

紫光展锐科技有限公司CEO 楚庆:

## 没有5G的AI是搁浅的航船



“没有AI的5G是一条空船,没有5G的AI是搁浅的航船。基于此,紫光展锐向世界推出5G与AI的完美组合虎贲T710+春藤510。”

本报记者 李佳师

9月3日,在工业和信息化部、上海市人民政府指导下,中国半导体行业协会、中国电子信息产业发展研究院联合主办的第二届全球IC企业家大会暨第十七届中国国际半导体博览会(IC China 2019)在上海举行。紫光展锐科技有限公司CEO 楚庆出席会议并发表了主题演讲。

楚庆在主题演讲中表示,5G是人类历史上最野心勃勃的网络连接计划,它目标是让所有的石头都上网。而AI,它是人类历史上最野心勃勃的科技革命,目标是让所有的石头都说话,这是革自己的命。关于5G与AI的关系,楚庆说:“没有AI的5G是一条空船,没有5G的AI是搁浅的航船。基于此,紫光展锐向世界推出5G与AI的完美组合虎贲T710+春藤510。”

## 5G将是

## 未来智能社会基础

在这次论坛上,楚庆演讲的内容是关于5G以及相关通信技术的前世今生。5G到底来了没有?目

前已经有5G手机开始销售,而且一些公司的5G手机相当热销,甚至是刚问世,今年的产能就全部订光。各种各样的信息都在显示,5G非常近了。但要想享受到5G的服务还需要一点时间,因为现在只是商用试验网。试商用的目标是商用,技术水平是商用,但仍然需要试验、试用。移动通信领域的人都知道,移动通信最重要的是频点,没有无线信号移动通信就不存在,只有有了网络才能有服务,而这个频点一旦改变,就会对网络的布局产生非常深入的影响,包括网络如何规划,蜂窝怎么布,频段应该怎么划分,网络模型如何调整等,只有解决这些问题,才能够成为具有服务价值的网络。要解决这些问题,要达到目标,需要经过多轮网络优化、重整、调整、试用。

5G真正商用之后会带来哪些好处?楚庆表示,5G带来的第一个好处是网络速度快,速度提升7到10倍。很多网红已经用他们的视频验证了5G带来的网络速度的提升。人们常说“由奢入简难,由简入奢易”,当你使用惯了高速的网络,就很难忍受网络速度变慢,所以说

速度提升是5G带给用户的第一个“鸦片”。第二个好处是密度高。比如浦东嘉里酒店这样一个区域,可以有100万个接入,这样的接入显而易见不是为人而设置,因为这样大小的地方无法聚集这么多人,所以它是为物连接而来的,是构成未来智能社会尤其是智能制造的基础。第三个好处是实时。实时就是有承诺的延时,就是说如果你承诺延时是5毫秒,那么所有的动作都必须在5毫秒内完成,不能一会儿延时长一会延时长,当你承诺了5毫秒的延时,就是保证了所有延时不能高于这个数值。第四是无缝升级。未来的升级是纯柔性的升级。

## 5G与AI

## 互相成就

另一个有趣的话题是关于WiFi和5G。WiFi和5G是一对相爱相杀的伴侣,它代表了两大无线通信的体制,一个是PLMN公众陆地移动通信网,我们所说的1G、2G、3G、4G、5G都是PLMN,手机支持PLMN的通信设备,

PLMN的目标是在人类所有可达的地方都要提供网络,无论是珠穆朗玛峰的峰顶还是在南极洲,只要人类可及就希望提供网络。WiFi英文翻译成汉语是无线本地环路,就是无线的本地接入,不需要连接任何数据线可以进行通信,这是它的目标,目标小,但不需要申请频点。现在已经到处都有WiFi。WiFi是放任的技术体制,而PLMN背后是非常严谨的标准体制。WiFi是在2G时代出现的,那时候彼此互补,3G时代这两者和平相处,4G时硝烟起来,这两者之间开始有了冲突。到了5G时代,这两者开始互相进攻对方的领地,WiFi也开始希望漫游,能够切换,甚至能大规模的组网,把世界上所有的WiFi都连接起来了,拿着WiFi6的手机到处可以享受数据通信。

楚庆谈到的另一个话题是5G和AI。楚庆表示,如果没有AI,5G装什么呢?它就是空船,没有5G的AI,再大的船也是搁浅在沙滩上。这两者谁也离不开谁,所以紫光展锐决定向世界推出AI+5G的完美组合,即T710和510。T710三周之前在欧洲苏黎世理工组成的测试中,获得了冠军。510就是展锐在巴塞罗那展发布的5G处理器,这个处理器是5G全功能处理器,包含所有十种通信技术,而且有强大的CPU和自我电源管理能力。我们相信这个组合将使人类愿望更美好。

楚庆最后表示,5G既然来了,6G也不会远了,6G速度一定比5G更快,那么它还会有什么特征呢?一定会包含5G所有具有的特征,还会有更多的新特征出现,究竟是什么?人类打开的永远是一个潘多拉的魔盒,我们不知道里面释放出什么,也许是天使,也许是魔鬼,但是人类会永远往前走,科技进步不会停息。

英伟达公司全球副总裁潘迪:

从边缘到云  
人工智能将无处不在

“有数据显示,目前传统企业AI的渗透率不到3%或4%。AI还需要成长,且发展潜力巨大,AI需要5G技术等多方面的‘拥抱’,需要多一点‘爱’。”

本报记者 卢梦琪

9月3日,在工业和信息化部、上海市人民政府指导下,中国半导体行业协会、中国电子信息产业发展研究院联合主办的第二届全球IC企业家大会暨第十七届中国国际半导体博览会(IC China 2019)在上海举行。英伟达公司全球副总裁潘迪出席会议并发表了主题演讲,潘迪在主题演讲中表示,未来从边缘到云,人工智能将无处不在。

## AI改变半导体企业

## 半导体也在改变AI

AI和半导体都是上世纪50年代的产物,之后的大部分时间都是在分别走自己的路,直到四五年前基于加速计算技术,二者才走向融合。“AI+5G”时代,传统的串型计算模式难以满足需求,在深度学习、机器学习、加速计算之下,AI正在改变半导体企业,半导体也在改变AI。

上世纪80年代以来,关于云架构与端架构,集成式与分布式,一直没有绝对的孰好孰坏的定论,而是要根据具体应用场景来评判,其不同领域都有各自该解决的问题。潘迪表示,未来从边缘到云,人工智能将无处不在。人工智能正在经历AI计算、云AI、企业级AI到自主AI的进阶,然而其真正的落地和商业化还有一定距离。有数据显示,目前传统企业AI的渗透率不到3%或4%。“AI还需要成长,且发展潜力巨大,AI需要5G技术等多方面的‘拥抱’,需要多一点‘爱’。”潘迪表示。

## 人工智能越来越复杂

## 应用场景将进一步拓宽

英伟达业务集中在三个核心市场,分别为图形、HPC和人工智能。图形处理包括游戏、专业可视化和渲染,HPC包括高性能计算,人工智能包括AI模型的训练,云端、数据中心端、边缘推理。英伟达在26年前开展了图形方面的早期工作,创建了GPU的并行处理架构,这种架构至今仍在继续,这为开发早期高性能计算应用奠定了基础,现代计算机图形学已经整合了人工智能,使计算机图形学更加逼真。现在,图形、HPC和人工智能这三个方面并非相互独立,而是关联交叠的,通过AI、算法、算力的协同来针对不同场景形成解决方案。

人工智能软件正在变得越来越复杂,经历语音识别、语音建模、从文本到语音转换再到语音编码器的过程,一旦整体实现突破,应用场景将被拓宽。潘迪表示,对话型AI是人工智能最大的创新应用。目前,已经有众多巨头企业基于

BERT模型进行优化,从中做了许多工作。从2017年谷歌的Transformer、2018年谷歌的BERT,到2019年微软的用于学习通用语言嵌入的多任务深度神经网络模型MT-DNN、阿里巴巴的Enriched BERT base、优步的Plato、百度的ERNIE等,已经取得了发展。在GLUE(通用语言理解)基准测试中,Facebook AI的RoBERTa排名第一,XLNet Team的XLNet-Large(ensemble)排名第二,微软D365 AI & MSR AI的MT-DNN-ensemble排名第三,GLUE Human Baselines的GLUE Human Baselines排名第四。潘迪相信,未来会有更多中国公司进入这个跑道。

而这一切的实现都伴随较高的工作负荷,对于AI的算力有较高要求。对话型AI不仅需要超级计算机,而且运行的人工智能程序只能在GPU上完成,传统的CPU速度过慢难以满足。事实证明,人工智能是GPU最理想的工作负载,并且需要高性能计算程序类的数据中心。

英伟达在生产世界级GPU之外,还为高性能计算和人工智能生产世界级的GPU加速系统和GPU超级计算机。其可进行大规模的AI训练的DGX SuperPOD平台是AI性能评测标准MLPerf的赢家,而这个平台可从中国英伟达合作伙伴网络获得,OEM版本也想通过NGC认证的服务器提供。

## 打造开源开放环境

## 释放AI巨大潜力

“AI一定要进入企业,并也要满足每一个企业的不同需求。目前,AI的发展体量较小,因此需要更多力量加入,协同向前,开发未来的潜力市场,这是英伟达的理念和目标所在,而这也离不开行业、政府、社区的支持。”潘迪表示。英伟达秉承开放和标准的理念,为众多企业和研发者打造一个开放、开源的环境,为行业的大前景共同努力。英伟达的开源人工智能软件平台包括DGX服务器、超级计算机DGX SuperPOD、企业级服务器和云,覆盖医疗保健、消费者互联网和金融服务等领域。

英伟达的软件是开放、开源的,因此可以很容易地与世界各地的研究人员合作。在中国,作为NVAAIL合作的一部分,英伟达正在与排名领先的大学合作开发上述所有软件。英伟达人工智能实验室项目目前正在与清华大学、北京大学等中国领先大学合作,通过METROPOLIS、CLARA、METROPOLIS、ISAAC平台进行人工智能城市、人工智能医疗、人工智能零售、人工智能制造的开发。“英伟达致力于给大家提供开放、开源、标准的环境,与同业一起,把AI事业做起来。”潘迪表示。

高通公司全球副总裁雷纳·克莱门特:

## 5G 将把工业发展推向新阶段



“未来会有更为先进的5G网络,更多可拓展的能力和可靠、灵活的系统,人工智能、增强现实、边缘计算,所有的这些都可以通过5G实现。”

本报记者 刘晶

9月3日,在工业和信息化部、上海市人民政府指导下,中国半导体行业协会、中国电子信息产业发展研究院联合主办的第二届全球IC企业家大会暨第十七届中国国际半导体博览会(IC China 2019)在上海举行。高通公司全球副总裁雷纳·克莱门特出席会议并发表了主题演讲,雷纳·克莱门特在主题演讲中表示,未来会有更为先进的5G网络,更多可拓展的能力和可靠、灵活的系统,人工智能以及增强现实、边缘计算,所有的这些都可以通过5G实现。

## 5G网络能力还会拓展

## 适用范围不断扩大

雷纳·克莱门特表示,5G将会影响到我们生活的方方面面,它会帮助我们实现未来工厂、自动化交通以及更为可靠的远程医疗。在咨询机构iHS的研究当中,5G经济时代将会在全球范围内,到2035年实现在更多的行业产生12万亿美元

设备、产品、服务,这都是5G技术推动的,在未来,整个行业都有非常庞大的增长机会。

雷纳·克莱门特说,5G将会把工业推向工业4.0,推向下一个阶段,达到新的水平。未来会有更为先进的5G网络,更多可拓展的能力和可靠、灵活的系统,处理器、人工智能以及增强现实、边缘计算,所有的这些都可以通过5G实现。高通有一个案例,在工业物联网中使用专用5G网络,通过对技术的优化,5G网络将这些设备连接在一起,就可以解决设备同步化问题。高通寻求在5G技术里与行业进行合作,在德国汉诺威的工业展上,高通展示了基于5G的物联网技术,并与博世展

开合作。

## 发展AI需要边缘云

## 低时延成就工业控制

高通在AI方面已经有十年的研究,2009年,通过合作方式开始了第一个AI项目,之后和越来越多的合作伙伴建立了AI生态系统,投入了研发、知识产权、人才的培训,在2015年生产了商业化的产品,即骁龙820移动平台,其中高通提供了首个边缘云解决方案。

雷纳·克莱门特说,我们一直关注AI,特别是在终端设备上实现AI,我们希望打造更多工具,可以适用于更多细分市场。在设备优化上,从软件和硬件两个方向展开,也创立了5G的SoC。

雷纳·克莱门特认为,在工业自动化的实现中,智能控制的要求是要做到及时的控制,这就必须使用边缘设备——边缘云。通过这样的技术,可以为客户提供定制化业务,也可以在边缘云上提供更好的网络服务,为公司、工厂或者私人网络提供帮助。但要把客户的设备能力和边缘云结

合在一起,需要完全发挥5G的潜能,通过提升现在的通信服务,面向产业应用转型,创新行业应用,积累新经验。

“不仅在工业自动化控制中需要低时延,在很多其他应用上也需要实现低时延。”雷纳·克莱门特说,“我们希望与生态系统的各个参与方合作,让这些企业为无线的边缘转型做重要的贡献,而且希望和整个行业合作,形成端到端的系统解决方案,高通可以在设备的处理能力、安全性、传感器等方面提供系统解决方案,完全释放5G的可能性。”

雷纳·克莱门特现场展示了在混合现实应用中,边缘云的增强能力。当一个人在终端设备上摇头,通过边缘云的渲染,最后呈现出来新效果,这个过程可以变得时延很低,从而实现5G对高质量低时延业务的支撑。考虑到网络传输中产生的延迟以及为保障网络安全牺牲的时延,5G实现低时延确实要解决很多问题,但带来的沉浸式体验非常好。

雷纳·克莱门特说,新的领域可以帮助我们更好地推动5G技术的发展,例如通过分布式技术,可以帮助我们更好地进行设备的互联互通;通过XR头部设备的改进,能够使穿戴的感觉更为舒适。新的领域也可以帮助我们挑战更多限制,例如通过分裂式的渲染解决方案,特别是对中档客户而言有更好的总体性能和性价比。

在智能交通中,利用现有蜂窝技术可以解决交通方面的问题,网络沟通能够更好、更快,特别是在一些远程的控制技术中,5G能够提高道路安全性和有效性。物联网将无处不在,例如道路、桥梁、基础设施也都可以连接在一起,5G是推动这一切转型的基础力量。“所以,5G只是一个开端,欢迎大家一起来到创新的新时代。”雷纳·克莱门特说。