

## <<GPU高性能运算之CUDA>>

### 图书基本信息

书名：<<GPU高性能运算之CUDA>>

13位ISBN编号：9787508465432

10位ISBN编号：7508465431

出版时间：2009-10

出版时间：水利水电出版社

作者：张舒，褚艳利 主编

页数：276

字数：438000

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

## &lt;&lt;GPU高性能运算之CUDA&gt;&gt;

## 前言

CUDA是NVIDIA公司于2007年推出的GPU通用计算产品。

CUDA能够有效利用GPU强劲的处理能力和巨大的存储器带宽进行图形渲染以外的计算，广泛应用于图像处理、视频播放、信号处理、人工智能、模式识别、金融分析、数值计算、石油勘探、天文计算、流体力学计算、生物计算、分子动力学计算、数据库管理、编码加密等领域，并在这些应用中对CPU获得了一到两个数量级的加速，取得了令人瞩目的成绩。

CUDA在产业界和学术界引发了巨大反响，目前已有Adobe、Autodesk、HP、联想、浪潮等国内外著名企业推出了基于CUDA的产品；哈佛、伯克利、清华等大批国内外高校也开设了相关课程，将CUDA作为一种典型的并行系统来帮助学生学习高性能计算的原理与应用。

多核的CPU和众核的GPU已经成为目前大多数计算机中最重要的两种处理器。

传统上，GPU只用于处理3D图像渲染任务，而其他大多数的任务都交给了CPU。

作为一种通用处理器，CPU的设计必须兼顾各种任务的需要。

因此，CPU中大多数的晶体管都被用于制造庞大的缓存和复杂的控制逻辑，而运算单元占用的面积则并不多。

近几年来，CPU生产厂商已经认识到依靠增加单个核心的缓存尺寸和控制逻辑的复杂程度对提高处理器性能已经没有太大帮助，因此纷纷在一块芯片中集成更多的CPU核心。

不过，这样做只是能够用更多的晶体管制造一个处理器，并没有提高晶体管的利用率。

与此同时，由于图形渲染的并行特性，GPU与生俱来就是一种拥有大量运算单元的并行处理器。

近几年来年的发展经验表明，一块高端GPU的单精度浮点处理性能可以达到一块同时期高端桌面CPU的10倍，其显存带宽也达同时期桌面平台的5倍左右。

并且，由GPU提供相同的计算能力，所需要的成本和功耗都要小于基于CPU的系统。

由于传统GPU硬件架构的限制，很难有效利用其资源进行通用计算。

NVIDIA推出的CUDA产品则完全扭转了这一局面。

与传统GPU通用计算开发方式相比，CUDA编程更简单，功能更强大，应用领域更广泛，支持CUDA的硬件性能也更强。

CUDA完全改变了PC中的一些游戏规则，使用GPU计算无需投入额外的成本，却可以在一些应用中获得一个数量级的加速。

在科学计算中，一些过去必须由集群处理的任务，现在也可以由每个研究人员的桌面Pc完成了，这使得科研人员能够更加自由地将时间投入到自己的研究中。

在超级计算机与集群中，过去往往拆除“多余的”显卡以节省功耗，现在的设计却开始采用大量GPU来获得更加绿色的计算能力。

CUDA的强大性能引发了一场通用计算革命，这场革命将极大地改变计算机的面貌。

我们很幸运地参与了这场GPU通用计算革命。

在2007~2009年间，我们看到国内有越来越多的开发者尝试将CUDA应用到自己感兴趣的领域中，在学术和产业领域都取得了令人欣喜的成绩。

与此同时，我们也注意到很多程序员在由CPU程序开发转移到CUDA程序开发的过程中遇到了一些问题。

特别是出于CUDA程序的性能往往和并行算法以及硬件架构有很大关系，初次接触GPU的程序员往往需要较长的时间才能完全掌握CUDA程序的编写与优化。

## <<GPU高性能运算之CUDA>>

### 内容概要

本书是全国第一本全面介绍CUDA软硬件体系架构的书籍。

全面介绍使用CUDA进行通用计算所需要的语法、硬件架构、程序优化技巧等知识，是进行GPU通用计算程序开发的入门教材和参考书。

本书共分5章。

第1章介绍GPU通用计算的发展历程，介绍并行计算的历史、现状以及面临的问题；第2章介绍CUDA的使用方法，帮助读者理解CUDA的编程模型、存储器模型和执行模型，掌握CUDA程序的编写方法；第3章探讨CUDA硬件架构，深入分析Tesla GPU架构与CUDA通用计算的相互作用；第4章总结CUDA的高级优化方法，对任务划分、存储器访问、指令流效率等课题进行探讨；第5章以丰富的实例展示如何使用CUDA的强大性能解决实际问题。

本书可作为CUDA的学习入门和编程参考书，主要面向从事高性能计算的程序员与工程师，使用GPU加速专业领域计算的科研人员，以及对GPU通用计算感兴趣的程序员。

开设相关课程的高等院校与科研机构也可选用本书作为教材。

## <<GPU高性能运算之CUDA>>

### 作者简介

张舒，电子科技大学信息与通信工程专业硕士，现任NVIDIA深圳有限公司系统设计验证工程师，CUDA技术顾问。

曾实现基于CUDA的神经网络、聚类分析、主分量分析等模式识别算法，以及信号仿真、密码破解、字符串匹配等应用。

## &lt;&lt;GPU高性能运算之CUDA&gt;&gt;

## 书籍目录

前言第1章 GPU通用计算 1.1 多核计算的发展 1.1.1 CPU多核并行 1.1.2 超级计算机、集群与分布式计算 1.1.3 CPU+GPU异构并行 1.2 GPU发展简介 1.2.1 GPU渲染流水线 1.2.2 着色器模型 1.2.3 NVIDIA GPU发展简介 1.3 从GPGPU到CUDA 1.3.1 传统GPGPU开发 1.3.2 CUDA开发第2章 CUDA基础 2.1 CUDA编程模型 2.1.1 主机与设备 2.1.2 Kernel函数的定义与调用 2.1.3 线程结构 2.1.4 硬件映射 2.1.5 deviceQuery示例 2.1.6 matrixAssign示例 2.2 CUDA软件体系 2.2.1 CUDA C语言 2.2.2 nvcc编译器 2.2.3 运行时API与驱动API 2.2.4 CUDA函数库 2.3 CUDA存储器模型 2.3.1 寄存器 2.3.2 局部存储器 2.3.3 共享存储器 2.3.4 全局存储器 2.3.5 主机端内存 2.3.6 主机端页锁定内存 2.3.7 常数存储器 2.3.8 纹理存储器 2.4 CUDA通信机制 2.4.1 同步函数 2.4.2 Volatile关键字 2.4.3 ATOM操作 2.4.4 VOTE操作 2.5 异步并行执行 2.5.1 流 2.5.2 事件 2.6 CUDA与图形学API互操作 2.6.1 CUDA与OpenGL的互操作 2.6.2 CUDA与Direct3D互操作 2.7 多设备与设备集群 2.7.1 CUDA设备控制 2.7.2 CUDA与openMP 2.7.3 CUDA与集群第3章 CUDA硬件架构 3.1 NVIDIA显卡构造简介 3.1.1 图形显卡概览 3.1.2 PCI—E总线 3.1.3 显存 3.1.4 GPU芯片 3.2 Tesla图形与计算架构 3.2.1 SPA—TPC—SM 3.2.2 主流GPU架构 3.3 Tesla通用计算模型 3.3.1 数据与指令的加载 3.3.2 warp指令的发射与执行 3.3.3 纹理、渲染和存储器流水线第4章 CUDA程序的优化 4.1 CUDA程序优化概述 4.2 测量程序运行时间 4.2.1 设备端测时 4.2.2 主机端测时 4.3 任务划分 .....第5章 综合应用附录A 安装、配置、编译及调试附录B 常见问题与解答附录C 技术规范附录D C扩展附录E 数学函数附录F 纹理拾取附录G 着色器模型

## <<GPU高性能运算之CUDA>>

### 章节摘录

插图：第1章 GPU通用计算目前，主流计算机中的处理器主要是中央处理器CPU和图形处理器GPU。传统上，GPU只负责图形渲染，而大部分的处理都交给了CPU。

21世纪人类所面临的重要科技问题，如卫星成像数据的处理、基因工程、全球气候准确预报、核爆炸模拟等，数据规模已经达到TB甚至PB量级，没有万亿次以上的计算能力是无法解决的。

与此同时，我们在日常应用中（如游戏、高清视频播放）面临的图形和数据计算也越来越复杂，对计算速度提出了严峻挑战。

GPU在处理能力和存储器带宽上相对CPU有明显优势，在成本和功耗上也不需要付出太大代价，从而为这些问题提供了新的解决方案。

由于图形渲染的高度并行性，使得GPU可以通过增加并行处理单元和存储器控制单元的方式提高处理能力和存储器带宽。

GPU设计者将更多的晶体管用作执行单元，而不是像CPU那样用作复杂的控制单元和缓存并以此来提高少量执行单元的执行效率。

图1-1对CPU与GPU中晶体管的数量以及用途进行了比较。

## <<GPU高性能运算之CUDA>>

### 编辑推荐

《GPU高性能运算之CUDA》精选典型实用例程，详解CUDA使用细节，重视理论结合实际，介绍并行程序设计方法，深入分析硬件架构，揭示模型与底层映射关系，精心总结优化经验，解析高性能编程技巧。

技术支持：论坛（<http://bbs.hpctech.com>）

## <<GPU高性能运算之CUDA>>

### 版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>