

基于Hadoop的LHAASO数据分析平台

李强, 黄秋兰, 孙功星
计算中心/高能所
2017.1.18



提纲

- 计算需求与挑战
- Hadoop平台介绍
- 平台运行情况
- 下一步工作计划



计算需求与挑战

- 未来实验规模越来越大，数据量越来越多，对传统的计算集群提出了新的挑战。
- 随着数据量的增加，需要更多的资金购买昂贵的网络设备和存储设备。
- 在传统集群中，网络IO已成为数据密集型作业的瓶颈。



Hadoop平台介绍

■ Apache Hadoop

- 一个能够对大量数据进行分布式处理的开源软件框架，被设计部署在廉价的服务器上，具有高可靠性、高扩展性、高效性、高容错性和低成本等特点。
- ✓ 一个适合海量数据存储的分布式文件系统—HDFS。
- ✓ 一个适合大数据集并行处理的计算框架—MapReduce。
- ✓ 有众多相关工具：Hbase, Hive, Pig, Flume等。
- ✓ 互联网大数据处理的标准，被众多公司使用。



Hadoop优势

- 节省系统造价

 - 不需要昂贵的网络设备和磁盘阵列

- 扩展性好

 - 可以很容易的扩展到上千、上万节点

- 提升性能

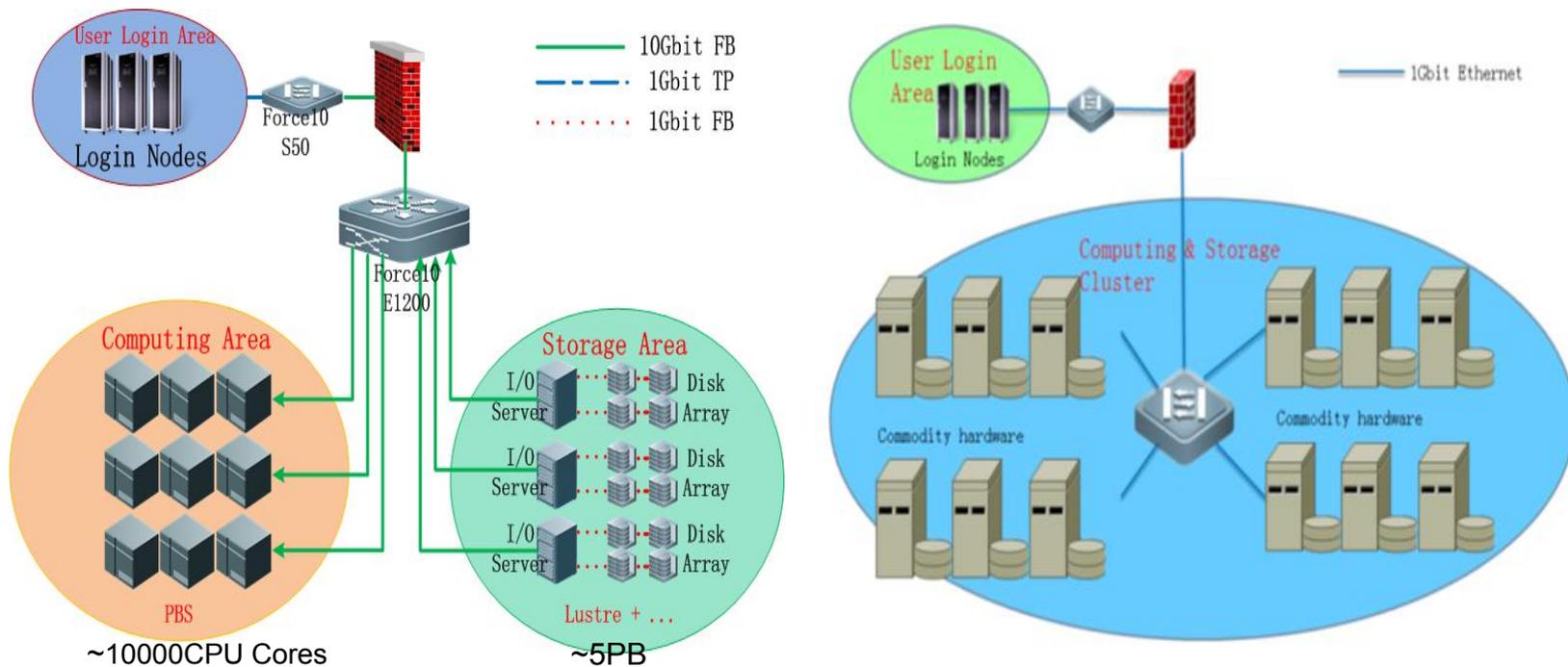
 - 本地数据读写，无网络延迟。

 - 不受网络限制，可以充分利用本地磁盘的IO性能。

- 应用广泛

 - 互联网领域广泛应用，众多相关工具可以使用。

当前计算集群架构 vs Hadoop架构



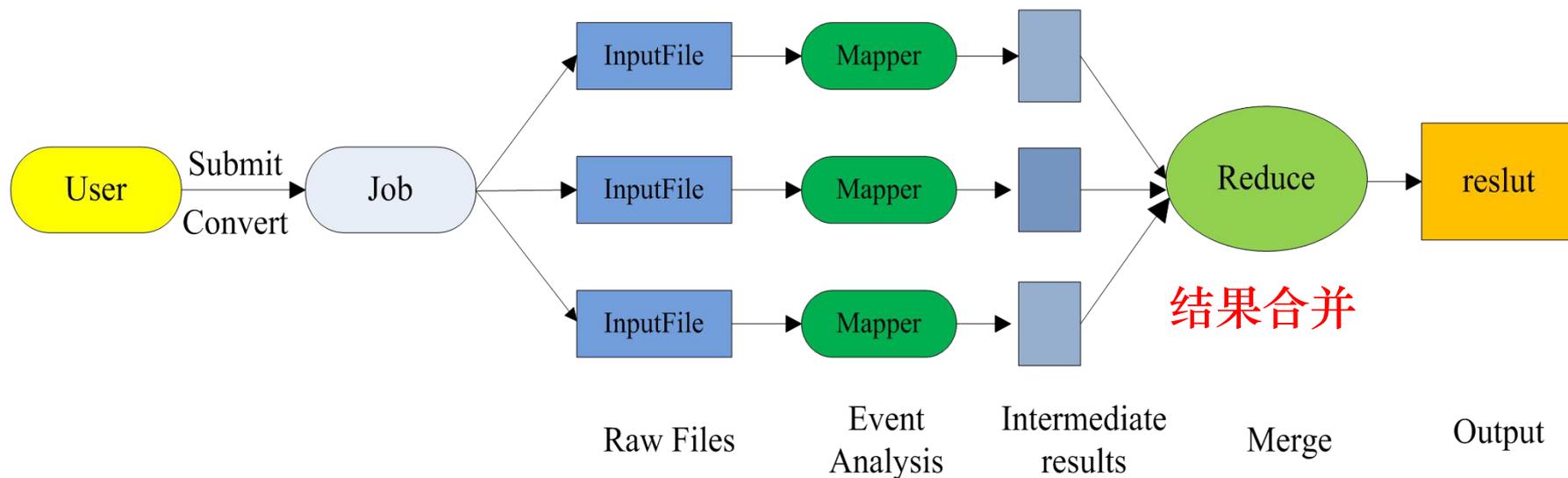
	当前计算集群	Hadoop集群
网络	万兆网络	千兆网络
存储	磁盘阵列	本地磁盘
数据读写	网络传输，受网络限制	本地磁盘读写



Hadoop作业执行

- Hadoop作业的执行流程如下：

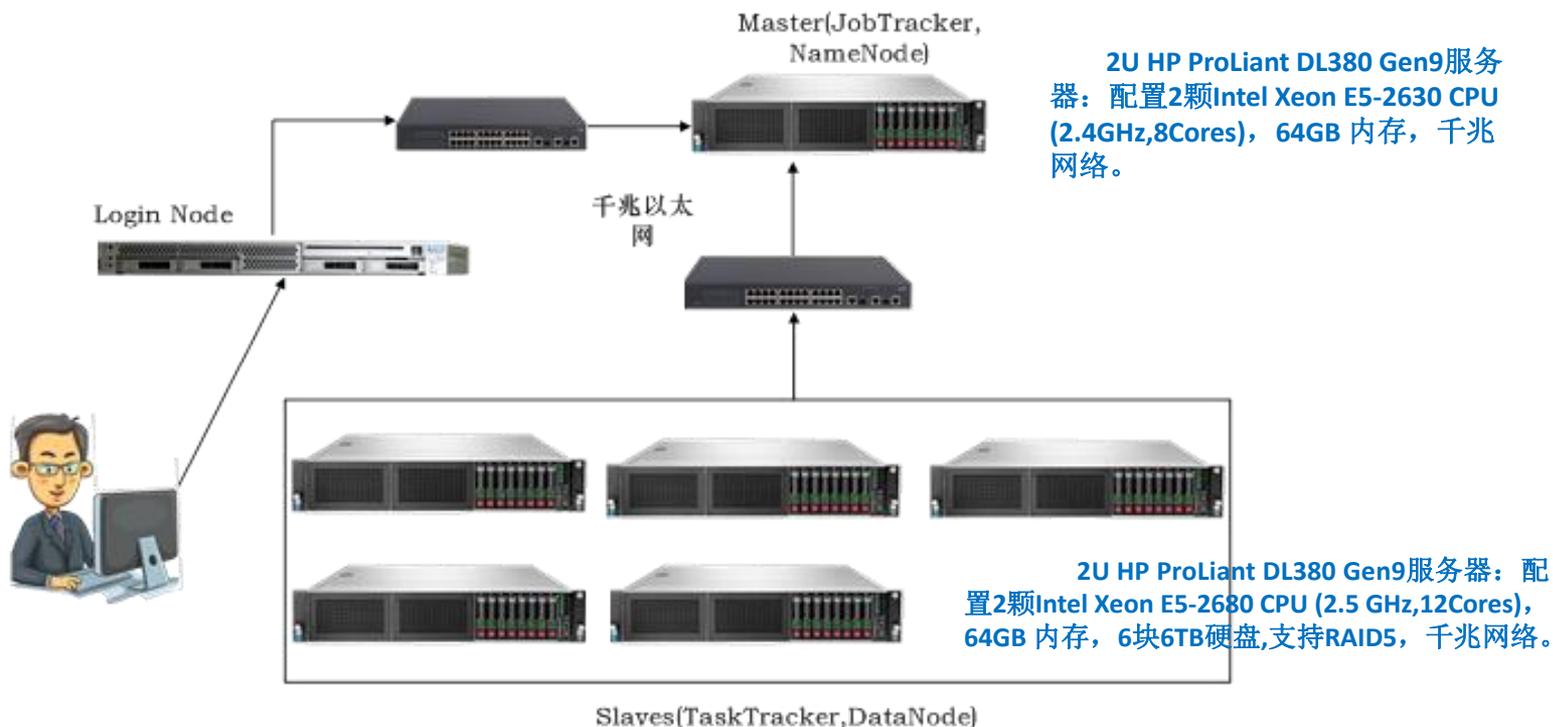
作业自动拆分，并行处理





平台建设情况

目前hadoop集群由1个登陆节点，1个主节点和5个从节点组成，网络配置为千兆以太网。可用计算资源为**120 CPU核**，**140TB**存储空间。目前支持**宇宙线模拟 (corsika)**、**ARGO探测器模拟 (Geant4)** 以及**数据分析 (KM2A)** ,有8个用户正在使用，活跃用户3个。





平台运行情况

■ 存储资源使用情况

- 已存放140万个文件，使用50TB存储空间（36%）

■ 作业运行情况（11月~12月）

- 提交作业数2513（任务数23191），失败作业365
- 使用66875个CPU hours，12月资源使用率63.7%

■ 作业失败原因

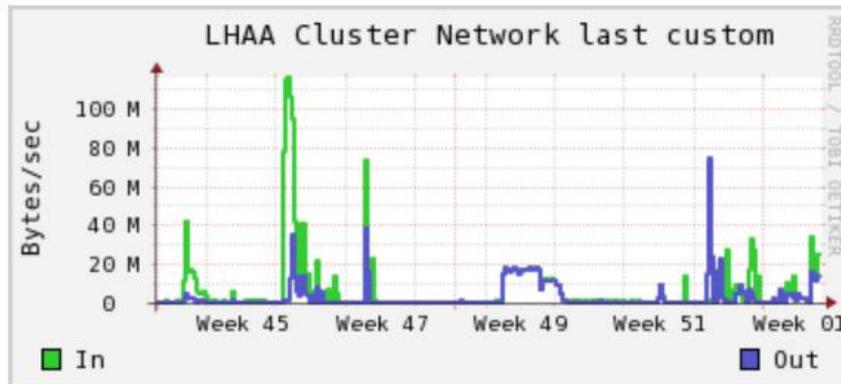
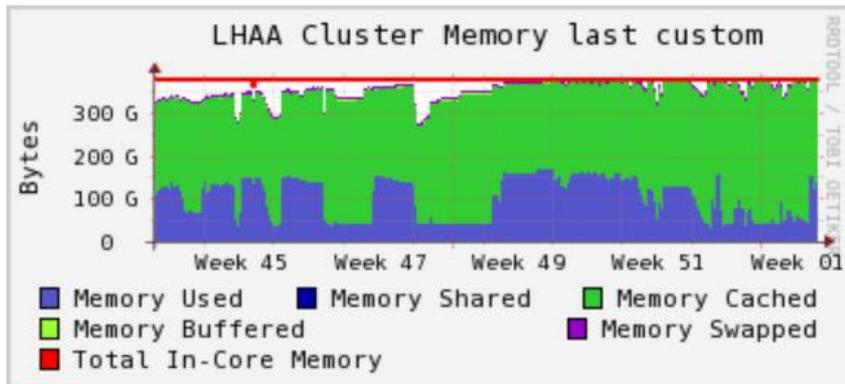
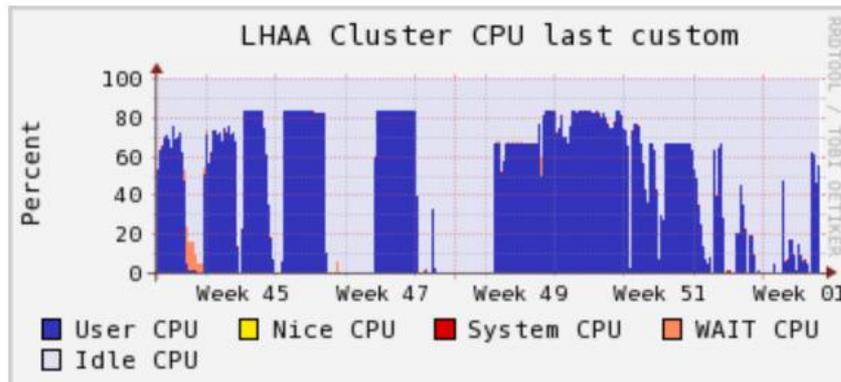
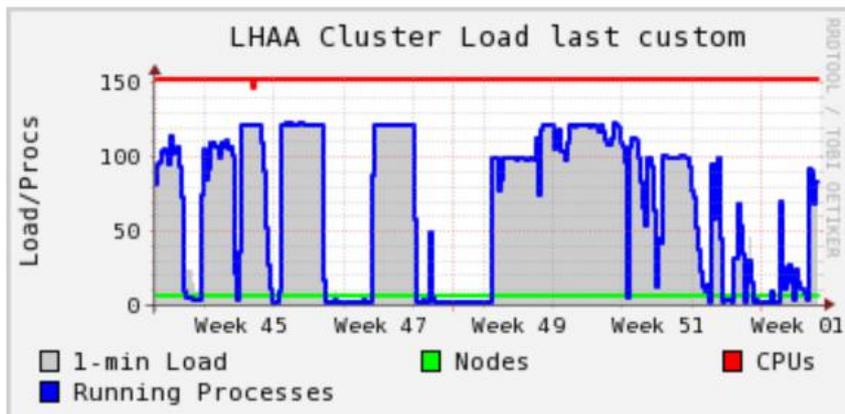
- 作业脚本格式问题
- kill作业
- 操作失误



集群监控 (ganglia)

Overview of LHAA

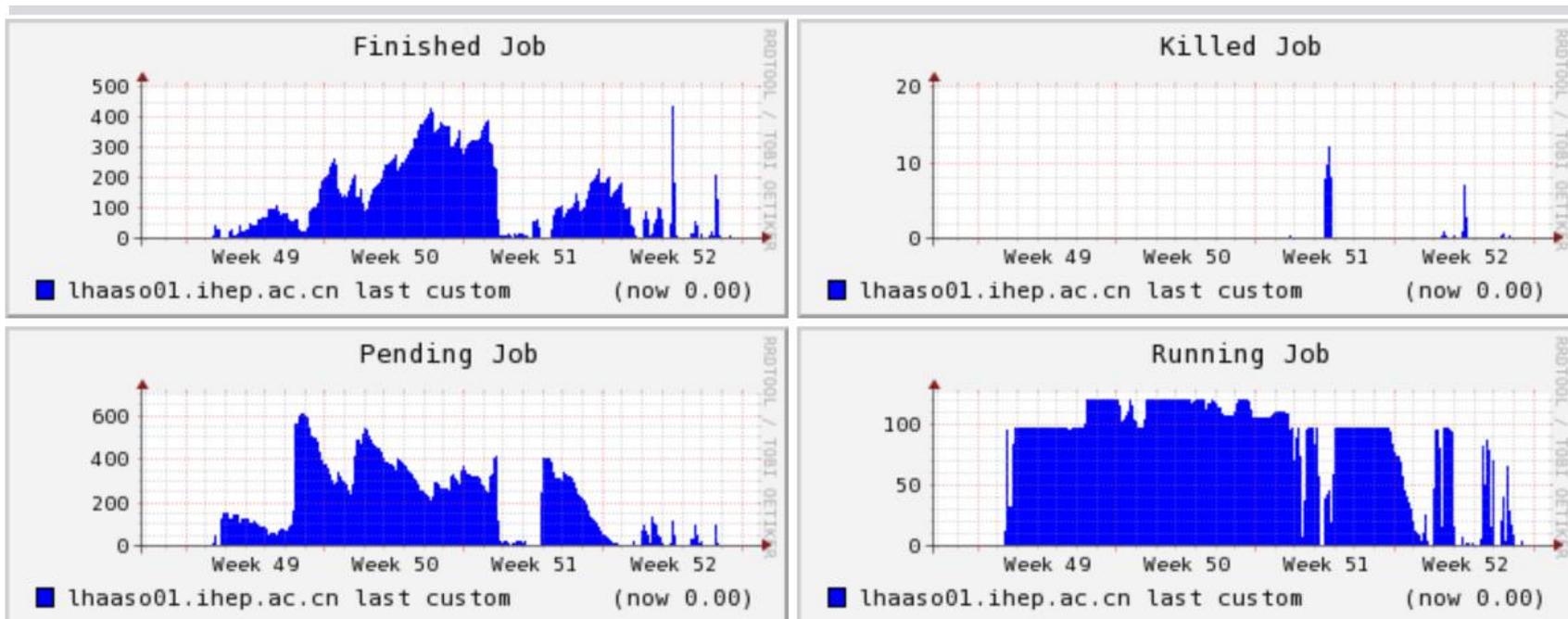
2016.10.10--2017.1.7





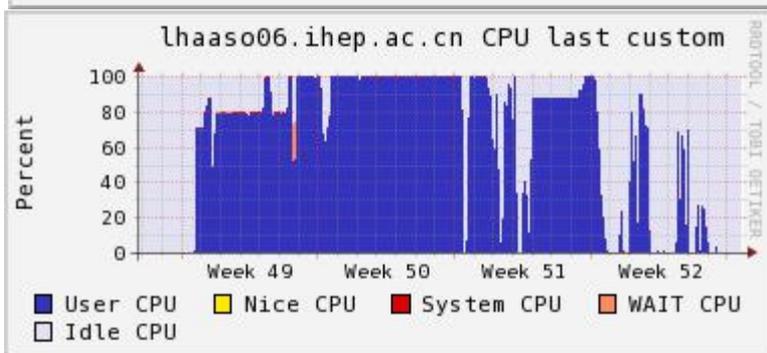
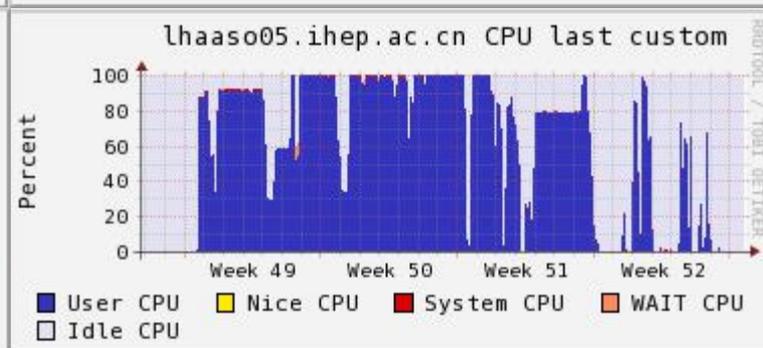
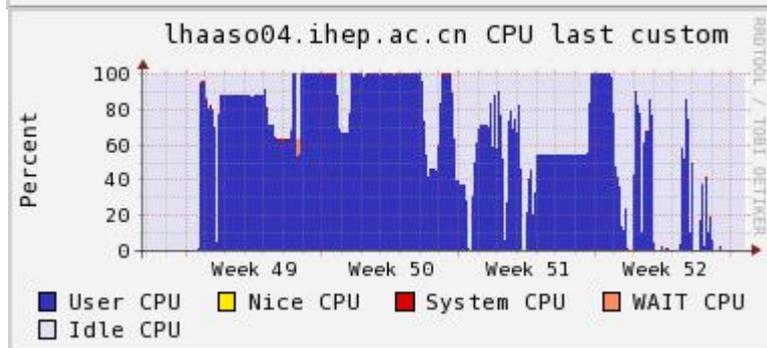
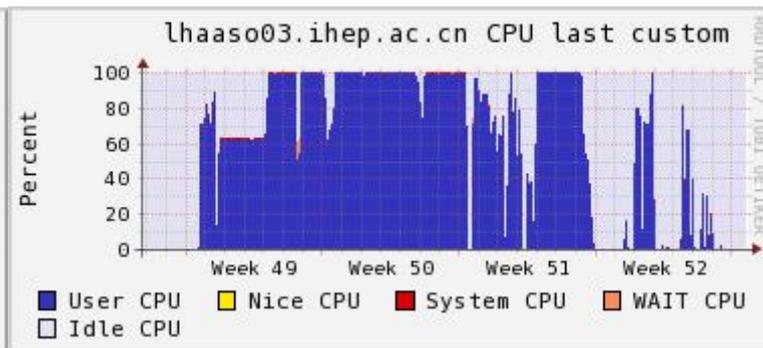
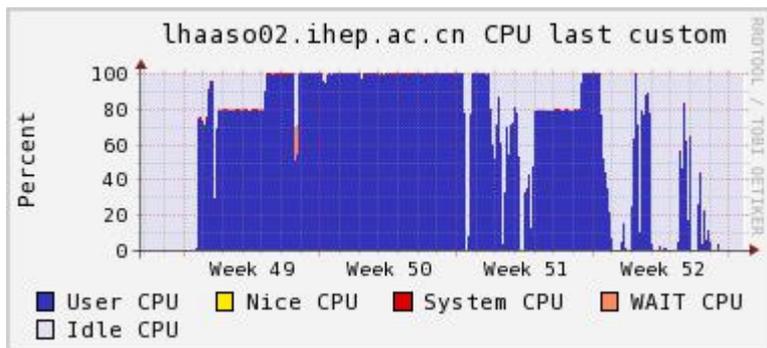
作业执行情况实时监控

(2016年12月)





各个计算节点CPU监控



(2016年12月)



下一步工作计划

■ HDFS负载均衡研究

为了充分利用本地磁盘的IO资源，实现了作业完全本地化执行。在数据分布不均衡的情况下，导致集群负载不均衡。因此，我们根据高能物理数据处理的特点，实现了基于目录的数据负载均衡策略。

节点	均衡前		均衡后	
	文件大小	百分比	文件大小	百分比
1	21	11.3	34.8	18.7
2	56	30	39.6	21.3
3	38	20.4	38	20.4
4	43	23.2	38.5	20.7
5	28	15.1	35.1	18.9

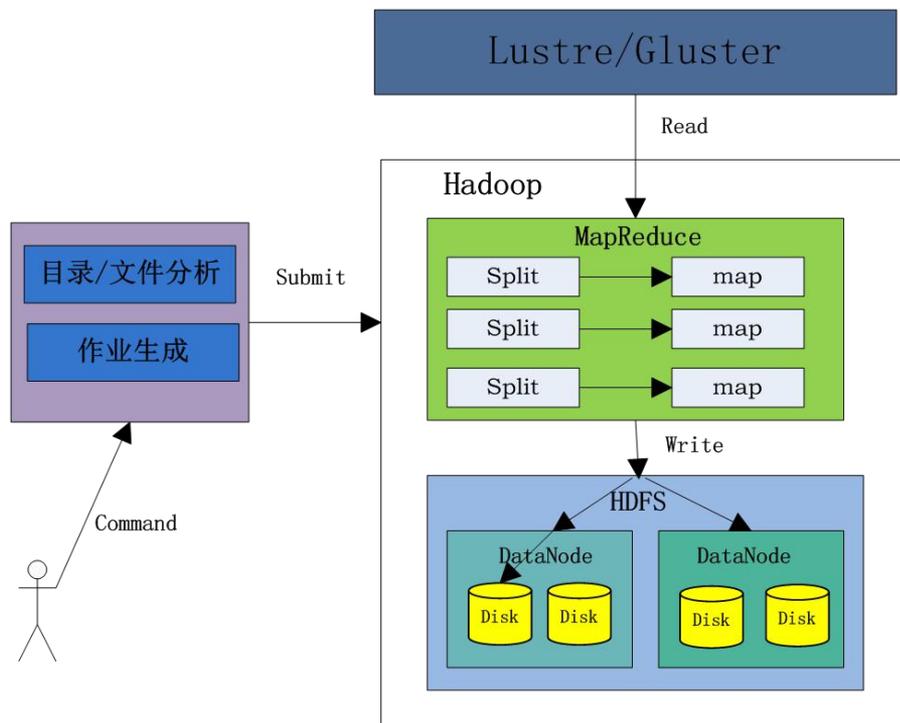


下一步工作计划

■ 数据传输系统

- 已实现基于MapReduce的数据导入系统

针对数据在lustre (/workfs)、Gluster (/ybjgfs) 上的数据。



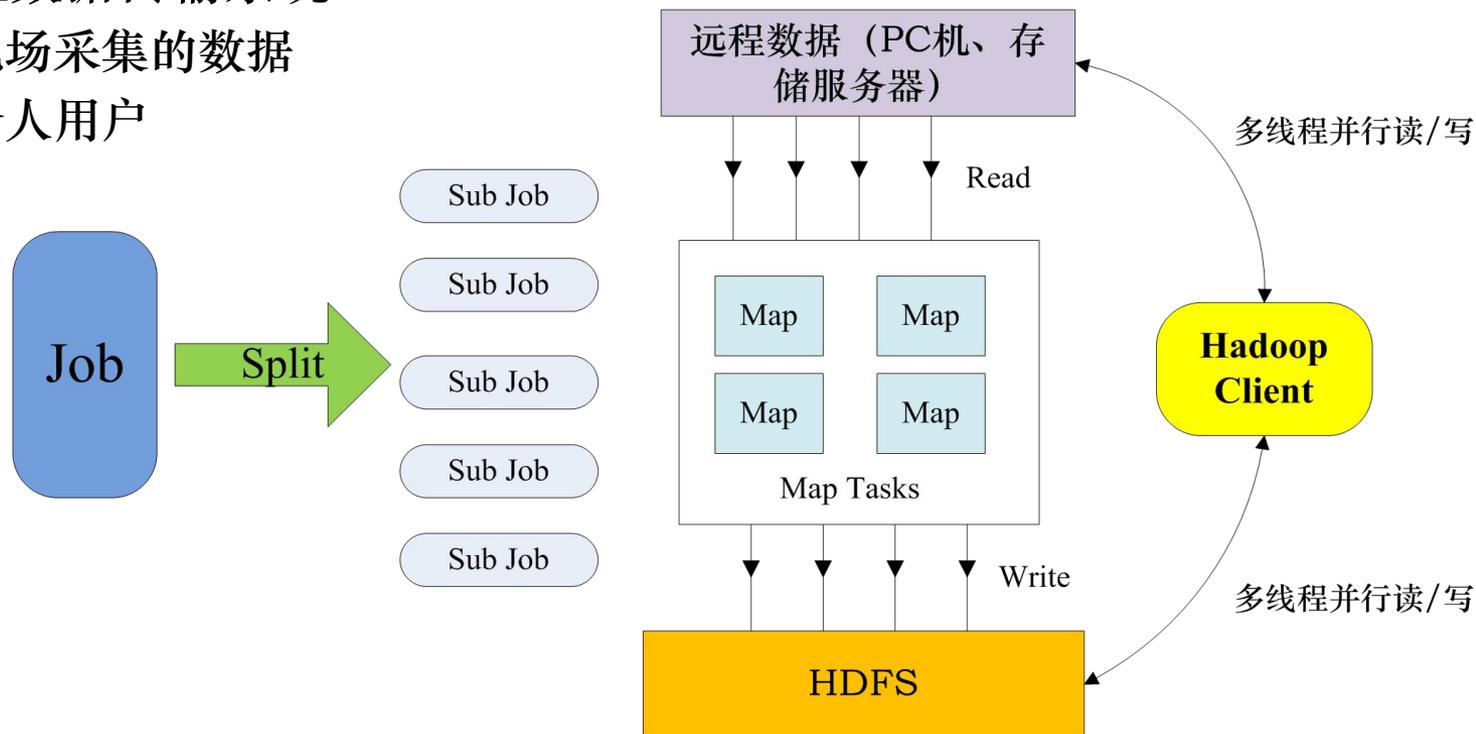


下一步工作计划

■ 数据传输系统

• 远程数据传输系统

- 现场采集的数据
- 个人用户





下一步工作计划

■ 作业提交优化

规范作业脚本格式，作业提交方式。

■ 作业调度管理研究

-合理划分集群资源（CPU、存储资源）

-合理配置用户使用额度（CPU、存储资源）

-作业优先级设置



小结

- 计算平台已面向用户提供服务
- 基本作业类型都可以正常运行
- 已提供集群监控及记账
- 一些细节需要进一步完善
- 新的功能模块正在研究与开发
- 希望有更多的用户使用和提供意见，帮助我们完善系统

联系我们：

黄秋兰： huangql@ihep.ac.cn

李强： liqiang88@ihep.ac.cn



谢谢!