

人工智能及其对数据中心的影响

这可能有助于发现新的效率水平，伴随的是带宽需求的大幅增加

Tony Robinson, Corning

电影制作人如何能够将那些在当时看来与现实相去甚远的概念引入我们的日常生活中，这一点一直让人感到惊奇。1990年，阿诺德施瓦辛格 (Arnold Schwarzenegger) 的电影《Total Recall》向我们展示了“Johnny Cab”，这是一种无人驾驶汽车，可以载着我们去任何想去的地方。现在，大多数大型汽车公司都在投资数百万美元将这项技术推广到大众。多亏了《回到未来2》(Back to the Future II)，马蒂·麦克弗莱 (Marty McFly) 在磁浮滑板上躲避暴徒，可以让我们的孩子们站在类似于1989年我们看到的东墙上撞向家具 (以及彼此)。

回到1968年 (我们有些人还能记得)，我们用 HAL 9000 (电影《2001 太空漫游》中发现号宇宙飞船上的一台有感知的计算机) 接触到了人工智能 (AI)。HAL 能够说话和面部识别，自然语言处理，唇读，艺术欣赏，解释情感行为，自动推理，甚至可以下棋。

快进到过去几年，您可以非常快速的确定 AI 已经成为我们日常生活中不可或缺的一部分。您可以向智能手机询问下一个旅行目的地的天气情况，虚拟助理可以播放您喜欢的音乐，您的社交媒体账户将根据您的个人喜好更新新闻和广告。没有冒犯科技公司的意思，这就是 AI 101。

但是在这个背景下发生的事情太多了，我们看不到这有助于改善，甚至拯救生活。语言翻译、新闻推送、面部识别、复杂疾病的更准确诊断以及药物研发速度的加快，只是企业开发和部署人工智能的部分应用。据 Gartner 预测，人工智能带来的商业价值到2022年将达到3.9万亿美元。

缜密思考的服务器

那么人工智能是如何影响数据中心的呢？早在2014年，Google就在他们的一个数据中心部署了 Deepmind AI (使用人工智能的一种应用——机器学习)。结果呢？他们能够持续地减少40%用于冷却的能量，这相当于在考虑到电气损耗和其他非冷却效率之后，总 PUE 开销减少15%。这产生了该数据中心见过的最低 PUE。基于这些显著的节省，Google 希望在他们的其他数据中心里部署这项技术，并建议其他公司也这样做。

Facebook 的使命是“赋予人们建立社区的力量，让世界更紧密地联系在一起”，他们在白皮书《Facebook 应用机器学习：数据中心基础设施的视角》(Applied Machine Learning at Facebook: A Datacenter Infrastructure Perspective) 中概述了这一点。它描述了在全球范围内支持机器学习的硬件和软件基础架构。

为了让你对 AI 和 ML 需要多少计算能力有个基本概念，百度硅谷实验室的首席科学家 Andrew Ng



表示，训练一个百度的中文语音识别模型不仅需要 4 terabytes 的训练数据，还需要 20 个计算机的 exaflops 计算量，也就是整个培训周期内需要 200 亿亿次数学运算。

但是我们的数据中心基础设施呢？人工智能是如何影响不同规模和类型基础设施的设计和部署的，我们正在建造、租用或升级数据中心基础设施，以适应这种创新的、节省成本的和更为高效的技术。

ML 可以在一台机器上运行，但由于难以置信的数据吞吐量通常会在多台机器上运行。所有机器都互相连接，在训练和数据处理阶段，所有设备都相互连接，以确保持续的通信和低延迟，绝对不会中断在我们指尖、屏幕或音频设备上的服务。作为一个人类，我们对越来越多数据的渴望正在推动带宽指数的增长，以满足我们最简单的想法。

这种带宽需要在设备内部和跨多个设备分布，使用更复杂的架构设计 (spine-and-leaf)。我们说的 super-spine 和 super-leaf 它们为所有复杂的算法提供了一条高速公路，以便数据在不同的设备之间传输，最终回到我们的接收器。

数据中心中的技术部署选项

这就是光纤在将您的特殊（或搞怪）时刻的图片或视频播放给全世界观看，分享和评论方面发挥关键作用的地方。与铜缆相比，光纤具有高速和超高密度的性能，已成为我们数据中心基础设施中的实际传输介质。随着我们向更高的网络速度迁移，还在混合部署中引入了全新的复杂性，也就是将采用哪种技术？

传统的 3 层网络架构使用核心、聚合和边缘交换来连接数据中心内的不同服务器，服务器间的通信通过光收发器以南北向的方式彼此通信。然而现在，非常感谢 AI 和 ML 为游戏带来的高计算要求和相互依赖性，更多的网络使用 2 层的 spine-and-leaf 网络，由于生产和培训网络所需的超低延迟，服务器之间以东西方向进行通信。

自 2010 年 IEEE 批准 40G 和 100G 网络传输技术标准以来，出现了许多相互竞争的解决方案，这些解决方案在一定程度上影响了用户的判断，因为他们不确定应该采用哪种方式。解释一下，在 40G 之前，我们使用的是 SR（即多模的短距离传输）和 LR（即单模的长距离传输）。两者都使用一对光纤在两个设备之间传输信息。无论您使用哪种设备或在该设备中安装了哪种收发器，这都是通过两芯光纤进行的简单数据交换。

但是随着 IEEE 批准了 40G 及更高速率的解决方案后，游戏规则发生了改变。现在我们正在研究使用标准批准的或专有的两种方案，专有的 WDM 技术，和标准批准的、或多源协议 (MSAs) 认可的并行光学技术，并行光学技术使用 8 芯光纤 (4 芯发送和 4 芯接收) 或 20 芯光纤 (10 芯发送和 10 芯接收)。

- 如果您想使用标准认可的解决方案并降低光学器件成本，并且在不需要单模光纤的长距离能力的情况下，您可以选择多模并行光学技术。它还可以让您将高速 40 或 100G 的交换机端口拆分为更灵活的 10 或 25G 服务器端口。我将在本文中更详细的介绍这一点。
- 如果您希望延长已安装的双工光纤系统的使用寿命，并且不介意在没有通用性选项的情况下继续使用单一的硬件供应商，也不需要更长的传输距离，那么您可以选择多模 WDM 解决方案。

现在我要告诉你的是，大多数大规模部署人工智能的科技公司都在为今天和明天的网络设计…单模并行光学系统。这里有三个简单的原因。

1. 成本和距离

目前的市场趋势是并行光学解决方案首先开发和发布，几年后 WDM 解决方案也随之发布，因此使用并行光学解决方案的数量大大增加，从而降低了制造成本。与 2 km 和 10 km WDM 解决方案

相比，并行光学解决方案支持更短的距离，因此您不需要太多复杂的组件来冷却激光器，并对两端信号进行复用和解复用。虽然我们已看到这些“超大型”设施的规模已经激增到 3-4 个足球场的大小，但我们的数据显示，单模光纤的平均部署长度尚未超过 165 米，因此无需使用更昂贵的 WDM 收发器来满足他们不需要支持的距离。

2. 灵活性

部署并行光学设备的一个主要优势是能够使用高速交换机端口，比如 40G，并将其分解为 4x10G 服务器端口。端口拆分提供了巨大的经济性，因为拆分低速端口可以显著的将电子设备的机箱或机架安装单元数量减少到 1/3（数据中心资产并不便宜），并且使用更少的电源，需要更少的冷却，从而进一步降低能源消耗。我们的数据显示这相当于在单模解决方案上节省 30%。光纤收发器供应商还确认，在所有销售的并行光收发器中，有很大一部分是为了利用这个端口拆分能力而部署的。

3. 简单清晰的迁移

主要交换机和收发器厂商的技术路线图部署并行光学的客户提供了非常清晰和简单的迁移路径。我之前提到过，大多数科技公司都遵循这条路径，所以当系统从 100G 迁移到 200G 或 400G 时，光纤基础设施仍然保持不变，不需要任何升级。那些决定使用双工，2 芯光纤为布线基础的公司可能会发现他们想要升级到 100G 以上，WDM 光学技术可能无法在其迁移计划的时间范围内完成。

对数据中心设计的影响



从连接性的角度来看，这些网络是高度网状的光纤基础设施，以确保没有任何一台服务器彼此之间的网络跳数超过两个。但是带宽需求如此之大，以至于从 spine 交换机到 leaf 交换机的 3:1 的收敛比配置是不够的，因此更常用的是从不同数据核心之间的超级 spine 进行分布式计算。

由于交换机 IO 速度的显著提高，网络运营商正在努力提高利用率，我们通过使用从 spine 到 leaf 的 1:1 收敛比例设计系统，达到了更高的效率和超低的延迟，在当今的人工智能环境中，这是一个昂贵但必要的需求。

此外，在谷歌最近宣布推出最新的人工智能硬件后，我们又一次改变了传统的数据中心设计，这是一种定制的专用集成电路，称为张量处理单元（TPU 3.0），在其巨大 pod 设计中，其功率将是过去的 TPU 的 8 倍，超过 100 petaflops。但是，在芯片中加入更多的计算能力也会增加驱动它的能量，因此也会增加热量，这也是为什么大量的声明声称，由于 TPU 3.0 产生的热量已经超过了以前的数据中心冷却解决方案的限制，将转向对芯片进行液体冷却的原因。

最后

人工智能是下一波商业创新。它所带来的优势来自运营成本的节约、额外的收入流、简化以及更高效的客户交互，数据驱动的工作方式带来的优势太有吸引力了——不仅对 CFO 和股东如此，对客户也是如此。这一点在最近的一次组内讨论中得到了证实，专家当时表示使用聊天机器人的网站，如果效率不高且客户关注度不够，客户就会放弃对话，并且很难再次合作。

因此，我们必须接受这项技术并从中受益，这也意味着采用一种不同的方式来思考数据中心的设计和实施。由于 ASIC 的性能显着提高，我们最终会看到 IO 速度的提高，从而进一步提高系统的联通性。概括来讲，您的数据中心在支持 ML 训练的同时，需要超高效率，高密度，超低延迟，东西向的 spine-and-leaf 光纤网络，以适应您的日常流量需求。

我们已经看到，主要的科技公司是如何接受人工智能的，以及采用并行单模技术如何帮助它们比传统的双工模式获得更好的资本和运营成本的，后者承诺从第一天起就降低成本。但是，数据中心的运营从第二天就开始了，随着我们个人及行业内交流习惯和方式的不断变化、速度的提高和复杂性的增加，数据中心的运营也在不断发展。现在部署正确的布线基础设施解决方案，将使您的企业从一开始就获得更大的经济效益，留住和吸引更多的客户，并使您的数据中心设施能够更灵活的蓬勃发展。

康宁光通信中国

上海市漕河泾高科技开发区桂箐路 111 号立明大厦 3 楼

电话: 86 21 5450 4888

传真: 86 21 5427 7898

www.corning.com