

# 谷歌 TPU 及强化学习

## 谷歌 TPU 以时间换吞吐量，加速云端 AI 帝国；AlphaGo 从 Lee 到零，探索强化学习新起点

### 谷歌 TPU：以时间换吞吐量，软硬兼施，冲入云端

人工智能立夏将至的大趋势下，芯片市场蛋糕越做越大，足以让拥有不同功能和定位的芯片和平共存，百花齐放。后摩尔定律时代，我们强调 AI 芯片市场不是零和博弈。我们认为在 3-5 年内深度学习对 GPU 的需求是当仁不让的市场主流。行业由上至下传导形成明显的价值扩张，英伟达和 AMD 最为受益。在深度学习上游训练端(主要用在云计算数据中心里)，GPU 是当仁不让的第一选择，但以 ASIC 为底芯片的包括谷歌的 TPU、寒武纪的 MLU 等，也如雨后春笋。而下游推理端更接近终端应用，需求更加细分，我们认为除了 GPU 为主流芯片之外，包括 CPU/FPGA/ASIC 等也会在这个领域发挥各自的优势特点。

但我们需要强调，包括 TPU 在内的 ASIC 仍然面临通用性较弱，以及开发成本高企等局限。TPU 虽然理论上支持所有深度学习开发框架，但目前只针对 TensorFlow 进行了深度优化。另外 ASIC 芯片开发周期长和成本非常高，在开发调试过程中复杂的设计花费有时甚至会超过亿美元，因此需要谷歌这样的计算需求部署量才能将成本分摊到大量使用中。同时 ASIC 开发周期长，也可能会出现硬件开发无法匹配软件更新换代而失效的情况。

TPU 是针对自身产品的人工智能负载打造的张量处理单元 TPU。第一代主要应用于在下游推理端 TPU。本质上沿用了脉动阵列机架构(systolic array computers)，让推理阶段以时间换吞吐量。第二代 TPU 除了在推理端应用，还可以进行深度学习上游训练环节。将 TPU 部署在云计算中以云服务形式进行销售共享，在为数据中心加速市场带来全新的需求体验的同时，可进一步激活中小企业的云计算需求市场，另辟 AWS、Azure 之外蹊径。

### AlphaGo 的“终点”，强化学习的起点

我们在今年 1 月的报告《2017 MIT 人工智能 5 大趋势预测：寒梅傲香春寒料峭，人工智能立夏将至》中提到第一大趋势预测：正向强化学习(Positive Reinforcement)正成为深度学习(Deep Learning)后研究应用的最新热点。

强化学习(Reinforcement Learning)的目的是尝试解决对人类标注样本的依赖，并打破特定板块和领域里的学习局限，向无监督、通用型人工智能拓展。强化学习的灵感来自于动物的学习方式。动物能够学会某些特定行为所导致的正面或负面结果(a positive or negative outcome)。按照这种方法，计算机可以通过试错法(trial and error)来与训练环境互动，包括 sensory perception 和 rewards，来决定这一结果的行为相关联。这使得计算机可以不通过具体指示或范例(explicit examples)去学习。

当前人工智能主流应用还是基于深度学习神经网络，从针对特定任务的标记数据中学习，训练过程需要消耗大量人类标注样本。而在很多现实场景下，特定垂直领域的数据并不足以支持系统建构，我们认为强化学习有机会成为下一个机器学习商业成功的驱动力。

### 谷歌重申买入：人工智能巨头新征途：云+YouTube+硬件

我们早在年初已经开始强调，AI 巨头谷歌的新征途——云+YouTube+硬件。YouTube 和云计算的巨大增长动力将是谷歌持续转型的助推器：YouTube 百般武艺冲劲十足，Q3 广告业务净营收增速回升至 21%，移动端积极转型执行力坚决。长期眼光看 AI 和 Other Bets 创新业务厚积薄发：谷歌是人工智能的龙头标的，我们长期看好语音识别和无人驾驶的发力。公司 20 日收盘价 1035 美元，根据彭博一致预期 2018 年 EPS 40.15 美元，给予 30x PE，目标价 1200 美元，重申“买入”评级。

**风险提示：芯片市场竞争加剧，研发速度不及预期等。**

### 作者

何翩翩	分析师
SAC 执业证书编号：S1110516080002 hepianpian@tfzq.com	
雷俊成	联系人
leijuncheng@tfzq.com	
马赫	联系人
mahe@tfzq.com	
董可心	联系人
dongkexin@tfzq.com	

### 相关报告

- 1 《谷歌(GOOG.US) 3Q17 点评：营收盈利超预期，YouTube 照耀转型路，人工智能巨头新征途：云+YouTube+硬件》2017-10-27
- 2 《谷歌(GOOG.US) 2Q17 点评：欧盟处罚争议难阻营收盈利超预期，云计算继续发力，人工智能巨头百般武艺扎实前进》2017-07-25
- 3 《人工智能芯片行业点评：英伟达 GPU 王者风范，Google TPU 破局科技；人工智能冲入云霄，看好 GPU、ASIC 各领风骚》2017-05-31
- 4 《2017 MIT 人工智能 5 大趋势预测：寒梅傲香春寒料峭，人工智能立夏将至》2017-01-25
- 5 《谷歌人工智能深度解剖：从 HAL 的太空漫游到 AlphaGo，AI 的春天来了》2017-01-05

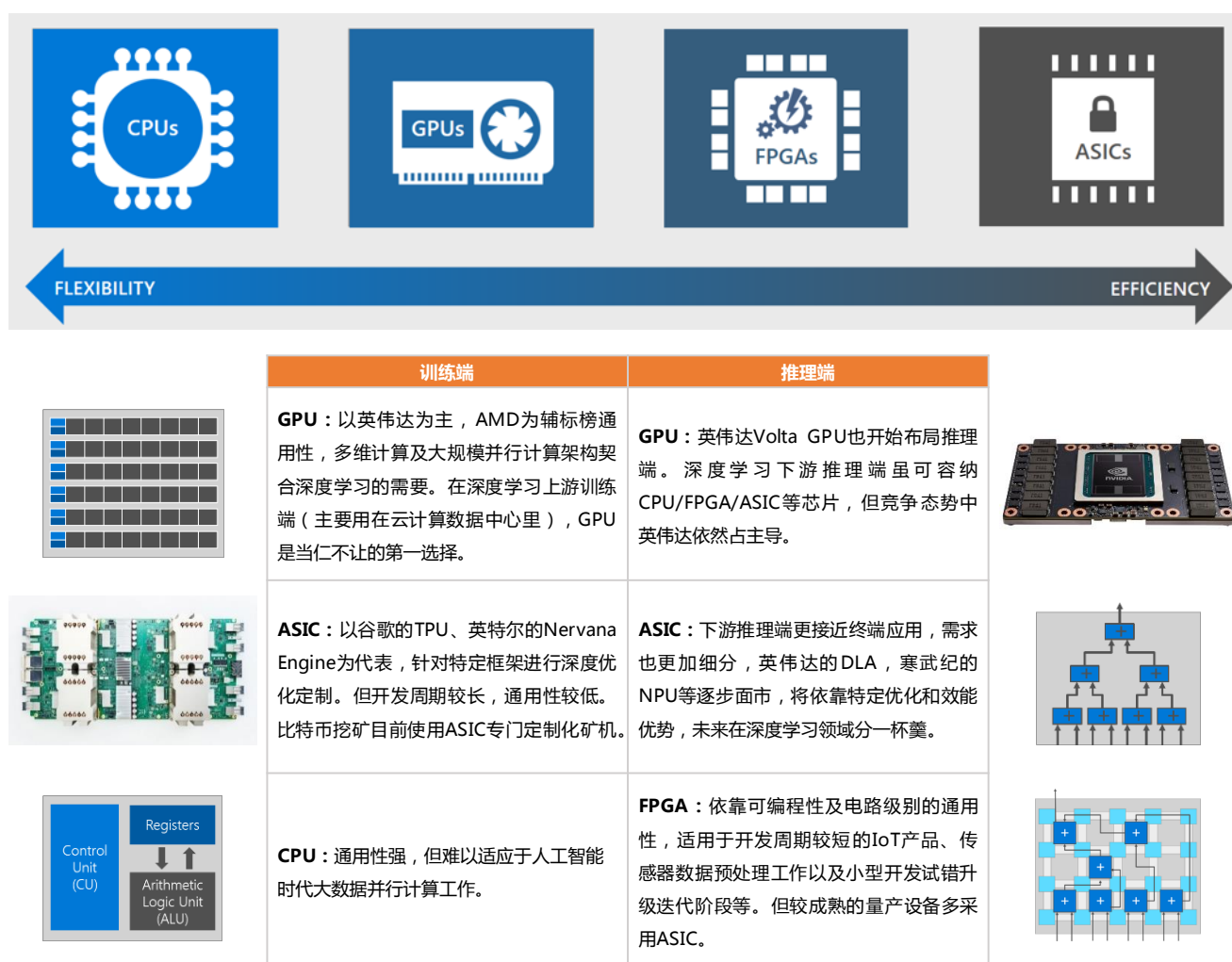
## 1. 谷歌 TPU：以时间换吞吐量，软硬兼施，冲入云端

AI 芯片市场蛋糕越做越大，足以让拥有不同功能和定位的芯片和平共存，百家争鸣非零和博弈。“通用性和功耗的平衡”——在深度学习上游训练端(主要用在云计算数据中心里)，GPU 是当仁不让的第一选择，ASIC 包括谷歌 TPU、寒武纪 MLU 等也如雨后春笋。而下游推理端更接近终端应用，需求更加细分，GPU 主流芯片之外，包括 CPU/FPGA/ASIC 也会在这个领域发挥各自的优势特点。

但我们需要强调，包括 TPU 在内的 ASIC 仍然面临通用性弱、开发成本高企等局限。TPU 虽然理论上支持所有深度学习开发框架，但目前只针对 TensorFlow 进行了深度优化。另外 ASIC 芯片开发成本非常高，在开发调试过程中复杂的设计花费有时甚至会超过 1 亿美元，因此需要谷歌这样的计算需求部署量才能将成本分摊到大量使用中。同时 ASIC 开发周期长，会出现硬件开发无法匹配软件更新换代而失效的情况。

ASIC (Application Specific Integrated Circuit, 专用集成电路): 细分市场确定后, 以 TPU 为代表的 ASIC 定制化芯片 (或者说针对特定算法深度优化和加速的 DSA, Domain-Specific-Architecture), 在确定性执行模型(deterministic execution model)的应用需求中发挥作用。我们认为深度学习 ASIC 包括英特尔的 Nervana Engine、Wave Computing 的数据流处理单元、英伟达的 DLA、寒武纪的 NPU 等逐步面市, 将依靠特定优化和效能优势, 未来在深度学习领域分一杯羹。

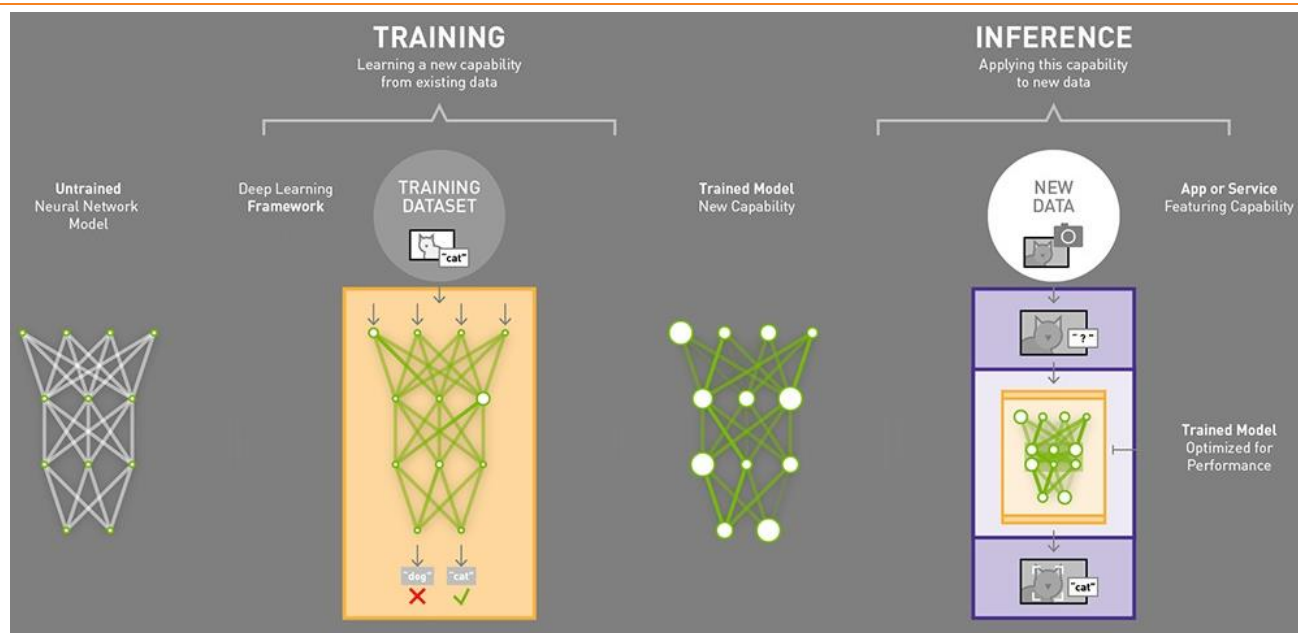
图 1：目前深度学习领域常用的四大芯片类型，“通用性和功耗的平衡”



资料来源：微软 Build，天风证券研究所

神经网络的两个主要阶段是训练( Training 和 Learning )和推理( Inference 和 Prediction )。当前几乎所有的训练阶段都是基于浮点运算的，需要进行大规模并行张量或多维向量计算，GPU 依靠优秀的通用型和并行计算优势成为广为使用的芯片。

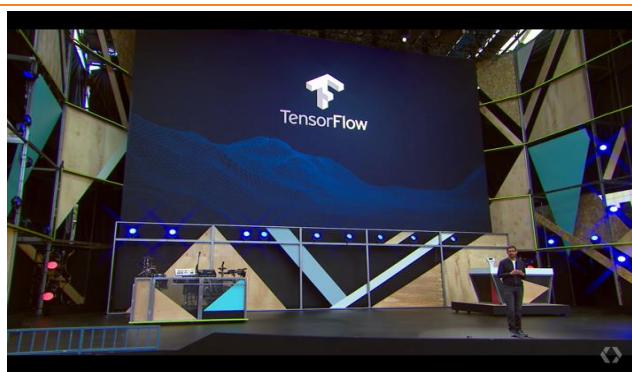
图 2：深度学习在神经网络模型的应用中主要分为上游训练端和下游推理端



资料来源：英伟达官网，天风证券研究所

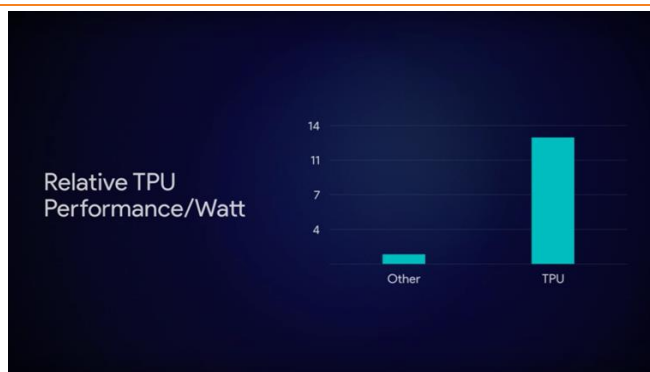
在推理阶段，由于更接近终端应用需求，更关注响应时间而不是吞吐率。由于 CPU 和 GPU 结构设计更注重平均吞吐量(throughput)的 time-varying 优化方式，而非确保延迟性能。谷歌设计了一款为人工智能运算定制的硬件设备，张量处理单元(Tensor Processing Unit, TPU)芯片，并在 2016 年 5 月的 I/O 大会上正式展示。

图 3：皮查伊在 2016 I/O 大会上介绍 TensorFlow



资料来源：谷歌 2016 I/O 大会现场照片，天风证券研究所

图 4：皮查伊介绍 TPU 性能对比



资料来源：谷歌 2016 I/O 大会现场照片，天风证券研究所

第一代 TPU 的确定性执行模型(deterministic execution model)针对特定推理应用工作，更好的匹配了谷歌神经网络在推理应用 99%的响应时间需求。第一代 TPU 是在一颗 ASIC 芯片上建立的专门为机器学习和 TensorFlow 量身打造的集成芯片。该芯片从 2015 年开始就已经在谷歌云平台数据中心使用，谷歌表示 TPU 能让机器学习每瓦特性提高一个数量级，相当于摩尔定律中芯片效能往前推进了七年或者三代。

谷歌表示，这款芯片目前不会开放给其他公司使用，而是专门为 TensorFlow 所准备。TPU 的主要特点是：

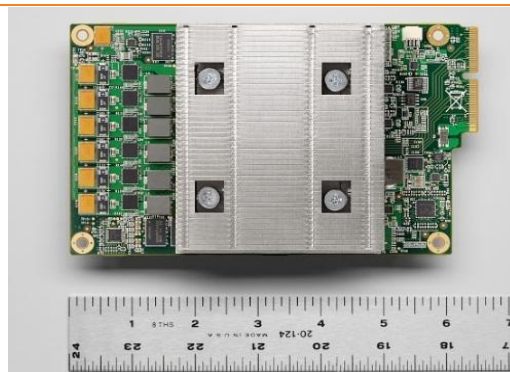
- 1、从硬件层面适配 TensorFlow 深度学习系统，是一款定制的 ASIC 芯片，谷歌将 TPU 插入其数据中心机柜的硬盘驱动器插槽里来使用；
- 2、数据的本地化，减少了从存储器中读取指令与数据耗费的大量时间；
- 3、芯片针对机器学习专门优化，尤其对低运算精度的容忍度较高，这就使得每次运算所动用的晶体管数量更少，在同时间内通过芯片完成的运算操作也会更多。研究人员就可以使用更为强大的机器学习模型来完成快速计算。

图 5：谷歌第一代 TPU 电路板



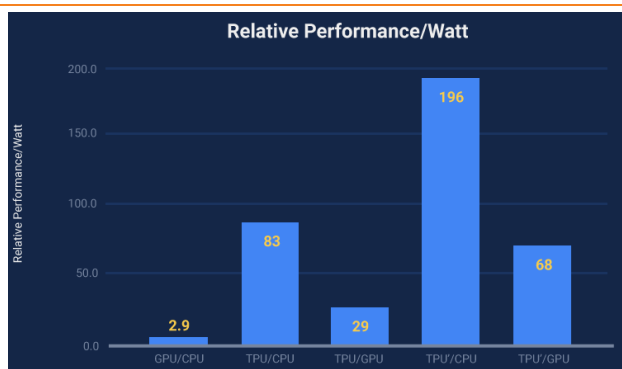
资料来源：谷歌研究所官方博客，天风证券研究所

图 6：谷歌第一代 TPU 尺寸示意图



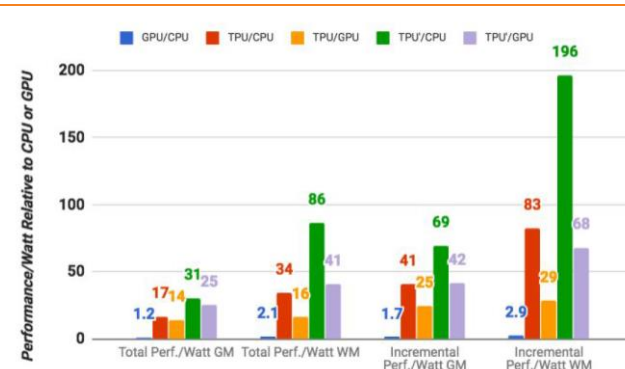
资料来源：谷歌研究所官方博客，天风证券研究所

图 7：TPU 的性能/功耗比较优势



资料来源：公司官网，天风证券研究所

图 8：TPU 的性能/功耗比较优势



资料来源：公司官网，天风证券研究所

自 2016 年以来，TPU 运用在人工智能搜索算法 RankBrain、搜索结果相关性的提高、街景 Street View 地图导航准确度提高等方面。在 I/O 大会上，皮查伊顺带提到了 16 年 3 月份行的举世瞩目人机大战里，在最终以 4:1 击败围棋世界冠军李世石的 AlphaGo 身上，谷歌也使用了 TPU 芯片。

谷歌把：

- 1、2015 年击败初代击败樊麾的版本命名为 AlphaGo Fan，这个版本的 AlphaGo 运行于谷歌云，分布式机器使用了 1202 个 CPU 和 176 个 GPU。
- 2、去年击败李世石的版本 AlphaGo Lee 则同样运行于云端，但处理芯片已经简化为 48 个第一代 TPU。
- 3、今年击败柯洁的 Master 以及最新版本 Zero 则通过单机运行，只在一个物理服务器上部署了 4 个第一代 TPU。（AlphaGo 的背后算法详解，可参见我们此前的深度报告《谷歌人工智能：从 HAL 的太空漫游到 AlphaGo，AI 的春天来了》）

图 9：AlphaGo 版本进化







版本	AlphaGo Fan	AlphaGo Lee	AlphaGo Master	AlphaGo Zero
时间	2015年10月	2016年3月	2017年5月	2017年10月
使用芯片	运行于谷歌云，分布式机器使用1202个CPU和176个GPU。	运行于谷歌云，使用48个第一代TPU。	单机运行，只在一个物理服务器上部署4个第一代TPU。	单机运行，只在一个物理服务器上部署4个第一代TPU。

资料来源：DeepMind 官网，天风证券研究所

### 1.1. 谷歌 TPU 软硬兼施，加速云端 AI 帝国

AI 芯片领域数据中心市场空间巨大，我们看到市场主流 GPU 之外，谷歌破局者之态依靠 TPU 2.0 的浮点运算升级自下而上进入云计算服务。谷歌当下不直接销售硬件，但将 TPU 部署在云计算中以云服务形式进行销售共享，在为数据中心加速市场带来全新的需求体验的同时，可进一步激活中小企业的云计算需求市场，另辟 AWS、Azure 之外蹊径。我们长期看好谷歌基于公司 AI First 战略规划打造 AI 开发软硬件一体化开发帝国。

不过 TPU 虽然理论上支持所有深度学习开发框架，但目前只针对 TensorFlow 进行了深度优化。而英伟达 GPU 支持包括 TensorFlow、Caffe 等在内所有主流 AI 框架。因此谷歌还在云计算平台上提供基于英伟达 Tesla V100 GPU 加速的云服务。在开发生态方面，TensorFlow 团队公布了 TensorFlow Research Cloud 云开发平台，向研究人员提供一个具有 1000 个云 TPU 的服务器集群，用来服务各种计算密集的研究项目。

图 10：TPU Pod 由 64 台第二代 TPU 组成



资料来源：公司官网，天风证券研究所

图 11：TensorFlow Research Cloud 云开发平台



资料来源：公司官网，天风证券研究所

### 1.2. 第一代 TPU：脉动阵列“获新生”，以时间换吞吐量

第一代 TPU 面向的推理阶段，由于更接近终端应用需求，更关注响应时间而不是吞吐量。相对于 CPU 和 GPU 结构设计更注重平均吞吐量(throughput)的 time-varying 优化方式，而非确保延迟性能。第一代 TPU 的确定性执行模型(deterministic execution model)针对特定推理应用工作，更好的匹配了谷歌神经网络在推理应用上 99%的响应时间需求。由于 TPU 没有任何存储程序，仅执行从主机发送的指令，这些功能的精简让 TPU 有效减小芯片面积并降低功耗。

谷歌在今年 4 月的体系结构顶会 ISCA 2017 上面，发布了一篇介绍 TPU 相关技术以及与其

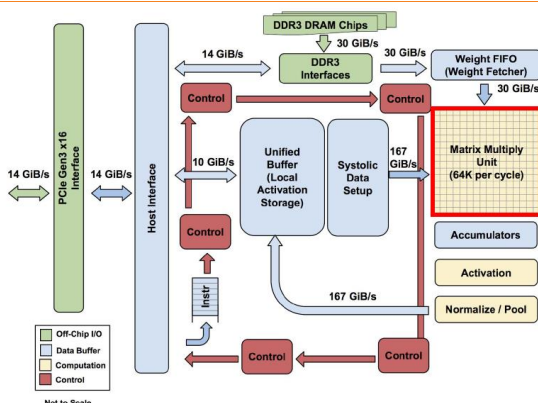
它硬件比较的论文，并被评为最佳论文。我们通过论文得以看到第一代 TPU 的设计思路以及性能比较。

第一代 TPU 从 2015 年开始就被使用在谷歌云计算数据中心的机器学习应用中，面向的是推理阶段。首先看性能比较（鉴于 2016 年以前大部分机器学习公司主要使用 CPU 进行推理，谷歌在论文中 TPU 的比较对象产品为英特尔服务器级 Haswell CPU 和英伟达 Tesla K80 GPU），谷歌表示：

- 1、针对自身产品的人工智能负载，推理阶段，TPU 处理速度比 CPU 和 GPU 快 15-30 倍；
- 2、TPU 的功耗效率（TOPS/Watt，万亿次运算/瓦特）也较传统芯片提升了 30-80 倍；
- 3、基于 TPU 和 TensorFlow 框架的神经网络应用代码仅需 100-1500 行。

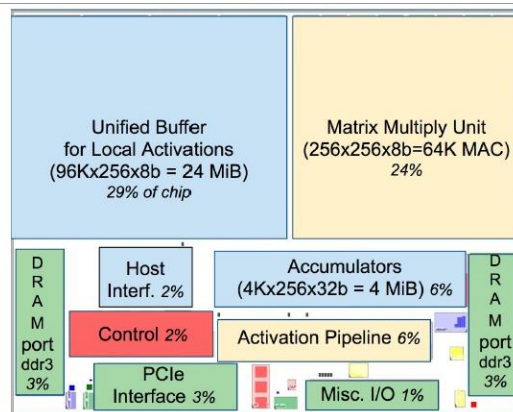
基于在成本-能耗-性能(cost-energy-performance)上的提升目标，TPU 的设计核心是一个 65,536(256x256)个 8 位 MAC 组成的矩阵乘法单元(MAC matrix multiply unit)，可提供峰值达到 92 TOPS 的运算性能和一个高达 28 MiB 的软件管理片上内存。TPU 的主要设计者 Norman Jouppi 表示，谷歌硬件工程团队最开始考虑过 FPGA 的方案，实现廉价、高效和高性能的推理解决方案。但是 FPGA 的可编程性带来的是与 ASIC 相比在性能和每瓦特性能的巨大差异。

图 12：第一代 TPU 各模块的框图，红框为核心矩阵乘法单元



资料来源：公司官网，天风证券研究所

图 13：第一代 TPU 的芯片布局图



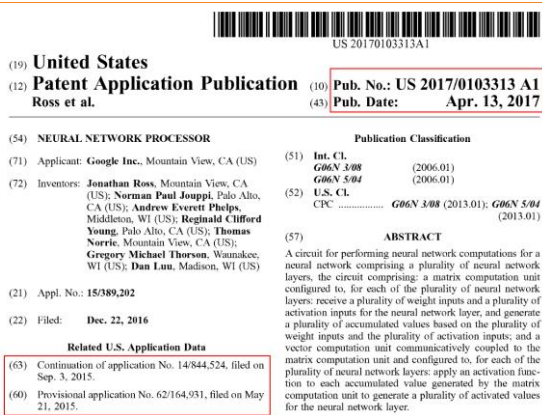
资料来源：公司官网，天风证券研究所

从上图我们看到，TPU 的核心计算部分是右上方的黄色矩阵乘法单元(Matrix Multiply unit)，输入部分是蓝色的加权 FIFO 和一致缓冲区(Unified Buffer，输出部分是蓝色的累加器(Accumulators)。在芯片布局图中我们看到，蓝色的缓存的面积占 37%，黄色的计算部分占 30%，红色的控制区域只占 2%，一般 CPU、GPU 的控制部分会更大而且难以设计。

我们深挖谷歌 TPU 论文，在参考文献中提及了谷歌申请的专利，核心的专利 Neural Network Processor 作为总构架在 2015 年就已提交，并在 2016 年公开（后续专利在 2017 年 4 月公开，专利号：US 2017/0103313，即下图 12 所示），同时还包括了几个后续专利：如何在该构架上进行卷积运算、矢量处理单元的实现、权重的处理、数据旋转方法以及 Batch 处理等。

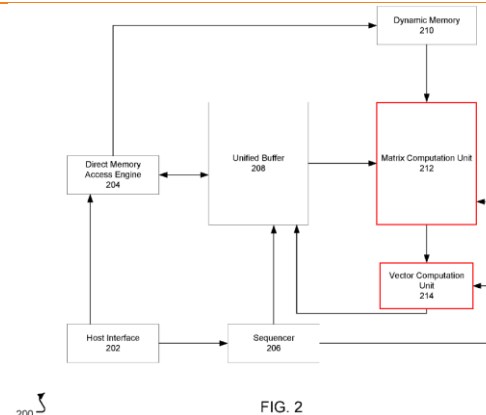
专利摘要概述：一种可以在多网络层神经网络中执行神经网络计算的电路，包括一个矩阵运算单元(matrix computation unit)：对多个神经网络层中的每一层，可以被配置为接收多个 weights 输入和多个 activation 输入，并对应生成多个累积值；以及矢量运算单元(vector computation unit)，其通信耦合到所述矩阵运算单元。

图 14: TPU 论文核心专利: Neural Network Processor



资料来源: Google TPU 专利, 天风证券研究所

图 15: 第一代 TPU 各模块设计原理专利, 核心计算单元为红框中的矩阵运算单元和矢量运算单元

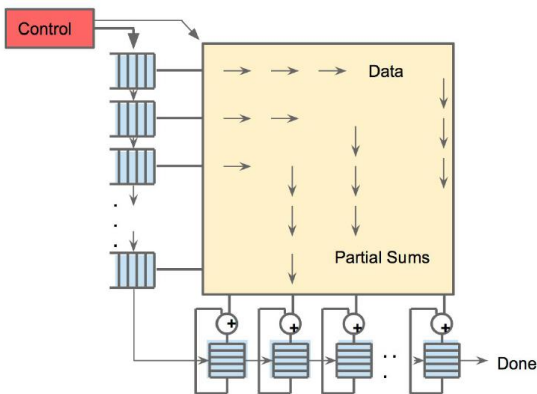


资料来源: Google TPU 专利, 天风证券研究所

TPU 的设计思路比 GPU 更接近一个浮点运算单元, 是一个直接连接到服务器主板的简单矩阵乘法协处理器。TPU 上的 DRAM 是作为一个独立的并行单元, TPU 类似 CPU、GPU 一样是可编程的, 并不针对某一特定神经网络设计的, 而能在包括 CNN、LSTM 和大规模全连接网络(large, fully connected models)上都执行 CISC 指令。只是在编程性上 TPU 使用矩阵作为 primitive 对象, 而不是向量或标量。TPU 通过两个 PCI-E 3.0 x8 边缘连接器连接协处理器, 总共有 16 GB/s 的双向带宽。

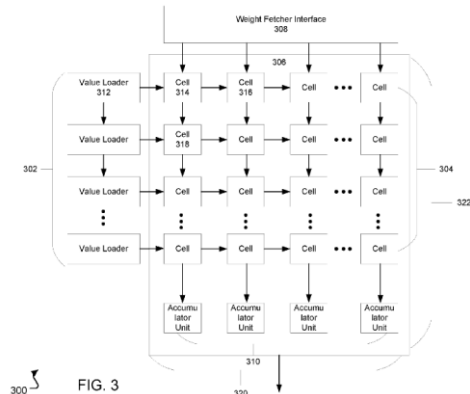
我们看到, TPU 的 matrix 单元就是一个典型的脉动阵列架构(systolic array computers)。weight 由上向下流动, activation 数据从左向右流动。控制单元实际上就是把指令翻译成控制信号, 控制 weight 和 activation 如何传入脉动阵列以及如何如何在脉动阵列中进行处理和流动。由于指令比较简单, 相应的控制也是比较简单的。

图 16: 矩阵乘法单元的脉动数据流(Systolic data flow)



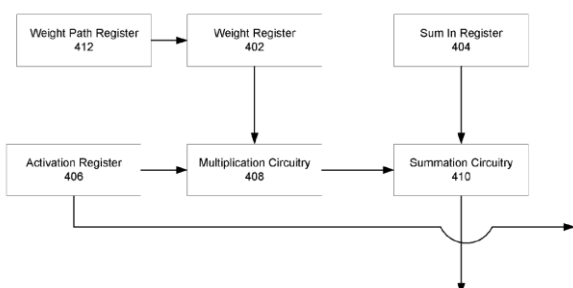
资料来源: 公司官网, 天风证券研究所

图 17: 矩阵运算单元的架构原理图



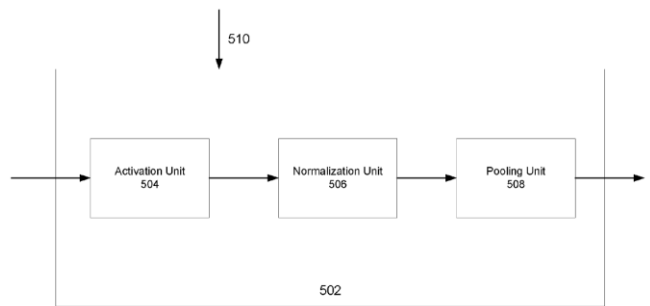
资料来源: Google TPU 专利, 天风证券研究所

图 18: 矩阵运算单元中一个 Cell 的架构



资料来源: Google TPU 专利, 天风证券研究所

图 19: 矢量运算单元的架构原理图



资料来源: Google TPU 专利, 天风证券研究所

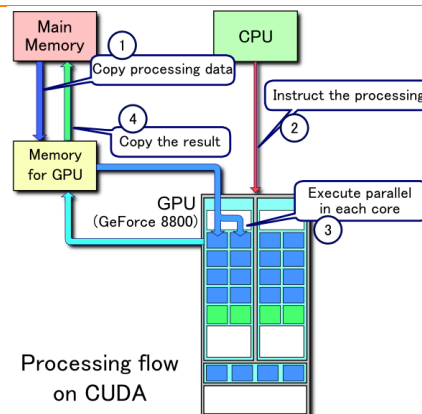
从性能上，脉动阵列架构在大多数 CNN 卷积操作上效率很好，但在部分其他类型的神经网络操作上，效率不是太高。另外脉动阵列架构在上世纪 80 年代就已经被提出，Simple and regular design 是脉动阵列的一个重要原则，通过简单而规则的硬件架构，提高芯片的设计和实现的能力，从而尽量发挥软件的能力，并平衡运算和 I/O 的速度。脉动阵列解决了传统计算系统：数据存取速度往往大大低于数据处理速度的问题，通过让一系列在网格中规律布置的处理单元(Processing Elements, PE)，进行多次重用输入数据来在消耗较小的带宽的情况下实现较高的运算吞吐率。但是脉动阵列需要带宽的成比例的增加来维持所需的加速倍数，所以可扩展性问题仍待解决。

图 20：英伟达 GeForce GTX 1070 Ti 模块框图



资料来源：公司官网，天风证券研究所

图 21：CUDA 核心计算处理流程图



资料来源：Wikipedia，天风证券研究所

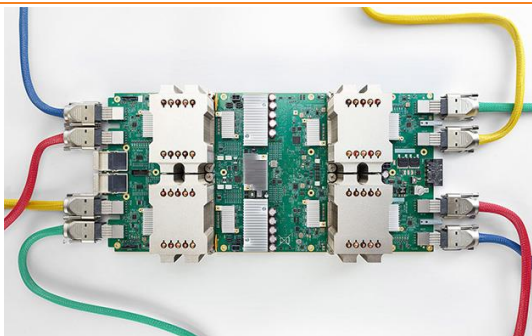
对比 GPU 的硬件架构，英伟达的游戏显卡 GeForce GTX 1070 Ti 使用的是 Pascal 架构 16 纳米制程，主频 1,607 MHz，拥有 2,432 个 CUDA 核心和 152 个纹理单元，2 MB L2 cache，功耗 180 W，8GB GDDR5 内存。英伟达 GPU 的核心计算单元 CUDA 核心专为同时处理多重任务而设计，数千个 CUDA 核心组成了 GPU 的大规模并行计算架构。而在计算过程中，主要计算流程为：1) 从主机内存将需要处理的数据 read 到 GPU 的内存；2) CPU 发送数据处理执行给 GPU；3) GPU 执行并行数据处理；4) 将结果从 GPU 内存 write 到主机内存。通过编译优化把计算并行化分配到 GPU 的多个 core 里面，大大提高了针对一般性通用需求的大规模并发编程模型的计算并行度。

### 1.3. 第二代 TPU：可进行深度学习上游训练计算

第二代 TPU，又名 Cloud TPU，能够同时应用于高性能计算和浮点计算，峰值性能达到 180 TFLOPS/s。与第一代 TPU 只能应用于推理不同，**第二代 TPU 还可以进行深度学习上游训练环节**。随着第二代 TPU 部署在 Google Compute Engine 云计算引擎平台上，谷歌将 TPU 真正带入云端。

谷歌在今年 5 月 17 日举办了 2017 年度 I/O 开发者大会。一场并未有太多亮点的大会上，谷歌 CEO 皮查伊继续强调公司 AI First 的传略规划。最为振奋人心的当属第二代 TPU——Cloud TPU 的发布。

图 22：第二代 TPU 包含 4 个芯片



资料来源：siliconangle，天风证券研究所

图 23：第二代 TPU 包含 4 个芯片



资料来源：siliconangle，天风证券研究所



谷歌同时发布了 TPU Pod，由 64 台第二代 TPU 组成，算力达 11.5 petaflops。谷歌表示 1/8 个 TPU Pod 在对一个大型机器翻译模型训练的只需要 6 个小时，训练速度是市面上 32 块性能最好的 GPU 的 4 倍。

谷歌此前强调，第一代 TPU 是一款推理芯片，并不用作神经网络模型训练阶段，训练学习阶段的工作仍需交由 GPU 完成。早在去年 I/O 大会上公布 TPU 之前，谷歌就已经将 TPU 应用在各领域任务中，包括：图像搜索、街景、谷歌云视觉 API、谷歌翻译、搜索结果优化以及 AlphaGo 的围棋系统中。

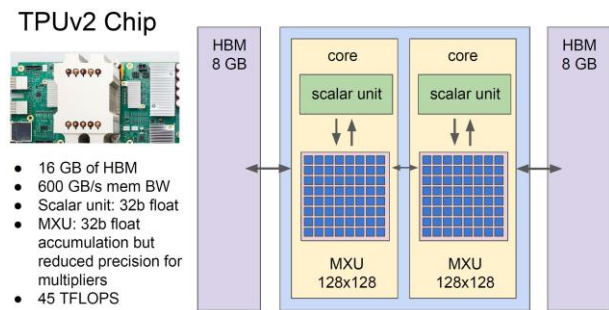
而这次第二代 TPU 的升级，自下而上的进入深度学习上游，应用在图像和语音识别，机器翻译和机器人等领域，加速对单个大型机器学习模型的训练。第二代 TPU 在左右两侧各有四个对外接口，左侧还有两个额外接口，未来可能允许 TPU 芯片直接连接存储器，或者是高速网络，实现更加复杂的运算以及更多的扩展功能。在半精度浮点数(FP16)情况下，第二代 TPU 的单芯片可以达到 45 Teraflops（每秒万亿次的浮点运算），4 芯片的设计能达到 180 Teraflops。（对比第一代 TPU 算力：8 位整数运算达 92 TOPS，16 位整数运算达 23 TOPS）

图 24：TPU Pod，由 64 台 TPU 组成，算力达 11.5 petaflops



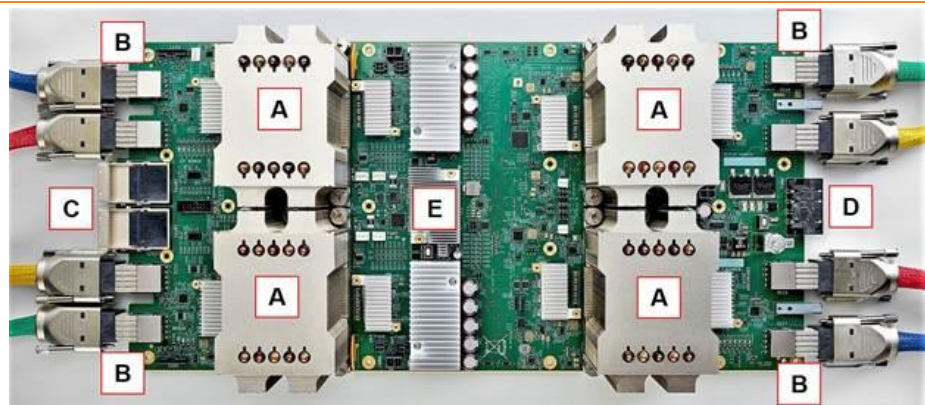
资料来源：公司官网，天风证券研究所

图 25：第二代 TPU 使用了 16 GB HBM 内存



资料来源：servethehome，天风证券研究所

图 26：A 是第二代 TPU 及散热片，B 是每块 TPU 的 2 根 BlueLink 25GB/s 电缆，C 是 Omni-Path 架构(OPA) 电缆接口，D 是电源连接器背面，E 可能为网络交换机

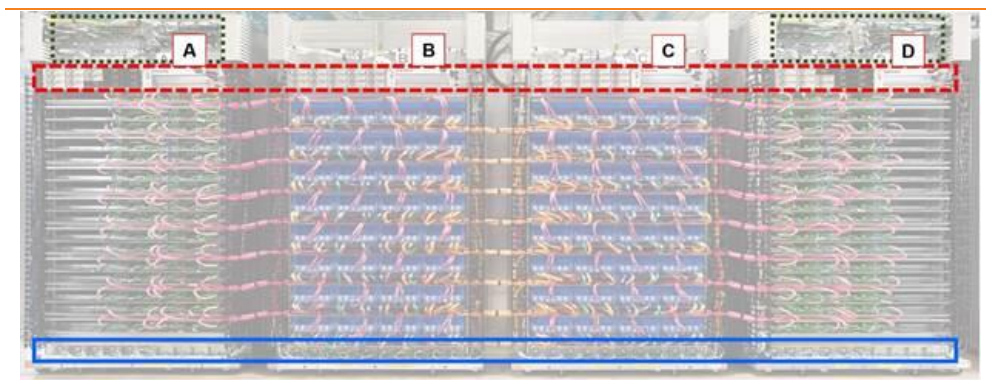


资料来源：The Next Platform，天风证券研究所

对 TPU Pod 的结构进行简要分析，四机架的镜像结构包含 64 个 CPU 板和 64 个第二代 TPU 板，The Next Platform 推测 CPU 板是标配英特尔 Xeon 双插槽主板，因此整个 Pod 机柜包括 128 个 CPU 芯片和 256 个 TPU 芯片。

The Next Platform 认为，谷歌使用两条 OPA 线缆将每块 CPU 板一一对应连接至 TPU 板，使得 TPU 与 CPU 的使用比例为 2:1，这种 TPU 加速器与处理器之间高度耦合的结构，与典型的深度学习加速结构中 GPU 加速器 4:1 或 6:1 的比例不太一样，更强调了 TPU 作为协处理器的设计理念——CPU 处理器还是需要完成大量的计算工作，只是把矩阵计算的的任务卸载到 TPU 中完成。

图 27: A 和 D 是 CPU 机架, B 和 C 是 TPU 机架, 蓝色方框为不间断电源(UPS), 红色方框为电源, 右上角绿色方框为网络交换机顶部



资料来源: The Next Platform, 天风证券研究所

#### 1.4. 谷歌重申买入: 人工智能巨头新征途: 云+YouTube+硬件

我们早在年初已经开始强调, 人工智能巨头新征途——云+YouTube+硬件。YouTube & 云计算的巨大增长动力将是谷歌持续转型的助推器, 长期看好 AI 和 Other Bets 创新业务厚积薄发。

3Q17 营收 277.7 亿美元, 同比涨 24%, 高于华尔街预期 219 亿美元, 主要鉴于移动端广告搜索业务和 YouTube 的增长。EPS 9.57 美元, 高于预期 8.31 美元。广告业务营收 240.7 亿美元, 同比涨 21%, 其他业务包括云计算和硬件销售达 34.1 亿美元, 同比大涨 40% (尚未囊括 10 月发布的 Pixel 2 等新产品销售收入)。新兴业务 Other Bets 营收同比涨 53% 至 3.02 亿美元, 但亏损环比略涨至 8.12 亿美元。

核心广告指标 Cost per click 实现环比转正, 移动端转型之势给予市场极大信心。谷歌股价 3 季度跑输大盘, 外部压力包括欧盟审查、美国选举操控等舆论监管压力。我们认为虽然在情绪面上承压, 但对公司业绩基本面影响有限。根据彭博一致预期 2018 年 EPS 40.15 美元, 给予 30x PE, 目标价 1200 美元, 重申“买入”评级。

##### YouTube 百般武艺冲劲十足, 移动端积极转型执行力坚决

YouTube 成长继续保持蓬勃动力, Pichai 表示用户通过电视观看 YouTube 的总时长达到 1 亿小时/日, 同比剧增 70%。YouTube TV 网络电视服务超过 30 个城市, 包括 40 个电视台节目的打包订阅费 35 美元/月, 仅为有线电视订阅均价的一半。根据 eMarketer 预测, 2017 年美国视频广告市场增速强劲, 整体规模预计增长 23.7% 至 132.3 亿美元, YouTube 作为龙头将贡献 21.7% 约 28.7 亿美元。

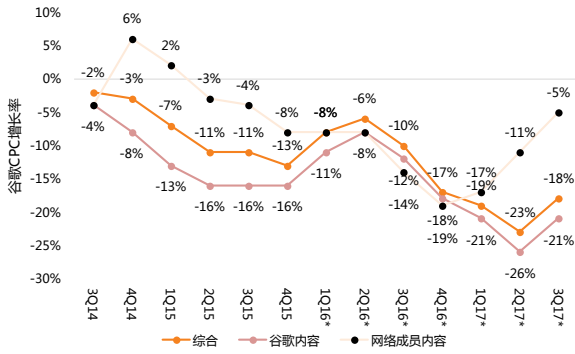
广告营收向移动设备转移步伐扎实, 广告业务净营收增速回升至 21%, 广告业务指标 Cost per click 同比降 -18%, 对比 Q2 的 -23% 和 Q1 的 -19%, 但 16 年以来环比首现转正。Paid clicks 同比涨 47%, 对比 Q2 的 52% 和 Q1 的 44%, 自由网站尤其是 YouTube 极大拉动用户点击意愿。我们强调, 在移动端获取搜索流量的成本会高于 PC 端, 谷歌需要向包括 iPhone 在内的合作伙伴支付更多的流量获取成本和收入分成, 谷歌已证明在移动广告上拥有不逊于 Facebook 的市场执行力。

##### 谷歌是人工智能的龙头标的

我们长期看好人工智能, 发力语音识别和无人驾驶: 我们认为语音识别技术已经足够进入普及。DeepMind 成为谷歌 AI 的标签门面, 看好进一步实现前瞻 AI 技术与现有业务的有效整合。

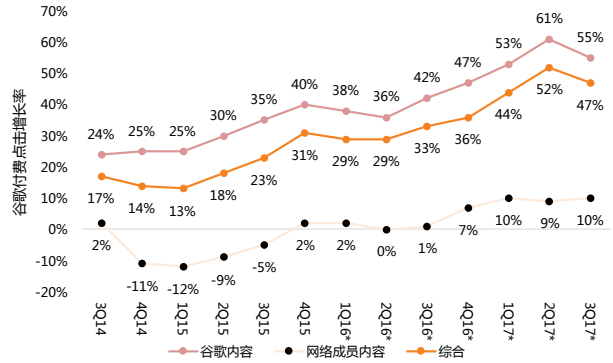
C 端谷歌软硬兼施, Pixel 手机+Home 音箱+Assistant AI 助理打造 AI 生态圈, 探索人机交互便捷方式和广告业务协同效应。9 月以 11 亿美元收购 HTC 打造 Pixel 手机的团队。无人驾驶业务 Waymo 初试共享经济, 领投 Lyft 把握用户入口将成为未来布局关键。

图 28: 谷歌 Cost-per-click 增长率 (\*号为算法调整后)



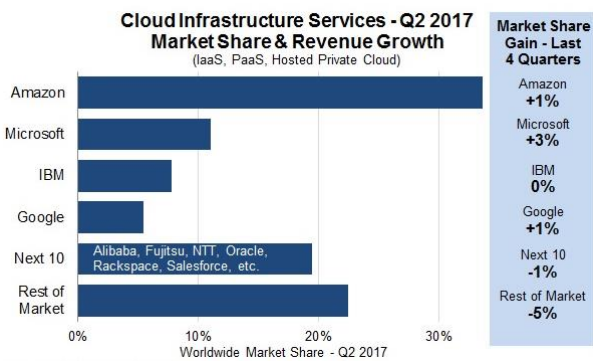
资料来源: 谷歌财报, 天风证券研究所整理

图 29: 谷歌 Paid clicks (\*号为算法调整后)



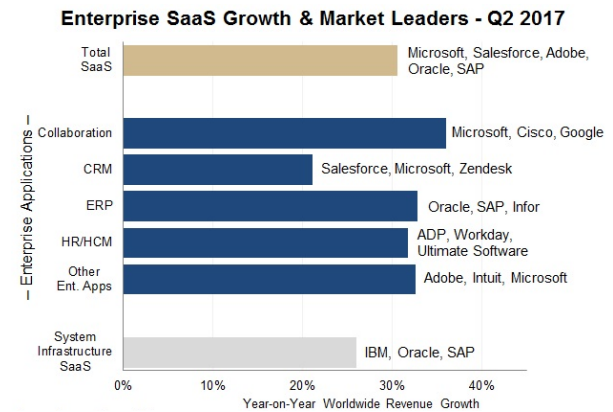
资料来源: 谷歌财报, 天风证券研究所整理

图 30: 全球云计算市场竞争格局



资料来源: Synergy Research, 天风证券研究所

图 31: 全球云计算企业 SaaS 市场格局

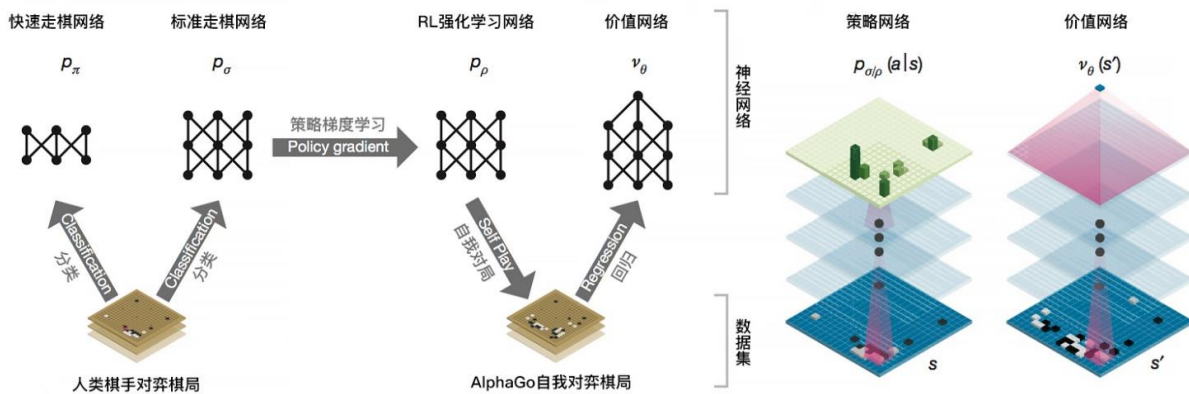


资料来源: Synergy Research, 天风证券研究所

## 2. 初代 AlphaGo 原理简介

我们在今年 1 月的《谷歌人工智能深度解剖》报告中详细介绍过 AlphaGo 的程序原理。简单来说, AlphaGo 的算法基于两个不同的部分: **蒙特卡洛树搜索和**指导树搜索的**卷积神经网络**。与以前的蒙特卡洛程序不同, AlphaGo 使用了深度神经网络来指导它的树搜索。**卷积神经网络**分为“**策略网络**”(这个网络又分为“**监督学习**”和“**强化学习**”两种模式)和“**价值网络**”。这两个神经网络以当前围棋盘面为初始值, 以图片的形式输入系统中。

图 32: AlphaGo 的神经网络

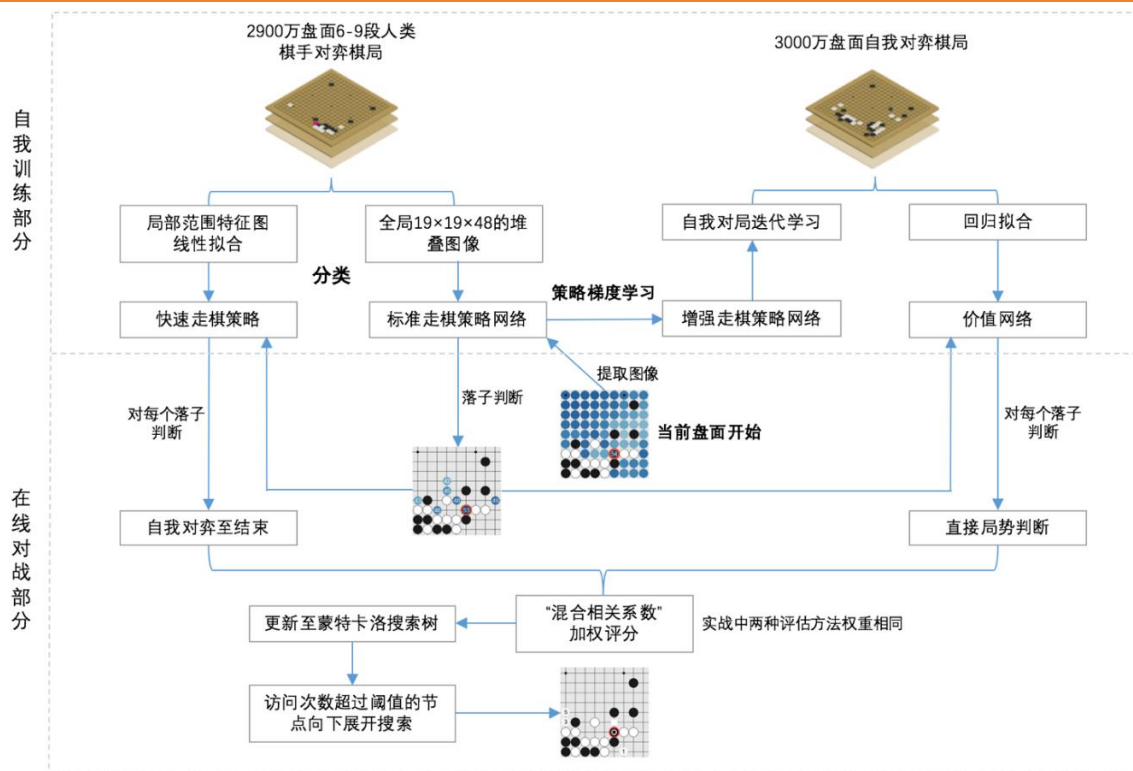


资料来源: Nature, 天风证券研究所

这里面的“策略网络”用来预测下一步落子并缩小搜索范围至最有可能获胜的落子选择。“价值网络”则用来减少搜索树的深度——对每一步棋局模拟预测至结束来判断是否获胜。与此前的蒙特卡洛模拟程序不同的是，AlphaGo 使用了深度神经网络来指导搜索。在每一次模拟棋局中，策略网络提供落子选择，而价值网络则实时判断当前局势，综合后选择最有可能获胜的落子。

下图是我们经过仔细研究 DeepMind 团队在学术杂志《自然》上发表的论文原文，精心制作的 AlphaGo 系统原理解图。

图 33：AlphaGo 系统原理解图



资料来源：参考 ACM 数据挖掘中国分会研究资料，天风证券研究所整理

### 3. AlphaGo Zero 的进化：强化学习，万物归零

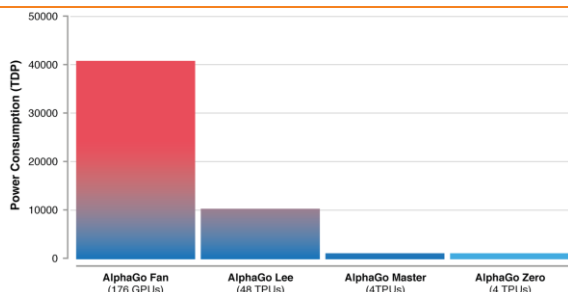
谷歌 DeepMind 最近在 Nature 上发布 AlphaGo 最新论文，介绍了目前最强版本 AlphaGo Zero，在没有先验知识的前提下，仅通过强化学习(Reinforcement Learning)自我训练，将先前的两个神经网络：价值网络和策略网络整合为一个框架，仅通过 3 天的训练就以 100:0 的成绩击败了曾经战胜李世石的 AlphaGo 版本。

我们认为 AlphaGo Zero 的进化是在围棋这个非常细分的特定问题上做的算法极致优化：

- 1、AlphaGo Zero 只使用棋盘上的黑白棋子作为输入，训练完全从随机落子开始，而此前的 AlphaGo 版本均使用了少量人工标注的特征。
- 2、Zero 背后的神经网络精简为一个，此前 AlphaGo 使用的两个神经网络：价值网络(Value Network)和策略网络(Policy Network)被整合为一个框架。在这个神经网络中，Zero 应用了更简单的 Monte-Carlo 树搜索，实现更高效的训练和评估。
- 3、Zero 没有使用“走棋策略”(Rollouts)，这是其他围棋程序以及 AlphaGo 最初版本中使用的快速走棋策略(Fast Rollout Policy)：进行快速判断较优落子选择，适当权衡走棋质量精准度与速度。相反 Zero 引入了全新的强化学习算法来在训练回路(training loop)中做前向搜索(lookahead search)。极大地提高了棋力以及学习稳定性。

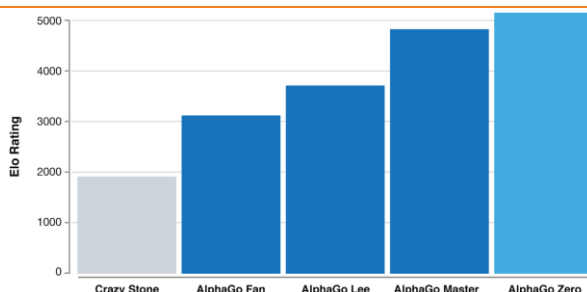
算法的改善提升了 AlphaGo 系统的性能以及通用性，并极大地降低了系统能耗。谷歌把 2015 年初代击败樊麾的版本命名为 AlphaGo Fan，这个版本的 AlphaGo 运行于谷歌云，分布式机器使用了 1202 个 CPU 和 176 个 GPU。去年击败李世石的版本 AlphaGo Lee 则同样运行于云端，但处理芯片已经简化为 48 个第一代 TPU。今年战胜柯洁的 AlphaGo Master，运算量只有上一代 AlphaGo Lee 的十分之一，在单机上运行，只在一个物理服务器上部署了 4 个第一代 TPU。

图 34: AlphaGo 4 个版本的能耗对比



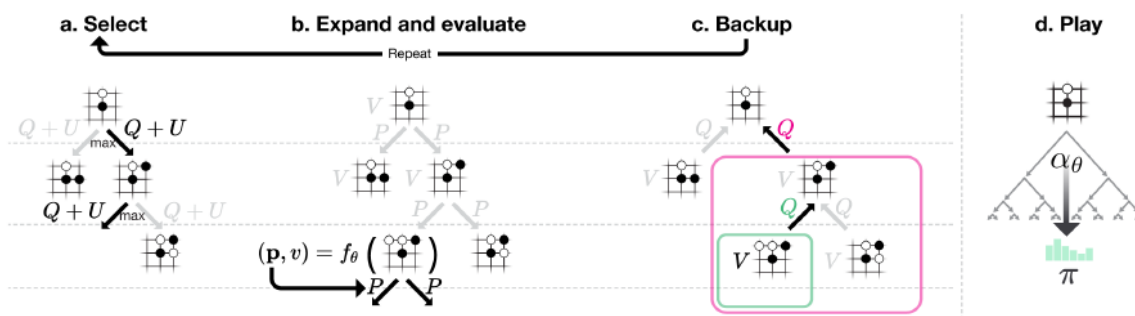
资料来源: DeepMind, 天风证券研究所

图 35: AlphaGo 4 个版本的围棋棋力对比



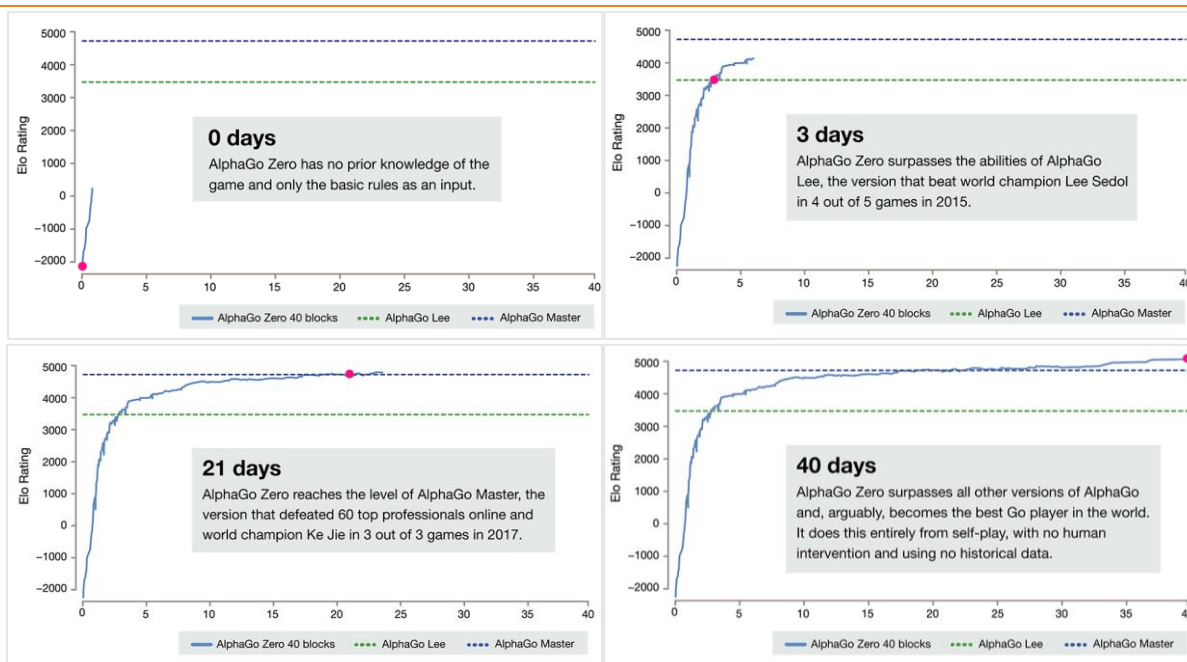
资料来源: DeepMind, 天风证券研究所

图 36: AlphaGo Zero 中的蒙特卡洛树搜索示意图



资料来源: Nature, 天风证券研究所

图 37: AlphaGo Zero 通过不同训练时间实现的棋力提升



资料来源: DeepMind, 天风证券研究所

## 4. 强化学习：减少对数据样本的依赖，里程碑后的新起点

AlphaGo 与李世石的世纪大战,是人工智能领域,尤其是深度强化学习技术的一个里程碑。

当前人工智能主流应用还是基于深度学习神经网络,从针对特定任务的标记数据中学习,训练过程需要消耗大量人类标注样本。在很多现实场景下,特定垂直领域的的数据并不足以支持系统建构,因此尝试解决对人类标注样本的依赖,包括强化学习、迁移学习(Transfer Learning)、多任务学习(Multi-task Learning)、零样本学习(zero-shot learning)有机会成为下一个机器学习商业成功的驱动力。

强化学习(Reinforcement Learning),就是智能系统从环境到行为映射的学习。人工智能的终极目标是模仿人类大脑的思考操作,而强化学习的灵感则来自于动物的学习方式。动物能够学会某些特定行为所导致的正面或负面结果(a positive or negative outcome)。按照这种方法,计算机可以通过试错法(trial and error)来与训练环境互动,包括 sensory perception 和 rewards,来决定这一结果的行为相关联。这使得计算机可以不通过具体指示或范例(explicit examples)去学习。

其实强化学习理论已经存在了数十年,但通过与大型深度神经网络的结合,让我们真正获得了解决复杂问题(如下围棋)所需的能力。通过不懈的训练与测试,以及对以前比赛的分析,AlphaGo 能够为自己找出了如何以职业棋手下棋的道路。

我们在今年 1 月的报告《2017 MIT 人工智能 5 大趋势预测:寒梅傲香春寒料峭,人工智能立夏将至》中提到第一大趋势预测:正向强化学习(Positive Reinforcement)正在成为深度学习(Deep Learning)之后研究应用的最新热点。虽然随着 Master 战胜柯洁,DeepMind 宣布了 AlphaGo 的退役,但我们认为 AlphaGo 的“终点”,正开启了强化学习的起点。

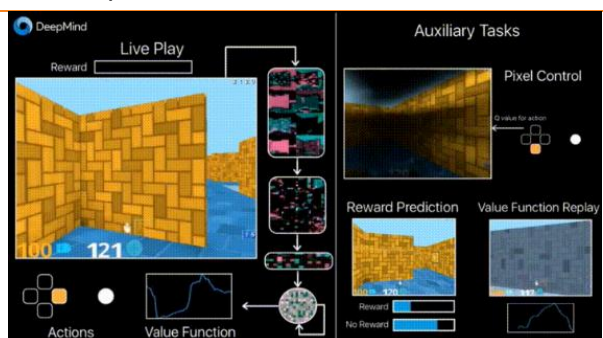
### 4.1. 游戏中的人机互动

从 Atari 到 Labyrinth,从连续控制到移动操作到围棋博弈,DeepMind 的深度强化学习智能系统在许多领域都表现出优异的成绩。人工智能的一大难题就是局限于在特定的板块和领域里学习。DeepMind 这个板块和领域中性的学习算法能够帮助不同的研究团队处理大规模的复杂数据,在气候环境、物理、医药和基因学研究领域推动新的发现,甚至能够反过来辅助科学家更好的了解人类大脑的学习机制。

可以预期的是,强化学习将能够在现实世界情景中得到更多的实用证明。过去一年中我们看到一些模拟环境(simulated environments)的推出,包括 DeepMind 的 DQN、OpenAI 的 Universe 以及著名的沙盒游戏《Minecraft》。这个游戏在 2014 年被微软收购,目前微软剑桥研究院的研究人员正通过这款游戏进行游戏内人与 AI 的互动协作研究。

研究人员开发了一个 Malmo 项目,通过这个平台使用人工智能控制 Minecraft 游戏里面的角色完成任务。这个项目被视为有效的强化学习训练平台。通过特定的任务奖励,人工智能能够完成研究人员布置的游戏任务,例如控制角色从一个布满障碍物的房间的一头走到另一头。同时,人工控制的角色还会在旁边提供建议,进行人与 AI 协作的测试。

图 38: DeepMind 打造的 3D 训练虚拟世界



资料来源: DeepMind, 天风证券研究所

图 39: Minecraft 训练界面



A character in Minecraft controlled by an algorithm and another operated by a person.

资料来源: 微软, 天风证券研究所

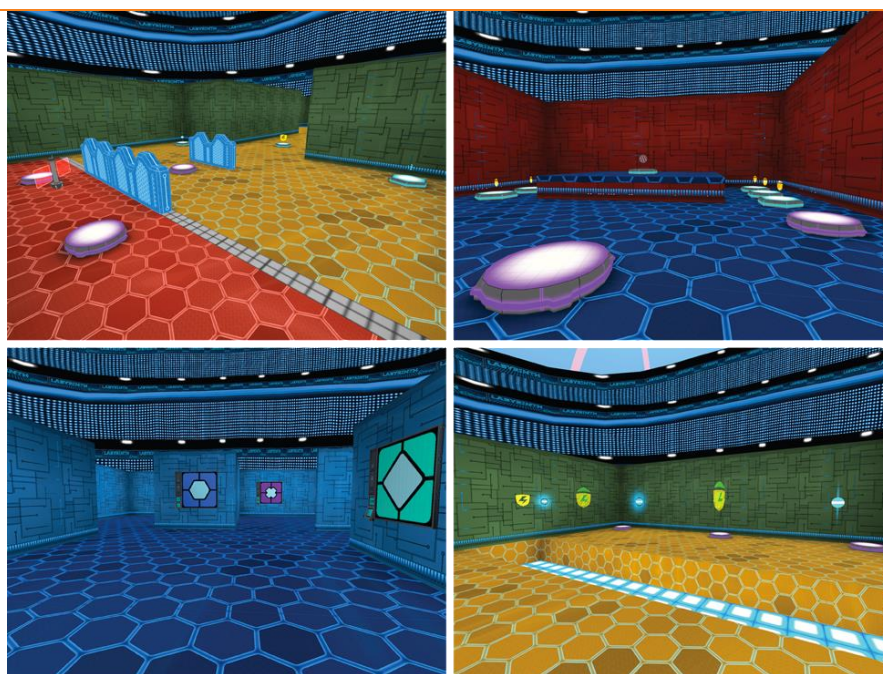
#### 4.1.1. DeepMind 之 Deep Q-Network (DQN)

DeepMind 在 2015 年 2 月于《自然》上发表了一篇《人类控制水平的深度强化学习》的论文，描述了其开发的深度神经网络 Deep Q-Network (DQN)将深度神经网络(Deep Neural Networks)与强化学习(Reinforcement Learning)相结合的深度强化学习系统(Deep Reinforcement Learning System)。Q-Network 是脱离模型(model-free)的强化学习方法，常被用来对有限马尔科夫决策过程(Markov decision process)进行最优动作选择决策。

谷歌设计的这个神经网络能够完成雅达利(Atari)游戏机 2600 上一共 49 个游戏，从滚屏射击游戏 River Raid，拳击游戏 Boxing 到 3D 赛车游戏 Enduro 等。令他们惊喜的是，DQN 在所有游戏过程都可以使用同一套神经网络模型和参数设置，研究人员仅仅向神经网络提供了屏幕像素、具体游戏动作以及游戏分数，不包含任何关于游戏规则的先验知识。

游戏结果显示，DQN 在一共 49 个游戏中的 43 个都战胜了以往任何一个机器学习系统，并且在超过半数的游戏中，达到了职业玩家水平 75%的分数水平。在个别游戏中，DQN 甚至展现了强大的游戏策略并拿到了游戏设定的最高分数。

图 40：DeepMind 开发的 3D 迷宫游戏 Labyrinth 界面



资料来源：DeepMind 官网，天风证券研究所

虽然 Atari 游戏为深度学习系统提供了多样性，但它们都还是二维动画层面的游戏。DeepMind 最近开始把研究重心放在 3D 游戏中并开发了一套 3D 迷宫游戏 Labyrinth 进行深度学习系统的训练。与之前类似，智能系统只获得了在视场(field-of-view)中观察到的即时像素输入，需要找到迷宫地图的正确宝藏路径。

DQN 作为第一个深度学习系统，通过进行端到端训练完成一系列有难度的任务。这样的技术能够有效运用到谷歌的产品服务中，我们可以想象一下，以后用户可以直接发出指令要求谷歌为他制定一个欧洲背包旅行计划。

#### 4.1.2. OpenAI Universe：通用 vs 具体

人工智能非牟利组织 OpenAI 去年推出了 Universe。这是一个用于训练解决通用问题的 AI 基础架构，能在几乎所有的游戏、网站和其他应用中衡量和训练 AI 通用智能水平的开源平台。这是继去年 12 月 OpenAI 发布可以用来开发强化学习算法的开发工具 Gym 之后，向通用型人工智能进一步扩展的新动作。

OpenAI Universe 的目标是开发一个单一的 AI 智能体，使其能够灵活运用过去在 Universe 中的经验，快速在陌生和困难的环境中学习并获得技能，这也是通往通用型人工智能的重

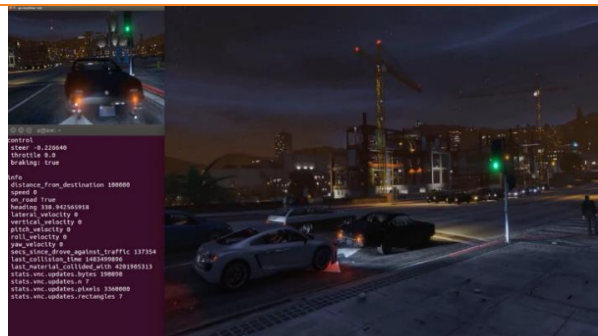
要一步。目前，Universe 包括了大约 2600 种 Atari 游戏，1000 种 flash 游戏和 80 种浏览器环境，可供所有人用于训练人工智能系统。最近，Universe 加入了游戏大作《GTA 5》。用户只需购买正版游戏，即可使用 Universe 中的人工智能在游戏中的 3D 环境中纵横驰骋。在 Universe 的新测试环境中，人工智能获取视频信息的帧数被限制在了 8fps，环境信息和视角管理齐备。此次开源 GTA 5 让普林斯顿大学开发的自动驾驶模拟测试平台 DeepDrive 在 GTA 世界中进行测试变得更加简便易行。

图 41: Universe 游戏环境范例



资料来源: OpenAI, 天风证券研究所

图 42: Universe 环境下的 GTA 5 自动驾驶测试界面



资料来源: GitHub, 天风证券研究所

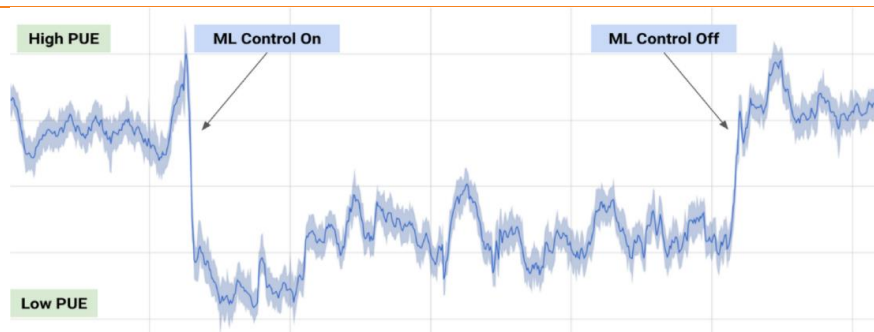
## 4.2. 更瘦、更绿的云计算数据中心

在 2017 年,我们预计看到强化学习更多的出现在自动驾驶系统和工业机器人控制等方面。谷歌一直在致力于更瘦、更绿的云计算数据平台。早在 2014 年,谷歌通过安装智能温度和照明控制以及采用先进的冷却技术而非机械冷却器,最小化能量损失,使其数据中心的耗电量比全球数据中心平均水平的低 50%。而且对比自己的数据中心,谷歌现在的数据处理性能是 5 年前的 3.5 倍,但能耗却没有提高。

如今,坐拥 DeepMind 的谷歌在很大程度上又走在世界前沿。DeepMind 将强化学习神经网络技术应用到云计算数据中心的能源控制方面,通过获取数据中心内的传感器收集的大量历史数据(如温度、功率、泵速、设定点等),首先在将来平均 PUE (Power Usage Effectiveness, 电力使用效率)值上训练神经网络系统。PUE 是总建筑能源使用量与 IT 能源使用量的比率,是衡量数据中心能源效率的标准指标,而每一部仪器可以受到几十个变量的影响。通过不断的模拟调整模型与参数,使其接近最准确预测的配置,提高设施的实际性能。团队训练两个额外的深层神经网络集合,以预测未来一小时内数据中心的温度和压力,模拟来自 PUE 模型的推荐行为。

通过 18 个月的模型研发与测试,DeepMind 联合谷歌云的研发团队成功为数据中心节省了 40% 的冷却能耗以及 15% 的总能耗,其中一个试点已经达到了 PUE 的最低点,未来该技术的可能应用于提高发电转换效率、减少半导体生产的能量和用水量,或帮助提高生产设施的产量。机器学习为数据中心节省能源,减少了更多的碳排放。DeepMind 和谷歌云计算团队计划将这项成果开源出来,造福全世界的数据中心、工厂和大型建筑等,打造一个更绿色的世界。

图 43: 数据中心 PUE 的机器学习测试结果



资料来源: DeepMind, 天风证券研究所整理



## 分析师声明

本报告署名分析师在此声明：我们具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，本报告所表述的所有观点均准确地反映了我们对标的证券和发行人的个人看法。我们所得报酬的任何部分不曾与，不与，也将不会与本报告中的具体投资建议或观点有直接或间接联系。

## 一般声明

除非另有规定，本报告中的所有材料版权均属天风证券股份有限公司（已获中国证监会许可的证券投资咨询业务资格）及其附属机构（以下统称“天风证券”）。未经天风证券事先书面授权，不得以任何方式修改、发送或者复制本报告及其所包含的材料、内容。所有本报告中使用的商标、服务标识及标记均为天风证券的商标、服务标识及标记。

本报告是机密的，仅供我们的客户使用，天风证券不因收件人收到本报告而视其为天风证券的客户。本报告中的信息均来源于我们认为可靠的已公开资料，但天风证券对这些信息的准确性及完整性不作任何保证。本报告中的信息、意见等均仅供客户参考，不构成所述证券买卖的出价或征价邀请或要约。该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。客户应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专家的意见。对依据或者使用本报告所造成的一切后果，天风证券及/或其关联人员均不承担任何法律责任。

本报告所载的意见、评估及预测仅为本报告出具日的观点和判断。该等意见、评估及预测无需通知即可随时更改。过往的表现亦不应作为日后表现的预示和担保。在不同时期，天风证券可能会发出与本报告所载意见、评估及预测不一致的研究报告。天风证券的销售人员、交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。天风证券没有将此意见及建议向报告所有接收者进行更新的义务。天风证券的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

## 特别声明

在法律许可的情况下，天风证券可能会持有本报告中提及公司所发行的证券并进行交易，也可能为这些公司提供或争取提供投资银行、财务顾问和金融产品等各种金融服务。因此，投资者应当考虑到天风证券及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突，投资者请勿将本报告视为投资或其他决定的唯一参考依据。

## 投资评级声明

类别	说明	评级	体系
股票投资评级	自报告日后的 6 个月内，相对同期标普 500 指数的涨跌幅	买入	预期股价相对收益 20%以上
		增持	预期股价相对收益 10%-20%
		持有	预期股价相对收益 -10%-10%
		卖出	预期股价相对收益 -10%以下
行业投资评级	自报告日后的 6 个月内，相对同期标普 500 指数的涨跌幅	强于大市	预期行业指数涨幅 5%以上
		中性	预期行业指数涨幅 -5%-5%
		弱于大市	预期行业指数涨幅 -5%以下

## 天风证券研究

北京	武汉	上海	深圳
北京市西城区佟麟阁路 36 号 邮编：100031 邮箱：research@tfzq.com	湖北武汉市武昌区中南路 99 号保利广场 A 座 37 楼 邮编：430071 电话：(8627)-87618889 传真：(8627)-87618863 邮箱：research@tfzq.com	上海市浦东新区兰花路 333 号 333 世纪大厦 20 楼 邮编：201204 电话：(8621)-68815388 传真：(8621)-68812910 邮箱：research@tfzq.com	深圳市福田区益田路 4068 号卓越时代广场 36 楼 邮编：518017 电话：(86755)-82566970 传真：(86755)-23913441 邮箱：research@tfzq.com