

阿里Q2财报：菜鸟、云业务求稳，电商主业仍难有起色

8月4日晚间，阿里巴巴发布了其2022年六月底止季度报告。

财报显示，本季度阿里巴巴营业收入为人民币2055.55亿元，同比保持稳定。营业利润为人民币249.43亿元，同比下降19%。自由现金流为人民币221.73亿元，相较2021年同期的人民币206.83亿元增加7%。

财报指出，营业利润的减少主要是受到疫情特殊形势的冲击，使得部分业务收入例如电商、本地生活业务、数字媒体及娱乐等相比去年同期有所减少。

从各项业务线上看，截至2022年6月30日止季度，阿里中国商业合计收入人民币1419.35亿元，与去年同期保持稳定。

受疫情负面影响，本季度服饰和消费电子等重要品类的支付GMV下降。尽管面临短期挑战，淘宝和天猫的消费者继续实现高留存率，尤其是消费能力较高的消费者。

截至2022年6月30日止的12个月，有超过1.23亿名年度活跃消费者每人在淘宝和天猫上消费超过人民币10000元。截至2022年6月30日，天猫有共计2500万名88VIP会员，每名会员的年平均消费额超过人民币57000元。

在6月份季度，淘宝及淘特上M2C商品产生的支付GMV同比增长超过40%。本季度，淘特通过优化用户获取的投入，以及提高活跃消费者的平均消费，使其其亏损同比及环比均大幅收窄。

报告期内，淘菜菜GMV同比增长超过

200%，亏损相较去年同一季度仅适度增加。此外，而在优化定价策略、提高采购能力和减少运营成本的推动下，淘菜菜亏损环比大幅收窄。

菜鸟持续透过加强端到端的物流能力来拓展其国际物流基建，其中包括eHubs、干线、分拣中心以及最后一公里派送网络。

在抵消跨分部交易的影响后，菜鸟的收入同比增长5%至人民币121.42亿元。主要增长来自国内强劲的业务量和增值服务的收入增长。

在国内市场中，菜鸟持续拓展增值服务，包括送货上门服务，以及通过菜鸟驿站来提升消费者体验，与中国商业业务相辅相成。截至2022年6月30日，约有70%的菜鸟驿站能够为消费者提供送货上门服务。

国际市场上，今年7月，菜鸟有一个新国际分拣中心在以色列投入服务，使得菜鸟的海外分拣中心总数增至十个。

本季度菜鸟通过进一步铺设智能柜，增强在欧洲最后一公里的配送能力。截至2022年6月30日，菜鸟在欧洲有超过7700个智能柜投入服务。

本地生活服务业务方面，主要包括“到家”及“到目的地”业务，如饿了么、高德及飞猪等。

财报显示，截至2022年6月30日止三个月，本地生活服务收入为人民币106.32亿元，相较2021年同期的人民币100.99亿元增长5%。

其中“到家”业务方面，受疫情防控政策影响，饿了么餐饮配送订单减少，本地生活服务整体订单同比减少5%，本季度整体分部GMV同比下降。但随着疫情的影响消退，饿了么的单位经济效益在6月份季度为正数，促使“到家”业务的亏损继续收窄。

“到目的地”业务仍受到新冠疫情的负面影响。不过在防疫措施放宽后，高德在6月份业务实现增长。6月高德日均活跃用户达到超过1.2亿的新高，主要由于疫情的影响有所舒缓，以及不断丰富本地内容和服务所带动，让高德用户能够发掘、联系和到访本地商家。

云业务由阿里云及钉钉组成。到2022年6月30日止季度，在抵消跨分部交易的影响后，阿里的云业务分部总收入（包括来自其他阿里巴巴业务提供服务的收入）为人民币176.85亿元，同比增加10%。

收入同比增长主要由金融服务、公共服务，以及电讯行业所驱动。截至2022年6月30日止季度，来自非互联网行业的客户收入占抵消跨分部交易的影响后阿里云总收入的53%，较去年同期上升超5%。

截至2022年6月30日止，云业务经调整EBITA为人民币2.47亿元，与2021年同期为人民币3.40亿元相比有所下降。下降原因财报表示，主要由于加大对技术的投入以及业务需求增大推动主机托管费和带宽费增加所致。云业务依旧是阿里聚焦的重点。

此前7月举办的阿里云合作伙伴大会



便强调了加强生态体系的举措，包括能力建设及培训项目，以及面向生态伙伴的奖励项目。阿里预期来自生态伙伴的收入贡献将通过此等举措进一步扩大。

数字媒体及娱乐方面，本季度收入为人民币72.31亿元，相较2021年同期的80.73亿元下降10%，主要由于阿里影业、优酷和其他娱乐业务的收入下降。

值得注意的是，受到优质内容及来自88VIP会员计划的持续贡献所带动，优酷的日均付费用户规模同比增长15%，从而达致本季度亏损为连续五个季度同比有所收窄。

阿里巴巴集团董事会主席兼首席执行官张勇表示：“过去一个季度，我们始终聚焦长期战略。虽然我们的业务在四、五月增速相对放缓，但进入六月已经看到向好迹象。基于我们拥有的高质量消费群体，以及满足客户不同需求的多元化业务所呈现出的韧性，我们对未来长期增长充满信心。”

财报公布后，阿里美股开盘涨幅超7.5%，截至发稿回落至4.4%左右浮动。

你身体里的DNA，能存下整个宇宙的数据

恐龙灭绝6000多万年后，科学家们获得了一块有史前蚊子的琥珀，从蚊子血中获得了恐龙的基因，从而让遥远的生物复活。讲这个故事的《侏罗纪公园》，至今仍位列全球电影票房前十。这个系列故事的原理很简单：DNA存储了恐龙的生物信息，科技让它重新表达。

现在，用DNA想象另一个故事：在宇宙长河中，“人类世纪”也寂灭了。另外一种智慧生物出现，TA们去探究远古的“人类文明”。有什么能承载人类文明的记忆？气温异变，地球上的庞大数据中心徒留遗迹。

而冻土中有一份DNA，它很轻，只有1公斤，看起来是一些被封装在胶囊里的白色粉末。读取后，里面却记载了地球上曾有过的巨量信息。视频、文字、代码展现了人类历史进程中的无数发明和文艺作品。于是那个遥远文明的痕迹在宇宙间再次展开。

这是另一个科幻设定了。背后的技术正是目前被关注的一个前沿方向：DNA存储信息。在大自然里，DNA负责存储遗传信息。单个人体细胞的平均直径是5到200微米，这其中的DNA可以包含一个人全部的遗传信息：30亿对碱基。

那为什么不能用碱基存储别的信息？这个科幻般的设想，正在走出实验室，被当作信息存储的未来方案。

01 基因组数据太多了，怎么办？

本来是生物学家想解决生物学发展的问题。

11年前，一群生物信息学家在德国的一家酒店里讨论「数据存储问题」。Nick Goldman也在其中，那是他在欧洲生物信息所(EBI)担任高级科学家的第二年。

大规模的基因组测序正在进行，随之产生的数据规模快速增长。存储、压缩这些数据是个麻烦事，现有的技术方案看起来不太行。据估计：人类基因组需要高达2-40EB的存储容量。这可能超过一个世界级科技公司的云存储量——全世界苹果用户存储在谷歌云上的数据总量大约是8EB。这8EB数据，每月存储费需要2.18亿美元。(1EB=1024GB)

生物学家们陷入了沮丧。

Nick Goldman拿着存储了莎士比亚所有十四行诗、一张照片和「我有一个梦想」演讲片段的DNA | 来源：EBI

有人灵光乍现：是什么东西阻止了我们用DNA来存储数据呢？

看起来是一句玩笑话，但是生物学家们意识到了这不仅仅是个玩笑，他们拿起手边的餐巾纸，用圆珠笔认真计算起可行性。

DNA存储遗传信息的原理并不复杂，它由四种核苷酸A、T、G、C组成，彼此两两对应，组成双螺旋结构。核苷酸的序列，记录

了遗传信息。

在数字世界，所有的信息本质上是0和1组成的数据串。想要DNA存储数字信息，简单理解，原就是将0和1的编码序列转换成核苷酸的序列。DNA存储的优势在于密度大，大约在你眼前逗号这么大小，1立方毫米的DNA，就可以容纳9TB(1TB=1024GB)的信息。

用DNA存储数据，也并不是完全新的想法，之前就有科学家尝试过。不过属于科学和艺术的先锋跨界实验。

1988年，艺术家Joe Davis和哈佛大学的研究员，将一副名为「小维纳斯」(Micro Venus)的图案存储到DNA短链中。

这个图案编码简单，白色的地方标记为0，黑色的线条部分标记为1，文件大小只有35bits，用了28个核苷酸长度的DNA链条来存储。

在那次酒店讨论的2年之后，2013年，Goldman团队发表了研究成果。这次，他们存储了5种不同格式的文件，一共有0.75MB。为了确保信息读取不出错，科学家存储的时候，每份信息按照四倍冗余的量来存储。

五个文件分别是：

- 154首莎士比亚的14行诗(ASCII编码格式)

- 提出DNA双螺旋结构的论文(PDF版)

- 一张照片(JPEG格式)

- 马丁·路德金「我有一个梦想」演讲其中26秒片段(MP3格式)

- 一串霍夫曼密码

这些年，DNA存储容量的上限不断被突破。2019年，美国一家创业公司Catalog在DNA中存储了16GB的维基百科。这个公司表示自己正在建设世界上第一个基于DNA的大规模数字数据存储和计算平台。

02 编码和解码，要处理的事情很多

在一些生物学家看来，用DNA来存储是一件非常「顺滑」的事。「大自然的编码语言非常类似于我们在计算机领域使用的二进制语言。在硬盘上我们使用0和1来代表数据，而DNA中，我们拥有4种形式的核苷酸，A、C、T和G」。在瑞士联邦理工学院的生物学家Robert Grass说。

DNA存储的关键之一是用四个核苷酸去映射0和1两个数字。方案可以很简单。比如：A对应00，C对应01，G对应10，T对应11。然后再按照所需要的核苷酸序列，像串珠子一样，把核苷酸们串成一串。(这就是DNA合成)需要读取信息的时候，再运用基因测序技术，把这一串核苷酸序列读取出来，再翻译成0和1的字符串。这个流程就是编码—DNA合成—测序—解码。

这个听起来像是「把大象装进冰箱」的流程，操作起来需要考虑的问题还有很多。不然科学家就不必一直研究新的编码方案了。

在自然界存在的DNA中，A和T，C与G两两配对，在一条DNA中，CG与AT的存在比例基本均匀，为50%左右。如果C和G的含量过高，可能会让DNA链产生一些复杂的物理结构。这就会让DNA测序(解码)变得复杂。

而且在「串珠子」(也就是合成DNA链条)的过程中，错误率不可避免。目前大约每合成100个碱基就会出现一个错误。这是由目前的化学合成技术带来的瓶颈，每合成一个碱基，有99.9%以上的正确率。但是当碱基串变长，0.01%的概率相乘，错误就难以避免。目前人工合成DNA的单链的长度一般不超过100个碱基，极限在300个碱基左右。而在自然界的DNA动辄有几千个碱基对。

也就是说，虽然DNA的存储能力很强，但它们不得以很多条短链的方式存在。如果存储的信息量比较大，这些DNA短链就像一本散装的书。它可以存储很多信息，存在形式却是一张张标着页码的纸。当然，可以将一条条DNA短链拼接成长链。这就意味着增加了一道工序。在测序的过程中，又需要把长链打断成短链。这是因为目前技术还不能一次性读取长链。

在测序的过程中，也存在错误率。尽管目前的错误率已经低至10⁻³数量级，比起商业硬盘的读写错误率，仍相差至少9个数量级。

正确率受到合成和测序这两项技术的影响，科学家想到设计编码方案来避免：在编码中增加纠错机制。这样，哪怕碱基合成和测序中出现了错误，依旧能够保证被存储进DNA的内容能够被正确读取出来。

03 走出实验室，还要考虑速度和成本

DNA存储也正在尝试走出实验室。

2020年10月，微软、西部数据和基因测序巨头illumina、DNA合成初创公司Twist Bioscience等联合成立了DNA数据存储联盟。

这是世界上第一个该领域的学术和产业链联盟。这个联盟希望制定技术和格式标准，最终建立一个可以通用的商业系统。

微软研究院在2015年就成立DNA存储的项目，并聘请了华盛顿大学的计算机科学与工程学院的副教授Karin Strauss担任高级首席研究经理(Senior Principal Research Manager)。

2013年，她和同事去英国EBI访问，了解到Goldman和同事们关于DNA存储的研究，就对这个方向产生了很大的兴趣。

Strauss说，「DNA的密度、稳定性和成熟度让我们兴奋。」

在他们的研究中，想开发的是另一个功能：随机读取。常见的DNA测序技术中，必须要把所有的碱基串一次性读取完，才能够获得信息。要么不读取，要么全读。如果只想要数据中的某一个片段，就会非常麻烦。

2016年，他们发表了一项研究，可以在DNA已经存储的信息中搜索到指定的图像，定位后，用酶来复制所需的DNA片段，然后只需读取这一小段即可。

要让DNA存储离商用更进一步，还需要解决合成速度和成本。现在合成速度是每秒存储上千个字节(KB)，成熟的云存储方案已经有每秒兆字节(GB)以上。

这意味着，编写DNA的速度还需要提升6个数量级。如何让提升数据处理量？就像并行计算能够提升数据处理速度，科学家希望DNA在合成时也可以并行多条，同时处理。

2021年，微软开发出首个纳米级DNA存储器，能够在每个平方厘米的区域上，同时合成25X106(2650)条碱基序列。这个新的技术把原来同时合成碱基序列的数字从个位提升到了千位。这个吞吐量，让DNA合成速度变成了每秒兆字节(MB)。

新的方法让DNA合成的阵列数量大大增加 | 来源：微软研究院

更大的吞吐量，也就意味着更低的成本。现在DNA存储的成本是每万亿字节(TB)8亿美元。而磁带存储成本已经降到了每万亿字节16美元以下。这样比起来似乎毫无竞争力。但现实生活中的大型数据中心的维护成本极高，还要定期更新硬件；DNA存储密度大、体积小、可以长时间不变质的优势就变成了降维打击。

所以量大、读取频率低的「冷数据」，被认为是DNA存储最近的应用场景。Twist Bioscience最近在一份市场报告中强调，这种技术能够帮助科技企业进入「大规模、低功耗」情况下更有效地部署。

另外一些乐观的科学家，更相信技术的进步。

自2003年人类基因组计划完成以来，测序成本降低了200万倍。2016年时，面对每秒千字节的速度，Goldman说，「(读写的速度提升)6个数量级对基因组学来说没什么大不了的。你只需要再等一会儿。」

那这「一会儿」是多久呢？这个领域似乎到了临门一脚，仍在等待突破。