

高能物理计算环境

杜然，计算小组

中国科学院高能物理研究所计算中心

2017-Jun-05

大纲

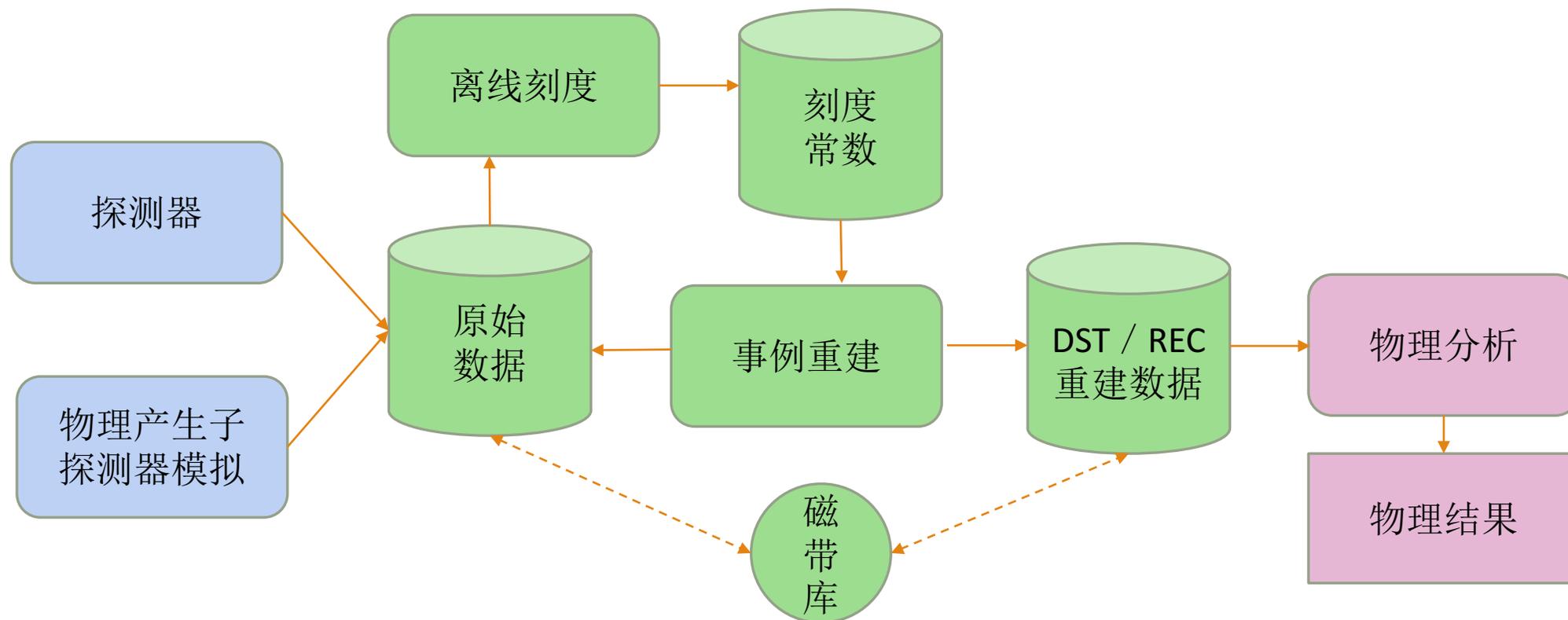
- 背景
- 集群计算
- 网格计算
- 云计算
- 分布式计算
- 志愿计算
- 总结

高能物理离线处理过程

○ 数据获取

○ 数据处理

○ 数据分析



大纲

- 背景
- 集群计算
- 网格计算
- 云计算
- 分布式计算
- 志愿计算
- 总结

计算集群

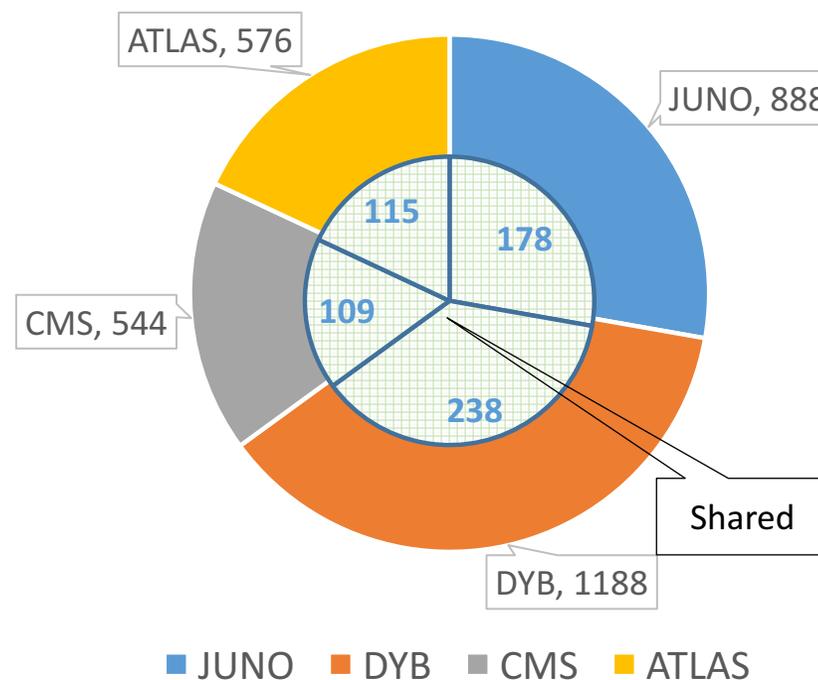
- 计算集群：把一组计算机通过高速网络连接在一起，构成一个整体，为用户提供计算服务。
- 利用计算集群进行数据处理是高能物理计算的主要手段。
- 计算集群中的最核心部件：资源管理和调度系统。
- 为满足高能物理实验不同需求，构建两种计算集群：
 - HTCondor集群
 - SLURM集群



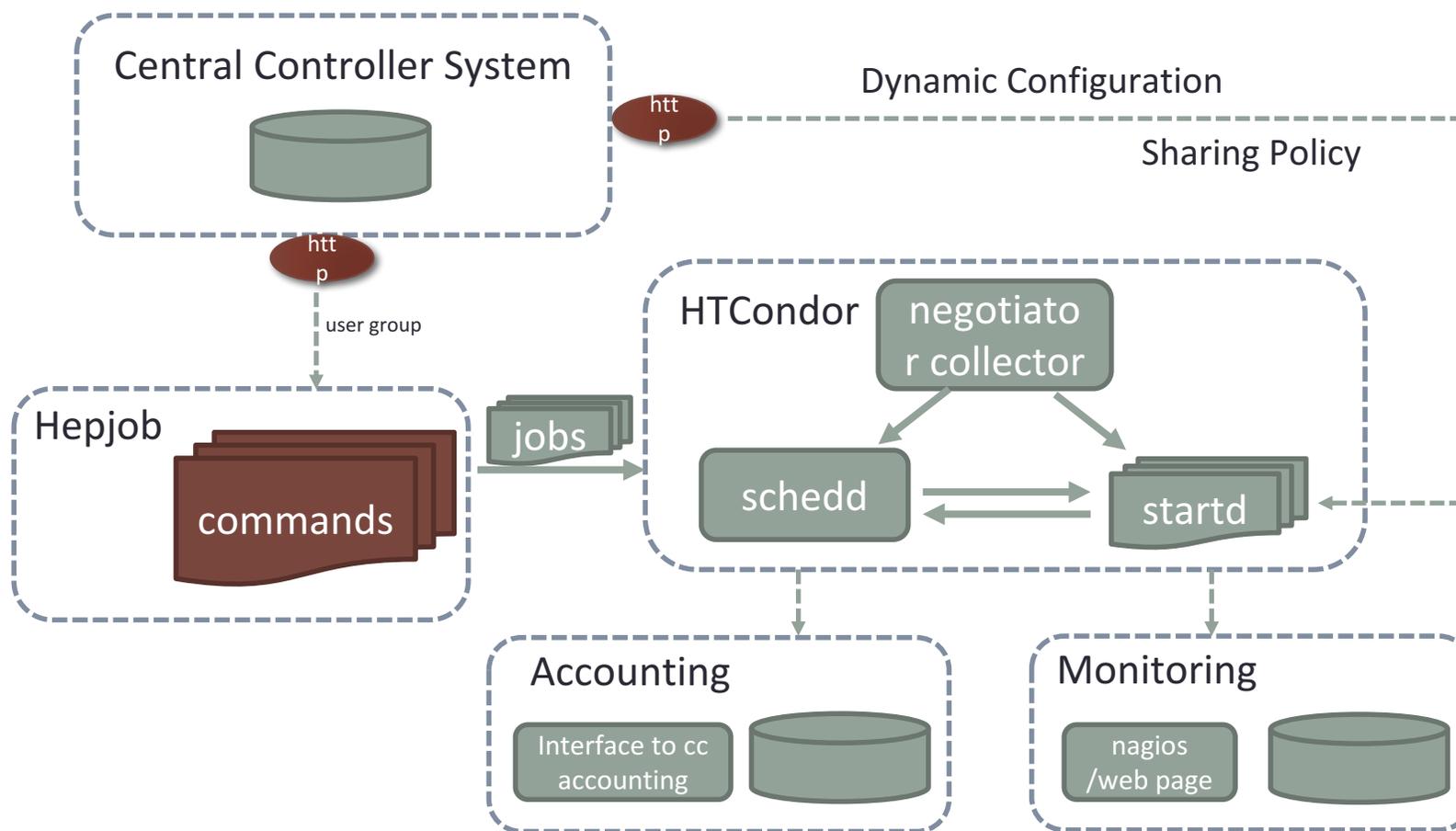
IHEP HTCondor集群

- 需求：
 - 多以单核、串行作业为主
 - 注重作业高吞吐量
 - 用户群复杂
 - 资源隔离
 - PBS: 可扩展性较差
- HTCondor:
 - 管理大型集群具备更好的性能
 - 活跃的社区支持
- 在HTCondor基础上，制定了组间共享策略，实现了资源共享

HTCondor Cluster Sharing Policy



IHEP HTCondor 集群架构



IHEP HTCondor 集群现状

➤ 硬件

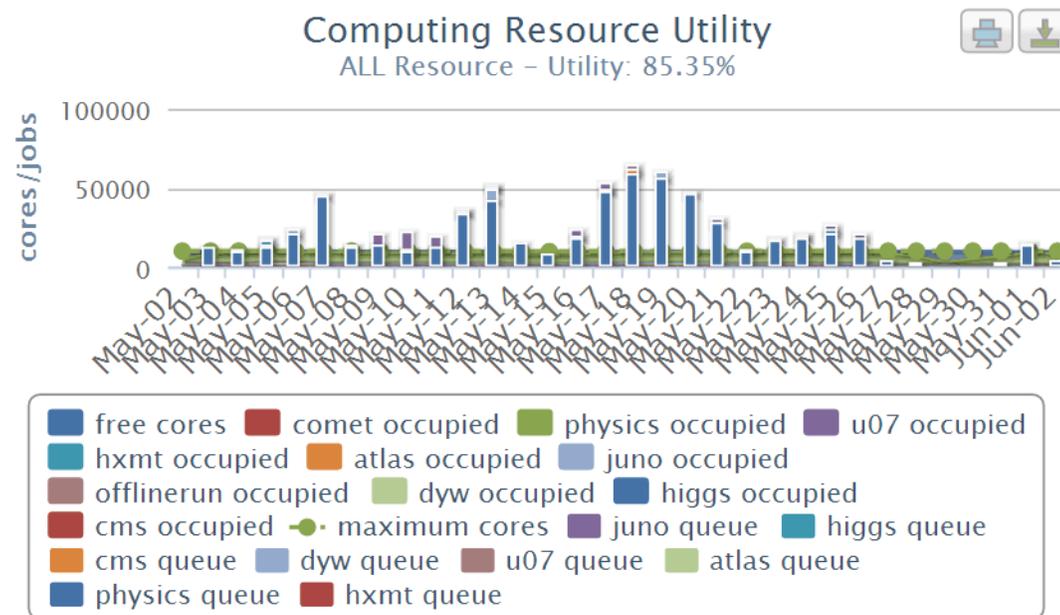
- 28个提交节点
- 3个队列服务器（local, virtual, MPI）
- 3个中央管理服务器（local, virtual, MPI）

➤ 资源

- ~11,000 CPU cores, 以及部分动态虚拟核
- 利用率提升, 由50%左右提升至90%左右
- 支持BES、JUNO、DYW、LHAASO、CMS等实验

➤ 作业

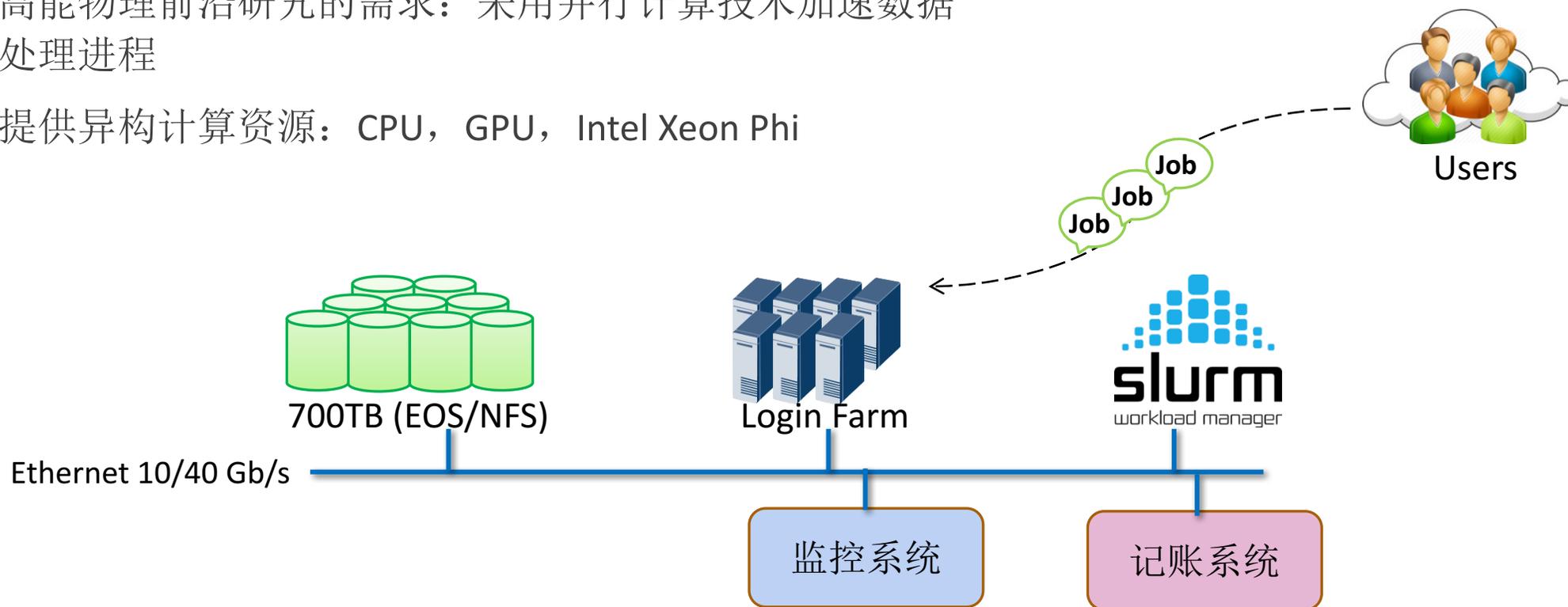
- 150,000 作业 / 天
- 213,642 CPU时 / 天



IHEP SLURM集群

需求

- 高能物理前沿研究的需求：采用并行计算技术加速数据处理进程
- 提供异构计算资源：CPU，GPU，Intel Xeon Phi



IHEP SLURM集群现状

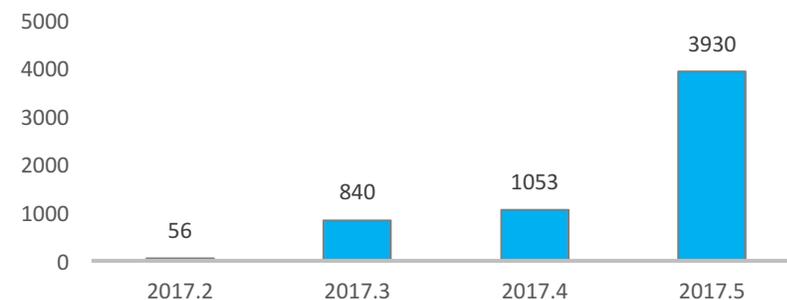
➤ 资源

- 1个主控节点，1个记账节点，16个登录节点，70个计算节点
- ~1,500 CPU cores, 8 GPU cards

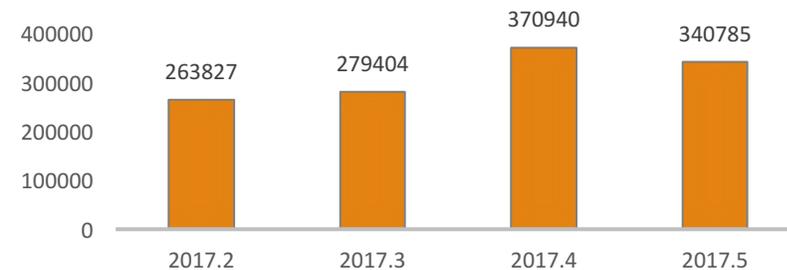
➤ 作业（2017.2~2017.5）

- 总共运行作业个数：5,879
- 运行总CPU时：1,254,956

SLURM集群每月作业运行个数



SLURM集群每月运行CPU时



大纲

- 背景
- 集群计算
- 网格计算
- 云计算
- 分布式计算
- 志愿计算
- 总结

网格计算

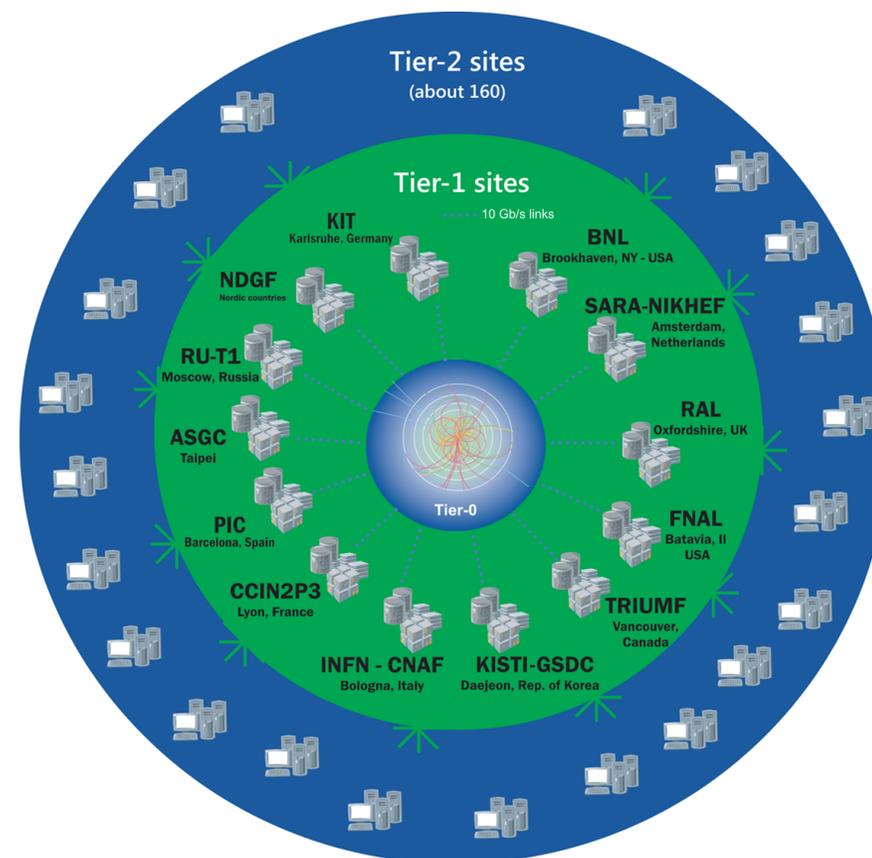
➤ 网格计算

- 高能物理实验大数据时代：单一的计算集群已经远远不能满足需求
- 网格计算应需而生：将全球地理上分散的计算资源有机的整合起来协同工作，能够完成单个集群无法完成的大规模计算任务，是“集群的集群”。

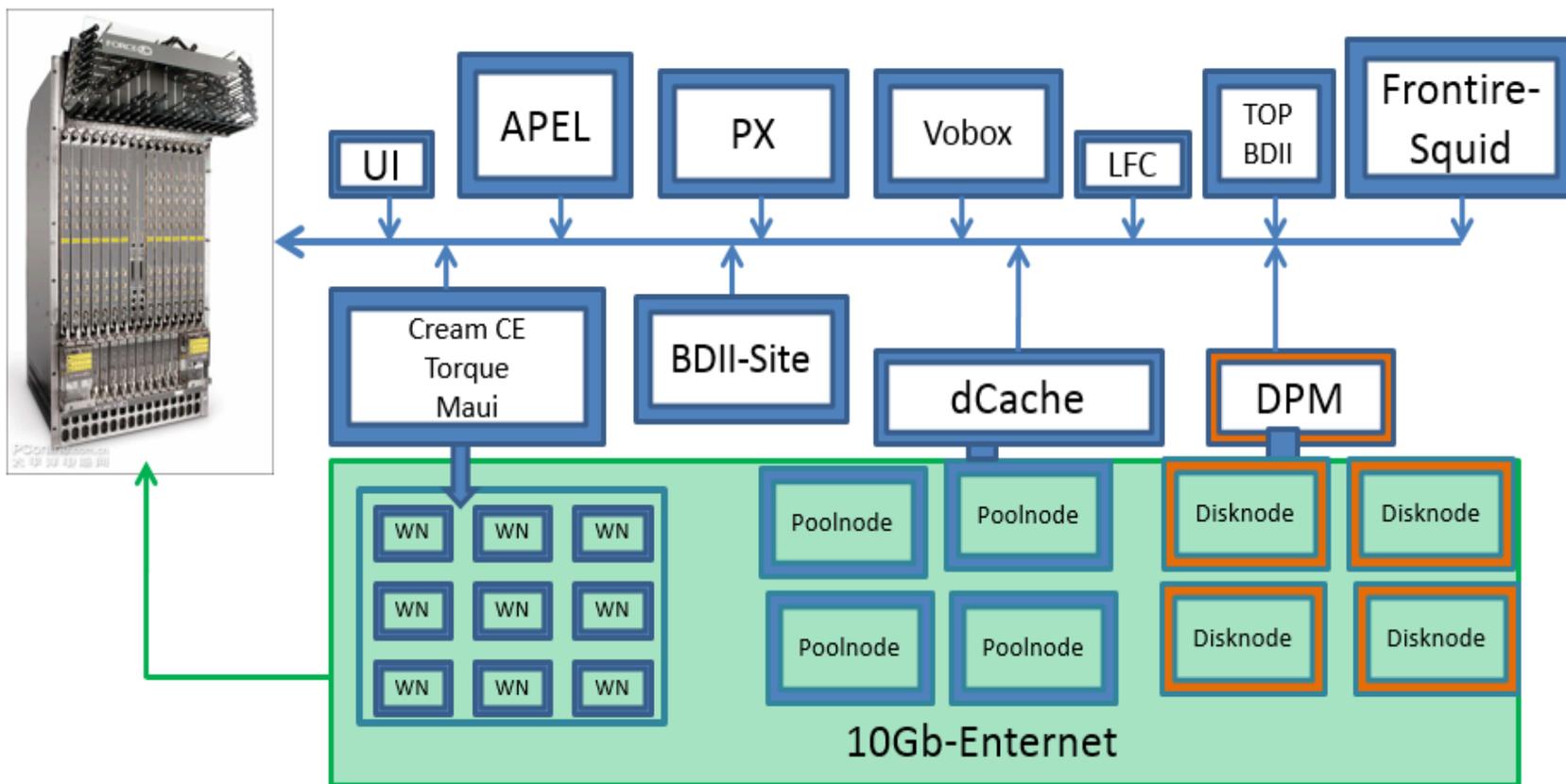
➤ 影响最大的网格：由CERN主导、多国参与的WLCG网格计算系统

➤ 高能所是WLCG网格二级站点

- 站点名称：BEIJING-LCG Tier 2
- 支持实验：CMS、ATLAS



WLCG网格计算系统架构



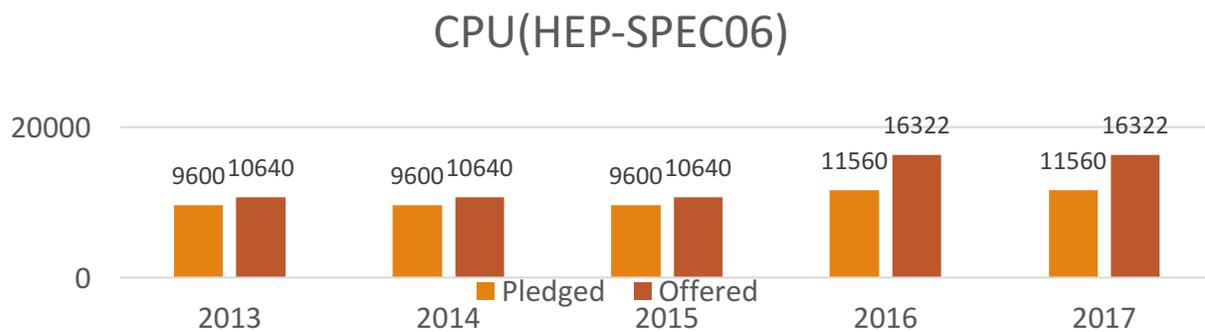
BEIJING-LCG Tier2现状

➤ 计算资源

- CPU: 888 Cores
 - Intel E2680V3: 696 Cores
 - Intel X5650 192 Cores
- 计算能力比合约高出约41% (2017年)

➤ 调度系统: Torque

- Beijing-LCG Tier 2站点
- 100% 可靠性 (2017.5)
- 共78个二级站点



大纲

- 背景
- 集群计算
- 网格计算
- 云计算
- 分布式计算
- 志愿计算
- 总结

云计算环境

➤ 云计算技术:

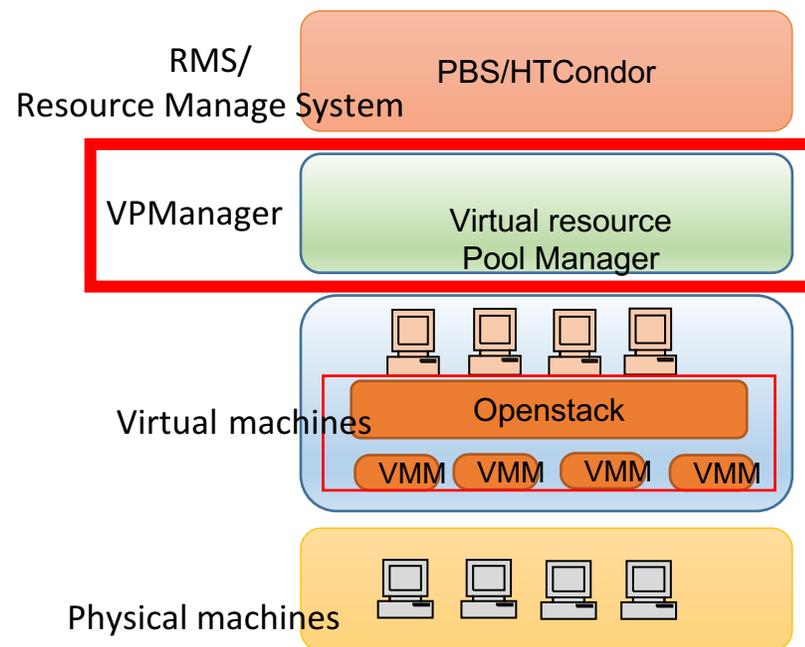
- 通过对所有资源进行整合、抽象后以新的业务模式提供高性能、低成本的持续计算
- 特征: 资源池化、弹性可伸缩、按需自助服务
- 优点: 灵活性、可靠性、可扩展性

➤ 个人云计算需求

- 用户临时性、多样性的测试环境需求
- 提供物理机维护成本较高
- 提高资源利用率

➤ 虚拟计算集群需求

- 多样性计算环境
- 按需弹性提供计算资源
- 提高资源利用率



IHEP 个人云计算 & 虚拟计算集群

➤ 个人云计算服务

- 按需申请个人测试机，拥有root权限
- 在线自助申请，无需审批
- <http://cloud.ihep.ac.cn>



HEP Cloud
Powered by OpenStack

登录

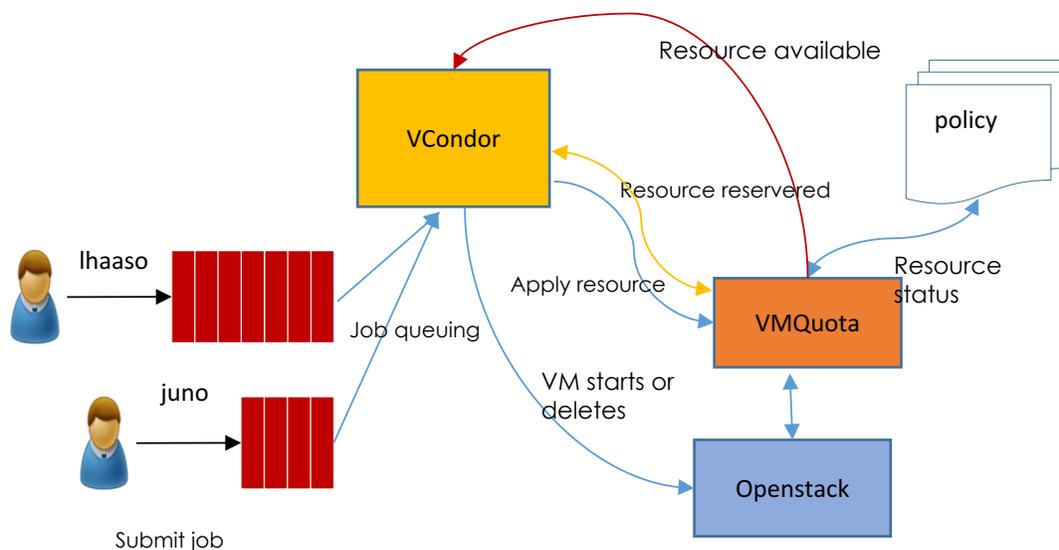
用户名

密码

帮助 登入 高能所统一认证登录

➤ 虚拟计算集群

- 共享不同实验/组织的计算资源，实现资源整合
- 按需分配资源，提高资源利用率，满足峰值需求
- 支持LHAASO、JUNO实验



IHEP 个人云计算&虚拟计算集群现状

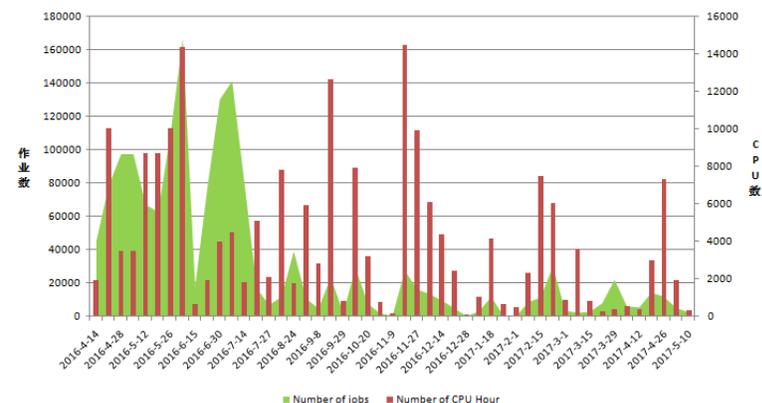
➤ 个人云计算

- 352个虚拟cpu cores
- 205个用户，运行235个虚拟机，约330个虚拟机核，资源使用率93.75%

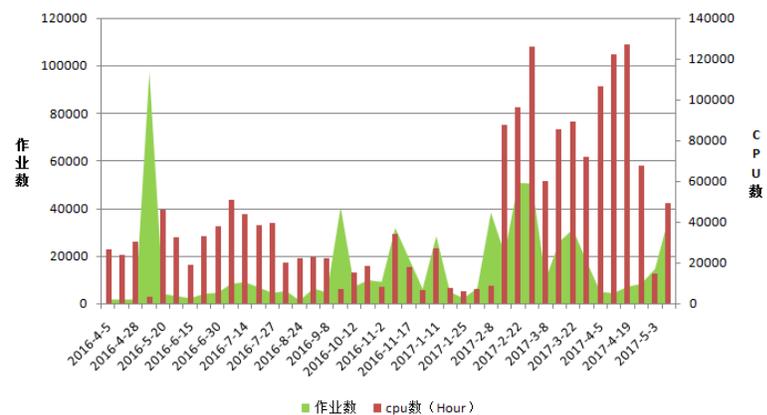
➤ 虚拟计算集群

- 1041个虚拟CPU cores
- LHAASO已运行65万个作业，累计177万cpu时
- JUNO已运行154万个作业，累计20万cpu时

JUNO 作业统计 (2016.4-2017.5)



LHAASO作业统计 (2016.4-2017.5)

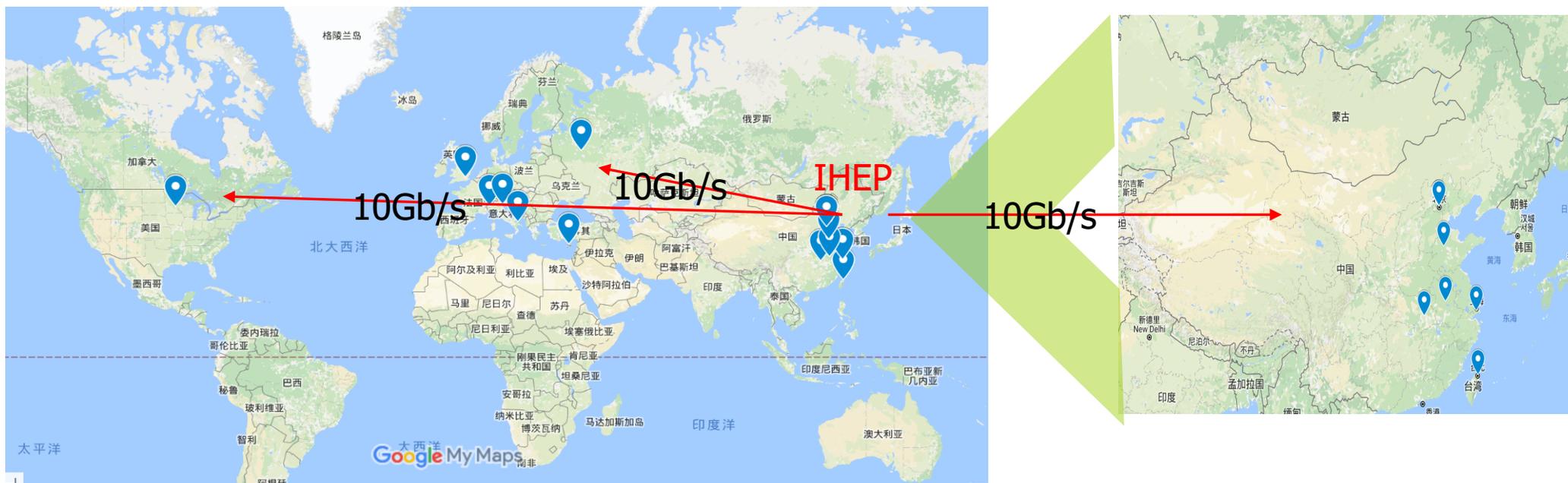


大纲

- 背景
- 网格计算
- 集群计算
- 云计算
- 分布式计算
- 志愿计算
- 总结

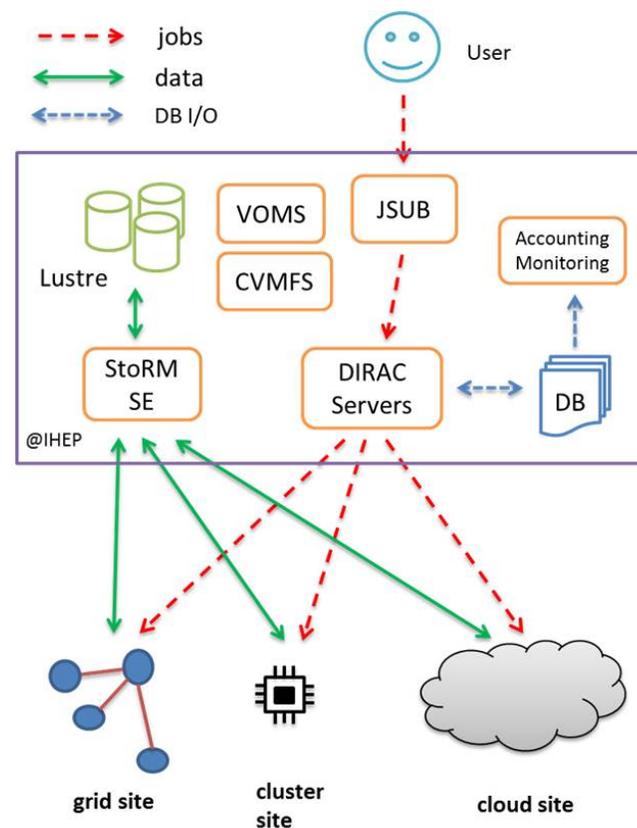
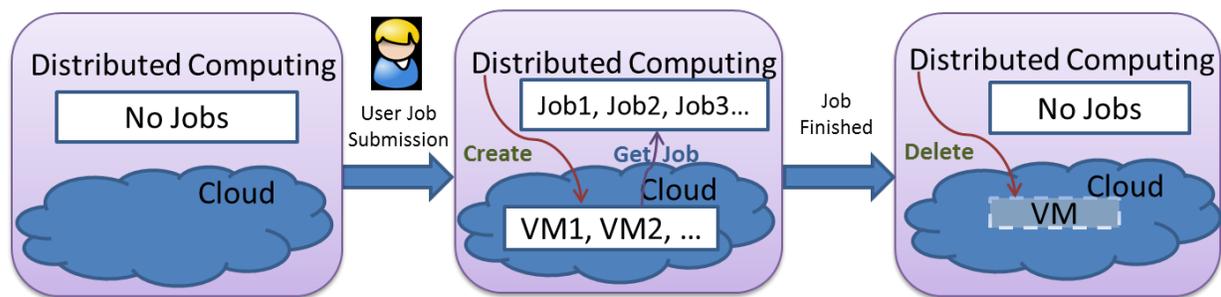
分布式计算需求

- 解决一些实验（如CEPC）本地计算资源不足的问题
- 一些突发的资源需求
- 利用合作单位的资源，提高资源利用率
- 对异构资源的整合（集群、网格和云）



分布式计算架构

- 基于DIRAC实现
- 整合了网格、集群和云
- 支持实验： BESIII、 CEPC、 JUNO
- 为用户提供统一的平台利用异地站点资源



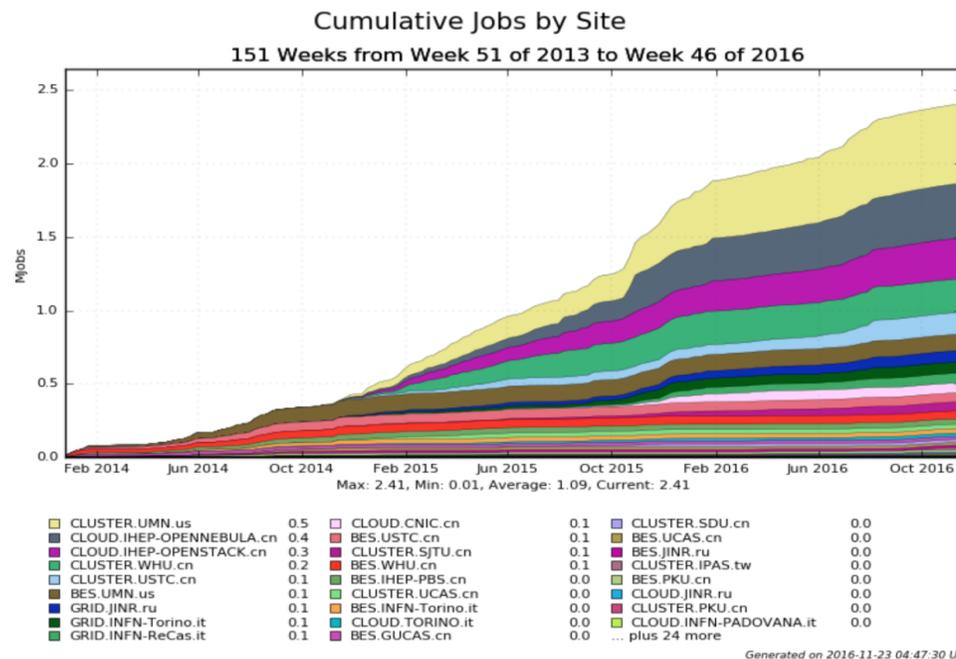
分布式计算现状

资源

- 资源量超过3500核
- 15个来自国内外合作单位加入

作业

- CEPC: 探测器设计的模拟, 总事例数~9600 万
- BESIII: 完成了近30亿事例的官方MC产生
- JUNO: 主要用于现阶段的探测器设计和重建算法的研究



大纲

- 背景
- 集群计算
- 网格计算
- 云计算
- 分布式计算
- 志愿计算
- 总结

志愿计算需求&架构

➤需求

- 高能物理实验资源短缺
- 志愿计算资源零散，但整合后很可观
- 适合运行周期长、CPU密集型模拟作业

➤架构

- 基于BOINC中间件
- BOINC服务器端分发作业并收集结果
- BOINC客户端运行作业，并可以自主设定资源使用参数



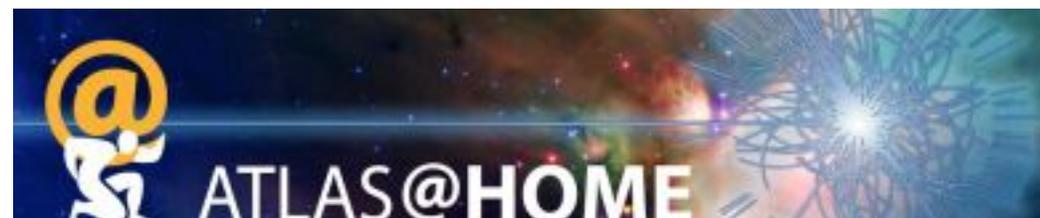
志愿计算环境现状

➤ 计算项目:

- CAS@HOME: 动态分子模拟, 蛋白质结构预测
- ATLAS@HOME: ATLAS作业

➤ 计算规模:

- CAS@HOME: 2000台主机, 约4TeraFLOPS的计算能力, 一百万CPU时/月免费计算资源
- ATLAS@HOME: 运行4年, 5000万CPU时/年, 相当于1万CPU核专用网格站点的计算能力



大纲

- 背景
- 网格计算
- 集群计算
- 云计算
- 志愿计算
- 分布式计算
- 总结

总结

- 不同计算技术满足不同高能物理计算需求
- 根据高能所计算资源现状与实验具体需求，计算中心不断调整与优化相应计算环境

谢谢！