

阿里巴巴集团

2026 年 6 月份季度

业绩发布电话会

记录稿

2026 年 8 月 20 日（星期四）

介 绍

Lydia Liu 刘冰楠

阿里巴巴集团投资者关系主管

（翻译）

大家好，欢迎参加阿里巴巴集团 2026 年 6 月份季度业绩电话会议。

出席本次电话会议的有：阿里巴巴集团董事会主席蔡崇信、首席执行官吴泳铭、首席财务官徐宏、阿里电商事业群首席执行官蒋凡。

在此提示您，本次电话会议正在集团网站上进行网络直播。今天晚些时候，本次电话会议的录音将上传到我们的网站上。

现在我来简要介绍免责声明。今天的讨论可能包含前瞻性陈述，尤其是关于我们业务前景和预期财务结果的陈述，该等陈述受各种风险和不确定因素的影响，可能导致实际业绩大幅偏离前瞻性陈述中的内容。详情请参阅我们于今日发布的新闻稿和投资者演示材料中的安全港声明。请注意，电话会议中将使用某些非公认会计准则财务指标，公司按公认会计准则编制的财务指标以及公认会计准则指标和非公认会计准则指标之间有关调节的更多详情，请参阅我们的季度业绩新闻稿和投资者演示材料。

除非另有说明，我们的所有表述均为同比比较。

现在，请 Eddie 发言。

阿里巴巴集团业务进展

吴泳铭

阿里巴巴集团首席执行官

(原话)

各位投资人朋友，欢迎参加阿里巴巴集团 2027 财年第一季度业绩电话会。

过去一个季度，阿里巴巴对 AI 的战略投入转化为强劲的业务成果，集团整体收入增长 9%。同时，AI 商业化全面提速，阿里云外部商业化收入强劲增长 45%，经调整 EBITA 同比增长 133%，持续兑现和加速增长的承诺。AI 相关产品收入连续十二个季度保持三位数的增长，AI 相关产品年化收入已经突破 495 亿元人民币，接近 73 亿美金，是阿里云加速增长的核心引擎。

接下来，我将从 AI＋云商业化、AI 全栈能力、AI 应用生态以及大消费四个方面来详细解读。

第一，AI＋云商业化全面提速，预计未来将继续保持高速增长。

本季度，阿里云外部商业化收入增速提升至 45%，创二十二个季度新高。经调整 EBITA 利润率 11.6%。值得注意的是，45%这一增速是来自算力、存储、MaaS、AI 应用的全面增长。同时我们主动收缩低毛利的一些网络流量业务，增长质量持续提升。

本季度，AI 相关产品年化收入突破 495 亿元人民币，在阿里云外部商业化收入中的占比提升至 35%。AI 相关产品毛利率显著高于云产品的平均水平。我们的 AI 产品经常性收入，覆盖 AI 算力、模型服务、AI 应用等多个层面，这种多层次的 AI 营收来源和盈利模式，意味着客户在算力、模型、应用任何一层的需求增长，都会转化为我们的商业机会。这一结构性优势，将支撑 AI 产品经常性收入的持续高速增长。

AI Agent 爆发，一方面直接推动了 Token 和 GPU 算力需求，同时也大幅拉动了 CPU 相关的算力、存储、数据库、网络等传统云产品需求。阿里云正在向 Agentic Cloud 全面升级。根据八月份的最新数据，包括 Maas 在内的模型与应用服务收入 ARR 已突破 160 亿元人民币。

从当前市场反馈及合同订单来看，算力需求会持续大于供给。随着我们供应能力的不断提升，未来几个季度，AI＋云的收入增长还将持续加速，盈利能力也将继续提升。

第二，AI 全栈能力持续提升，自研芯片规模化商用，模型迭代加速，开源生态繁荣。

平头哥的自研芯片和我们的自研模型的深度协同，在本季度进一步提升了 AI 的商业化效率。

平头哥已建立覆盖 GPU、CPU、网络芯片的全栈自研芯片组合。真武芯片到 8 月初已经服务了超过 650 家客户。基于全国产的新一代平头哥芯片真武 M890 的超节点实例，近日已上线阿里云进行规模化销售，下半年将持续放量以满足客户的旺盛需求。阿里云真武 M890 超节点可高效运行超 2 万亿参数的大模型推理，Kimi K3 和 Qwen3.8-Max 均已通过该超节点实例对外提供 MaaS 服务。

在数据中心层，阿里云已将大型 AIDC 数据中心的交付工期缩短至 100 天，处于全球最领先水平。这将显著提升我们的全球算力布局速度。

在模型层，最近一个月，我们的模型发布全面提速，大语言模型、图像模型、语音模型、视频模型和音乐模型都实现了重要版本迭代，性能均位于国际第一梯队。

上周，我们刚开源了 2.4 万亿参数的 Qwen3.8-Max 和 Qwen3.8-27B 模型。Qwen 系列模型的全球下载总量已超过 30 亿次，衍生模型数超 30 万个。开源模型生态的繁荣，将有利于促进更多的云服务需求，这是一个正向的商业循环。

第三，AI 原生应用 B 端、C 端双布局，带动 Token 调用快速增长。

在企业端，我们推出了全新的 AI 生产力产品千问办公，主打企业办公场景，规模化输出大模型和 Agent 能力。我们判断，生产力 Agent 将成为 ARR 增长的又一引擎。

在消费端，千问 app 用户规模持续保持稳定增长，并正在增加多元化的付费增值服务。

在 ATH 和云的紧密协同下，“算力—模型—Token—应用—收入”的完整商业链路正在高效运转。

第四、本季度，阿里电商整体稳健。即时零售在大幅减亏的同时业务规模保持 45%的高增长，而且单位经济效益（UE）环比持续改善。

从上季度 AI 商业化跨越拐点，到本季度我们看到增长在提速，利润率在攀升。AI 业务自身的造血和支撑能力变得更强，让我们更有信心持续投入。面向未来，AI 已经成为阿里最确定性的增长引擎，我们将保持战略定力，以全栈 AI 驱动长期增长。

接下来请 Toby 分享财务进展。谢谢大家。

财务业绩

徐宏

阿里巴巴集团首席财务官

（翻译）

谢谢 Eddie。

本季度，我们在 AI+云和消费领域的战略重点持续推进，依托纪律性的投入，取得了强劲业绩。

云分部收入同比增长加速至 45%，EBITA 利润率环比提升至 12%。AI 相关产品收入继续推动增长势头，实现连续第十二个季度的三位数同比增长，占云外部收入的 35%。这一强劲表现反映越来越多的客户采用我们的全栈 AI 产品，包括 AI 智能体、模型、云基础设施和自研芯片，也体现了我们增长的规模效应，以及在供给紧张的市场环境下的高定价能力。

在消费方面，淘宝闪购在保持市场份额的同时，UE 持续改善。整体电商业务 EBITA 同比保持相对稳定。

为实现各电商平台间的协同效应，并进一步强化全栈 AI 能力，我们对财报中的业务进行了战略性调整。自本季度起，我们的分部报告将呈现如下：（1）阿里巴巴电商集团，（2）AI 云与算力服务，（3）AI 实验室与应用，以及（4）所有其他。

现在看一下本季度的财务业绩：

总收入同比增长 9%至人民币 2,690 亿元，主要由云和即时零售业务的强劲增长所带动。

总经调整 EBITA 同比下降 30%至人民币 273 亿元，主要归因于对科技的投入，部分被云业务的经营业绩改善，以及多项业务运营效率的提升所抵消。

我们的 GAAP 净利润为人民币 104 亿元，同比下降 75%，主要归因于经营利润的减少、处置投资以及股权投资按市值计价变动的净收益减少。

本季度的经营现金流为人民币 229 亿元的流入，同比增长 11%，去年同期为人民币 207 亿元。

自由现金流为人民币 447 亿元的流出，去年同期为人民币 188 亿元流出，主要归因于对云基础设施的投入。

本季度，资本支出为人民币 677 亿元，源于我们持续投入 AI 基础设施，以满足强劲且不断增长的客户需求。显著的同比增长受多项因素驱动，包括交付节奏波动性、预期客户对 AI 智能体采用量提升所带动的 CPU 采购增加，以及各类芯片组件价格上涨。

截至 2026 年 6 月 30 日，我们持有约 307 亿美元的净现金。若剔除期限超五年的债务，我们的净现金约为 465 亿美元。稳健的资产负债表使我们有信心投资以实现强劲增长。

AI+云的投资具有实现可观 ROIC 的清晰路径。搭载芯片的服务器通常在不超过 3 年收回成本。按 5 年使用年限计算，我们预计其在实现盈亏平衡后的至少 2 年将持续产生正向自由现金流。

截至 2026 年 6 月 30 日止季度，我们以 1.62 亿美元的总价回购了股份。

我们持续致力于实现长期股东回报的最大化，通过有纪律性的资本配置，进行 AI+云业务投入驱动业务增长、股份回购和派息。随着市场环境及战略变化，我们会相应调整优先级。

现在，我们先看电商业务。

新的阿里巴巴电商集团体现我们释放境内与跨境电商业务之间显著协同的战略重心。

从本季度起，我们将按照如下呈现阿里巴巴电商集团的收入：（1）中国电商，（2）中国即时零售，（3）国际电商，以及（4）全球批发。

阿里巴巴电商集团收入为人民币 2,059 亿元，同比增长 4%。

客户管理收入（CMR）同比下降 7%。若不考虑新的营销发展计划所产生的视作收入冲减项的影响，同口径 CMR 同比增长 1%。

中国即时零售业务收入为人民币 533 亿元，同比增长 45%，由盒马和淘宝闪购带动。

阿里巴巴电商集团经调整 EBITA 为人民币 397 亿元，同比保持相对稳定，体现了我们在增加用户体验和技术投入的同时，有效控制了成本。

淘宝闪购在保持市场份额的同时，UE 环比持续改善，主要得益于笔单价与履约效率的提升。

此外，本季度速卖通录得经营利润。

我们的目标是保持传统电商业务利润稳定，同时推动即时零售业务利润率持续改善。

接下来，我们来看 AI 云与算力服务的业务进展和业绩，该分部包括云智能集团和平头哥。

总收入及外部客户收入同比增长均加速至 45%。阿里云收入增速也进一步加快，同比增长 45%。我们有信心收入在未来几个季度持续加速增长。

本季度，AI 相关产品收入达到人民币 124 亿元，年化收入为人民币 495 亿元。其连续第十二个季度实现三位数同比增长，占云外部收入的 35%。

经调整 EBITA 利润率扩大至 12%，主要得益于规模效应提升，以及在市场供给紧张的环境下我们更强的产品定价能力。

我们预期接下来几个季度经调整 EBITA 利润率会继续稳步增长。通过提升资源利用率、优化模型组合并创新应用场景，我们正在加速 AI+云业务增长，进一步释放规模效应。

AI 实验室与应用包括 AI 模型实验室、千问 2C 事业部及千问办公。

经调整 EBITA 为亏损人民币 139 亿元，主要是由于对 AI 能力投入以及千问 app 推理成本的增加。由于千问 app 市场费用的减少，亏损环比显著收窄。由于模型训练和千问 app 营销效率的提升，我们预计今后几个季度的亏损会改善。

我们推出了前沿的语言、编程、视频、语音、图像、音乐模型，性能均位列第一梯队。

自千问 app 上线以来，已有 2.5 亿用户通过其智能体功能，在日益丰富的电商及其他服务场景中首次实现 AI 驱动的购物体验。

所有其他分部收入保持稳定，为人民币 288 亿元。

所有其他经调整 EBITA 为亏损人民币 33 亿元，主要是由于对技术的投入增加。

AI 已经从孵化阶段进入到规模商业化阶段。随着我们扩大市场份额、强化 AI 领先地位并提升运营效率，我们在战略和财务方面拥有更大的灵活性，能够有纪律性地、可持续地投入全栈 AI 和消费领域，驱动长期结构性增长并为股东创造更大价值。

谢谢，我们的发言部分到此结束，现在开始问答环节。

问答

刘冰楠：

（翻译）

感谢 Toby。我们现在开始问答环节。欢迎大家用中文或英文提问。第三方工作人员将会提供交替传译。如有任何疑义，请以管理层原始语言所作的陈述为准。

请接线员现在开始问答环节，谢谢。

Alicia Yap (Citigroup):

(翻译)

祝贺公司在云业务上取得了扎实的业绩。

可否请管理层谈一下：这一季度 CapEx 显著增加的原因和驱动因素是什么？对于接下来几个季度的资本支出趋势有何预期？此前提到的三年人民币 3800 亿元的资本支出预算，目前是否有任何更新？

此外，可否请管理层也介绍一下资本支出在不同业务（如模型训练等）上的分配细节？最后，管理层对这些投入的 ROIC 预期如何？

吴泳铭：

（原话）

谢谢您的问题，这些问题挺重要的，我也想借这个问题详细解释一下我们 AI 方面整个的商业模型，以及对未来 CapEx 投入的预期。

我们去年 2 月份公布了三年 3800 亿的投资计划，到今年 6 月季度末，已经累计投入了 1900 亿，进度还是符合预期的。本季度的 677 亿确实高了一些，但是因为硬件交付有一个周期，并不是每个季度均匀的，所以更多是一个设备交付的波动性。同时，本季度我们还增加了对于 CPU 的采购，因为 Agent 时代我们看到对 CPU 的需求在大幅增长。当然，也有芯片组件价格上涨的因素构成。

总体来看，我们不应该强调本年度的 CapEx 是 677 亿乘以 4 这样匀速投入的建设。但总体来说，我们希望今年的投入建设维持更积极的态度。

我也想借此机会和各位投资人解释一下全栈式 AI 平台服务的商业模式。它本质上是一个重资产的商业模式，我们现在看到所有的 AI 变现方式，无论是 AI 软件订阅，还是大模型的 API 推理，还是 GPU 的 IaaS 服务，以及 AI 相关的训练或者推理的软件服务，本质上所有的这些变现都需要在 AI 的算力中心基础上。而 AI 算力中心的建设，我们才能够获得市场份额的高速增长。所以这是一个非常明确的重资产的商业模式，而重资产商业模式要获得高速的增长，前提需要在 CapEx 投入上进行前置。

所以会看到，我们从 2025 年开始进入了一个非常积极的硬件方面的投资建设周期。从这个角度上来说，我们要获得一个高速增长，就必须在前面几年去投入更多的 AI 算力中心的建设。所以这是我们商业模式上决定的，必须是 CapEx 先行，才能获得后面业务的增长。

我再来解释一下，为什么在现在这个阶段进行 AI 算力相关 CapEx 投资的回报，确定性非常高。行业共识是在 2030 年之前，我们看不到 AI 算力紧缺的情况会有非常大的改变。所以在这个行业背景下，我们现在看到我们 AI 算力的 CapEx 投资回报确定性非常高。从数据上来看，按照我们现在 AI 产品的平均毛利水平，我们的 CapEx 可以实现在三年之内回本。而我们 AI 产品的毛利率水平还处于持续提升过程之中，这个回报周期未来还会持续缩短，比如往 2.5 年回本的方向去推进。

按照我们现在的投资回报周期，我们的 AI 算力在投入三年回本之后，还会非常长时间地、持续不断地贡献正向的现金流。从我们现在看到的实际情况来看，一个最直接的例子，在我们的数据中心里面，2020 年购买的 A100、2018 年购买的 V100，到现在还是近乎满载地被客户使用。AI 算力资产的实际使用寿命周期，远远长于理论上的折旧周期。

同时，我们在提升 AI 产品的毛利率水平上和投资回报率上有三个重要的手段：

第一，我们可以持续提升 AI 产品的毛利率，包括我们现在持续投入，推动我们的 SOTA 模型建设，拓展高毛利的 MaaS 业务。同时还可以使用我们更完善的产品组合，提升我们的 AI IaaS 产品或者 AI 相关的软件产品的竞争力，从而提升毛利率水平。我们看到 AI 产品的毛利率还在持续提升，本季度已经带动了阿里云整体板块的利润率提升 4.4 个百分点，来到 11.6%，这是初步的验证。

第二个重要的措施是我们的自研芯片，也可以说是在这个过程当中最重要的一个方案。我们平头哥的自研芯片路径覆盖 GPU、CPU 和网络芯片，这是新一代 AI 数据中心当中最核心的芯片组合。而在 AI 数据中心里面，最昂贵的成本就是芯片以及存储，自研芯片是我们长期重要的方向。随着平头哥自研芯片未来产能的持续提升，在我们数据中心的自研芯片比例还会持续提升，替代我们更多的商业化采购的芯片。我们知道，在现在算力紧缺的时代，商业化芯片本身毛利率就很高。所以自研芯片的比例大幅提升，会大幅度提升我们的产品竞争力以及毛利水平。

第三点也很重要，我们还有多种商业化手段，来提升我们现金流的使用效率。比如，我们和合作伙伴共建 AI 算力中心；比如，我们现在有一些 AI 算力合同已经开始采取预付的方式。所以通过这些商业上的方式，也可以大幅提升我们的投资回报率。

通过这三条路径，我们可以持续地把 AI CapEx 的回本周期推向更短的时间，比如推向 2.5 年甚至接近 2 年。

我觉得可以给大家算一个简单的框架，如果按我们现在 AI 产品毛利水平，CapEx 三年回本的条件下，理论上把我们的增速控制在 33%以内，就可以实现正向的现金流。但这现在不是我们的战略选择，由于 AI 还处于整个行业的非常早期阶段，对我们的战略选择，我们会坚定地投入 CapEx 积极的建设，去推动业务的高速增长。

而后面随着我们的产品毛利率的提升，自研芯片的替代比例提高，我们的回本周期将进入 2.5 年甚至更短的时候，我们可以看到，在追求 40%以上增速的同时，就能够维持正向的现金流，这是我们的长期方向。

Charlene Liu (HSBC):

（原话）

我想问一个跟电商相关的问题，首先我想问一下闪购的进展情况，另外公司把很多业务进行了梳理，组成了阿里巴巴电商集团，这些业务未来的各自的一些战略重点是什么呢？

蒋凡：

(原话)

新财年我们对电商板块进行了重新的梳理。未来我们会围绕着中国电商、即时零售，还有国际电商、全球 B2B 四个核心板块，向大家持续同步我们的进展。今天我想简要分享一下这四个板块的进展和思考。

首先是中国电商。国内电商短期宏观环境面临挑战。放眼长期，我们聚焦核心供给，同时希望通过 AI 提升电商的体验跟整体经营效率。

首先在供给方面，从去年开始，淘宝天猫就聚焦对平台包括品牌商家在内的原创商家进行扶持，同时挖掘优质产业带的白牌供给的潜力。

我们会继续加强和头部品牌商的合作关系，帮助品牌客户实现稳定且可持续性的生意发展，天猫依然是品牌商跟很多原创商家最核心的经营阵地。同时我们也会深入产业带，挖掘源头好货，支持更多制造业工厂直接到平台上开店经营。

也利用平台 AI 能力，帮助更多白牌商家进行更加简单、更加高效地托管经营，产业带托管经营模式在平台的交易比例在持续提升。在刚过去的 618，尽管宏观环境面临一定挑战，但从结果上看符合我们的预期，尤其是核心商家的经营结果，还是取得了不错的成绩。

同时我们也看到，AI 在电商供需两侧的机会。用户侧，我们会持续推出更多基于 AI 能力的新体验、新场景，例如多模态的搜索、AI 试衣等。另一方面，利用 AI 技术提升现有购物场景下的体验效率，

例如我们看到 AI 对商品推荐带来了非常显著地提升。

商家侧，我们看到商家已经在经营中非常普遍地使用 AI，我们在经营各个环节尝试通过 AI 帮助商家提升能力，尤其是数据分析、广告营销、客服等环节，商家可以明显受益。后面我们也会跟千问办公合作，推出更加适配电商场景的 AI 智能体。

回到即时零售的这个问题，经过一年多的发展跟投入，淘宝闪购的规模跟市场份额都已经发生了实质性的变化。无论是用户心智、供给丰富性、物流体验还是订单规模，都取得了巨大的提升。上个季度（6 月季度）我们在用户跟订单规模都保持上涨的前提下，实现了 UE 的大幅度优化，亏损规模显著缩小。

在今天这样一个基础上，接下来我们会加速整合盒马、天猫超市等相关板块，发展非餐饮品类的即时零售，尤其是加速发展前置仓。盒马过去一年加快前置仓的布局，带动了整个交易额的同比加速。

同时闪购也会持续拓展更多品类覆盖，在一些重点品类上持续创新，提升消费者体验。我们认为非餐饮品类的即时零售交易额，规模将在下个财年内超过餐饮品类，会带动电商大盘很多实物品类。即时零售板块预计在 29 财年实现整体盈利。长期看我们认为即时零售有望贡献平台整体交易额的 30%，成为电商板块的第二曲线。

关于国际电商板块，短期来看，我们的海外电商确实受到了国际税收政策和地缘环境的影响，面临一定的压力。但我们也看到，尽管市场环境很复杂，跨境电商依然保持交易规模上涨的同时，盈利水平也有显著提升，不管是交易规模还是盈利能力，我们都认为跨境业务具备长期潜力。另外，我们在土耳其、中东等区域的电商平台也发展很快，东南亚等市场的经营效率也在持续改善。

最后关于全球的 B2B 业务。B2B 业务包括我们 1688 平台跟 Alibaba.com，过去 20 年，这个业务一直在持续地发展，我们看到，AI 技术会给 B2B 交易平台带来非常大的变化，甚至可能从根本上改变原有的商业模式，尤其是 Agent 代理模式，会在 B2B 交易中扮演越来越重要的角色。我们面向跨境电商商家推出的 AI 智能体 Accio Work，上线后很快就拥有了 5 万个付费用户在使用。

AI 正在全面改变 B2B 商家，尤其是跨境商家做生意的方式。我们认为，基于过去 20 多年在这个领域的积累，我们有机会在 AI 时代的 B2B 跟跨境贸易领域创造出全新的业务模式跟商业机会。

总体来看，过去几年我们在电商板块几个领域分别完成了阶段性布局。接下来我们希望继续发挥从供给协同到 AI 技术的优势，让各个业务在 AI 时代释放更多的潜力。同时也实现更加多元化的收入利润结构，推动整个板块更加稳健的发展。

Yang Bai (CICC):

(原话)

我的问题是关于云与 AI 的业务，我们观察到阿里云收入增速逐季提升，本季度已经来到了 45%，公司此前也提出过未来五年外部云收入突破千亿美元的长期目标，这次也提出了未来几个季度仍会加速。

我想请教两点，第一，如果展望未来几个季度云业务的增长节奏，支撑云计算进一步加速增长的核心驱动因素是什么？第二，当前行业还是处于算力供给吃紧的阶段，虽然管理层刚才也提到，可能在 2030 年前这个供需格局都未必会有变化。但也想请教一下，如果是站在更长期的维度去看，云业务长期增长的核心驱动因素有哪些？是否会和短期有所不同？

吴泳铭：

(原话)

我可以展开分三部分：先讲业务现状和数据，再讲驱动 AI 相关业务增长的动能，最后讲远期展望。

第一部分，先来讲一下现状和数据。

AI+云的外部收入已经连续 9 个季度加速增长，本季度已经持续加速到 45%。我们看到客户的需求强劲，而我们的供给相对别的云厂商来说，也具备非常强的优势。所以未来几个季度，我们判断收入增速还会持续加速。我们看到 AI 相关的产品本季度的收入来到 124 亿元人民币，对应年化我们可以换算成约 73 亿美元。在我们现在的数据业务预测情况下，我们在下个季度的 AI 产品年化收入将接近 100 亿美元，所以可以看到我们的增速还是非常之高的。

同时我们也判断，未来几个季度我们的 EBITA 利润率每个季度都会有逐步的提升。还有对于我们云业务相当重要的 MaaS 业务，在本季度 AI 需求的增长以及推理效率的提升共同推动下，八月份 MaaS 业务的 ARR 已经突破 160 亿元人民币。

然后，来说一下我们商业模式的增长动能。

阿里巴巴在 AI 方面的投入模式和单一的 AI 公司最大的区别在于我们是在 AI 全栈技术上进行饱和式的投入，尤其是在芯片、AI 云基础设施和模型侧的饱和投入，保证了我们在这三个最重要的技术层面，都处于行业的领先地位。

而且我们判断 AI 行业的技术发展还在早期，在不同的技术发展阶段，AI 行业的核心商业价值会在芯片、云、模型和应用等不同层次之间流动。而我们全栈式的投入，可以保障我们具备最佳的服务能力和最好的性价比，也使得我们在未来行业技术发展的不同阶段，都能够保持竞争力和长期的增长动能。

下面讲一下最近一两年短期内最重要的增长动能。

从 2025 年底开始爆发的大规模的商业化推理服务，已经彻底改变了 AI 云的商业模式。刚才我已经提到了，算力才是 AI 营收能力的核心资产，现在所有的 AI 营收模式都是围绕着 AI 算力所展开，而算力紧缺现在又是全行业的中期共识。同时由于 MaaS 推理服务厂商的高毛利，使得 AI 算力不再是一个原来意义上的成本中心，而是成为与收入正相关的生产资料。

所以，行业内高定价的算力供不应求，推动了各种不同的场景下都需要使用 GPU，但是不同场景下它们的定价模型开始越来越趋近于行业最高毛利的变现方式，共同推动了几乎所有 GPU 相关产品的定价模型。

而阿里巴巴同时具备全模态的模型能力，我们的大部分模型能力都处于行业前沿的一线水平，也使得我们在算力的价值实现上具备优势，使得在我们的算力定价上有一个非常强的锚点，让阿里云具备算力的差异化定价能力。所以我们看到，在旺盛的需求以及算力的商业价值正在被市场重新定价，而我们也能够以更健康的价格去签新的合同，以及包括客户老合同的续约，都是随着新的市场需求发生定价模式的变化。所以，我们看到阿里云的盈利能力，在今年开始处于显著提升过程当中。

AI 云商业模式长期的规模效应和网络效应，会带动长期的增长动能。前两年行业投资人一直在问 AI 的超级应用是什么，但其实我们认为，AI 最大的超级应用就是以 AI 算力为核心的 AI 云平台。

我们看到，长期来看大部分企业核心的 AI 工作负载，都需要全栈的 AI 云来提供服务，比如大规模的训练、大规模的推理、定制化的推理软件、企业 Agent 开发与运营，这些都需要在一个全栈式的云平台，提供覆盖 GPU、CPU、存储、数据库和虚拟化，包括 Harness 工具等全面的云端基础技术服务。

所以 AI 云就像是一个超级城市，训练、推理、企业系统，以及 AI Agent 就像是这个城市的居民，持续迭代提升全栈的 AI 云服务，就像城市的基础设施，源源不断地吸引更多的新居民，并在提升老居民的留存率。

所以，AI 云平台的规模效应和网络效应可以推动我们的 AI 业务的长期增长。

因此，鉴于我们拥有亚洲云厂商里最多的数据中心，具备最强的规模效应。同时我们自研的平头哥 AI 芯片的规模化部署扩大，避免了外购高端的商业化 GPU 的高额溢价对毛利的侵蚀。同时由于我们自研模型的行业 SOTA 的能力，具备极高的算力定价能力。

因此，我们从行业的发展趋势以及我们自身的产品优势上来看，AI+云长期的营收增长和利润率的增长动能非常强劲。所以，对于加速实现 2030 日历年云外部收入 1000 亿美元的目标，我们具备极强的信心，也认为对 AI 云业务实现 20%以上的利润率有很强的可见性。

Yuan Liao (CITIC Securities):

(原话)

恭喜公司这个季度在云和 AI 领域的积极进展。

我有一个 follow-up 的问题，关于我们的 MaaS 业务，刚才 Eddie 也讲到，我们 8 月份的 ARR 超过 160 亿人民币。那我想问，我们上一季度也有讲到，就是年底的目标要突破 300 亿 MaaS 的 ARR。基于现在的进展，我们年底的目标会不会有一些调整？

另外，在 MaaS 业务当中，我们 Qwen 模型和三方模型的占比是怎么样的？未来模型竞争加剧，以及有更多开源模型的涌现，这些将怎么影响到 MaaS 的毛利率以及盈利能力？

吴泳铭：

(原话)

我们现在百炼平台 MaaS 业务的增速确实非常迅速。8 月份的 MaaS ARR 已经突破了 160 亿人民币。从现在的增速以及后面我们持续还会上线的更多的新的模型来看，我们在年底突破 300 亿的 ARR 目标，我们觉得还是会比较确定性地实现。

在我们现在的 MaaS 业务中，我们的自研模型还是占大部分的比例，但是三方的开源模型比例，其实营收也是不小的。

关于模型竞争的长期看法。

一方面，客户有比较强的需求，在一个 AI 应用当中，他可能会使用多方的模型，因为每个模型都会有自己的能力特点。同时我们也会看到，未来更多的开源模型会涌现。从这个角度上来说，其实是有利于阿里云百炼这样的云厂商的推理平台。因为实际上在我们的毛利水平当中，托管开源的、免费的三方模型和托管我们的自研模型，从毛利水平上来说其实是相差不大的。

我们的自研模型，其实更多的是为了长期的去追求模型的智能水平上的提升和 AGI 方面的突破。但是在短期的 MaaS 服务上面，实际上整体的毛利水平和托管开源模型毛利水平是差不多的。所以，更加竞争或者更加开放、繁荣的生态，其实有利于阿里云这样的云厂商的。

Alex Yao (JP Morgan):

（原话）

刚才 Eddie 也花了很多篇幅讲全栈式的 AI 生态——这个全栈式的 AI 生态，它的这个价值最终是在每一层去沉淀，或者去变现。

因为我们看到公司在旗舰模型发布三个月之内就开源了 Qwen3.8-Max 的权重。同时咱们自研芯片的商业化开展很成功，现在落地 650 多家外部客户，这是不是意味着管理层判断价值将会最终沉淀在算力与调度编排层，而不是模型层？

还是管理层觉得 AI 在不同的发展阶段，它的全栈式生态的价值，可能会沉淀在不同的环节。如果是觉得长期的价值沉淀在硬件或者是算力这个层面的话，怎么考虑当政府主导的算力供给或者是同样拥有合规芯片的竞争对手扩大供给以后，在硬件或者算力这一层的回报靠什么来保护？

吴泳铭：

（原话）

这个涉及到长期的判断，我觉得其实还是有一些不确定性。但是总体来说，我们觉得，因为我们全栈式的 AI 投入，所以无论价值在哪一层，在不同阶段时每一层无论多还是少，我们都有机会捕获。

说一下我自己的短期判断，在我们这个生态当中，我们觉得大部分的价值还存在于芯片和云基础设施这一层，从国内外的更多公司当中也看到的情况。在技术的早期，尤其技术供给稀缺的情况下，大部分情况确实会是这样，大量的价值会存在提供基础设施以及最核心的硬件比如芯片、存储这样的供应商中。在阿里巴巴的设计中，我们把算力、云基础设施和 AI 推理这一层放在一个板块。

关于刚才说的模型的这一层的商业价值如何体现，我觉得在行业内的争论也很多，在我们公司内部也是大家各有不同的看法。

我说一下个人看法，AI 大模型的这一层，现在通过 API 来变现的商业模式，对于大模型来说是一个短期阶段性的商业模式，肯定不是最终的商业模式。我们投入这么大的算力在全栈模型上，目标肯定不是为了短期的 API 收入。因为我们觉得未来当 AI 大模型能够实现 AGI 或者接近 AGI 的能力，其实最终的商业模式应该是会直接创造产品或者直接创造客户所需要的结果，尤其是在商业产品研发或者产品创造上面，或者商业的直接运营上。到了那个阶段，其实 AI 模型公司或者 AI 模型的商业模式不应该是提供 API 服务。那个阶段的商业价值，这也是支撑全世界这么多模型公司在现阶段持续不断军备竞赛投入的原因。所以从这个角度上来看，我觉得 AI 大模型的商业价值现在还远远没有看到，未来的商业价值会比现在所提供的 API 商业服务高很多。

在硬件算力层面，这个问题我也想重点回答一下。实际上对于我们的自研芯片，平头哥芯片的布局，我们以前在投资人会议上说的比较少，今天稍微展开一下。

平头哥上一代芯片已经出货了 50 万片以上，而我们刚才所说的全国产化的第一代芯片，已经在八月开始以超节点的形式在阿里云上全面上架。这一代的芯片超节点上架，意味着我们是中国市场唯二的可以放量进行超节点上架的公司。这一代国产芯片的客户接受度，以及大规模的商业化程度上，我们的信心非常强。平头哥芯片在所有国产芯片中有一个特点——因为平头哥的 AI 芯片是在国内芯片中比较独特的以 GPGPU 架构作为核心的技术架构。在这个技术架构下，我们能够同时非常好的支持训练和推理的工作负载。所以我们支持的几百个中大型的客户，既在帮他们做推理，也在帮他们做模型训练，尤其是像众多的具身智能和自动驾驶公司以及大的模型公司。从这个角度上，我们可能是中国唯一具备大规模训练和推理商业化客户的公司。

另外，我也想说一下我们对平头哥的第二代国产芯片的期待。

平头哥的第二代国产芯片，今年下半年会开始逐渐流片和产出，我们这一代的芯片会有非常强的算力，以及非常强的互联带宽。我们觉得完全可以用于大规模的模型训练。从这个角度上，我觉得我们在国内芯片技术上的身位是非常独特的。另外我不觉得有所谓政府主导的算力供给市场。所以从这个角度上，我们觉得对于平头哥未来成为阿里云的主要核心竞争力，还是非常有信心的，尤其是阿里云现在已经有几百个中大型客户在平头哥的芯片上获得了比较好的使用价值。在国内，从我接触的工程师的角度上来说，其实我们的芯片接触的工程师范围也是非常广的。

总结来说，我觉得平头哥的芯片在支持多行业的训练和推理的多样性上，我觉得可以说是国产芯片第一。我们在行业未来的产能和供应链的布局上，我们可以说是前二的存在。从我们对于未来 AI 芯片对接客户的渠道能力上，由于阿里云长期在中国 AI 云上是市场份额最大的厂商，在未来的渠道能力上，我觉得会是非常强的存在。所以从这个角度上，我觉得我们平头哥的芯片，长期的商业价值，我觉得我们还是非常有信心的。

刘冰楠：

（翻译）

非常感谢大家。谢谢各位的支持，我们期待下个季度向大家报告最新进展。谢谢。

【记录稿结束】