

如今，云基础设施领域，几乎所有的大佬都在布大局、下大棋，而在这盘大棋里，“DPU”堪称非常关键的一步。

我这里所说的

“DPU”，是泛指各种具备
DPU类似功能和定位的加速卡或协处理器
，大家的叫法虽然不同，想干的事、想搅的局都差不多。

但是在实战布局中，不同出身的棋手，下棋的路数却大相径庭↓

传统处理器大厂，目标是全行业通吃，很注重产品的“通用性”，这是很合乎他们以往套路的正规下法
，就像他们多年以来一代代CPU/GPU
的迭代，盯住对家，一步一个脚印，如“挤牙膏”般稳健布局。

所以，这类大厂，下的是“**本手**”。

随着DPU的热炒，还有一大批独立公司涌现出来，纷纷入局入坑，由于势能有限，这类公司的产品往往盯住一些行业细分场景，比如数据库加速、统一通信加速。

这个棋路，初看合理，但长线会很艰辛，定制化多，市场规模也有限。

因此，这类玩家，下的是“**俗手**”。

那么，“DPU”的大棋局里，谁有机会下出“妙手”？

我认为只有CSP
云服务商，而且是头部的云大厂
，比如AWS和阿里云们，才有机会走出“妙手”。

第一，无论是叫做数据中心处理器（英伟达DPU），还是叫做基础设施处理器（英特尔IPU），最大的目标市场，都是CSP，也就是像“3A”这样的超大云服务商。

这说明云服务商使用DPU是必然趋势，但通用型的DPU能满足各大云商“变态”的需求吗？显然不能！

道理很简单，所有的硬件基础设施，都需要跟软件来相互加持。传统IT时代，跟CPU配合的是OS操作系统，当年Wintel组合就是用这种互相加持，取代了IBM的PC霸主地位。

而到了如今的云时代，DPU这样的新型基础设施，必须要靠CloudOS来加持，才能发挥最大作用。

那么云操作系统掌握在谁手里呢？当然是这些云大厂们。

第二，手握CloudOS、深谙云需求的云大厂们，硬件能力够格吗？能造出靠谱的DPU来吗？

完全不用担心！

其实这几年来他们一直在修炼硬件内功，AWS自研的ARM处理器已经进化到第三代了（Graviton3），而接近DPU形态的Nitro则打磨了更多年。

阿里云也是一样，自研的“神龙”架构其实就可以看做是DPU，去年阿里云还发布了吊炸天的5nm工艺倚天710处理器...

不止这“两A”，其他的云大厂也都在悄么声的搞硬件、搞芯片，大家都看到了重新定义云基础设施、把CloudOS与底层紧耦合的机会，所以云大厂的硬件

能力，不必担心。

再举个极端的例子，上个月发生了件大事，博通610亿美金收购了VMware，接下来，很有可能整出一个针对私有云环境深度优化的“DPU”或者什么怪东西出来。

博通大概率会用软硬件的深度垂直整合，来收割私有云市场。

私有云都敢这么下场，对于自主能力更强的公有云大厂来说，水到渠成。

云大厂亲手下场垂直整合，CloudOS与DPU软硬一家亲。

不过，这虽是妙手，还没有秒到极致，头部稍有点道行的云大厂都在这么琢磨。

所以，真正绝顶的棋手，要做到别人下棋看3步，他却能看到10步。如今，这样的超级妙手，还真让我们等到了。

在今天举办的阿里云峰会上，阿里云宣布推出CIPU：云基础设施处理器。

在认真扒了扒这个“CIPU”之后，我发现，阿里云这不仅仅是妙手，甚至是“掀桌”级别的，彻底换了玩法。在所有的云大厂中，阿里云抢先迈出了一大步！

为什么说这是“掀桌式”的玩法呢？

长期以来，IT基础架构，都是以CPU为中心的。不管是集中式的传统IT架构，还是分布式的云架构，CPU都是当仁不让的老大。

即便是曾经的阿里神龙或者AWS Nitro，虽然具备了强大的卸载加速能力，但它们依然是插在服务器里的一块卡。

负责充当CPU老大的打手，干一些脏活累活（存储卸载、网络卸载、虚拟化和
管理卸载）。

所以，从当前时间点往前看，
整个云基础设施，仍然以服务器为单位、以CPU为中心。

即便服务器上都插了各种加速卡，但CPU是“话事人”，CloudOS想要更直接、更流畅地操控基础设施，总有那么一点不顺滑，无法完成垂直整合。

但是，今天阿里云CIPU发布，云基础设施的架构被颠覆了，从以CPU为中心，变成了以CIPU为中心。

“DPU”从以前的CPU小跟班，摇身一变当家做主了。

CIPU就是这个升格版的主人，它成为数据中心里连接计算、存储、网络的中枢。

在以CIPU为中心的架构下，云基础设施即便物理外观看起来还是一台台服务器，但内部的逻辑结构和连接关系已经发生了巨变。

下图是阿里云给出的逻辑架构。

而我脑补的画面是这样的↓

在云数据中心的一组集群里，有台插满CIPU的“装置”，作为中枢，把周边的计算型节点、存储型节点统统“纳管”起来，变成了一台“超级服务器”。

以前，虚拟化技术把单台物理服务器化整为零，而如今，CIPU又把多台服务器化零为整，从“一虚多”到“多合一”，整个底层完全被重构。

新体系架构下，在一个集群里，CIPU是核心“话事人”，阿里飞天OS可以更直接的操控CIPU，把飞天的所有软实力，全部装载到CIPU里，管理平面更简洁，加速更彻底。

以前阿里云神龙也好，AWS Nitro也罢，加速能力相当于“外挂”，而CIPU新架构，这种加速和管控，相当于“原生”，形成了云操作系统的垂直一体。

因此，CIPU带来了更加吊炸天的加速体验，我不啰嗦了，直接列一下阿里云今天公布的官方数据吧↓

这里面最夸张的一项，莫过于eRDMA的延时低至 $5.5\mu\text{s}$ ，此前业界在公有云上能拿得出手的RDMA延时，只有 $15.5\mu\text{s}$ 。

5.5 μs 意味着什么？用户可以真正放心地把HPC

工作负

载放到云上去

跑了，通过云计算，高性能

计算得到了双重普惠：①云上超算低成本，性能无忧，更灵活的规模和弹性

②全应用场景的加速编程生态兼容。

而一些通用场景，同样可以考虑用eRDMA网络来提速，根据阿里云官方公布的数据，Redis场景提速130%，Spark场景提速30%，AI深度学习训练场景提速30%

其它主流通用计算场景，在CIPU的加持下，性能更是得到大幅提升，全面领先于友商云，以下数据来源于阿里云官方↓

好了，以上就捋完了CIPU的基本特征，那为什么说阿里云推出CIPU，是一招超乎寻常的“妙手”的呢？

我来总结下——

CIPU是阿里云为飞天操作系统量身订做的，是业界第一个云操作系统垂直软硬一体，All in One的全新基础设施体系架构。

CIPU将成为阿里云基础设施底座的中轴，对计算、存储、网络进行全面的云化加速，未来，CIPU Centric的体系不仅会颠覆数据中心的逻辑架构，也会对物理架构产生深远影响。

CIPU消除了90%以上的云税（云化开销）：计算虚拟化开销、存储虚拟化+存算分类转发开销、网络虚拟化+网络转发开销、内核TCP协议栈开销、RPC序列化&反序列化开销。

这就好比从蒸汽机进化到内燃机，本质上是能量转换效率的大幅提升。

对广大云用户来说，大家将会享受到云基础设施“内燃机时代”的红利，获得更高性能、更高安全、更加稳定、更普惠的云服务。

阿里云的这招“妙手”也绝非偶得，是通过无数次软硬件的“本手”历练才悟到的：飞天OS“本手”、神龙“本手”、倚天“本手”、中国公有云市场深度实践“本手”...

苦练“本手”，不恋“俗手”，擅出“妙手”，方能赢得云计算的大棋局！

最后一张图，我们再来品品阿里云的“妙手”↓