



服务器加速器

增强应用程序的性能



7/10 的

技术高管将 AI 确定为今后 5-10 年内三大最重要的技术之一¹

高达 100 倍

的 CPU 数量增加 — 单个 GPU 中²

超过 600 个应用程序

得到加速 — 全是深度学习框架³

加速获得见解和推动创新

对于数字化企业，成功的关键是以较快的速度利用大数据。但是随着数据集的增长，传统数据中心开始达到性能和规模上的限制 — 特别是在接收和查询实时数据源方面。

有些企业利用加速器来加速可视化、建模和模拟工作负载已经有了很长历史，而在今天，有比以往任何时候都更多的主流应用程序能够利用加速器来加速获得见解和推动创新。加速器，例如图形处理单元 (GPU) 和现场可编程门阵列 (FPGA) 可补充并加快 CPU，使用并行处理以更快的速度处理大量数据。加速的数据中心还能够带来更好的经济性，使用更少的服务器提供突破性的性能，从而更快获得见解并降低成本。

许多行业中的组织都在采用服务器加速器来超越竞争对手 — 利用从数据中获得的见解调整产品和服务，凭借更好的应用程序性能提高工作效率，利用快速、强大的分析功能优化运营，以及以卓越的开发速度加快产品上市。

Dell EMC 在 Dell EMC PowerEdge 服务器中提供了一系列服务器加速器，让您能够给您的应用程序提速。

加速器使用情形

- **机器学习和深度学习** — 通过帮助实现加速训练和推理工作负载所需的并行处理能力，加速器已经使 AI 从理论进入主流应用。
- **预测性分析** — 通过加速器实现的 AI 可以加强分析功能，从而实现动态关联，并以惊人的速度、准确性和规模提供预测结果。
- **更快的数据库** — 加速器可帮助加速聚合、排序和分组操作，解决使传统数据库过载的复杂的分析操作。
- **流数据** — 物联网 (IoT) 的出现产生了数据洪流。加速器支持流数据的同步接收、浏览和可视化，以实现实时分析。

¹“2019 Edelman AI 调查”，2019 年 3 月。

²“NVIDIA Tesla V100 Tensor 核心 GPU”，2019 年 5 月。

³“NVIDIA Tesla P100”，2019 年 5 月。

GPU

图形处理单元 (GPU) 是旨在加速计算性能的协处理器。一个 GPU 通常有数千个核心，用于实现数学函数的高效运算。工作负载的某些部分从 CPU 分流到 GPU，其余代码运行在 CPU 上，从而改善总体应用程序性能。

Dell EMC 以 PCIe 卡的形式提供了一系列 GPU，它们插入服务器的 PCIe 插槽中，还有一些作为 SXM2 模块安装到服务器主板上。

搭载 SXM2 模块的 PowerEdge C4140 服务器可利用 GPU 之间的 NVLINK 高速互连。

FPGA

现场可编程门阵列 (FPGA) 是一些具有特定技术特征的加速器，用于以比传统软件解决方案快 1000 倍的速度执行某些类型的算法。可对 FPGA 进行配置，以精确匹配特定任务或应用程序的要求。

Dell EMC 以 PCIe 卡的形式提供了一系列 FPGA，它们可插入服务器 PCIe 插槽中。

并行处理

并行处理是一种方法，就是将程序任务拆分并同时在多个微处理器上运行它们，从而缩短处理时间。并行处理可通过一台具有两个或更多处理器的计算机实现，或通过一个计算机网络实现。并行处理也称为并行计算。

优化代码

为充分利用服务器加速器的优势，可优化软件代码。

计算统一设备体系结构 (CUDA) 允许直接访问 GPU 虚拟指令集和并行计算元素，以完成计算内核的执行。

FPGA 配置通常是使用硬件描述语言 (HDL) 指定的。通过 HDL，可配置基本规则以匹配特定任务或应用程序的要求，实际上就是模仿了应用程序专用集成电路 (ASIC) 的原理。

- **可视化** — 加速器增强了诸如计算机辅助设计之类的 3D 可视化应用程序的性能，从而使软件能够在用户移动模型的过程中绘制模型。
- **建模和模拟** — 加速器可以提供超实时的建模模拟，以便及早评估设计、快速测试设计修改，并进行更多的设计迭代。
- **财务建模** — 加速的 HPC 和人工智能 (AI) 解决方案正在革新分析工具，使行业能够利用大规模的数据点来更好地了解风险和回报。
- **地震处理** — 石油和天然气公司正在寻求用新的、更好的办法来从庞大的地震数据存储库中提取信息，利用加速器加快获得结果并削减成本。
- **信号处理** — 加速器使提供商能够对来自计算机、收音机、视频设备和手机的大规模信号数据流进行实时建模和分析。

Dell EMC PowerEdge 服务器的 GPU 和 FPGA

部分 Dell EMC PowerEdge 塔式和机架式服务器中提供了性能加速器，可加速您的应用程序。下表显示了 [PowerEdge 服务器](#) 中配置的加速器 (GPU 和 FPGA) 的数量和型号。Xilinx® Alveo™ U200 和 Intel® Programmable Acceleration Card (PAC) with Arria® 10 GX 是 FPGA，其余的是 NVIDIA® GPU。

2S 塔式	1U 2S 机架式		2U 2S 机架式		4U 4S 机架式	
T640	R640	C4140	R740	R7425	R840	R940xa
4 个加速器	1 个加速器	4 个加速器	3 个加速器	3 个加速器	2 个加速器	4 个加速器
			P4			
	T4		T4	T4		
P40		P40	P40	P40	P40	
M10			M10	M10	M10	
V100 16 GB 或 32 GB		V100 16 GB 或 32 GB PCIe 或 SXM2	V100 16 GB 或 32 GB	V100 16 GB 或 32 GB	V100 16 GB 或 32 GB	V100 16 GB 或 32 GB
			U200	U200	U200	U200
	10GX		10GX		10GX	10GX

U = 机架单元, S = 插槽或 CPU
表中列出的加速器数量针对的是双宽全高卡。T4 是一个薄型半高卡。

GPU 因 CUDA 核心数量、内存量以及电源和冷却需求而异。例如，NVIDIA T4 GPU 有 2560 个 CUDA 核心、16 GB 内存，功耗最高 70 瓦。NVIDIA Tesla® V100 有 5120 个 CUDA 核心、最高 32 GB 内存，功耗最高 300 瓦。英特尔和 Xilinx FPGA 规格包括嵌入式内存和芯片外内存，容量从 35 MB 到 64 MB，功耗为 10 到 225 瓦。



增加 47 倍

的深度学习推理吞吐量 — 与一个 CPU 相比⁴



1.5 TFLOPs

(万亿浮点运算/秒)，能效高达 40 GFLOP/瓦⁵



高达 90 倍

的性能提升 — 与 CPU 相比，在运行机器学习推理、视频转码以及数据库搜索和分析工作负载时⁶

NVIDIA Tesla GPU

NVIDIA Tesla GPU 提供了比以往更快的速度运行可视化和其他工作负载所需的处理性能。另外，Tesla 提供了虚拟桌面基础架构 (VDI) 所需的高性能和用户密度。

- [T4 Tensor 核心 GPU](#)
- [Tesla V100 GPU 加速器](#)
- [NVIDIA NVLink™ 结构互连](#)
- [NVIDIA GPU CLOUD™ 容器](#)
- [软件应用程序目录](#)
- [NVIDIA 开发人员资源](#)

NVIDIA GPU 在 [Dell EMC PowerEdge 服务器](#) (包括 R640、R740、R740xd、R7425、R840、R940xa、C4140) 以及针对 HPC 和 AI 的 [Dell EMC 就绪型解决方案](#) 中提供。

英特尔 FPGA

英特尔 FPGA 可使用与您的工作负载 (如数据分析、图像推理、加密和压缩) 精确匹配的数据路径进行动态重新编程。

- [Intel PAC with Intel Arria 10 GX FPGA](#)
- [软件加速堆栈](#)
- [英特尔 FPGA 加速中心](#)

在 [Dell EMC PowerEdge 服务器](#) (包括 R640、R740、R740xd、R840、R940xa) 以及 [HPC 和 AI 解决方案](#) 中提供。下载 [配置指南](#)。

Xilinx FPGA

基于 Xilinx 16 纳米 UltraScale™ 体系结构而构建的 Xilinx Alveo 加速器卡可根据不断变化的加速要求和算法标准而调整，能够加速任何工作负载而不用更改硬件，并降低了总体拥有成本。

- [Alveo U200](#)
- [软件生态系统](#)
- [开发人员工具](#)

在 [Dell EMC PowerEdge 服务器](#) (包括 R740、R740xd、R7425、R840、R940xa) 以及 [HPC 和 AI 解决方案](#) 中提供。

加速的 Dell EMC 解决方案

利用配备有 GPU 和 FPGA 的 Dell EMC 及合作伙伴解决方案节约时间。

Dell EMC Ready Solutions

[Dell EMC Ready Solutions for HPC](#) 让客户能够更快、更简单地采用高级计算。Dell EMC 提供了一系列灵活而且可扩展的高性能计算解决方案，利用经过优化整合的服务器、网络、存储解决方案和服务适应众多行业中的使用情形。

[Dell EMC Ready Solutions for AI](#) 囊括了加快 AI 计划所需的一切要素。这些集成式系统允许轻松利用 AI，非常适合于运行机器学习和深度学习工作负载，让您能够更快获得关于客户和业务的更深入的见解。

⁴ “[NVIDIA Tesla V100 Tensor 核心 GPU](#)”，2019 年 5 月。

⁵ “[Intel Arria 10 FPGA 功能特性](#)”，2019 年 5 月。

⁶ “[Alveo U200 数据中心加速器卡](#)”，2019 年 5 月。

Dell EMC 合作伙伴提供的解决方案

[Amulet Hotkey® 虚拟桌面解决方案](#)将企业级服务器与虚拟 GPU 加速器结合起来，以实现针对数据中心而优化的低密度解决方案，从而简化向 Windows® 10 的过渡。另外，虚拟 GPU 可帮助满足不断增长的常用程序（如 Windows 10、Microsoft® Office 365®、YouTube® 等）的图形加速虚拟化需求，实现卓越的虚拟桌面体验。 [阅读 Amulet Hotkey 客户成功案例](#)。

[基于 BittWare Stratix® 10 的 FPGA 加速计算节点](#)让您能够使用一种针对 FPGA 加速器而优化的低密度机架式服务器（每单元最多 4 个 Intel Stratix 10 FPGA）来运行要求苛刻的数据中心工作负载。可直接从 BittWare 或 Dell EMC 购买此类系统。 [观看视频](#)。



[Kinetica®](#) 是一个见解发现引擎，它包括一个 GPU 加速数据库、直观发现和机器学习功能，以及加速的并行计算能力。在 Dell PowerEdge 服务器和 NVIDIA GPU 上运行的 Kinetica 可帮助组织解决大量复杂、不可预测的数据带来的难题。阅读此文章：[向您的首席执行官解释 GPU：产品化的威力](#)。



[Tracewell Systems®](#) 为企业、政府机构和 OEM 提供了功能强大而且现成的技术，适用情形包括：环境因素会带来独特计算难题的位置，例如空中、海上或地面上；固定和移动安装情形；或者要求集成专用硬件或软件的情形。 [获得产品介绍、视频和资源](#)。

Dell EMC 加速软件合作伙伴

[NVIDIA GRID™](#) Virtual PC 和 [GRID®](#) Virtual Apps 凭借基于 NVIDIA® Tesla® GPU 而构建的经验证的性能改善了虚拟桌面和应用程序。



[Kinetica](#) 软件通过使用 GPU 进行并行处理而大大加快了传统在线分析处理 (OLAP) 工作负载的速度。



[BitFusion](#) 软件能够拆分 GPU、FPGA 和/或 ASIC，并将它们动态附连到数据中心的任何位置。

[SQream Technologies®](#) 的 GPU 加速数据仓库能够从 TB 级扩展到 PB 级，从而适应任何规模和工作负载。



[FASTDATA.io](#) PlasmaENGINE® GPU 原生软件可借助多个 GPU 跨多个节点实现对海量移动数据的实时处理。



[RAPIDS](#) 是一系列基于 NVIDIA CUDA-X 而构建的数据科学库，用于在 NVIDIA GPU 中执行端到端数据科学训练管道。

成为 Dell Technologies 合作伙伴

加入 [Dell Technologies 合作伙伴计划](#)后，您将加入一个合作伙伴生态系统。该生态系统助力世界各地的组织实现数字化、IT、员工和安全转型 — 每一天都是如此。从边缘到核心再到云，我们有着一个阵容庞大的产品组合，而 Dell Technologies 合作伙伴计划是其关键支柱，它简单、可预测，可带来盈利。



资源

借助 PowerEdge 服务器让您的数据中心做好应对任何工作负载的准备

PowerEdge 塔式服务器采用专门设计，可按您的节奏随着您的组织增长。PowerEdge 机架服务器结合了高度可扩展的体系结构和计算与内存的优化平衡，以尽可能提高跨更广泛的应用程序的性能。

查看性能结果

获得工程部门提供的按工作负载分类的基准测试数据、参考体系结构以及博客 (hpcatdell.com) 和 [服务器解决方案支持资源](#)。阅读以下白皮书：

- [深度学习性能比较 — PowerEdge C4140 纵向扩展与横向扩展](#)
- [PowerEdge R940xa 上的 GPU 数据库加速](#)
- [PowerEdge R7425 上的深度学习推理](#)
- [使用基于 Dell EMC 基础架构的英特尔可编程加速卡运行深度学习推理](#)

利用教育服务

在 education.emc.com 网站上获得您需要的技能、培训和认证。在 Dell EMC 主办的深度学习学院 [了解如何利用深度学习解决问题](#)。

访问 HPC 和 AI 创新实验室

卓越的解决方案工程设计、技术测试和协作：dell EMC.com/zh-cn/innovationlab

通过卓越中心建立联系

全球创新和专业知识中心：dell EMC.com/zh-cn/coe

联系我们

如欲了解详情，请访问 dell EMC.com/zh-cn/accelerators、dell EMC.com/zh-cn/servers 或与您当地的客户代表或授权经销商联系。

资源中心

下载 [服务器解决方案资源](#)

技术文档

查看工程部门提供的性能结果、参考体系结构和博客：hpcatdell.com

虚拟机架

查看虚拟机架中的服务器和解决方案：esgvr.dell.com

加入 Dell EMC HPC 社区

促进思想交流的全球性技术论坛
dellhpc.org

版权所有 © 2019 Dell Inc. 或其子公司。保留所有权利。Dell、EMC、Dell EMC 和其他商标为 Dell Inc. 或其子公司的商标。Dell® / EMC® / Dell EMC® 等品牌商标将有可能同时出现在戴尔易安信相关企业级产品（包括硬件和软件）、和/或产品资料、戴尔易安信的官方网站。如果您有关戴尔易安信产品相关的任何疑问，欢迎联系您的指定客户经理。其他商标可能是其各自所有者的商标。参考编号：05/19 DELL-EMC-BRO-Accelerators-USLET-101

Intel® 是 Intel Corporation 在美国和/或其他国家/地区的商标。NVIDIA®、Tesla® 和 GRID™ 是 NVIDIA Corporation 的商标。Microsoft®、Windows® 和 Office 365® 是 Microsoft Corporation 在美国和/或其他国家/地区的注册商标或商标。YouTube® 是 Google Inc. 的注册商标。Arria® 和 Stratix® 是 Altera Corporation 的商标并已经在美国专利和商标办公室以及其他国家/地区进行了注册。Xilinx®、UltraScale™ 和 Alveo™ 是 Xilinx, Inc. 的商标。SQream Technologies® 是 SQream Technologies Ltd. 的注册商标。Kineticia® 是 Kineticia 及其子公司在美国和其他国家/地区的商标。FASTDATA.io® 是 FASTDATA.io 在美国和其他国家/地区的注册商标。Amulet Hotkey® 是 Amulet Hotkey Ltd. 的商标，在英国、美国和/或其他国家/地区进行了注册。Tracewell Systems 是 Tracewell Systems 在美国和其他国家地区的注册商标。

Dell EMC 确认本文档在发布之日内容的准确性。本出版物中的信息可随时更改而不另行通知。