

# 戴尔易安信为什么关注FPGA

发表于 2018-12-20 10:39 | 来源 互联网 | 作者 互联网

摘要： 2018年12月19日,英特尔FPGA中国创新中心落地重庆西永微电子产业园。

2018年12月19日,英特尔FPGA中国创新中心落地重庆西永微电子产业园。戴尔易安信作为英特尔FPGA创新中心服务器产品的独家赞助商,为创新中心提供了搭载FPGA的第14代PowerEdge系列服务器。戴尔易安信为什么支持英特尔FPGA中国创新中心?FPGA为什么值得关注?我们提供一些背景信息,供您解读。



图1:英特尔FPGA中国创新中心盛大揭幕

### FPGA技术优势:高性能、低功耗、灵活性强

随着深度学习和神经网络等技术的发展,FPGA加速技术在人工智能、5G、自动驾驶和数据中心等领域的应用越来越广泛。FPGA支持差异化定制,是一种极具竞争力的大数据、深度学习、无线通信加速器;FPGA支持底层硬件架构重新调整和软件自定义,因此可融合最新的行业创新技术;FPGA具有高效可重复编程特性,可实现定制性能、定制功耗、高吞吐量和低批量延迟,满足用户各种的规格要求。



图2:戴尔易安信全球资深副总裁、大中华区企业解决方案总经理 曹志平

出席2018 FPGA国际创新峰会并做主题演讲

FPGA具有以下技术优势:

#### (1) 高性能、低功耗

FPGA是一种高性能、低功耗的可编程芯片。可以根据客户定制,做针对性的算法设计,在处理海量数据的时候,FPGA相比于CPU和GPU,优势在于:FPGA计算效率更高,其高并行性的特性往往可以使业务性能得到量级的提升。

**CSDN官方微信**  
扫描二维码,向CSDN吐槽  
微信号: CSDNnews



**程序员移动端订阅下载**



### 每日资讯快速浏览

#### 微博关注

CSDN 北京 朝阳区  
**CSDN**  
加关注

让机器掌握阅读理解能力究竟意义何在? 搜狗一直坚持探索以语言为核心的人工智能战略,相信搜索的未来是问答,而机器阅读理解是现今问答技术发展的核心之一。搜索走向问答将会在医疗、法律等专业领域产生颠覆性效果。这也是搜狗AI团队一直坚持的方向。祝贺! [挤眼][赞]

搜狗 : #搜狗AI夺国际顶级挑战赛CoQA第

#### 相关热门文章

- 什么蓝牙耳机好? 选购必须注意的三大因素
- 罗振宇2018时间的朋友跨年演讲: 我们这一代人...
- 换手机号不做这些很麻烦 支付宝宣称被盗全赔
- 2019年度年终销量榜单: 蓝牙耳机畅销排行榜1...
- 2018戴尔易安信服务器年度总结
- MORE Health爱医传递携新产品亮相J.P.Morga...
- 盘点远程监控重要玩家,无人驾驶的“最高级别...
- 蚂蚁金服ATEC城市峰会上海举行,三大发布迎...
- 用货拉拉戒掉网瘾,司机重新找回自己
- 企业老板常见的5大难题怎么破?“好风网”电...

它是软硬件合一的器件,不采用指令和软件。FPGA编程使用硬件描述语言,硬件描述语言描述的逻辑可以直接被编译为晶体管电路的组合。所以,FPGA实际上直接用晶体管电路实现用户算法,没有通过指令系统的翻译。

## (2) 灵活性强

FPGA是现场可编程逻辑门阵列,是一堆逻辑门电路的组合,可以编程还可以重复编程。

FPGA是动态可重配的,在数据中心部署之后,可以根据业务形态来配置不同的逻辑,实现不同的硬件加速功能。举例来讲,当前服务器上的FPGA板卡部署的是图片压缩逻辑,服务于社交即时通讯;而此时广告实时预估需要扩容获得更多的FPGA计算资源,通过简单的FPGA重配流程,FPGA板卡即可以变身成“新”硬件来服务广告实时预估,非常适合批量部署。

未来,FPGA将会成为越来越重要的芯片,CPU与FPGA深度融合的异构计算已成为产业发展的必然选择。



图3:戴尔易安信FPGA解决方案展台

CPU+FPGA的异构多核计算创新涉及软硬件等深层次融合创新,为赶超者提供了新的跑道。同时,CPU+FPGA的异构计算作为新的重要技术趋势,在加速了计算架构融合的同时,也给下游整机系统及软件开发带来了重要的机遇。

在大数据时代,机器学习和深度学习是人工智能的主要推动力。最近几年大数据和计算力的快速增长使得深度学习技术有了质的飞跃,从而在计算机视觉、语音识别、自然语言处理等领域出现了重大突破。

然而,深度学习需要大量的并行计算,对硬件平台有极高的要求,传统的计算机是无法满足的。FPGA具有很高的性能功耗比,而且基于门级电路设计使得FPGA是一个超低延时和确定延时的方案,FPGA的可编程性及动态可重构可以适应深度学习未来算法的变化,IO可编程性可以满足更多业务需求(网络加速,边缘计算),因此基于FPGA的深度学习方案成为未来技术发展方向。

FPGA作为未来强大的计算加速器,不但影响主要企业的决策和市场趋势,而且加速企业中的工作负载,促进超大规模数据中心的内部搜索,提高HPC高性能计算模拟的地位。

### FPGA行业现状:未来五年增长近1倍

以物联网、智能制造、大数据、云计算及人工智能的技术为代表的新一轮革命方兴未艾。在此背景下,中国政府提出了“智能制造”和“十三五”科技创新”的发展战略,大力推动产业转型升级和结构化调整。面对技术更新迅速、迭代频繁的研发应用环境,FPGA(现场可编程门阵列)正成为越来越重要的芯片,并得到越来越广泛的应用,在泛人工智能、5G、无人驾驶、智能终端和数据中心等领域发挥重要作用。

近年来全球FPGA市场规模基本在50-60亿美元左右。根据研调机构Global Market Insights的最新报告显示,FPGA市场在2015~2022年间将出现8.4%的年复合增长率,到2022年规模可望超过99.8亿美元。

FPGA主要的应用领域介绍如下:

#### (1) 泛人工智能

当前的前沿研究方向包括机器人、语言识别、图像识别、自然语言处理和专家系统等,自AI诞生以来,理论和技术日益成熟,应用领域也不断扩大,但人工智能目前仍处于早期阶段,预计未来的主战场是在推理环节。专注于推理环节的FPGA必将在AI领域发挥巨大价值。FPGA内部拥有大量计算单元,进行神经网络运算的速度远大于CPU;FPGA同时具

有架构可优化的灵活性,当平均性能接近一块GPU时,功耗远小于后者,可大大降低最终用户的散热负担。应用方面,由于FPGA可编程的特性,使得它能够灵活地针对应用所用的算法修改电路,为软件与终端应用公司提供与其竞争对手不同的解决方案。

## (2) 5G通信

异构无线网络是未来无线网络的主要发展方向,是多样化技术的融合。其灵活性、多协议兼容性和智能化,以及对成本和功耗的要求,需要一个可编程的芯片平台去支持,通过随时修改芯片上局部功能,实现5G网络所需的灵活性。传统的ASSP和ASIC不可编程,成本昂贵,无法满足这些方面的需求。FPGA可编程,性价比高,使之与5G无线网络的灵活性、性价比和智能化的需求良好契合,新一代FPGA技术将在5G市场大放异彩。

## (3) 自动驾驶

自动驾驶对计算能力的要求非常高。相对于手机等移动终端的计算需求,自动驾驶大概要超过两个数量级,数据生成速度方面超过四个数量级。实时性方面,自动驾驶有非常明确的最低延迟要求。在功耗方面,它需要有很低的能耗,从而提升车辆使用性能。FPGA同时拥有流水线并行和数据并行,而GPU几乎只有数据并行。因此对流式计算的任务,FPGA比GPU天生有延迟方面的优势。在功耗方面,FPGA比GPU的功耗低,同时FPGA需要的空间小更加适合在自动驾驶的车辆上使用。

## (4) 数据中心

数据中心计算加速领域,由于计算任务灵活多变、多租户等应用特性,传统的ASIC大规模神经网络部署,在研发成本、交付周期等方面已现颓势。FPGA在此应用环境中只需几百毫秒,即可完成逻辑功能更新,其灵活性可以大力保护数据中心的投资,并且始终保持数据中心同构性。在数据中心应用中,FPGA相比GPU的核心优势在于延迟:FPGA同时拥有流水线并行和数据并行两种方式,而GPU几乎只有数据并行(流水线深度受限)。当任务逐个而非成批到达时,流水线并行比数据并行可实现更低延迟。可见,针对流式计算的任务,FPGA在延迟方面比GPU具有天生优势,更符合未来数据中心的业务需要。

一位AI芯片创业公司CEO接受采访举例说,在深度学习中,经常需要调整算法。如果采用ASIC的方式,有可能刚做完传统算法硬化,过了半年,算法团队又把算法换了。而采用FPGA的灵活性是可以很好地适应这种变化。

奥迪正在使用英特尔FPGA,以期实现奥迪A8轿车自动驾驶,汽车零部件制造商电装(Denso)已经采用英特尔技术驱动的立体视效系统。越来越多的汽车公司在其车辆上使用FPGA,提供诸如自适应巡航控制、防撞和驾驶员辅助等功能。

FPGA还被大量用于大型云数据中心实现人工智能应用。例如,微软已在Brainwave项目是一个建立在FPGA之上的加速深度学习平台,它采用英特尔Stratix FPGA,通过数据的尽快处理和传输,在云中提供实时人工智能。

Dell EMC正在使用英特尔FPGA芯片,进行基因分析和蛋白质结构研究等领域的应用验证。

深思考人工智能(iDeepWise Artificial Intelligence)是一家专注于类脑人工智能与深度学习核心科技的公司。深思考CEO杨志明在接受智东西采访时说,智慧工业机器人、智慧通信、智慧医疗、智慧装备等领域,都是FPGA以外的其他AI芯片难以介入的应用领域。

### 戴尔易安信PowerEdge服务器加速数据中心工作负载

戴尔易安信PowerEdge服务器是全球X86服务器的领先品牌。根据IDC公司全球服务器市场跟踪数据,戴尔易安信始终保持着全球服务器市场的领先地位。以IDC在12月11日公布的最新数据为例,在2018年第三季度,戴尔易安信无论销售量还是销售额都名列全球第一。如下表。

**Top 5 Companies, Worldwide Server Vendor Revenue, Market Share, and Growth, Third Quarter of 2018** (Revenues are in US\$ Millions)

Company	3Q18 Revenue	3Q18 Market Share	3Q17 Revenue	3Q17 Market Share	3Q18/3Q17 Revenue Growth
1. Dell Inc.	\$4,093.0	17.5%	\$3,070.3	18.1%	33.3%
2. HPE/New H3C Group *	\$3,809.9	16.3%	\$3,317.4	19.5%	14.8%
3. Inspur/Inspur Power Systems *	\$1,711.9	7.3%	\$667.3	3.9%	156.5%
4. Lenovo	\$1,437.9	6.2%	\$860.8	5.1%	67.0%
5. IBM*	\$1,196.2	5.1%	\$1,093.7	6.4%	9.4%
5. Huawei*	\$1,052.7	4.5%	\$599.4	3.5%	75.6%
5. Cisco*	\$1,040.5	4.5%	\$992.5	5.8%	4.8%
ODM Direct	\$6,257.2	26.8%	\$4,118.7	24.3%	51.9%
Others	\$2,767.5	11.8%	\$2,255.0	13.3%	22.7%
<b>Total</b>	<b>\$23,366.8</b>	<b>100%</b>	<b>\$16,975.2</b>	<b>100%</b>	<b>37.7%</b>

Source: IDC Worldwide Quarterly Server Tracker, December 11, 2018.

**Top 5 Companies, Worldwide Server Unit Shipments, Market Share, and Growth, Third Quarter of 2018** (Shipments are in thousands)

Company	3Q18 Unit Shipments	3Q18 Market Share	3Q17 Unit Shipments	3Q17 Market Share	3Q18/3Q17 Unit Growth
1. Dell Inc.	556.0	17.6%	503.0	18.8%	10.5%
2. HPE/New H3C Group *	456.2	14.4%	501.4	18.8%	-9.0%
3. Inspur/Inspur Power Systems *	283.6	9.0%	149.1	5.6%	90.2%
4. Lenovo	193.5	6.1%	151.8	5.7%	27.5%
5. Huawei*	187.9	5.9%	133.3	5.0%	40.9%
5. Super Micro*	169.3	5.4%	136.7	5.1%	23.9%
ODM Direct	871.5	27.6%	668.0	25.0%	30.5%
Others	443.5	14.0%	428.0	16.0%	3.6%
<b>Total</b>	<b>3,161.5</b>	<b>100%</b>	<b>2,671.3</b>	<b>100%</b>	<b>18.3%</b>

Source: IDC Worldwide Quarterly Server Tracker, December 11, 2018.

戴尔易安信与英特尔的结合,可以将FPGA技术的优势带入各种规模的数据中心,加速处理数据中心日益增长的数据。

2018年4月,英特尔宣布其FPGA已经正式应用于主流的数据中心OEM厂商。其中,在全球市场销量领先的戴尔易安信PowerEdge系列服务器率先支持英特尔FPGA加速器。

戴尔易安信在PowerEdge R640、R740、R740XD等多个型号支持英特尔可编程加速卡(PAC),从而加速数据中心的多种核心工作负载,例如数据分析、人工智能等,并通过与英特尔紧密合作,提供面向加速的软件生态系统支持。其中,PowerEdge R740和R740XD可支持多达4路FPGA,PowerEdge R640支持1路FPGA。



2018年5月,戴尔易安信发布新的第14代戴尔易安信PowerEdge四路服务器和戴尔易安信ML/DL就绪解决方案,进一步加强人工智能和机器学习功能。

其中,戴尔易安信新推出两款的四路服务器:PowerEdge R940xa和PowerEdge R840。凭借英特尔®至强™可扩展处理器(多达112个处理核心)和大容量内存(多达6TB内存和NVDIMM选件)的卓越性能密度,PowerEdge R940xa和PowerEdge R840服务器提供了强大的性能。两款四路服务器都支持FPGA,在数据密集型计算方面表现卓越。

PowerEdge R940xa服务器设计用于加速关键业务应用数据库,没有使用云的费用和安全风险。它将4个CPU与4个图形处理单元(GPU)以1:1的大比率相结合,提高了应用性能,并且通过直接连接的NVMe驱动器实现了低延迟。

PowerEdge R840服务器专为数据库内分析而设计。它采用了比市场上其他服务器产品更多的直连NVMe驱动器,让数据延迟最小化,并通过完全集成的超级通道互连总线(UPI),加速数据传输。

戴尔易安信还参加了英特尔AI构建者计划,这是一个针对企业级技术合作伙伴的生态系统,旨在开发基于英特尔人工智能产品的高质量解决方案。戴尔易安信跟英特尔通力合作,通过戴尔易安信创新实验室和戴尔易安信客户解决方案中心,帮助客户评估其在英特尔®至强™可扩展处理器和英特尔®FPGA上的大规模深度学习和高性能计算工作负载。

戴尔易安信全球资深副总裁、大中华区企业解决方案总经理曹志平表示,戴尔易安信独家赞助英特尔FPGA中国创新中心,缘起戴尔跟英特尔紧密的战略伙伴关系,也是出于戴尔“在中国,为中国”戴尔中国4.0+战略的考虑。

面对数字化浪潮对各行各业带来的颠覆性改变,戴尔科技集团作为数字化转型的实践者、领导者和前瞻者,将汇集集团的综合实力,进一步深耕中国市场,持续履行本土化承诺,以领先的科技力量,丰富的实战经验为客户保驾护航,在数字化转型的道路上与客户一同砥砺前行,成为中国客户值得信赖的合作伙伴,共赢数字化未来。英特尔FPGA中国创新中心,正好跟戴尔的中国战略方向一致。

### 在戴尔易安信PowerEdge服务器上使用FPGA的例子

2018年5月,戴尔易安信推出了一项人工智能挑战赛,鼓励参赛者展示人工智能的实际应用,这对企业、科研以及社会都有着革命性的影响。冠军得主可获得在戴尔易安信HPC与人工智能创新实验室Zenith集群上的20万个核心小时,其项目也在2018年超算大会戴尔易安信展台上作为亮点展出,并参与到一系列推广活动当中。

最终,佛罗里达大学空间、高性能和弹性计算中心(SHREC)的一支研究团队被戴尔易安信评为2018年戴尔易安信人工智能挑战赛冠军。

佛罗里达大学空间、高性能和弹性计算中心(SHREC)由30多个行业、政府和学术合作伙伴组成,共同解决在通过重新配置、可靠的高性能计算来驱动或改进任务和应用上所面临的科研难题。对于此次人工智能挑战赛,SHREC团队开发并展示了一个异构计算(HGC)系统,该系统可以支持完整的机器学习工作流,包含数据分析以及预处理、模型训练、部署和推理,并适用于任何利用机器学习的应用领域,包括医疗、商业、金融、科学探索等等。

该团队的学生代表Chao Jiang博士表示:“为了这次人工智能挑战赛,我们的团队利用了CERN OpenLab数据集来确定HGC工作流的性能,包括将CPU、GPU和FPGA用于机器学习。研究表明,性能提高了1.45到2.22倍。这些早期的研究成果为我们提供了很大的帮助,我们正计划继续测试基于更复杂的3D图像的技术,例如利用3D U-net技术进行立体分割来帮助提高性能,以及用于加速例子模拟的3D GAN。”

### 关于戴尔易安信

作为戴尔科技集团的重要业务组成之一,戴尔易安信拥有行业领先的融合基础架构、服务器、存储和数据保护技术,助力企业实现现代化、自动化以及数据中心转型,为通过建立混合云、开发云原生应用和大数据解决方案实现业务转型提供了值得信赖的基础。戴尔易安信为遍及180个国家不同规模的客户提供服务—从全球财富500强至中小型企业—并为客户提供业界全面的从边缘计算到数据中心、再到云计算的创新产品组合。

【免责声明：CSDN本栏目发布信息，目的在于传播更多信息，丰富网络文化，稿件仅代表作者个人观点，与CSDN无关。其原创性以及中文陈述文字和文字内容未经本网证实，对本文以及其中全部或者部分内容、文字的真实性、完整性、及时性本网不做任何保证或者承诺，请读者仅作参考，并请自行核实相关内容。凡注明为其他媒体来源的信息，均为转载自其他媒体，转载并不代表本网赞同其观点，也不代表本网对其真实性负责。您若对该稿件由任何怀疑或质疑，请即与CSDN联系，我们将迅速给您回应并做处理。】



顶 0 踩 0

关于我们 招聘 广告

QQ客服 kefu@csdn.net 客服论坛 400-660-0108 工

百度提供站内搜索 京

©1999-2019 江苏乐知网络技术有限公司 江苏知之为计算机有限公司 北京创新乐知信息技术

经营性网站备案信息 网络110报警服务 中国互联网举报中心 北京互联网违法和不良信息举报