等不断提升市场份额。

目前,微软正在加紧将 GPT 向 Office 套件、必应搜索业务中数字支出中获得最大价值并为下一代 AI 进行创新的首选平台。”

而在今年 3 月,谷歌人工智能

自研 AI 芯片的优势

大模型竞争之外可能还有芯片选择之争。

据芯片分销平台数据,国内 A100 芯片现货市场价已经从 7 万元涨至 9 万元。另据财新,英伟达 A800 在国内也已被炒到原价的 2 倍多,且紧俏到脱销。

“目前,很多互联网大厂如微软、谷歌、亚马逊以及国内 BAT 等都在进行 AI 芯片开发。”程前表示,虽然每家的生产工艺和版本迭代各不相同,但这些 AI 芯片的诞生,能为 AI 推理或训练提供比较及时的算力支持,避免对英伟达的依赖并降低运营成本。

在程前看来,自研 AI 芯片能为云服务企业带来多种优势。“一是云厂商可提供更多的服务器机型,为用户带来更多的选择,找到最优方案。二是多类型

聊天机器人 Bard 对外开启测试,但效果不如外界预期,4 月又发布了更新版本,Bard 在处理常识推理和编程问题等任务时能力变强。皮查伊表示,谷歌推出 Bard 实际上也是作为搜索业务的补充产品,未来也会将语言模型更多地引入搜索业务之中。

有消息称,谷歌正计划发布新的人工智能搜索工具,以对抗微软必应搜索引擎中 ChatGPT 的整合。据《纽约时报》报道,人工智能搜索功能由谷歌开发,代号为“Magi”,可能会在 5 月首次推出,并在秋季推出其他功能。

财报电话会上,谷歌还透露,AI 是公司的发展基石,未来将继续对 AI 进行投资,预计在今年第二季度,谷歌的资本支出将有所提高,且后续全年都会有所增加。

虽然谷歌 AI 的热度还无法与 ChatGPT 相比,但其至少已经意识到自身短板,并积极补强。两大巨头的竞争预计会延续很长一段时间,鹿死谁手尚是未知之数。

自研 AI 芯片的优势

大模型竞争之外可能还有芯片选择之争。

AI 芯片供货,不完全依赖英伟达,云资源算力储备充足。三是云厂商通过上架 AI 芯片,全球用户更快使用到 AI 芯片,能更好促进 AI 芯片的迭代发展。最后云厂商通过加速自家 AI 芯片的发展,形成自己的竞争力优势,吸引更多的客户上云。”程前分析道。

除谷歌 TPU 外,微软也在自研 AI 芯片。*The Information* 报道称,微软大约在 5 年前就开始秘密研发一种内部代号为“雅典娜(Athena)”的芯片,来节省算力开支。

此外,AI 开发初创企业 Mijourney3 月称其正在使用基于谷歌云的 TPU 来训练其机器学习模型。这意味着,大模型竞争之外可能还有芯片选择之争。

中国云市场大变局:腾讯掉出前三 运营商云“狂飙”

本报记者 谭伦 苏州报道

2023 年初的中国云市场,比以往任何时候都“卷”得更厉害。

4 月 23 日,国际市场研究机构 IDC 发布《中国公有云服务市场(2022 下半年)跟踪》报告。报告显示,2022 年下半年中国公有云服务整体市场规模(IaaS/PaaS/SaaS)达到 188.4 亿美元,增长 24.5%。虽仍保持增长,但过去 3 年公有云客户缩减支出,云计算厂商缩减投资预算,使得增速低于此前预期的 30%。

值得注意的是,增长放缓之际,竞争并未放松。报告调研的厂商份额显示,IaaS+PaaS 方面,阿里云、华为云、中国电信天翼云、腾讯云和 AWS 排名前五位,其份额分别为 31.9%、12.1%、10.3%、9.9%、8.6%。IaaS 则和 IaaS+PaaS 市场排位保持一致,五大厂商分别为 32.6%、13.0%、11.8%、9.2%、7.9%。

其中,与 2022 年上半年相比,前五大厂商的份额与排名出现大幅变化。腾讯首次跌出前三,其 IaaS+PaaS 份额从 11.1% 减少到 9.9%,排名从第二降至第四;取代其位置的华为云 IaaS+PaaS 市场份额从 10.8% 增加到 12.1%,排名从第三升至第二;天翼云 IaaS+PaaS 市场份额从 8.9% 增加到 10.3%,排名从第四升至第三;此外,阿里云市场份额排名虽然依旧保持第一,但其份额从 36.7% 减到 31.9%。

对于此次云市场格局的变化,IDC 方面表示,由于音视频、电商等互联网传统强势领域市场恢复不及预期,与之关联的公有云市场增长持续疲软。随着国家产业结构调整逐步稳定,在线教育、游戏市场有所回暖,需求将逐步恢复,但距离发展高峰期仍有较大差距。

腾讯掉队 阿里降价

让业内对此次云市场格局之变最为意外却又在意料之中的,当属腾讯云的掉队。

“按照腾讯云当初的规模,掉到如今的位置,还是有点意外的。”一位国内腾讯云的客户企业负责人向《中国经营报》记者表示,四五年前腾讯云的影响力很大,市场规模也不错,当时公司正是看中腾讯云的口碑而选择了其服务。

Synergy Research Group 发布的研报显示,2018 年,中国云基础设施服务市场中,腾讯云拥有 16.5% 的市场份额,收入环比增速为 34.3%,排在份额 40.5% 的阿里

运营商全线“狂飙”

与互联网云厂商业务收缩形成鲜明对比的是,以天翼云为增速代表的三大运营商云全线扩张,年增长率均超过 100%,远超行业平均水平,而前者更是首次进入了公有云厂商前三。

最新财报显示,2022 年中国电信天翼云收入达到 579 亿元,同比增长 107.5%。移动云收入达到 503 亿元,同比增长 108.1%,其中,行业云收入达到 412 亿元,同比增长 114.4%,移动云盘收入达到 55

挑战与机遇并存

云厂商的竞争加剧,在反映中国公有云市场充满活力的同时,也带来了更多的挑战。

IDC 研报指出,2022 年下半年,疫情冲击对企业用户在公有云的投入上带来负面影响。一方面,企业认可公有云在企业降本增效方面的积极作用,公有云存量业务稳定增长。但另一方面,企业在新增

云之后,位列中国市场第二。

但是,腾讯云国内第二的位置却并未维持太久。在 Canalis 发布的 2022 年第二季度中国云计算市场报告中,腾讯云的份额为 17%,虽与 2018 年相比略有上涨也未下跌,但已被华为云赶超退居第三。

能够佐证其云业务增长下滑的数据,来自其公开财报。记者注意到,由于腾讯集团年度财报没有单独披露云业务营收的惯例,后者与其他业务被共同归结在“金融科技与企业服务”下。财报数据显示,自 2021 年开始,这一服务同比增长幅度明显下降,从

亿元,同比增长 59.5%。联通云实现收入 361 亿元,同比增长 121%。尤其是在份额排名方面,除了天翼云首度进入前三外,移动云公有云(IaaS+PaaS)服务市场的份额也上升到第六,同比增速 104%,位列前十位云服务商第一。

而这已不是运营商云第一次实现破百增长。2021 年,彼时已跻身中国公有云市场份额前五的天翼云便曾达到 102% 的营收增长率,而移动云也以 114% 的增长率

需求的决策和建设周期普遍延长,这在一定程度上阻碍了公有云整体市场的发展。有些互联网公有云用户出现自建云资源池和能力反向输出的趋势,为公有云市场带来更多变数和挑战。

面对云厂商之间的激烈“厮杀”,中国工程院院士郑纬民则更为直接地表达了担忧。“云计算的互操

作性不足,以及由此导致的‘供应商锁定’是全球性问题,但与国外相比,我国的‘供应商锁定’问题更为严重,成为影响运营商生态,阻碍云计算业务发展创新的瓶颈。”他在 2023 移动云大会期间指出。

郑纬民表示,出于不同厂商的技术差异和商业利益考虑,造成了传统上的“供应商锁定”问题。但由

创下当年新高。

“天翼云的成绩是标志性的,说明运营商云已经正式进入了云市场第一梯队。”周桂军向记者表示,在政企上云的需求驱动下,三大运营商的云业务自 2021 年开始高速增长,而随着数字化转型业务成为运营商第二增长曲线,预计其云业务的市场份额还会有进一步的增长空间。

记者注意到,随着云业务强势扩张,运营商已有了更大的雄心。

于国内计算技术发展相对较晚,市场上服务供应商众多,且正处于激烈的竞争之中,这成为我国“供应商锁定”问题更为严重的首要原因。

“同时,我国信息化应用的基础相对薄弱,行业经验沉淀不足,再加上代码开源、开放应用编程接口等措施的滞后,进一步加大了跨云开发的难度,因而阻碍了跨云软件的

调 15%~50%,其中存储产品最高降幅达 50%。而虽未直言降价原因,但张勇表示,希望将技术红利更多回馈给客户和伙伴,持续降低用云成本,扩大云市场空间。

对于腾讯阿里的困境,业内人士向记者分析了其遭遇挑战的主要原因。“政企上云是这几年云市场的主要增量,但显然在这块,政企更加信赖以央企为主的国家云。”C114 主编周桂军告诉记者,这也是业内对其份额会下降有所预期的一大原因。

启信宝中公开的 2021 年全国招投标信息显示,在当年全国公示

继中国电信董事长柯瑞文在日前举行的 2022 年业绩说明会上宣布 2023 年天翼云目标收入为超千亿元后,在 4 月 24 日举行的 2023 移动云大会上,移动云能力中心总经理吴世俊称,未来三年,移动云将力争突破年营收 1500 亿元。

移动云的底气来自过去一年市场的认可。据中国移动董事长杨杰在本次移动云大会上透露,截至 2022 年 12 月底,移动云签约大单超 6200 个,拉动收入超 225

亿元,央企国企上云项目超 3700 个。在移动云带动下,中国移动 2022 年数字化转型收入达到 2076 亿元,同比增长 30.3%。

“布局建设云基础设施、提供算力服务,是中国移动数字化转型的战略重点。”杨杰强调,云服务本质上是算力服务,当前,算力网络正驶入发展快车道,拥有广阔的市场空间,中国移动将坚定不移做大算力网络生态圈,共同拓展数千亿市场空间。

的超 1 亿元的政企数字化大单中,三大运营商合力拿下了超过 200 个订单,其中中国电信 104 个、中国移动 68 个、中国联通 34 个,相比之下,阿里云与腾讯云只拿到 9 个。

对此,Omdia 资深首席分析师杨光此前向记者表示,以腾讯为代表的互联网云巨头,在度过扩张期后,已进入更为看重利润回报率高的阶段,而很多政企项目对于安全的要求高,单位利润低,从成本回报率上考虑,对这些云巨头来说并不划算。因此,这也逼迫互联网云厂商减少做集成、总包类项目,从而进一步缩减了份额。

开发互操作的实现。”他补充道。

对此,郑纬民呼吁,未来云服务之间的协作共赢,共同创造更大价值是必然趋势。“实现良好的云计算互操作性,才能形成统一的计算生态,促进技术的持续迭代和行业的整体创新,为消费者带来更大价值,进而推动国家经济和社会发展。”