



第十五届全国科学计算与信息化会议
暨现代物理信息化论坛

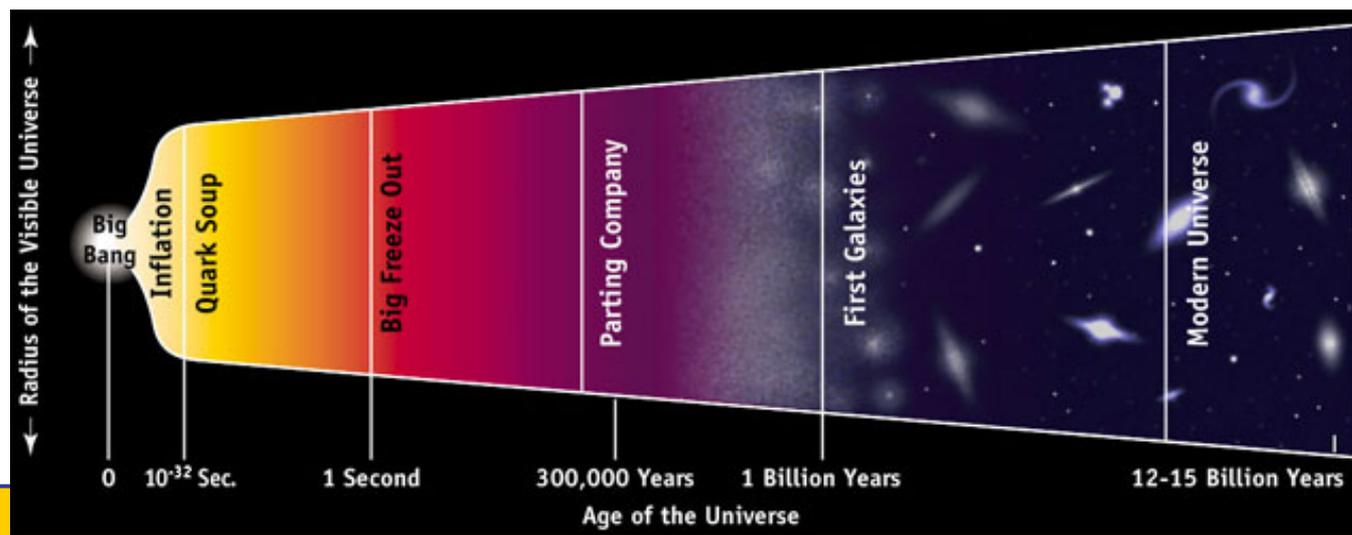
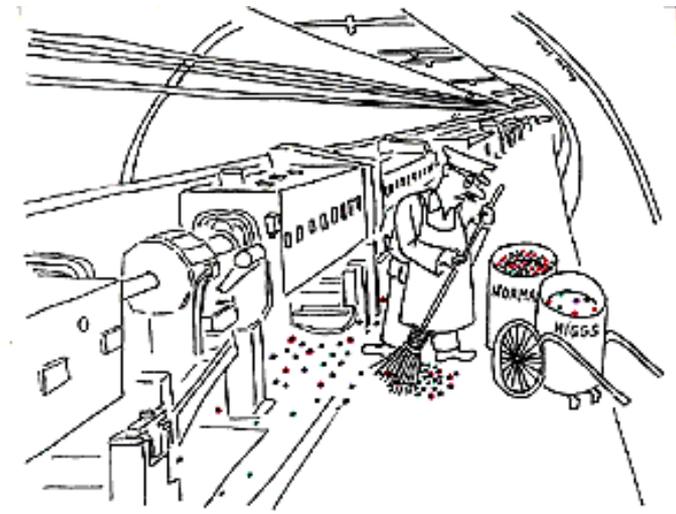
高能物理与信息化

陈刚

2011-07-04

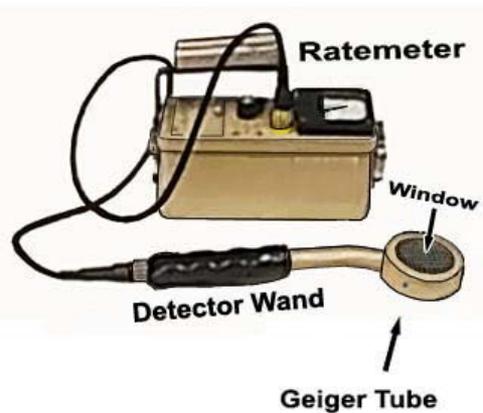
高能物理的目的

- 探索物质微观结构
- 探索宇宙起源等自然规律



高能物理学科特性

- 现代高能物理的特征：大科学装置和国际合作
 - 实验规模越来越大，参加人越来越多
 - 绝大多数是国际合作



一人



几十人



二千人



高能物理为什么要信息化

- 海量数据：数百PB的实验数据
- 大量计算：需要十几万台计算机
- 网络，数据传输，国际交流
-

信息化



(Borromean Rings)

理论

实验

从实验到数据处理分析

探测器

event filter
(selection & reconstruction)

事例重建

原始数据

event reprocessing

物理分析

batch physics analysis

重建数据

event simulation

物理模拟

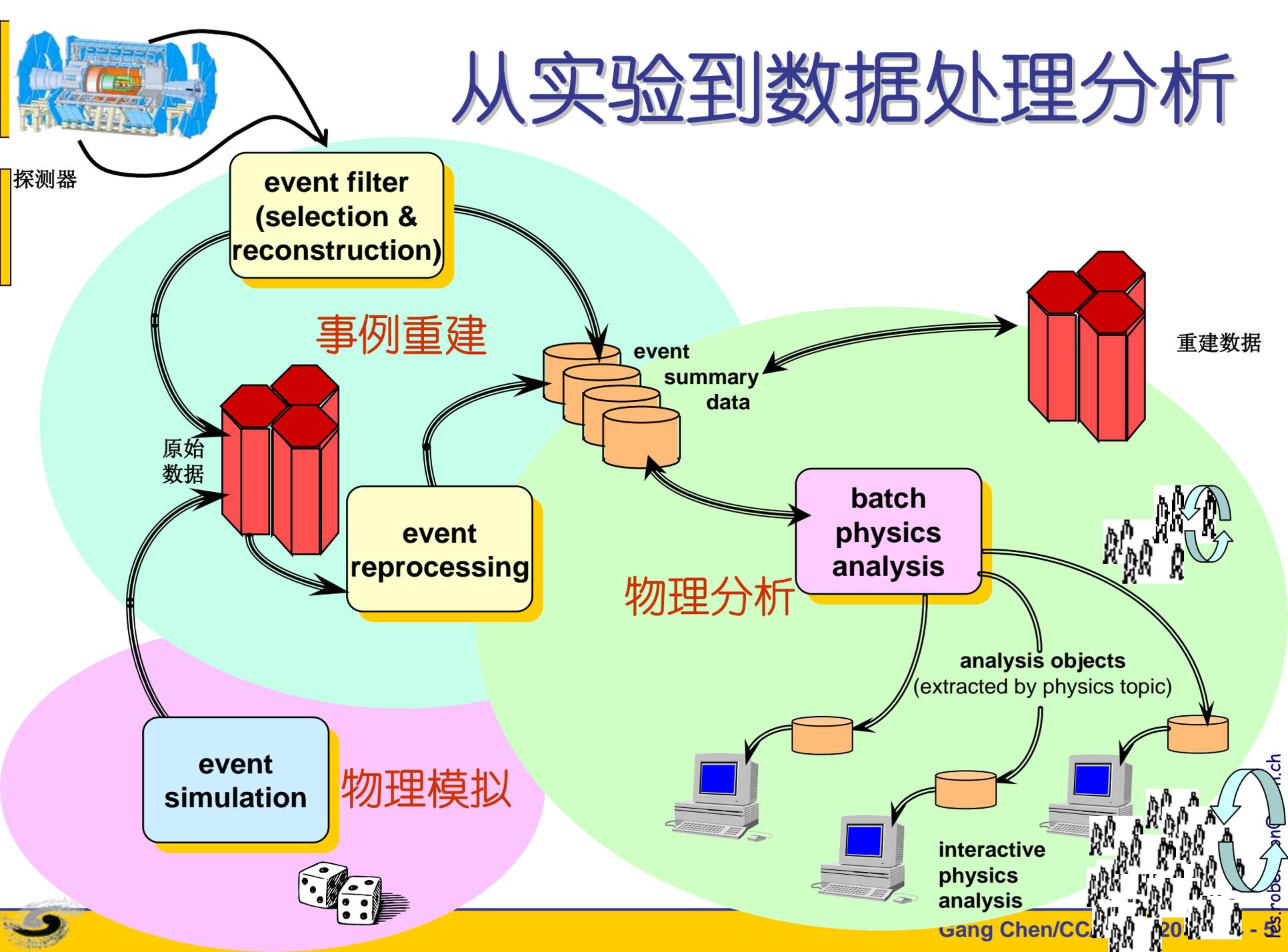
analysis objects
(extracted by physics topic)

interactive physics analysis

Gang Chen/CC

20

fb.in.ch



信息化的任务

- 实验装置设计
- 在线数据采集、实验装置控制
- 物理数据处理，事例重建、物理模拟、可视化
- 计算环境硬件基础，**benchmark**，虚拟化，硬件系统管理与控制
- 分布式及网格(云)计算
- 网络及应用
- 协同工具



国际合作与协调

- 国际合作与协调
 - 统一协调
 - 共同开发，形成全球可共享的平台
- **ICFA** 下属的相关技术委员会：
 - **IHEPCCC** - International High Energy Physics Computing Coordination Committee
 - **SCIC** - Standing Committee on Interregional Connectivity
 - **DPHEP** - Study Group on Data Preservation in High Energy Physics



IHEPCCC

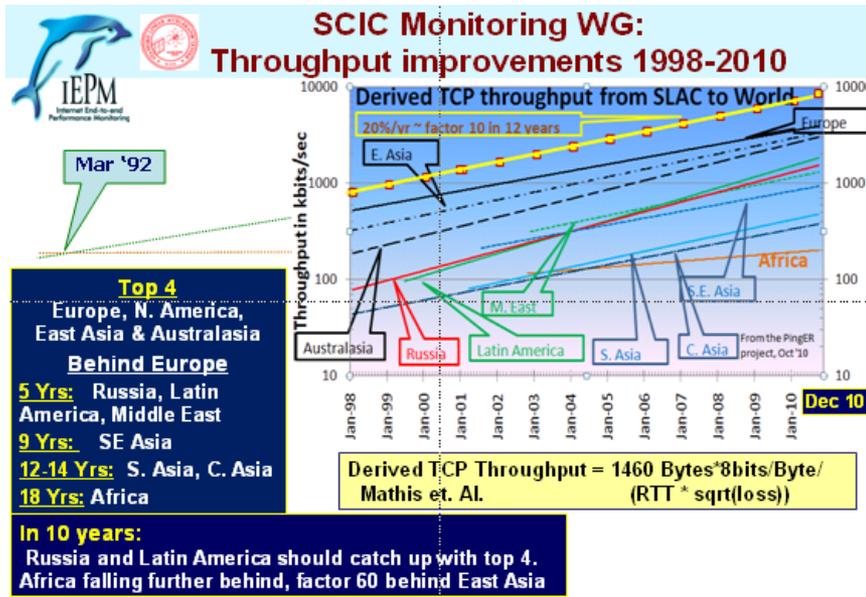
- **IHEPCCC**由负责高能物理计算的专家组成，对相关技术提出建议。
- 新技术与新趋势，计算中心战略，安全问题
- 高能物理计算新标准，硬件性能评价与比较，新建议
- 与其他科研领域的计算机相关工作进行沟通
- 与**SCIC**、高能物理网格等委员会或项目进行合作
- 向**ICFA**提交工作报告



Worker Node, cpu, clock, L2 cache, - Main Memory	SPECint200 gcc	SPEC CPU2006 int - gcc	HEP-SPEC	site measured
2 x Nocona/Irwindale 2.8 GHz/1 MB - 2GB	1501	11.06	10.24	CERN
2 x Nocona/Irwindale 2.8 GHz/2 MB - 4GB	1495	10.09	9.63	CERN
2 x Nocona/Irwindale 2.8 GHz/2 MB - 2GB	1673	11.87		CERN
2 x Nocona/Irwindale 2.8 GHz/2 MB - 2GB	1703	12.26		CERN
2 x Opteron 275 2.2 GHz/2 MB - 4GB	4133	28.76	28.03	CERN
2 x Woodcrest 2.66 GHz/4 MB - 8GB	5675	36.77	35.58	CERN
2 x Woodcrest 3.00 GHz/4 MB - 8GB	6181	39.39	38.21	CERN
2 x Opteron 2218 (Rev. F) 2.6 GHz/2 MB- 8GB	4569	31.4	31.67	CERN
2 x Clovertown 2.33 GHz/2x4MB - 16GB	9462	60.89	57.52	CERN
2 x Harpertown (E5410) 2.33 GHz/2x6 MB - 16GB	10556	64.78	60.76	CERN
2 x Harpertown (5440) 2.83 GHz/2x6MB - 16GB	11850	73.32		DESY
2 x Harpertown (5410) 2.33 GHz/2x6MB - 16GB	11164	65.93	62.12	INFN-PD
2 x Barcelona (2352) 2.10 GHz/2x4MB - 16GB	8488	56.23		INFN-PD
2 x Barcelona (2360) 2.50 GHz/2x4MB - 16GB	9939	63.75	63.19	INFN-PD
2 x Barcelona (2356) 2.30 GHz/4x512 KB - 16GB	9565	61.05	59.74	GridKa
2 x Shanghai (2376) 2.30 GHz/4x512KB+6MB L3 - 16GB	10962	66.88	65.85	GridKa
2 x Harpertown (E5430) 2.66 GHz/2x6MB -16GB	12122	72.14	68.04	GridKa
2 x Opteron (2216) 2.4GHz -8GB		21.86	22.22	RAL
2 x Barcelona (2354) 2.2 GHz -16GB		58.17	58.1	RAL
2 x Clovertown (E5335) 2.00 GHz/2x4 MB - 16GB		54.42	52.31	RAL
2 x Harpertown (L5410) 2.33 GHz/2x6MB -16GB		68.84	62.23	RAL
2 x Harpertown (E5410) 2.33 GHz/2x6MB -16GB		65.73	62.21	RAL
2 x Harpertown (L5420) 2.50 GHz/2x6MB - 16GB		68.73	62.11	RAL
2 x Harpertown (E5420) 2.50 GHz/2x6MB - 16GB		69.13	65.11	RAL
2 x Harpertown (E5420) 2.50 GHz/2x6MB, 16GB		66.83	64.85	RAL
2 x Harpertown (E5440) 2.83 GHz/2x6MB, 16GB		74.92	68.51	RAL

SCIC

- 监视和评价区域网络互联状态
- 分析评价高能物理的网络需求
- 提出网络优化与升级的建议



Summary and Conclusions Networks for HEP in the LHC Era

- The capacity and capability of HEP's networks continues to advance; we will soon be taking the next step to 40G/100G on major routes
- The experiments are building a new round of Computing Models, with greater reliance on networks
 - More intensive use of Tier2s & Tier3s; more complex flows
 - More agile, and more effective for discoveries
- The LHCOPN team has designed and is now developing a new architecture based on a global set of Open Exchange Points to meet the needs: **LHCONE**
 - Experiments and network providers need to work together now, to complete the Phase 1 plan and begin operations in 2011
- We must continue to work to resolve connectivity issues wherever they arise in the HEP community
 - This is especially important now that the 2011-12 LHC run is beginning
- We must continue to work on the Digital Divide in many regions: India, Pakistan, Mexico and elsewhere in Latin America, and Africa

DPHEP

- 评价与起草高能物理数据保存的物理目标
- 讨论数据分析模式，信息提取、软件、存档...
- 论述数据保存中的硬件和软件可持续问题
- 评价数据保存的经费问题以及国际合作
- 编制数据长期保存的技术蓝皮书



Data Preservation in High-Energy Physics

 Study Group for Data Preservation and
Long Term Analysis in High Energy Physics

<http://dphep.org>

Abstract

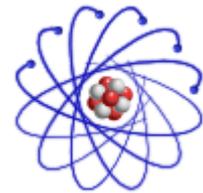
Data from high-energy physics (HEP) experiments are collected with significant financial and human effort and are mostly unique. At the same time, HEP has no coherent strategy for data preservation and re-use. An inter-experimental Study Group on HEP data preservation and long-term analysis was convened at the end of 2008 and held two workshops, at DESY (January 2009) and SLAC (May 2009). This document is an intermediate report to the International Committee for Future Accelerators (ICFA) of the reflections of this Study Group.



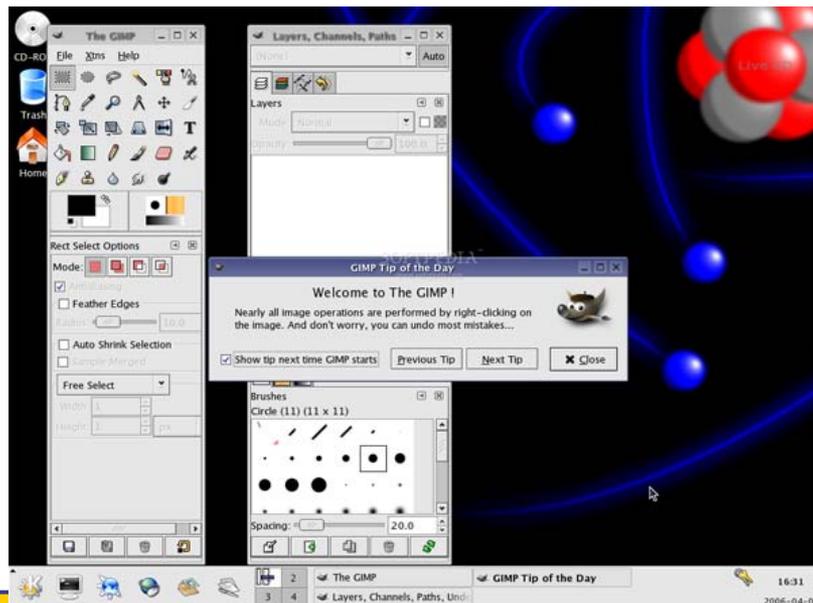
计算环境

- 面向需求，建立领域公用软件环境
- 操作系统：**Scientific Linux, ...**
- 支撑平台系统：**CASTOR, dCache, Quattor, ...**
- 通用物理及数值计算基础库：**Geant4, Root, ...**
- ...

Scientific Linux (SL)

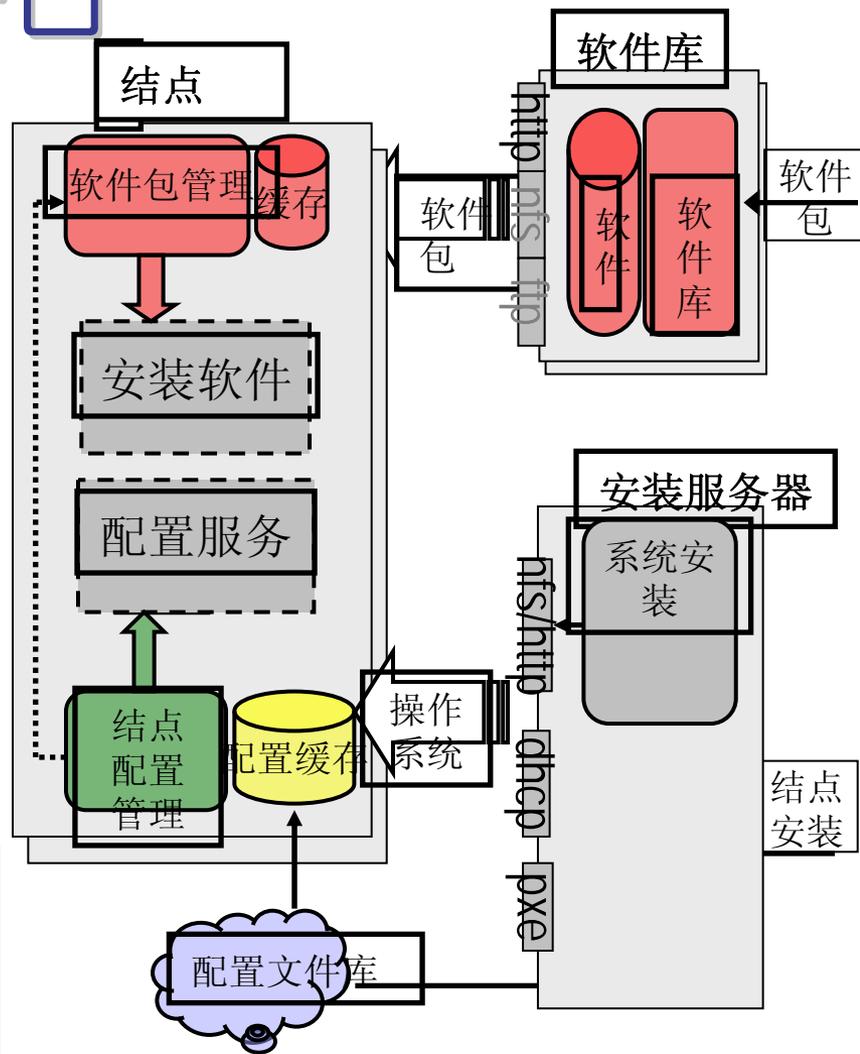
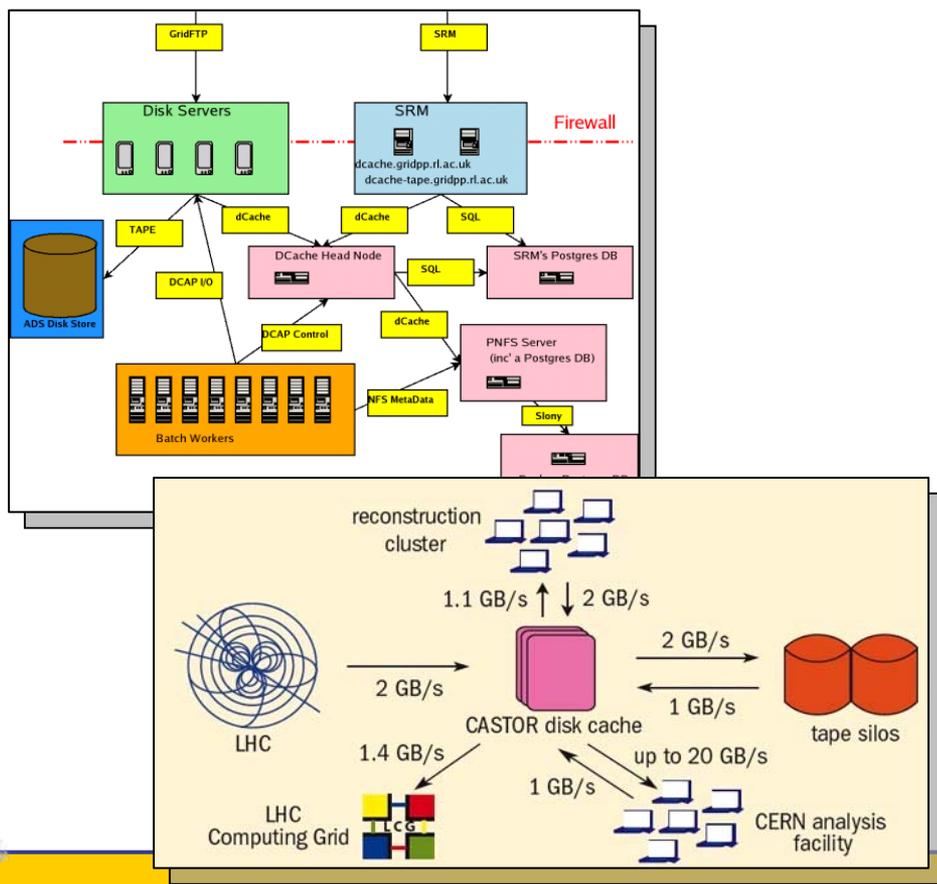


- 世界各地的科研院所联合制作，形成通用高能物理实验的基本软件包。减少重复劳动
 - 基于企业版Linux，
 - 可方便地按各自的需求加入组件，形成本地化的Linux发布版
- 成为高能物理计算标准操作系统



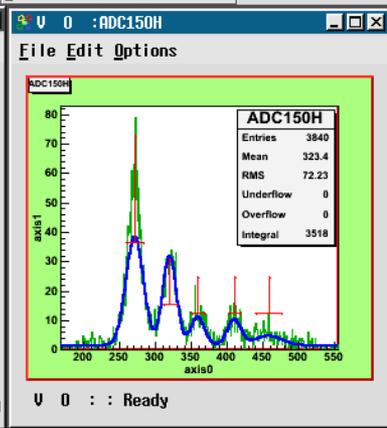
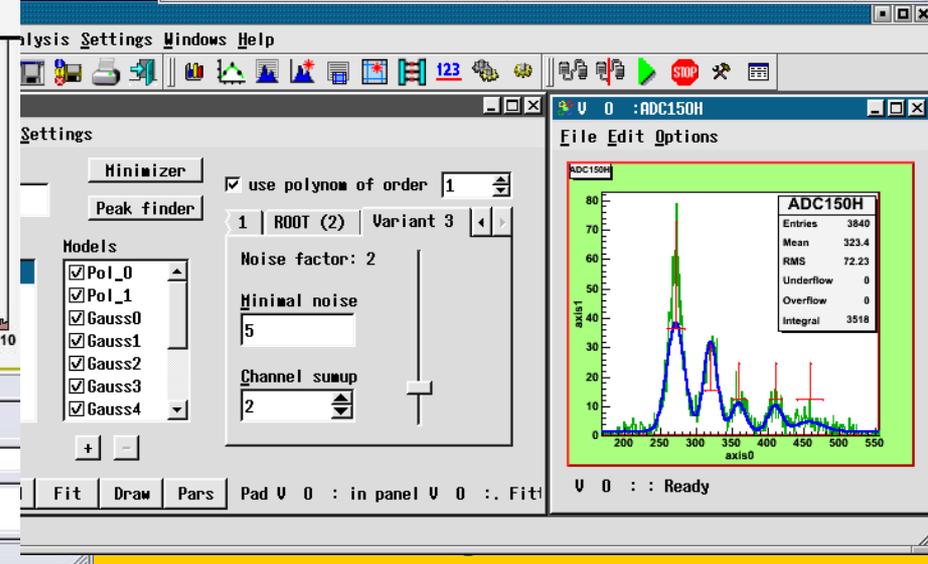
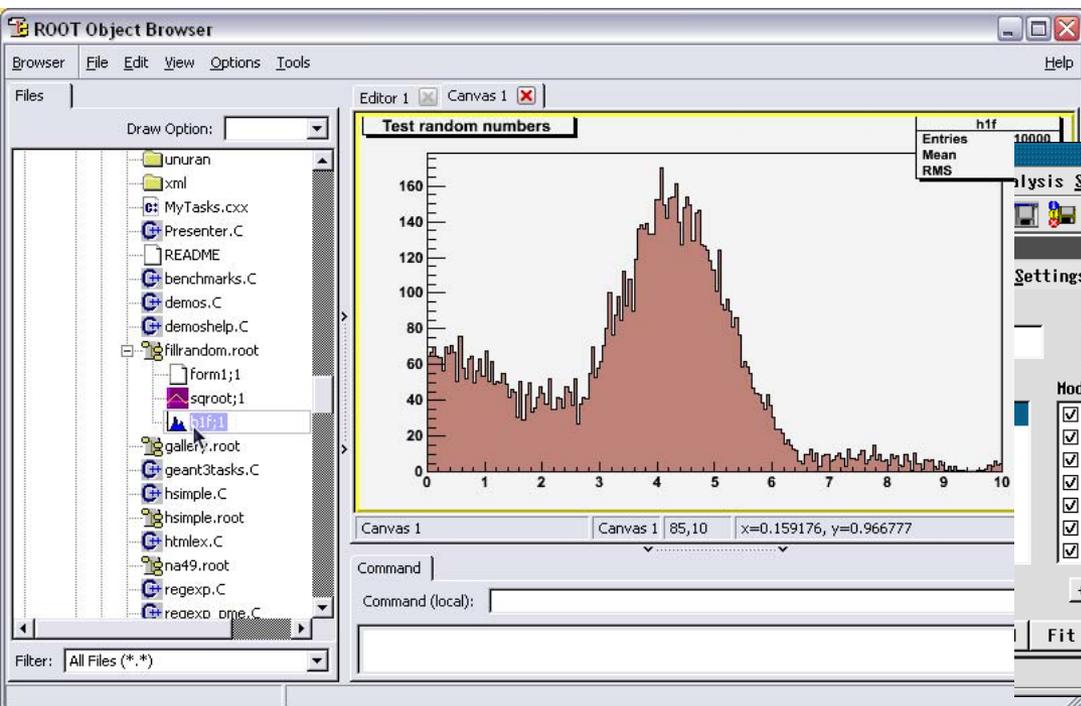
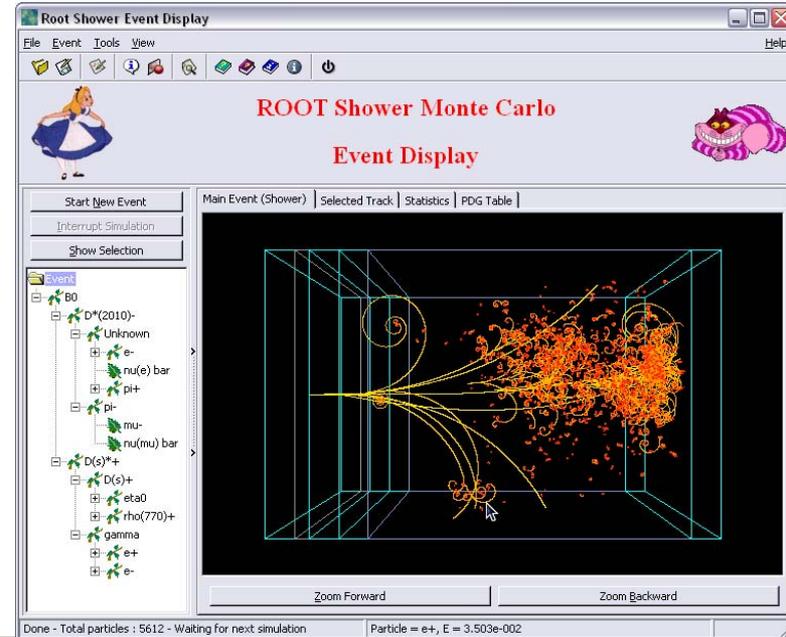
支撑平台

- 基础设施自动化及管理系统Quattor
- 海量存储管理系统CASTOR, dCache



Root

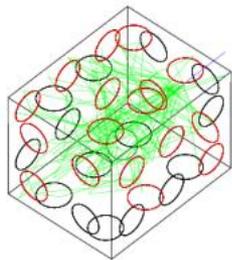
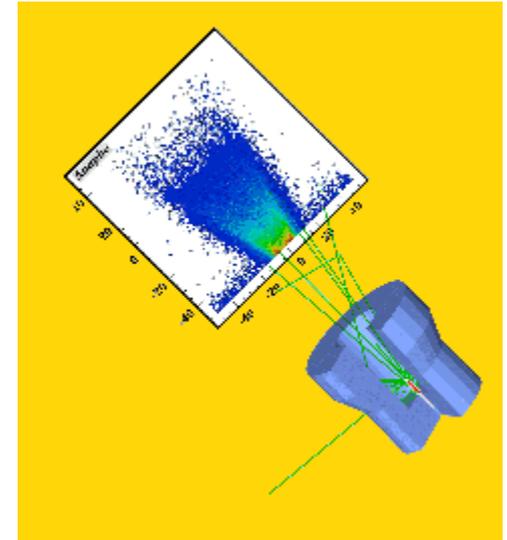
- 数值计算软件包
- 数据分析可视化
- 所有物理分析软件几乎都以Root为基础
 - GUI或者API方式



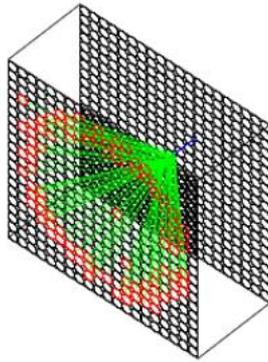
Geant4

- 功能强大的物理模拟软件
- 大量应用于“粒子与物质相互作用”的模拟计算

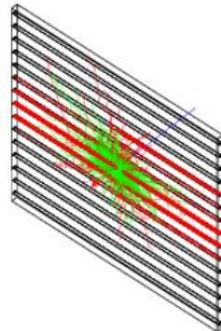
/examples/extended/optical/LXe



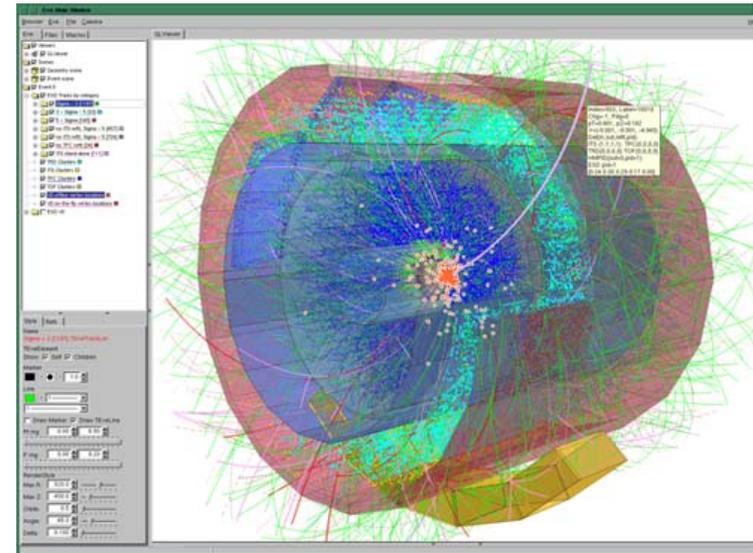
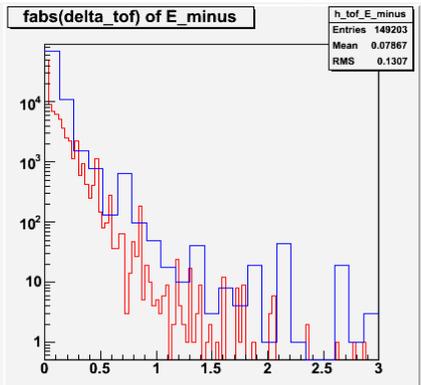
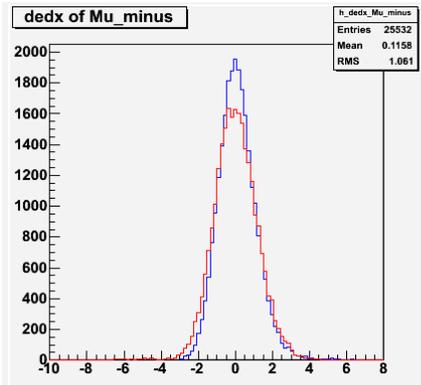
Scintillation



Cerenkov

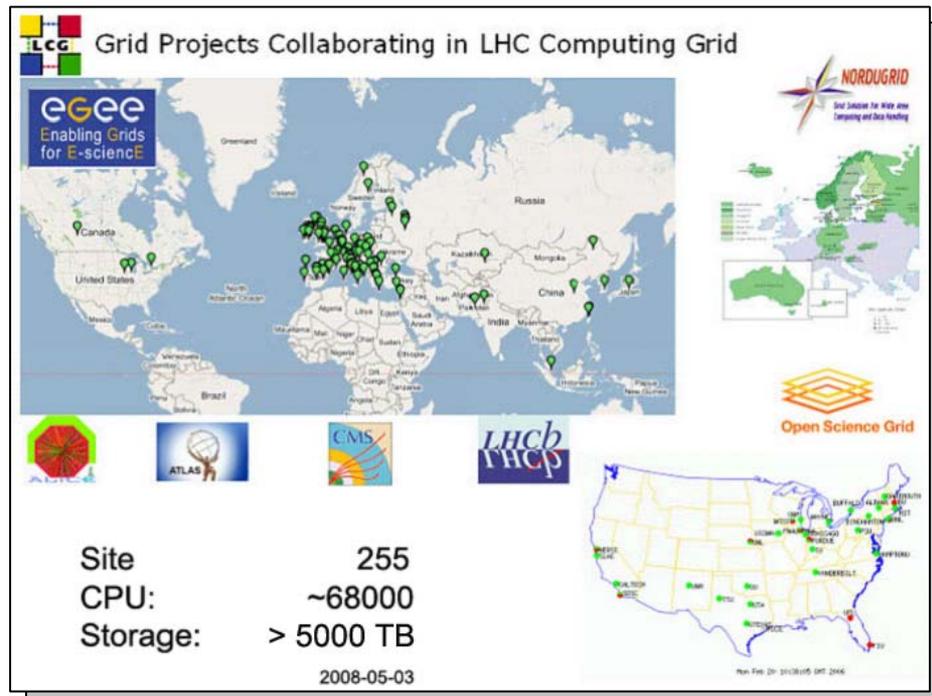
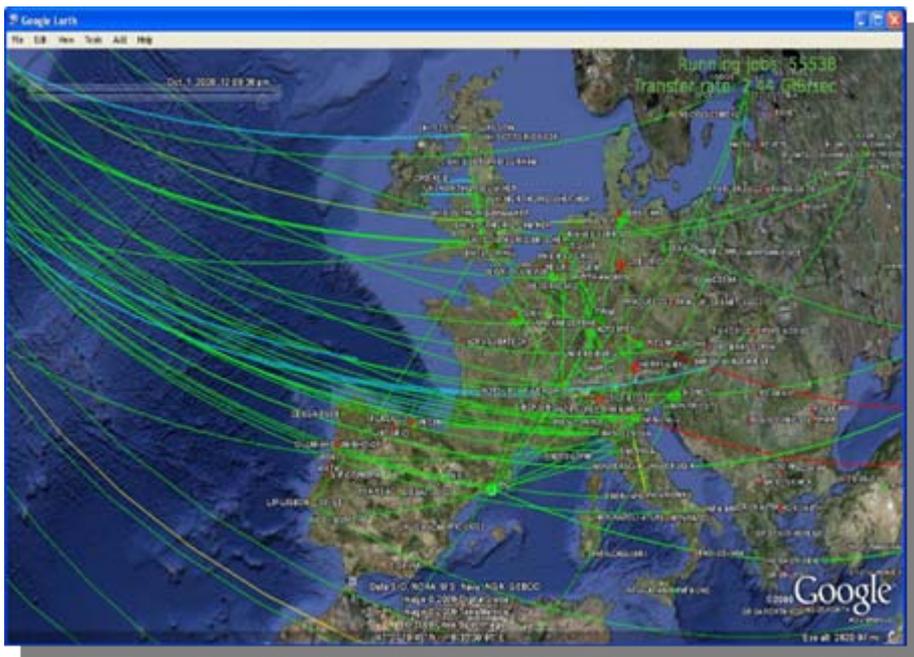


WaveLengthShifting



WLCG 网格平台

- gLite, OSG等中间件为基础
- 由EGEE, OSG, NorduGrid等网格系统互联互通
- 主要面向LHC的四个实验 (Alic, ATLAS, CMS, LHCb) 同时支持其它应用



协同工具

会议管理系统

第十五届全国科学计算与信息化会议暨现代物理信息化论坛 (04-07 July 2011) - Mozilla Firefox

http://indico.ihep.ac.cn/conferenceOtherViews.py?confId=2054&... - DFHEP

“第十五届全国科学计算与信息化会议暨现代物理信息化论坛” from Monday 04 July 2011 (08:00 to Thursday 07 July 2011 (18:00 at 成)

Description: 全国科学计算与信息化会议主要面向全国科学研究人员以及计算机专家与工程技术人员, 搭建一个沟通交流平台, 探讨科学研究需求、实际工作经验以及新技术发展与应用, 以更好的提升计算机与信息技术在现代科学研究中的应用水平。

Monday 04 July 2011 | Tuesday 05 July 2011 | Wednesday 06 July 2011 | Thursday 07 July 2011

Monday 04 July 2011 [top](#)

- 09:00->10:00 大会报告 (Convener: 陈刚) (Location: 三楼神光厅)
- 09:00 科学研究中的数据处理的 (30') 阎保平
- 09:30 次临界反应堆-核能可持续发展的重要途径 (30') 吴宜灿
- 10:00 茶歇 (30')
- 10:30->12:00 大会报告 (Convener: 吴宜灿) (Location: 三楼神光厅)
- 10:30 云时代的科学数据中心——架构与服务 (30') 吴开超
- 11:00 大型高功率固体激光驱动器的计算机集中控制 (30') 赵强
- 11:30 大亚湾反应堆中微子实验及其数据获取和探测器控制系统 (30') 李小勇
- 12:30 午餐 (1h00') (Location: 二楼中餐厅)
- 14:00->16:15 现代物理信息化论坛 (Convener: 阎保平) (Location: 三楼神光厅)
- 14:00 聚变能研究与信息化 (30') 吴宜灿
- 14:30 高能物理实验BESIII的数据处理和分析 (30') 李卫东
- 15:00 高能物理与信息化 (15') 陈刚
- 15:15 茶歇



视频会议系统

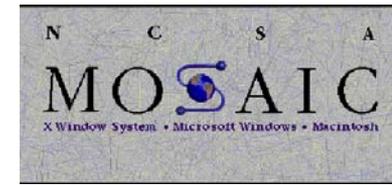
技术推广与转移

高能物理信息化形成的技术
几乎全部无偿开放给社会！

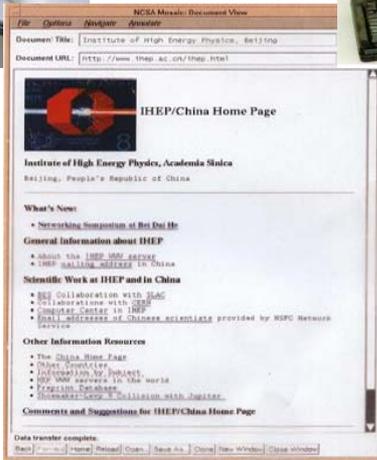




WWW



W3C

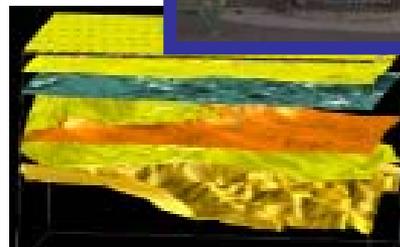


Google 谷歌

新浪微博

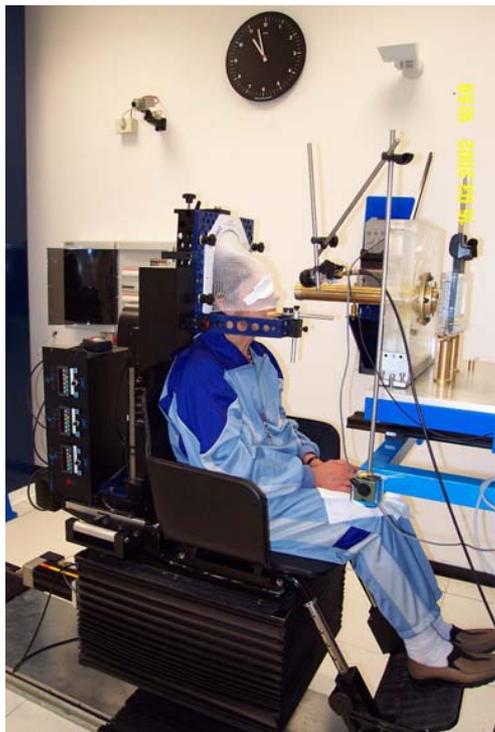
高能物理网格的应用

- 世界上最大最先进的网格平台
- 10大领域20多种应用
 - 高能物理 (引领性应用)
 - LHC 4个实验: **ATLAS, CMS, ALICE, LHCb**
 - 其他实验: 羊八井, 大亚湾, ...
 - 生物医学 (引领性应用)
 - 生物信息: **WISDOM, GPSA, CDSS, GridGRAMM, GDock**
 - 医学影像: **GATE, gPTM3D, SiMRI3D, ...**
 - 地球科学
 - 地球观测: **GOME**
 - 地球物理: **GeoCluster**
 - 气象学、水文学
 - 计算化学: **CHARON, CompChem**
 - 聚变: **ITER**
 - 天文学
 - 宇宙微波背景: **PLANK**
 - 伽马射线天文学: **MAGIC**
 - 工业应用
 -

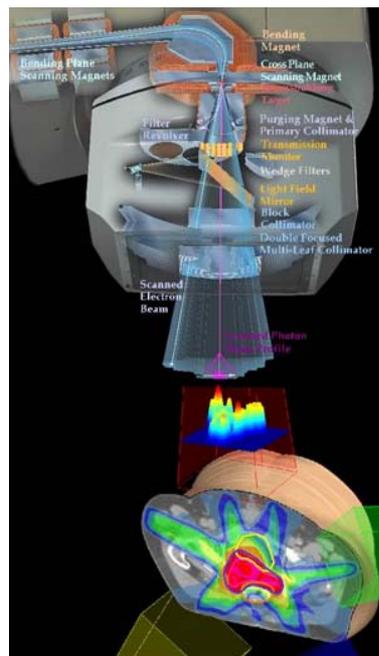


Geant4的推广应用

质子治疗



头部治疗的建模



人体解剖学结构重建



欧空局研究宇航员的辐射防护



小结

- 信息化是高能物理研究的必要条件和基础
— 没有信息化，就没有高能物理
- 高能物理是信息化发展的推动者
— 信息化推动高能物理发展，高能物理也推动了信息化的发展
- 信息化系统实现通用化和开放共享
— 避免重复建设，便于国际合作和交流



谢谢

