



北京交通大学  
BEIJING JIAOTONG UNIVERSITY



# 《大数据概论》

## 大数据架构与处理

鲍鹏  
软件学院





# 目录

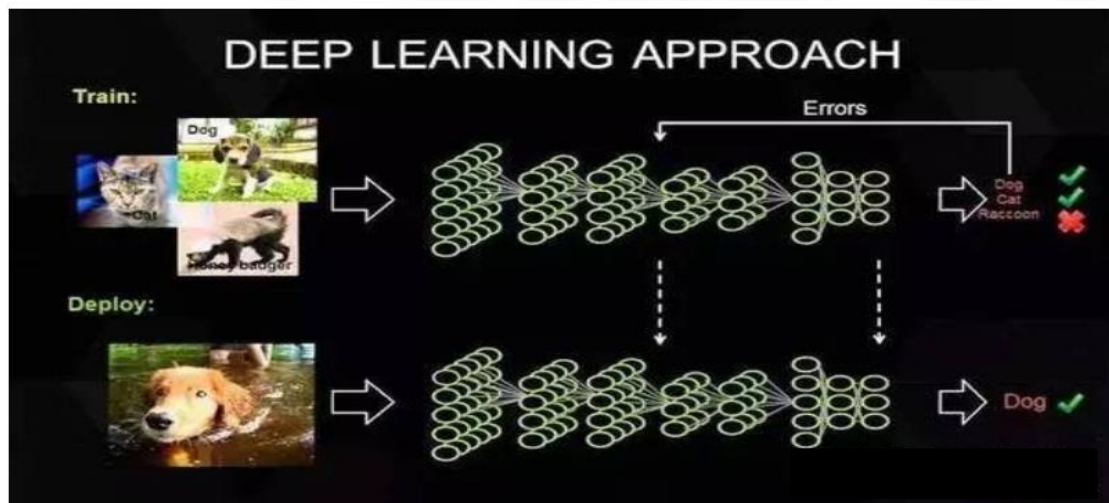
---

- 集中式分布架构
- 分布式计算架构
  - MapReduce
  - Spark
- 大数据计算处理加速技术
  - GPU
  - TPU
  - FPGA



# 深度学习的硬件方案

- 百度的硅谷人工智能实验室（SVAIL）已经为深度学习硬件提出了**DeepBench**基准，这一基准着重衡量的是**基本计算的硬件性能**。
- 现在的深度学习算法主要包括卷积神经网络（CNN）和循环神经网络（RNN）。





# 深度学习的硬件方案

- 基于深度学习算法，DeepBench提出以下四种基本运算：
  - **矩阵相乘**（Matrix Multiplication）——几乎所有的深度学习模型都包含这一运算，它的计算**十分密集**。
  - **卷积**（Convolution）——另一个常用的运算，占用了模型中大部分的每秒浮点运算（浮点 / 秒）。
  - **循环层**（Recurrent Layers）——模型中的**反馈层**，并且基本上是前两个运算的组合。
  - **All Reduce**——一个在优化前对学习到的参数进行传递或解析的运算序列。在跨硬件分布的深度学习网络上执行同步优化时（如AlphaGo），这一操作尤其有效。



# GPU

- **图形处理器**( graphics processing unit, GPU), 又称显示核心、视觉处理器、显示芯片, 是一种专门在个人电脑、工作站、游戏机和一些移动设备(如平板电脑、智能手机等)上进行图像运算工作的微处理器。
- **GPU加速计算**是指同时利用图形处理器(GPU)和CPU, 加快科学分析、工程、消费和企业应用程序的**运行速度**。





# GPU

- GPU加速计算可以提供非凡的应用程序性能，能将应用程序**计算密集部分**的工作负载转移到GPU，同时仍由CPU运行其余程序代码。
- 理解GPU和CPU之间区别的一种简单方式是**比较它们如何处理任务**。
- CPU由专为**顺序串行处理**而优化的几个核心组成，而GPU则拥有一个由数以千计的更小、更高效的核心(专为**同时处理多重任务**而设计)组成的大规模并行计算架构。



# GPU

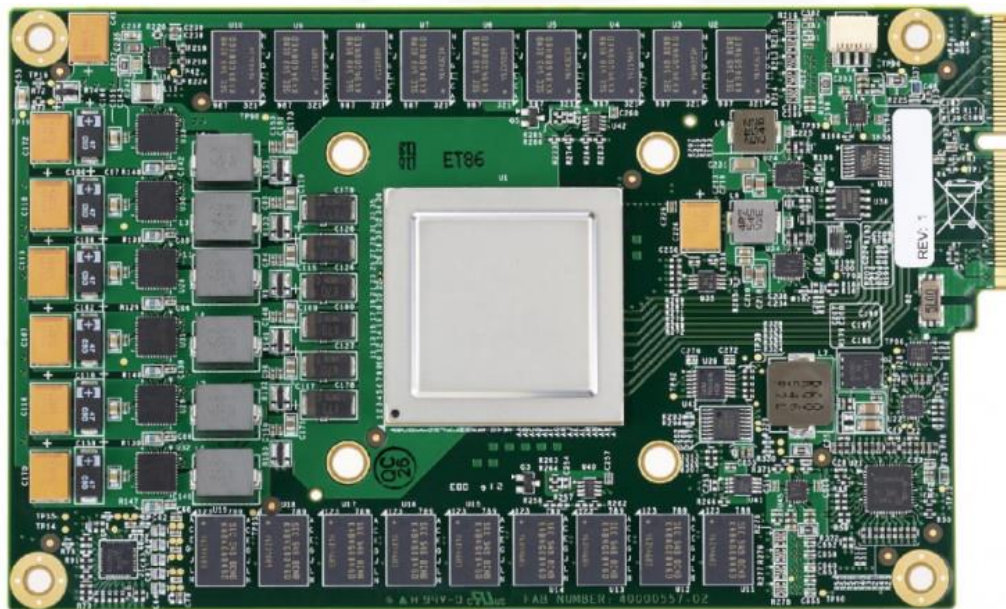
- GPU和CPU的浮点计算能力差异的原因:
- GPU是特别为计算密集、高并行度计算(如图像渲染)设计的, 因此将更多的晶体管用于数据处理而不是数据缓存和流控。
- 特别地, GPU非常适合处理那些能够表示为数据并行计算(同一程序在多个数据上并行执行)的问题。数据并行计算的算术计算密度(算术操作和存储器操作的比例)非常高。
- 由于同一程序在每个元素上执行, 因此对复杂流控的要求非常少, 更因为在多个元素上执行和高计算密度, 访存延迟可以被计算隐藏, 因此无需大的数据缓存。





# TPU

- 谷歌资深硬件工程师Norman Jouppi刊文表示，谷歌的专用机器学习芯片**TPU**处理速度要比**GPU**和**CPU**快**15-30倍**（和TPU对比的是英特尔Haswell CPU以及Nvidia Tesla K80 GPU），而在能效上，TPU更是提升了30到80倍。







# TPU

- **张量处理器**(tensor processing unit, TPU)是为机器学习定制的专用芯片，专为深度学习框架 **Tensorflow**而设计。
- 与图形处理器(GPU)相比，TPU采用**低精度(8位)**计算，以降低每步操作使用的晶体管数量。降低精度对深度学习准确度影响很小，但却可以**大幅降低功耗、加快运算速度**。同时，TPU使用了脉动阵列的设计，用来优化矩阵乘法与卷积运算，**减少I/O操作**。此外，TPU还采用了更大的片上内存，以此**减少对DRAM**的访问，从而更大程度地提升了性能。



# TPU

- 与CPU和GPU由于引入了**Cache**、乱序执行、多线程和预取等造成的行时间不确定相比，TPU的确定性执行模块能够满足神经网络应用上99%相应时间需求。
- CPU/GPU的结构特性对平均吞吐更有效，而TPU针对响应延迟设计。



# FPGA

- 现场可编程逻辑阵列( field programmable gate array, FPGA) 是在PAL、 GAL、 CPLD等可编程逻辑器件的基础上进一步发展的产物。
- 它是作为专用集成电路领域中的一种半定制电路而出现的, 既解决了定制电路的不足, 又克服了原有可编程器件门电路数有限的缺点。
- 目前以硬件描述语言( Verilog或VHDL)描述的逻辑电路, 可以利用逻辑综合和布局、布线工具软件, 快速地烧录至FPGA上进行测试, 这一过程是现代集成电路设计验证的技术主流。



# FPGA

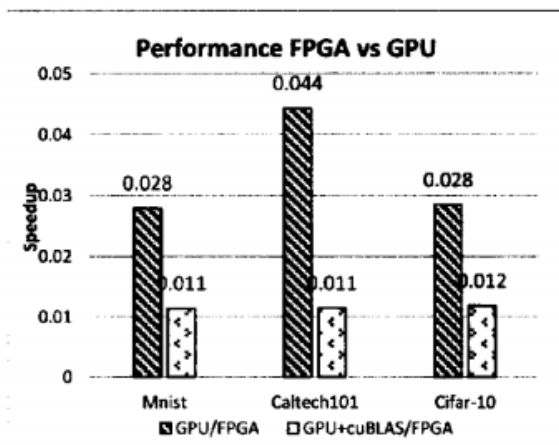
- **FPGA**一般来说比专用集成电路的**速度要慢**，无法完成更复杂的设计，并且会**消耗更多的电能**。
- 但**FPGA**具有很多优点，比如可以**快速成品**，而且其内部逻辑可以被设计者**反复修改**，从而改正程序中的错误，此外，使用**FPGA**进行除错的**成本较低**。



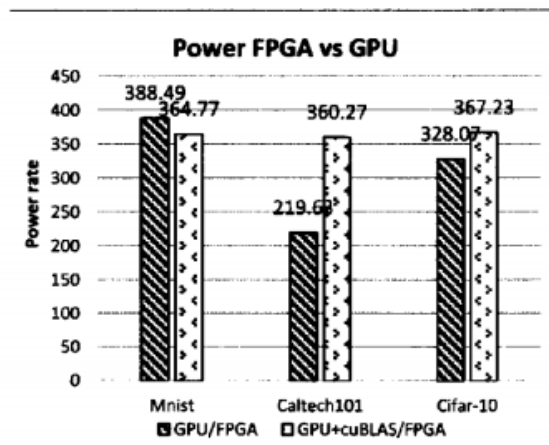


# GPU & FPGA

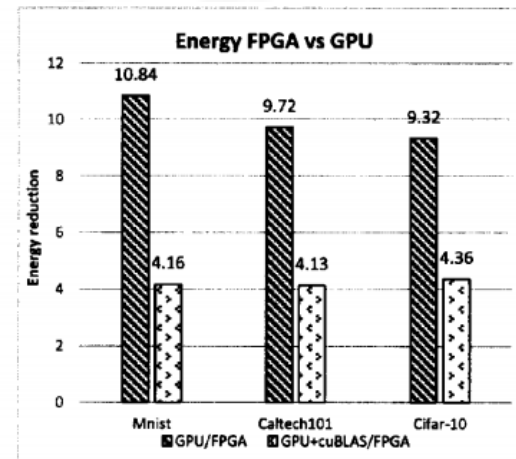
- 以NVIDIA Tesla K40c为基准，采用6种不同的神经网络结构分别测试了深度学习的预测过程、本地预训练过程和全局训练过程下FPGA(Zedboard)和GPU(K40c)的性能、功率和能耗数据。



(a) 性能加速比



(b) 功率比



(c) 能耗比



# GPU & FPGA

- 在异构处理器中，“**CPU+GPU**”是一个重要选项。**GPU**采用**SIMD**(单指令流多数据流)的方式让多个执行单元以同样的步调处理不同的数据，大大提升了**并行数据处理**的能力，在计算密集型任务中可堪重用。然而，**GPU**的“硬伤”在于**延迟比较高**。因为**GPU**虽可实现数据并行，但是其**流水线深度受限**，每个计算单元处理不同的数据包时，需要按照统一的步调做相同的事，使得**输入输出的延迟增加**，通常**GPU**的延迟会达到毫秒级。
- 从数据吞吐能力上看，新一代**FPGA**的数据处理加速能力理论上已经可以与**GPU**比肩。同时，由于半导体工艺的不断进步，**FPGA**器件的**功率**也控制得很好。**CPU+FPGA**这种异构处理器组合被越来越多的人所看好。