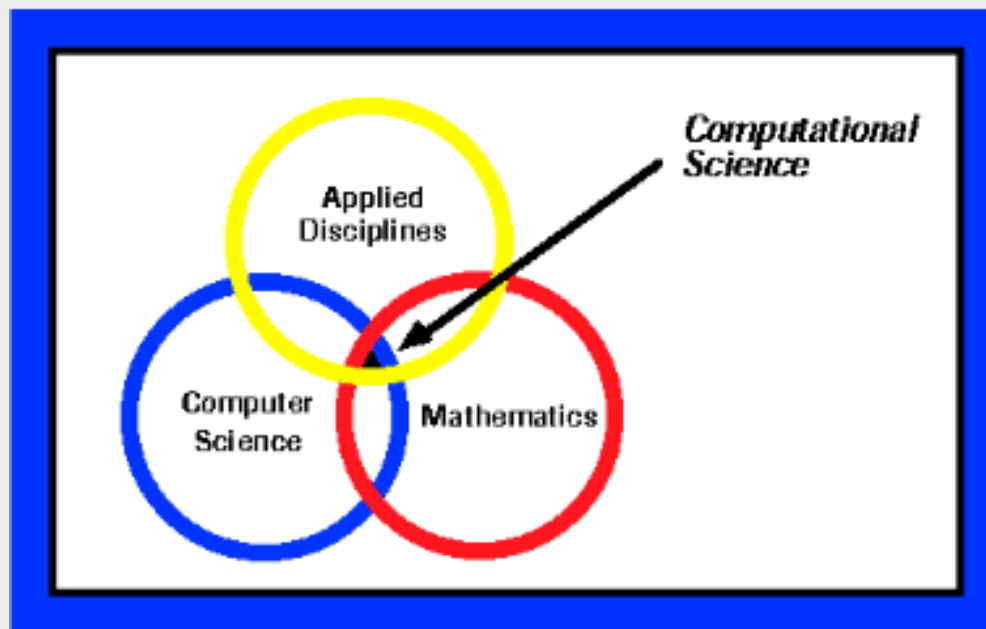


第一章 引言

计算科学 (computational science) 是以计算机为工具，计算机技术为手段，运用计算数学的方法，解决复杂的某科学问题的一门应用型学科。它总是与多（交叉）学科相关的。



授课人：董燎原（中国科学院高能物理研究所）

计算科学要求：

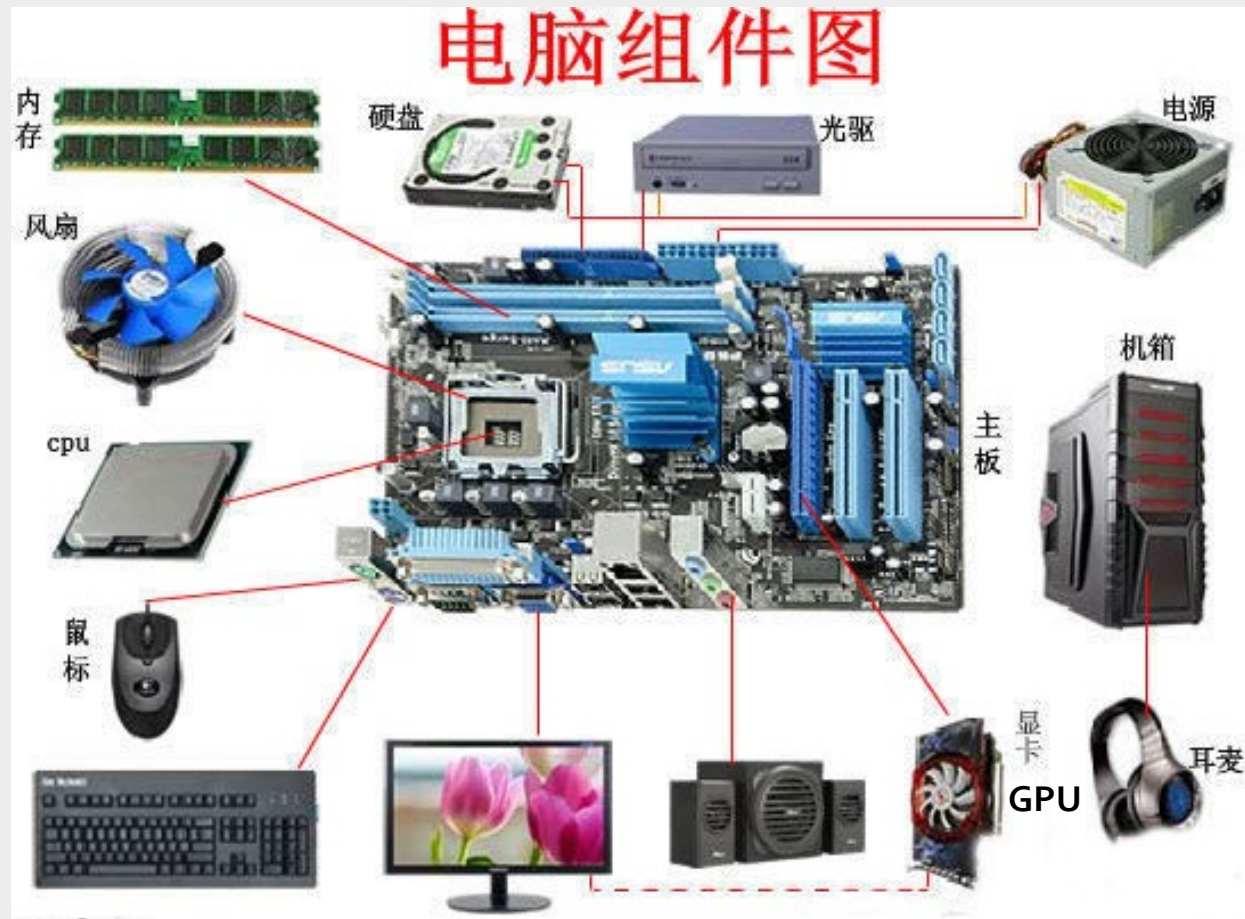
- 应用学科的专业知识
- 高性能的计算构架、存储和网络
- 算法、数值分析、数学方法
- 并行化
- 图形和可视化

**计算物理学与理论物理和实验物理有着密切的联系。
计算物理学的研究内容涉及到物理学的各个领域。**

计算机组件 (硬件+软件)

硬件:

- ✓ CPU
Central Processing Unit
- ✓ GPU
Graphic Processing Unit
- ✓ Memories
temporary and fast storage
- ✓ Permanent storage
hard disk
- ✓ IO devices
understandable input command and output results



软件: 包括操作系统和应用软件。

操作系统 是用户与计算机硬件的接口，负责控制和管理计算机系统的硬件和软件资源。

应用软件 是用户可以使用的各种程序设计语言编制的应用程序的集合

取下CPU散热器后的台式机电脑



配置:

CPU intel i9-7940X

显卡 GTX 1080Ti

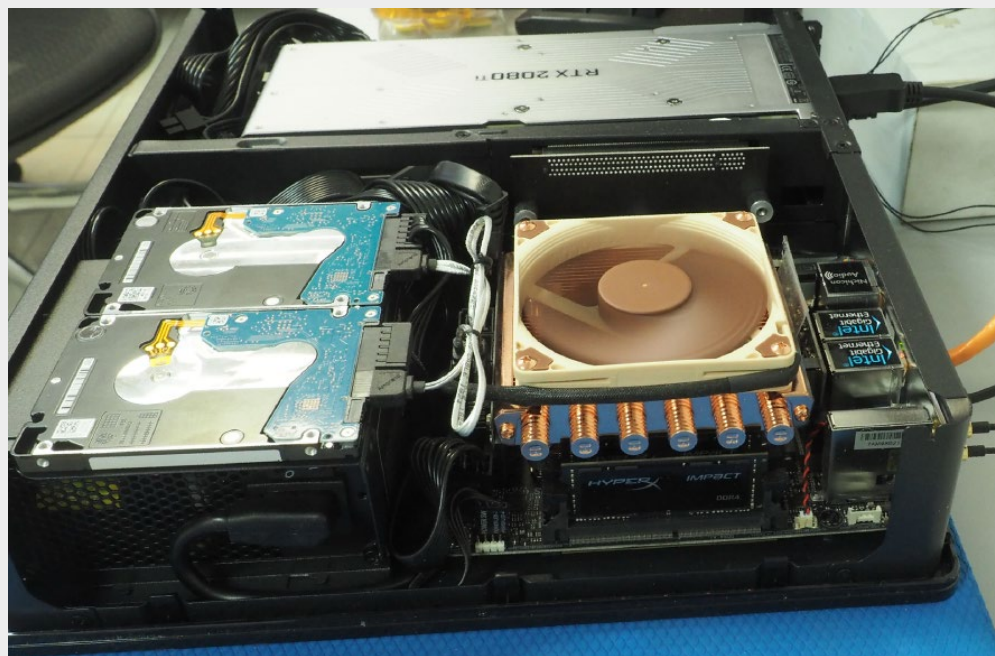
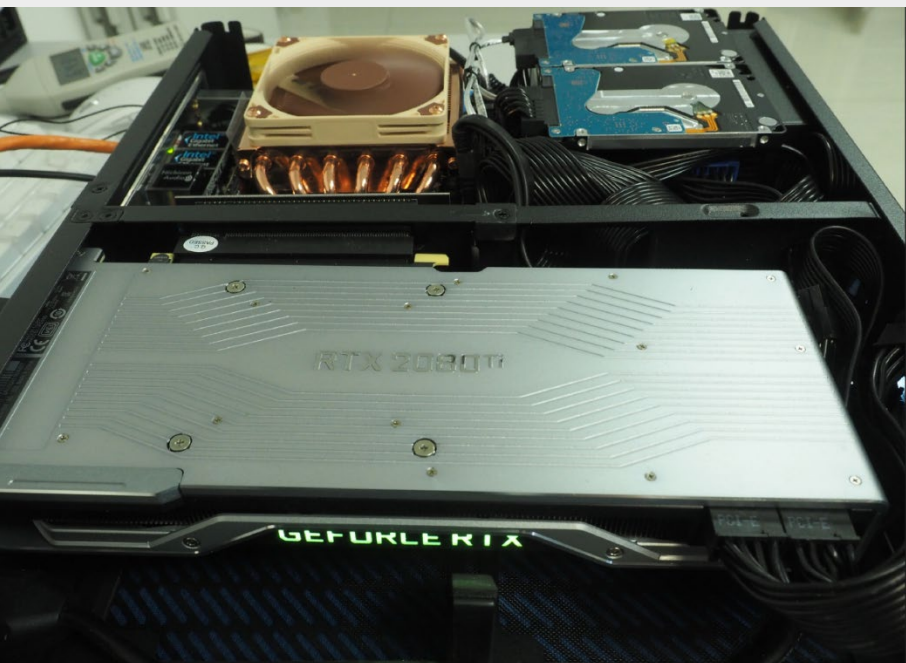
主板 PRIME X299-DELUXE



Double-precision (64-bit) Floating Point Performance
GeForce GTX 1080 Ti up to 0.355 TFLOPS

装上CPU散热器后开机

装上2080Ti的小型可移动电脑



配置:

Double-precision (64-bit) Floating Point Performance
GeForce GTX 2080 Ti **up to 0.44 TFLOPS**

CPU 英特尔 (Intel) i7 7820X 酷睿八核

显卡 英伟达 (NVIDIA) GeForce RTX 2080Ti Founders Edition

主板 华擎 (ASRock) X299E-ITX/ac主板 (Intel X299/LGA 2066)

电源 银欣 (SilverStone) 额定650W SX650-G SFX电源

机箱 分形工艺 (Fractal Design) Node 202 Slim机箱

操作系统 (双系统): Windows 10/Linux openSUSE 15.0

游戏显卡

ITX主板

宽8.2cm

高37.7cm

宽33.2cm

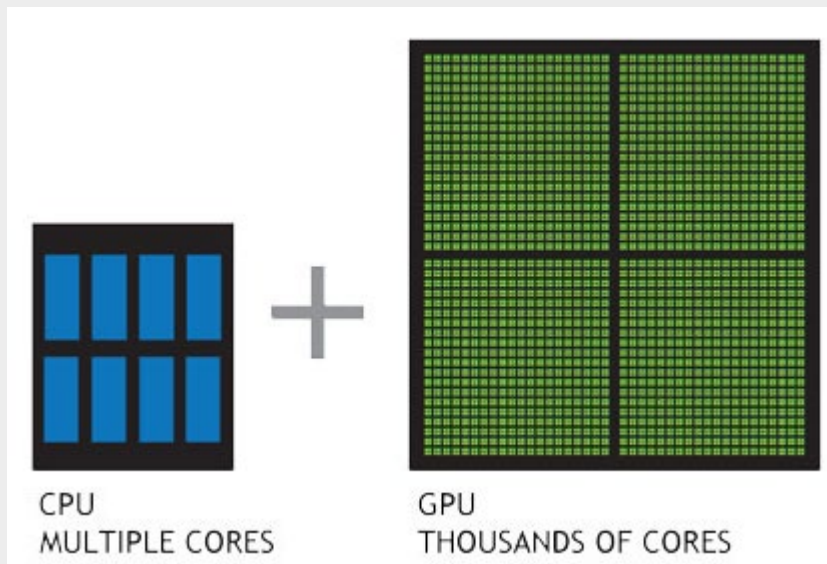


传统CPU 对阵 GPU

理解 CPU 与 GPU 之间区别：

CPU 由专为顺序串行处理而优化的几个核心组成。

GPU 则由数以千计的更小、更高效的核心组成，可并行处理多任务。

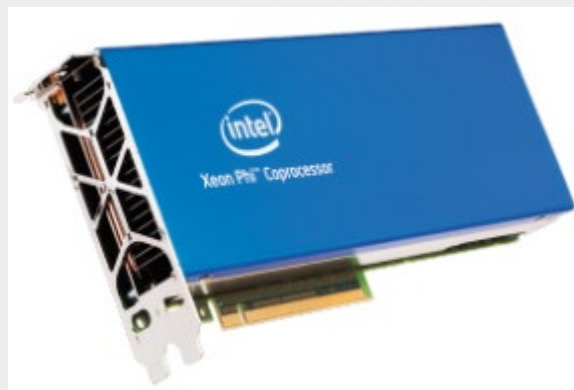


在浮点运算上，GPU可以提供数十倍乃至上百倍于CPU的性能！但移植代码到GPU是一个大工程。



INTEL 2012年11月12日正式推出 61核244线程处理器Xeon® Phi，它像与CPU协同工作的GPU。

因为Xeon Phi与传统CPU在架构上是一致的，因此原有代码可以不加修改就应用在Xeon Phi上。但是代码还需要优化。



最新的专业GPU

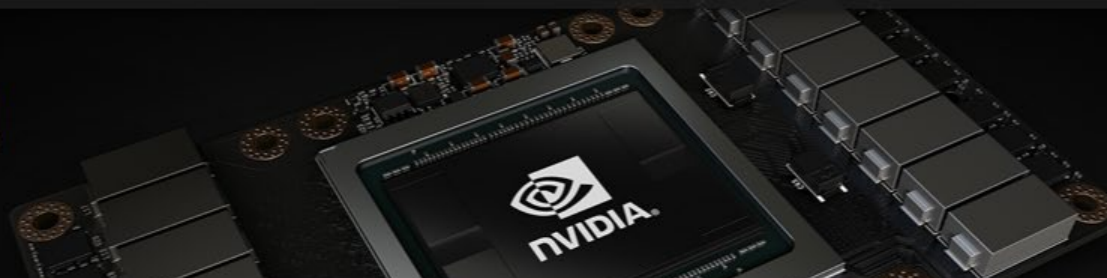
<https://www.nvidia.com/zh-cn/data-center/tesla-v100/>

Tesla V100

人工智能训练 人工智能推理 高性能计算 数据中心 GPUs

NVIDIA TESLA V100 TENSOR CORE GPU

当今市场上数据中心 GPU 中的精尖之作



NVIDIA® Tesla® V100 Tensor Core 是有史以来极其先进的数据中心 GPU，能加快 AI、高性能计算 (HPC) 和图形技术的发展。其采用 NVIDIA Volta 架构，并带有 16 GB 和 32GB 两种配置，在单个 GPU 中即可提供高达 100 个 CPU 的性能。如今，数据科学家、研究人员和工程师可以减少优化内存使用率的时间，从而将更多时间用于设计下一项 AI 突破性作品。

NVIDIA TESLA V100 规格

NVLink 版本 Tesla V100

PCIe 版本 Tesla V100

NVIDIA GPU 加速
带来高性能

双精度

7.8 TeraFLOPS

单精度

15.7 TeraFLOPS

深度学习

125 TeraFLOPS

双精度

7 TeraFLOPS

单精度

14 TeraFLOPS

深度学习

112 TeraFLOPS

双向互联带宽

NVLink

300 GB/s

PCIe

32 GB/s

```
lxslc6 - dongly@gpu012:~ - Xshell 6
1 lxslc6 2 lxslc6 3 lxslc6 4 gcy 5 cpcp 6 lsy 7 zhuz
[dongly@lxslc602 ~]$ cls
[dongly@lxslc602 ~]$ ssh gpu012
dongly@gpu012's password:
Last login: Mon Feb 25 17:05:23 2019 from lxslc602.ihep.ac.cn
mkdir: cannot create directory '/afs/ihep.ac.cn/users/d/dongly/.local': Permission denied
[dongly@gpu012 ~]$ /usr/bin/nvidia-smi -L
GPU 0: Tesla V100-SXM2-32GB (UUID: GPU-7a8b3850-b6bb-87a5-a6f9-14c77e4ece52)
GPU 1: Tesla V100-SXM2-32GB (UUID: GPU-87ba517c-ec19-8dfc-3ddf-e14837ac3af6)
GPU 2: Tesla V100-SXM2-32GB (UUID: GPU-cdf59d99-d308-98c1-d88a-606f0991bb93)
GPU 3: Tesla V100-SXM2-32GB (UUID: GPU-b4a35fc2-ff11-b9e9-de14-36f4bbcc0b10)
GPU 4: Tesla V100-SXM2-32GB (UUID: GPU-78c5b384-4986-bbf2-2ed8-8a0c200db7e7)
GPU 5: Tesla V100-SXM2-32GB (UUID: GPU-60200c52-cdda-6c0b-fe67-8e56785b0d8e)
GPU 6: Tesla V100-SXM2-32GB (UUID: GPU-c11fe911-4722-b3eb-2a08-290ee5f39555)
GPU 7: Tesla V100-SXM2-32GB (UUID: GPU-cd9abc09-aff5-4e9e-3c62-a98545d63aff)
[dongly@gpu012 ~]$
```

高能所计算中心购置了174块 V100
一个节点里面有8块。

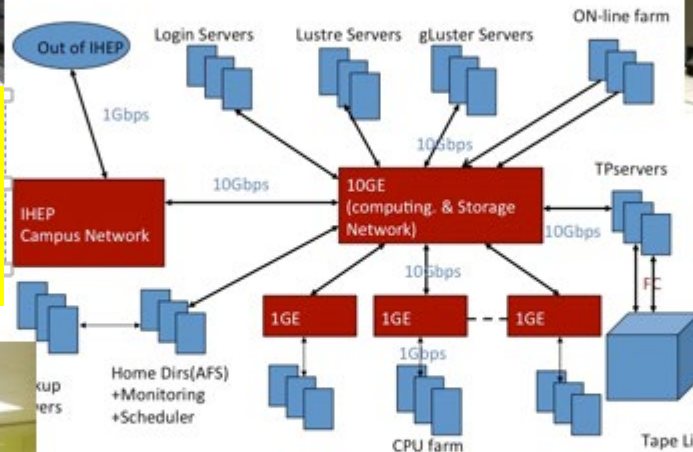
高能物理研究所的计算资源

支持的国际实验



~20000 CPU cores
其中13492 来自BESIII

174 GPU cards (PWA)



4.56PB+6PB(new) tape library



11.7 PB disk space

1. Local clusters



Power supply & cooling system

2. Grid site WLCG
1,500 CPU cores

3. BESIII DIRAC-based distributed
Computing system ~2,000 CPU cores

4. IHEPCloud based on
Openstack ~480 CPU cores

计算物理学在物理学研究中的应用

计算机数值分析

解决已经存在定量理论的物理问题

计算机符号处理

解析计算、公式的推导和高精度的数值计算

计算机模拟

计算机模拟实验

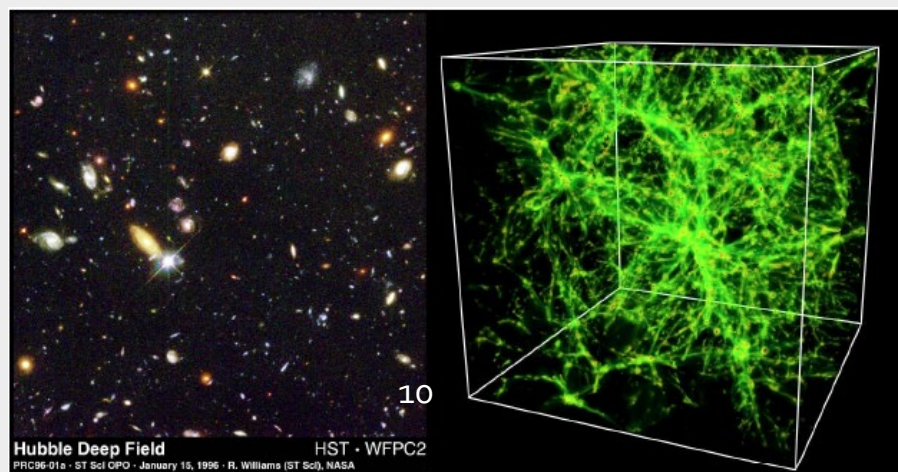
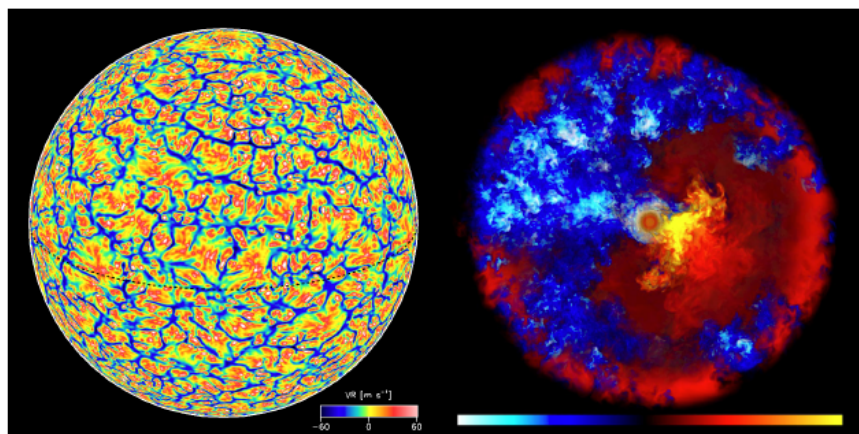
计算机实时控制

物理实验中的计算机控制

计算机模拟

利用现代电子计算机的大存储量和快速计算的有利条件，将物理学、力学、天文学和工程中复杂的多因素相互作用过程，通过计算机来模拟。

- 简单解析理论模型难以完全描述的复杂物理现象。
 - 克服实验物理中的困难，例如模拟实验上不能实现或技术条件要求很高、实验代价昂贵的物理系统等。
- 原子弹的爆炸、火箭的发射、代替风洞进行高速飞行的模拟试验、研究恒星特别是太阳的演化过程等。



实验的计算机实时控制

实验设备通过接口与控制计算机相连接，并结合在线数据获取和分析程序就可以对实验装置的整个实验进程做实时控制，使物理实验可以在没有人在场的情况下自己监测设备的正常运行，自动采集和分析实验数据。

- ✚ 在线分析

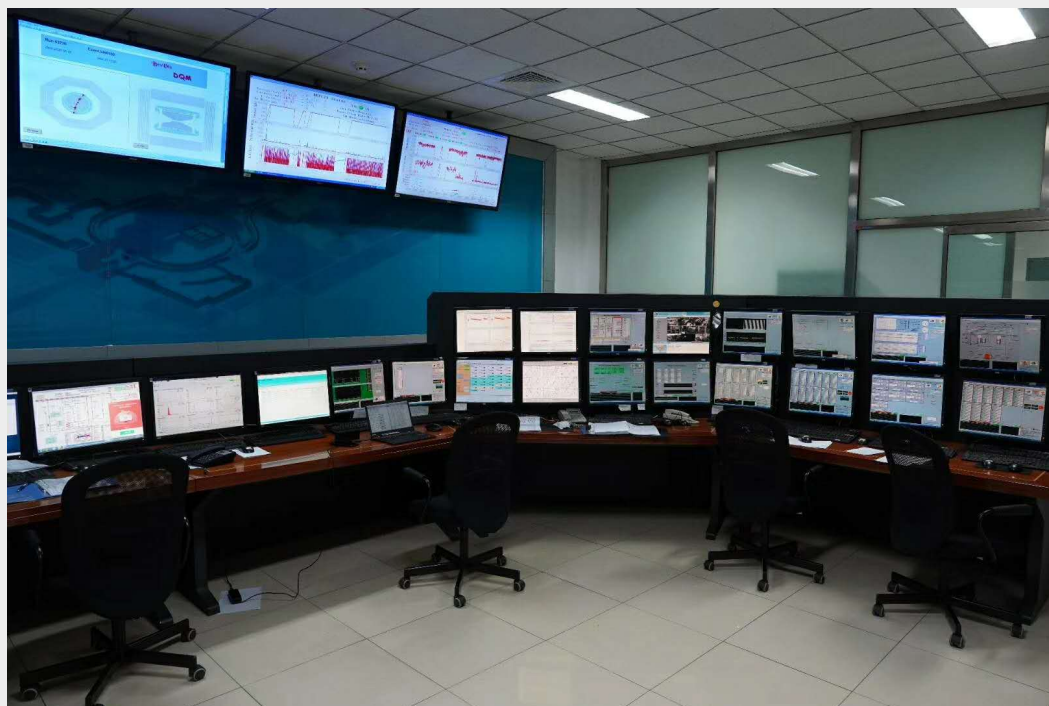
- ✚ 离线分析

实验上计算机在线分析

实验的在线分析：在实验进行过程中由计算机实现数据获取和数据分析。

以粒子物理实验为例，在线分析的任务包括四个方面：

- 控制系统运行
- 采集实验数据
- 监视仪器状态
- 数据在线分析

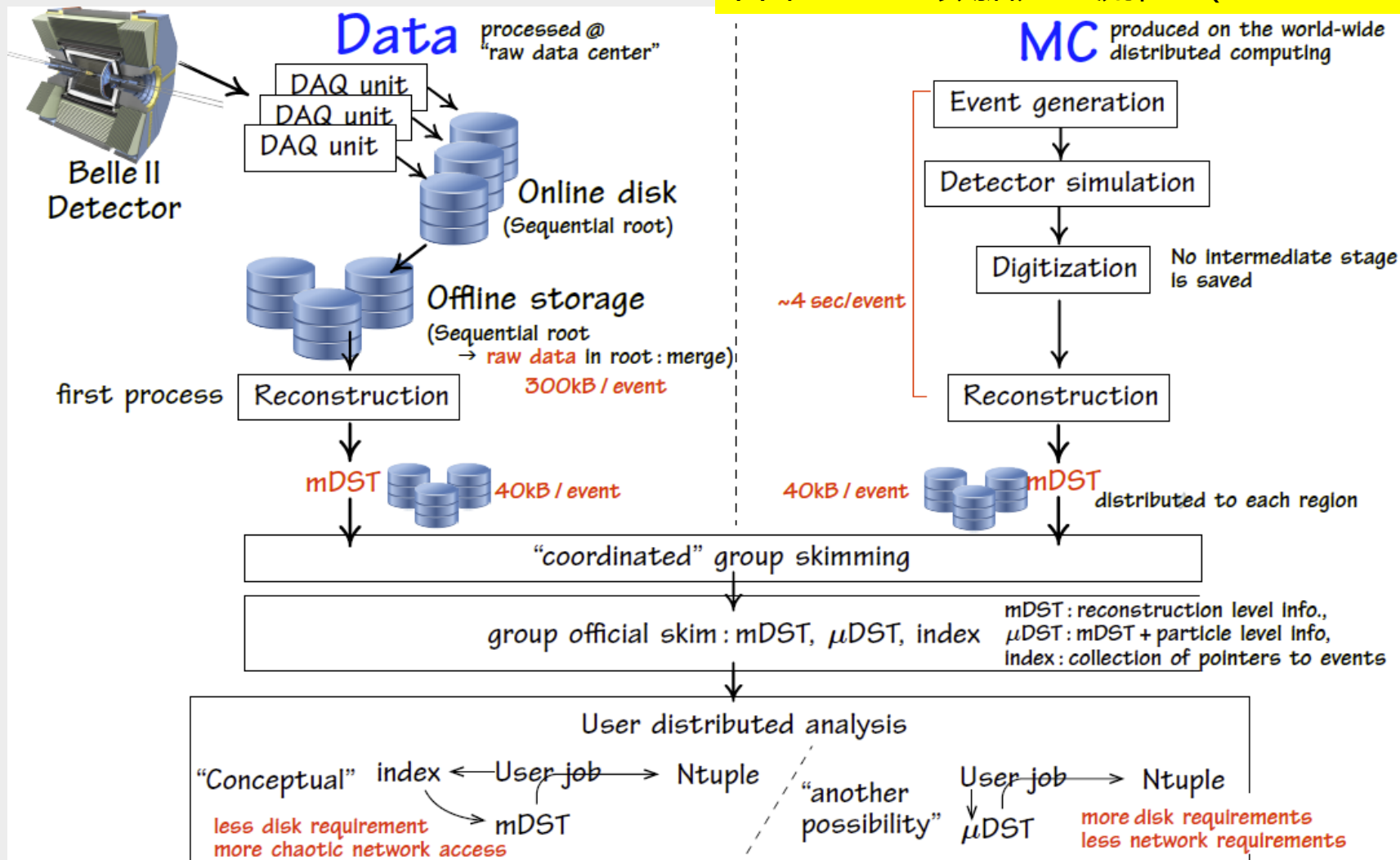


北京谱仪取数值班室，摄于2020年2月15日凌晨

实验上计算机离线分析

实验的**离线分析**：将实验数据送到计算中心做过滤和物理分析工作。包括对物理过程的MC模拟、探测器模拟、本底分析、理论和实验事例的分析对照等，形成物理结果。

日本Belle II 数据处理流程 (CHEP2015)



下面介绍的是：
实际工作中使用GPU加速科学计算的实例，
体会实现的思想。

GPU在Dalitz图分析中的应用

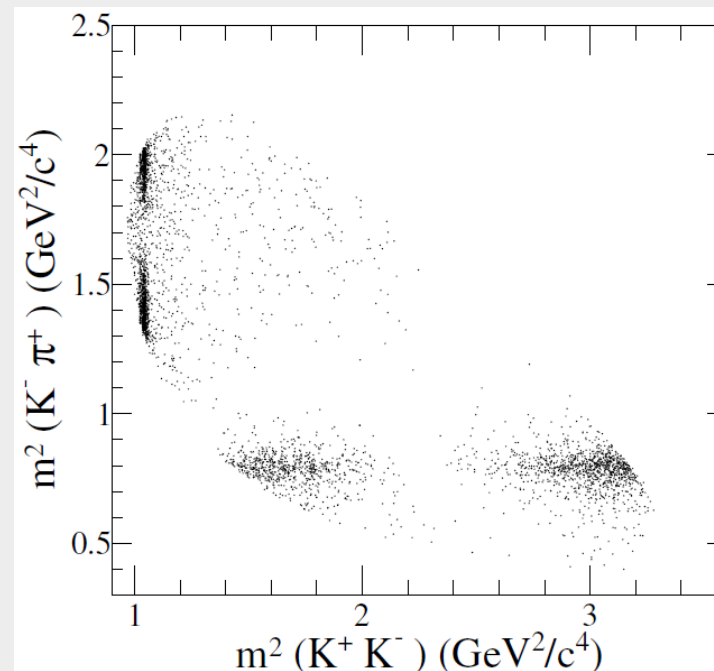
[核电子学与探测技术] 2014年第3期

$$D_s^+ \rightarrow K^+ K^- \pi^+$$

研究背景:

- **Dalitz图**是研究三体衰变过程的重要手段:

1. 在BESIII能区中, 需要对自旋为零的粒子(如: D或Ds)衰变到三个赝标量粒子的三体衰变过程(自由度为2)进行**Dalitz图**分析。
2. 对BESIII的大统计量数据, **Dalitz图**分析的计算开销很大, 使用中央处理单元CPU来计算, 已很难满足物理分析的需求。

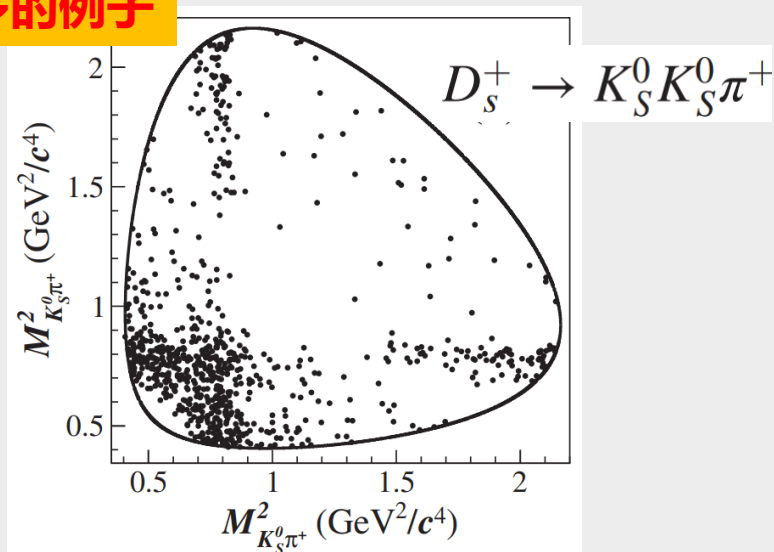


PRD 104, 012016 (2021)

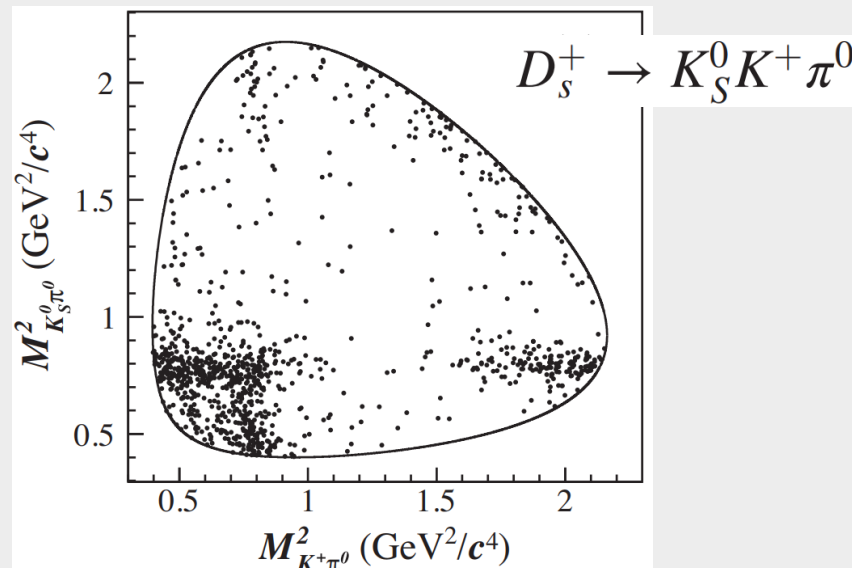
- 开发并行计算**Dalitz图**分析软件:

1. 在开放计算机语言OpenCL的编程标准下,
2. 利用图形处理单元GPU的多核心、适于单指令多线程并行计算的特性,
3. 开发利用GPU进行并行计算, 适合BESIII大统计量数据物理分析的工具。

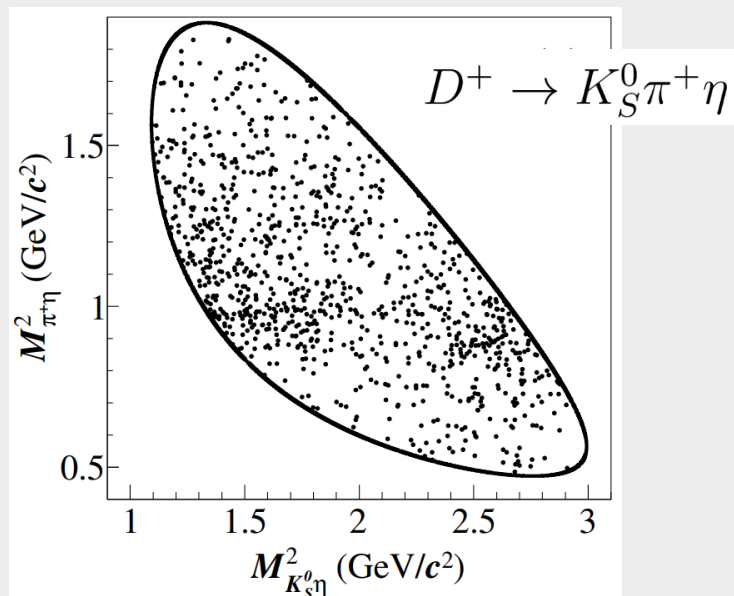
更多的例子



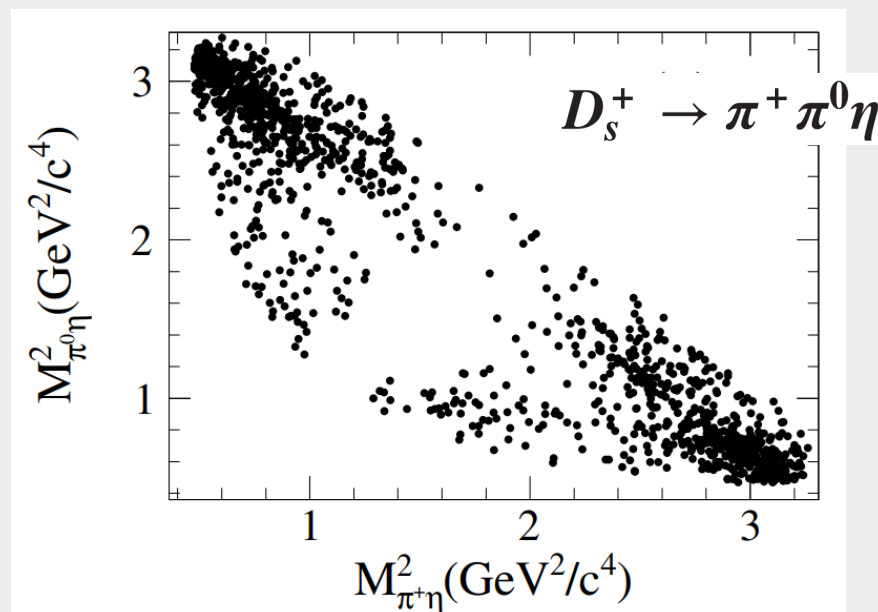
PRD 105, L051103 (2022)



PRL 129, 192001 (2022)



[arXiv:2309.05760](https://arxiv.org/abs/2309.05760) Accepted by PRL



PRL 123, 112001 (2019)

1. Dalitz图分析

对自旋为零的粒子衰变到三个赝标量末态粒子的过程，如D或 $B \rightarrow abc$ ，衰变率可写为：

$$d\Gamma = \frac{1}{(2\pi)^3 32\sqrt{s^3}} |M|^2 dx dy ,$$

式中： $x=m_{ab}^2$ ； $y=m_{bc}^2$ ；这里 m_{ij} 为粒子i和j组合的不变质量； x 和 y 的二维散点图称为Dalitz图， $|M|^2$ 包含动力学效应。Dalitz图上事例分布的任何变化由动力学而非运动学效应引起。

1.1 几率密度函数

几率密度函数（PDF）描述事例在Dalitz图的分布，它由信号和本底两部分组成：

$$\text{PDF}(x, y) = f_s S(x, y) + (1 - f_s) B(x, y) ,$$

式中， S 和 B 分别为信号和本底的PDF； f_s 为信号事例所占的比例。

信号PDF可写为：

$$S(x, y) = \frac{|M(x, y)|^2 \varepsilon(x, y)}{\int_{DP} |M(x, y)|^2 \varepsilon(x, y) dx dy} ,$$

式中， $\varepsilon(x, y)$ 为信号的效率函数，分母的积分是在整个Dalitz图(DP)进行的。

1.2 PDF中的归一化积分

Dalitz图分析常采用数值积分法：

1. 将Dalitz图分割成若干小块，并假定小块内任意点的 $|M(x, y)|^2 \varepsilon(x, y)$ 值均相同。
2. 取小块中心的 $|M(x, y)|^2 \varepsilon(x, y)$ 值乘小块的面积，即可求得单个小块的积分值。
3. 对Dalitz图内的所有小块的积分值求和，可得到近似的积分值：

$$\int_{DP} |M(x, y)|^2 \varepsilon(x, y) dx dy \approx \sum_{i=0}^{N_{bin}} |M(x_i, y_i)|^2 \varepsilon(x_i, y_i) \times S_i,$$

式中： N_{bin} 为分割的小块总数； x_i 和 y_i 为分别第 i 个小块中心点的 x 和 y 值； S_i 为第 i 个小块的面积。
 N_{bin} 由相空间大小、质量分辨和积分精度的要求决定，越大越好，通常取 >1000 。

1.3 极大似然法拟合

通过极大似然法拟合Dalitz图以确定PDF中的待拟合参数，似然函数 \mathcal{L} 定义为：

$$\mathcal{L} = \prod_{i=1}^{N_{event}} \text{PDF}(x_i, y_i),$$

式中： x_i 和 y_i 分别为第 i 个事例的 x 和 y 值； N_{event} 为事例总数。

在实际拟合中，最小化负的对数似然值(NLL)：

$$\text{NLL} = -\ln \mathcal{L}.$$

1.4 拟合过程和速度优化

1. 构造好PDF函数后，调用MINUIT即可对数据的Dalitz图进行拟合，经过反复的迭代，最终确定PDF中的待定参数。

2. 每次迭代都依次进行下面三个主要过程：

计算归一化积分；计算NLL值；找下一组参数值使NLL极小。

3. 速度的优化：

归一化积分的计算是耗时的过程， N_{bin} 越大积分值越准，但计算的时间越长。

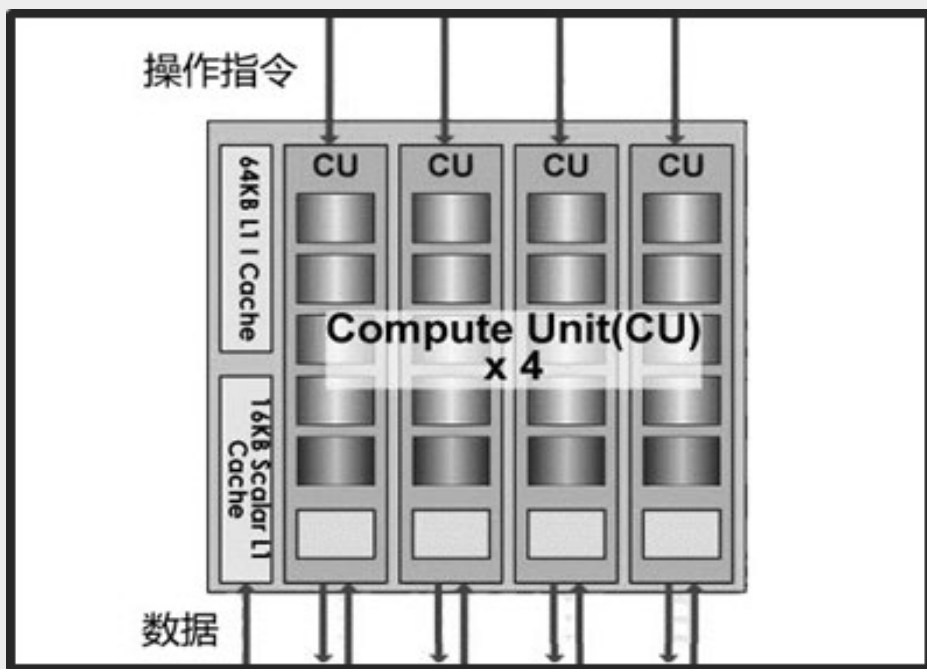
NLL的计算也是一个耗时的过程，待拟合的 N_{event} 越多，则计算的时间越长。

MINUIT找极小会消耗一定的时间，但其代码复杂不易并行化。

结论：归一化积分和NLL值的计算这两个过程都是基于相同代码的大数据量的计算，这一特性十分适合单程序多数据的GPU并行计算模式，故使用GPU来计算可极大地提高拟合的速度。

2. GPU构架

1. GPU的控制逻辑单元比CPU简单很多，有利于减少线程切换的开销；
2. GPU的每一个计算单元(CU)都拥有大量的算术逻辑计算单元(ALU)，以及少量用户可以管理的缓存，使GPU具有更强大的算术运算能力；
3. GPU的内存总线基于带宽优化，150GB/s以上的带宽可以使大量ALU同时进行内存操作，使GPU具有快速的数据处理能力。



GCN计算单元结构示意图

GPU 硬件构架（AMD HD 7970为例）

HD7970采用Graphics Core Next (GCN)架构。

- HD7970的2048个流处理器被划分成32个GCN并行计算单元。
- 而每个GCN内部都具有64个流处理器，这64个流处理器组成了4个计算单元(CU)，每个CU搭配64KB矢量寄存器。
- 每个计算单元都相当于一个运算核心，能同时领取、执行指令。
- GCN架构内，4个CU组成一个运算小组，配备了数据寄存器和一些辅助性功能模块，从而组成了一个完整的运算中枢。

单程序多数据并行计算在GPU上实现的方式:

- 对于单程序多数据的情况，输入的数据储存在各个CU的寄存器中，操作指令一致。
- GPU在工作时，同时执行各个CU的操作指令，完成对各组数据的处理。

3.在GPU上实现Dalitz图分析的并行计算

3.1 OpenCL简介

OpenCL是一个开放式并行计算标准，用于在各种不同的处理器上进行通用目的并行编程，使软件开发人员能够以一种高效可移植的方式使用这些异构处理平台。

OpenCL由语言规范、平台层API和运行时API构成：

- 语言规范描述在支持OpenCL的加速器（如GPU）上编写计算内核的语法和编程接口，所用语言均基于ISO C99子集。
开发人员可以选择预编译OpenCL计算内核，或者让OpenCL运行时根据需要编译内核。
- 平台层API使开发人员可以查询系统设备数目和类型，以便选择必要的计算设备并对其进行初始化，从而正确平衡工作负荷。
- 运行时API允许开发人员排队执行计算内核，并负责管理OpenCL系统中的计算和内存资源。

3.2 Dalitz图分析在GPU的实现

OpenCL平台包括一个宿主机及一个或多个OpenCL设备，设备是执行指令流（或内核）的地方，因此我们需要编写相应的宿主机程序和内核程序。

宿主机程序用来在初始阶段配置GPU资源、创建命令队列并执行、负责创建内存对象存放用于Dalitz图分析的数据、在整个拟合过程中管理内存与显存的数据传输、在计算结束后从GPU中取回计算结果等。

Dalitz图分析中，实验数据 (x_i, y_i) 以一维数组的形式存放，以方便在GPU中进行矢量操作。而少量的拟合参数和控制参数可以直接传递给GPU的私有内存，以提高计算速度。

内核程序在GPU上执行，完成具体工作，即将输入内存对象转换成输出内存对象。在Dalitz图分析中，内核程序主要用于计算 $|M(x, y)|^2 \varepsilon(x, y)$ 值，每个值的计算由相应的工作项来完成，它们有自己唯一的全球ID。对于每一个ID，都有对应的数据 (x_i, y_i) 。

Dalitz图分析的并行运算是数据的并行运算，内核在执行时，可近似地看成对所有的ID同时进行，即同时处理所有的实验数据得到各组 $|M(x, y)|^2 \varepsilon(x, y)$ 值，从而实现数据的并行计算，提高运算速度。

3.3 GPU程序与RooFit框架的整合

- RooFit基于C++开发，调用时遵循C++语言规范。故OpenCL可方便的与C++程序结合。
- 为了加速归一化积分和似然函数的计算过程，我们采用继承的方式：
由RooFit的计算归一化积分值的Integrator类和计算似然函数值的RooNLLVar类继承构造新的IntNLL类，既可计算归一化积分值，也可计算似然函数值NLL。
- 内核程序针对每一个事例同时计算 $|M(x, y)|^2 \varepsilon(x, y)$ 值。每一次拟合迭代中，IntNLL类第一次调用内核获得归一化积分值，第二次调用内核计算似然函数值。
- 归一化积分和似然函数的计算可通过调用同样的OpenCL内核程序实现加速，计算结果供MINUIT使用，寻找最佳的拟合结果。

4. GPU与CPU计算性能的对比

4.1 测试环境

测试采用的GPU为AMD Radeam HD 7970, CPU为Intel Core i7-2600k, 运行在Suse Linux 3.4.6-2.10(64位)操作系统下。

AMD Radeam HD 7970 GPU的参数

全局内存:	3 GB
局部内存:	32768 bit
核心频率:	925 Mhz
缓存大小:	16384 bit
常量缓存大小:	65536 bit
流处理器数量:	2048

测试用计算机配置

CPU:	Intel Core i7-2600k 运行在3.80GHz
内存大小:	16 GB
内存总线频率:	2.133 GHz
操作系统:	SUSE linux 3.4.6-2.10 64位

软件开发环境

AMD APP SDK:	AMD-APP(1113.2)
OpenCL版本:	OpenCL 1.2
Root版本:	5.34/01 (July13th , 2012)
gcc版本:	4.7.1 20120723

4.2 测试结果

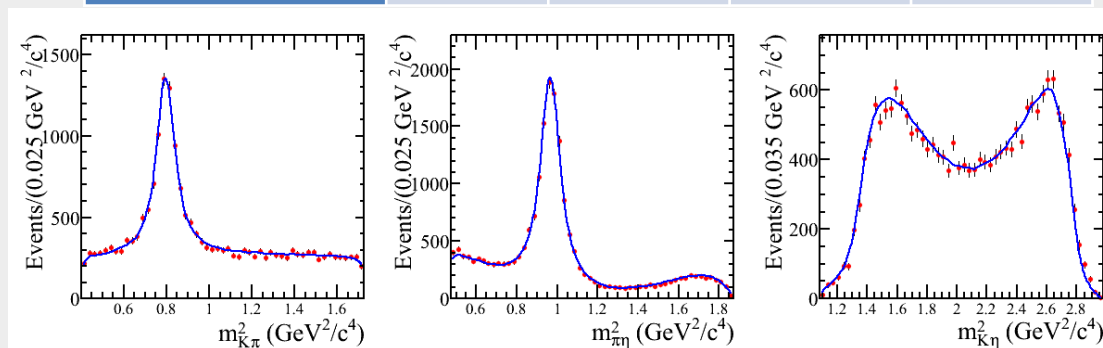
$D^0 \rightarrow K^- \pi^+ \eta$ 的Dalitz图分析

对BES-III MC样本的20335个事例，用开发出的CPU和GPU的Dalitz图分析程序分别进行拟合。

CPU与GPU的Dalitz图分析程序得到的拟合结果和投影图完全一致。

参数	CPU设备		GPU设备	
	拟合值	拟合误差	拟合值	拟合误差
$a_0(980)$ 质量(GeV)	0.9846	0.0004	0.9846	0.0004
$a_0(980)$ 宽度(GeV)	0.0591	0.0011	0.0591	0.0011
$a_0(980)$ 振幅的强度	0.5563	0.0115	0.5563	0.0115
$K^*(892)$ 质量(GeV)	0.8969	0.0005	0.8969	0.0005
$K^*(892)$ 宽度(GeV)	0.0564	0.0014	0.0564	0.0014
非共振态振幅的强度	1.286	0.077	1.286	0.077

拟合结果比较



拟合投影图

CPU程序的拟合时间 (T_{CPU}) 与GPU程序的拟合时间 (T_{GPU}) 的比较
(拟合时间的定义为程序开始到程序结束所消耗的真实时间(期间CPU的占用率大于99.5%))

N_{bin}	1000 × 1000	3000 × 3000	5000 × 5000
T_{CPU} : CPU拟合时间/s	2315	19074	54049
T_{GPU} : GPU拟合时间/s	15.7	122	333
$T_{\text{CPU}}/T_{\text{GPU}}$	147	156	162

GPU的拟合速度高于CPU 150倍左右。

5. 总结

- 建立了基于RooFit框架、应用图形处理单元GPU多核心单指令多线程的并行计算技术，实现Dalitz图分析并行计算的方法。
- 相应开发出的Dalitz图分析并行计算软件与使用CPU的Dalitz图分析软件测试比较表明：两种软件的拟合结果一致；而GPU的拟合速度是CPU的150倍以上。
- GPU在Dalitz图分析中的成功应用，极大减少了数据分析的时间开销，有效提高了物理分析的工作效率。