


# 基于dHTC的分布式计算

## - 在高能物理的应用



姜晓巍 (jiangxw@ihep.ac.cn)

高能所计算中心

2023开放科学计算联盟学术年会

2023/06/27

# OUTLINE

- ❖ 高能物理计算背景
- ❖ 基于dHTC的分布式计算资源共享方案
  - Factory和调度策略
  - Glidein Job
  - Tokens认证
- ❖ 站点接入情况：本地HPC、东莞区域中心、CENI算力网络&超算等
- ❖ 作业接口与作业运行环境
- ❖ dHTC分布式计算在高能物理的应用

# 背景 - 高能物理实验

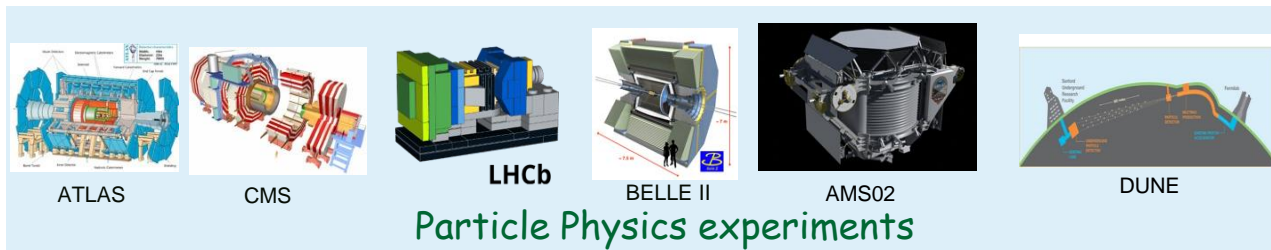
## ❖ 高能物理实验及相关应用:

- 实验粒子物理学
- 理论粒子物理学
- 天体物理学和宇宙射线
- 加速器技术和应用
- 同步辐射和应用
- 核分析技术
- 计算和网络应用

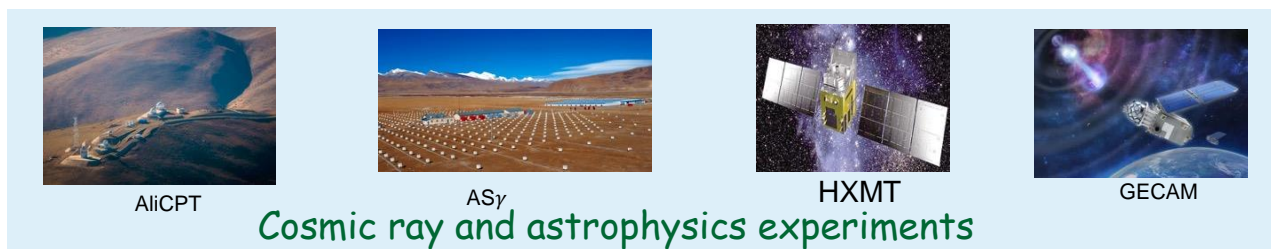
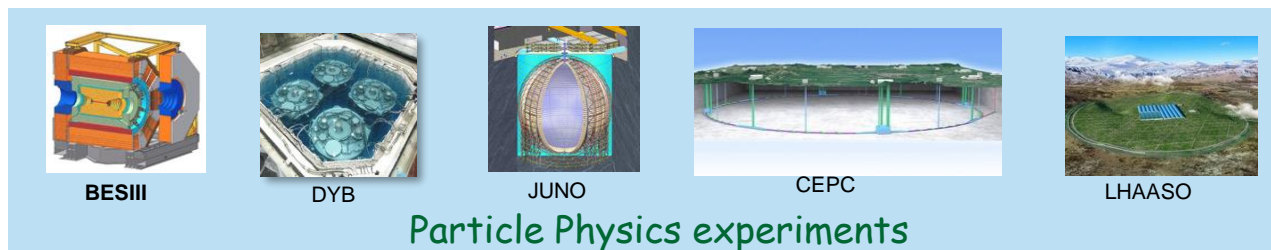
## ❖ 计算模式及特点

- 高通量计算: 日均百万计算任务
- 高性能计算: 并行&GPU等异构
- 分布式计算: 资源整合及共享

International collaboration

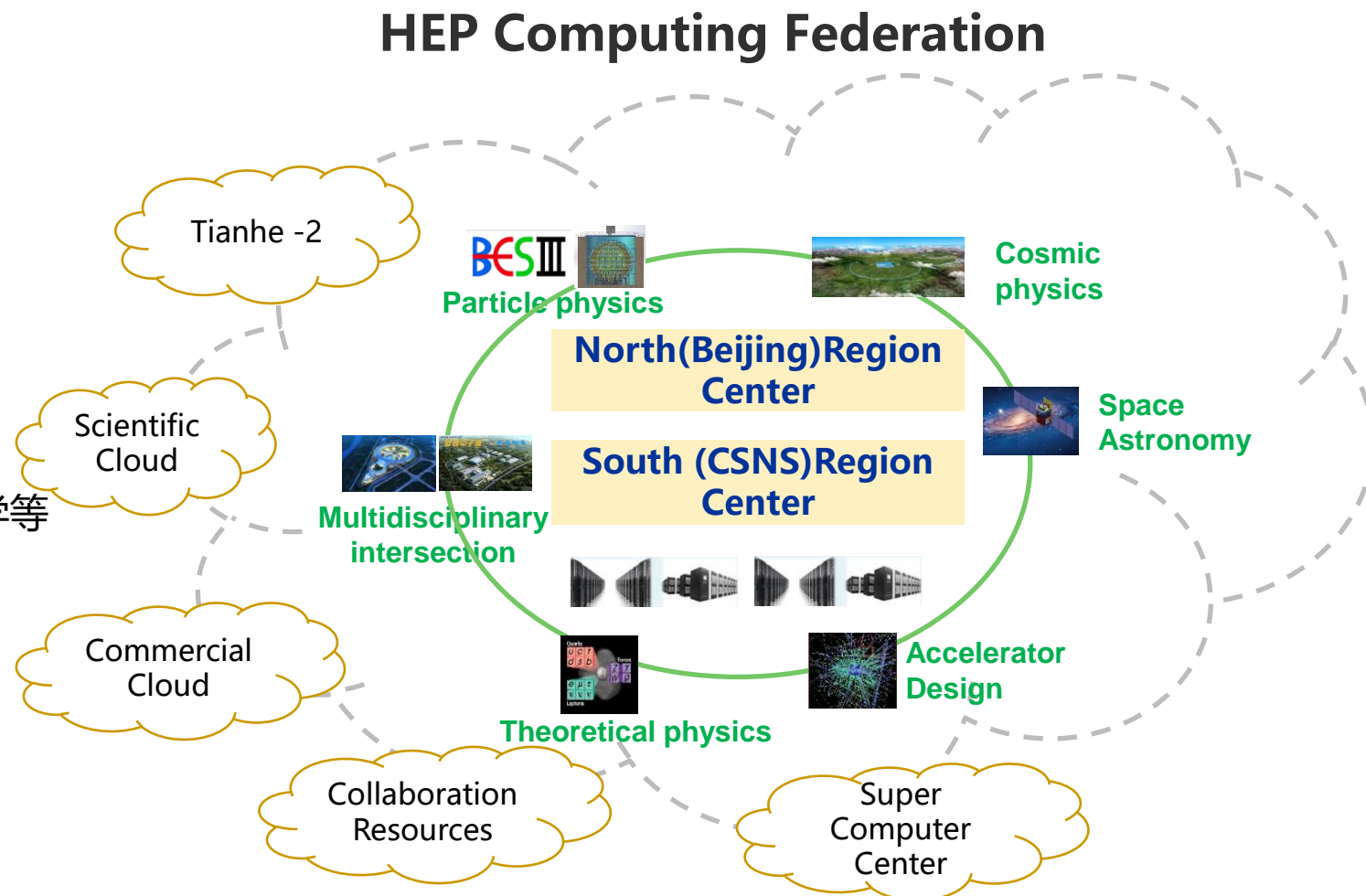


IHEP Leading



# 背景 - 高能物理计算联盟

- ❖ 跨区域的统一数据处理平台
  - 南北两个区域中心
- ❖ 整合多远程计算站点，兼容异构硬件
  - 合作组成员提供资源
  - 商业云
  - 超级计算中心等
- ❖ 平台支持多个学科的科学研究的
  - 粒子物理、理论物理、天体物理、光子科学等
  - 服务数千名物理学家
- ❖ 两种计算模型
  - 本地集群计算
  - 分布式计算



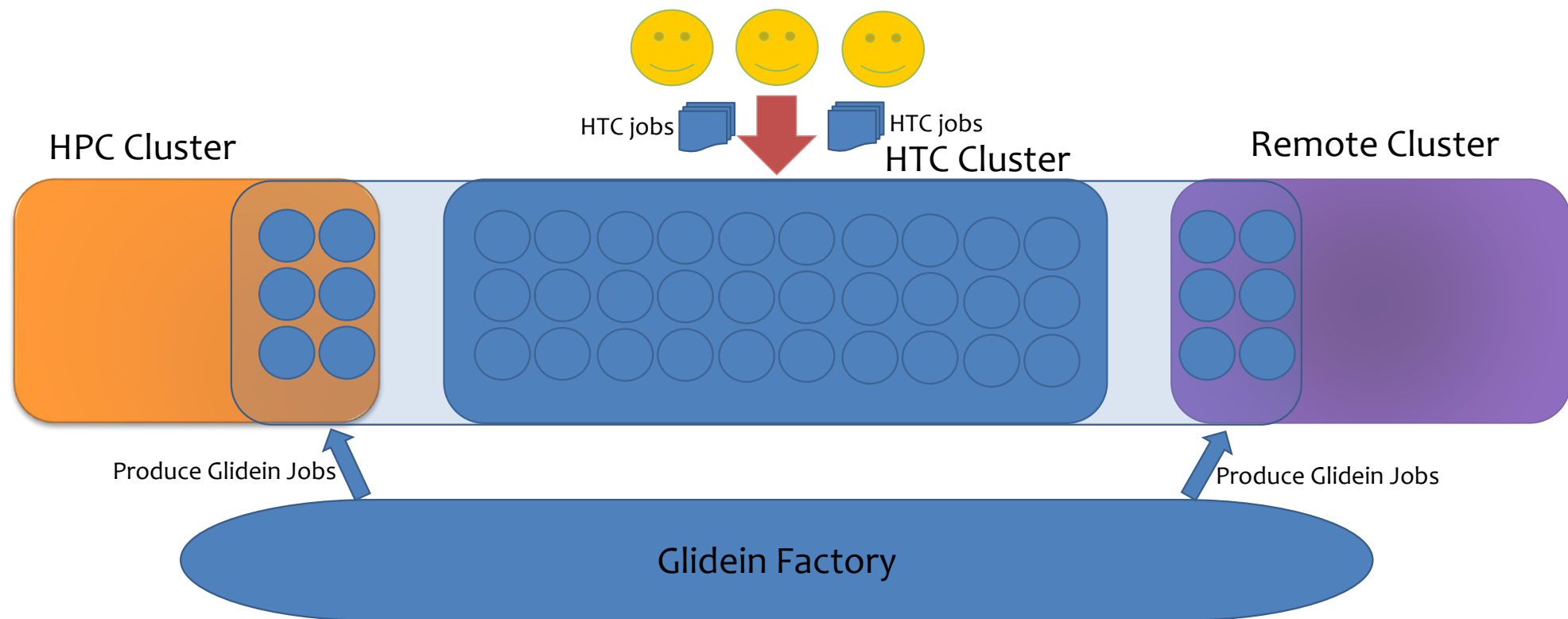
# 背景 - 区域数据中心与站点

- ❖ 为支撑高能物理数据处理，国内建设越来越多的数据中心或站点
- ❖ 区域中心（主要为高能所分支机构）
  - 提供长期资源，一般视为 HTC（高通量计算）池的一部分
- ❖ 实验合作组成员单位
  - 网络带宽和存储空间有限
- ❖ 合作研究机构、数据中心
  - 如超算、CENI算力网络等
  - 网络访问受限
- ❖ 基于不同站点条件产生不同的需求
  - CPU & GPU & ARM & Cloud
  - 站点计算系统架构不同
  - 作业调度策略不同

	CPU Cores	Disk Storage	Tape Storage	Network
IHEP-CC	>50000	Lustre, ~39 PB EOS, ~50 PB	41 PB	100G
IHEP-Dongguan Guangdong Province	~6000	Lustre 3 PB	1 PB	10G-100G to IHEP
IATAMS Shandong Province	~27400	Lustre 2.5 PB		1G-10G to IHEP
IHEP-Daocheng Sichuan Province	~3400	EOS ~6 PB		2G to IHEP
USTC, Anhui Province	~3200	Lustre 4 PB		Internet
LZU Gansu Province	~1700			2G to IHEP
Shandong Uni. Shandong Province	~1500			Internet
Dongguan DC Guangdong Province	~30000	6 PB		20G to Dongguan DC
IHEP Huairou Branch (ready by 2025)	~10000	Lustre 30 PB		100G to IHEP

# dHTC – 集群扩展模式

- ❖ 在本地集群基础上，基于集群扩展模式加入资源
- ❖ 用户可以保持原本使用本地集群的方式，仍然使用原有入口与集群交互



# 计算架构及关键问题

## ❖ 一种典型的分布式高通量计算架构

- 同时结合国内高能物理计算的特点

## ❖ 关键问题

### • Glidein Factory

- Glidein 工厂根据定义的策略生成 Glidein 作业。

### • Glidein Job

- 启动 htcondor startd 并将工作节点发布给本地 HTC资源池

### • Certification and Authentication

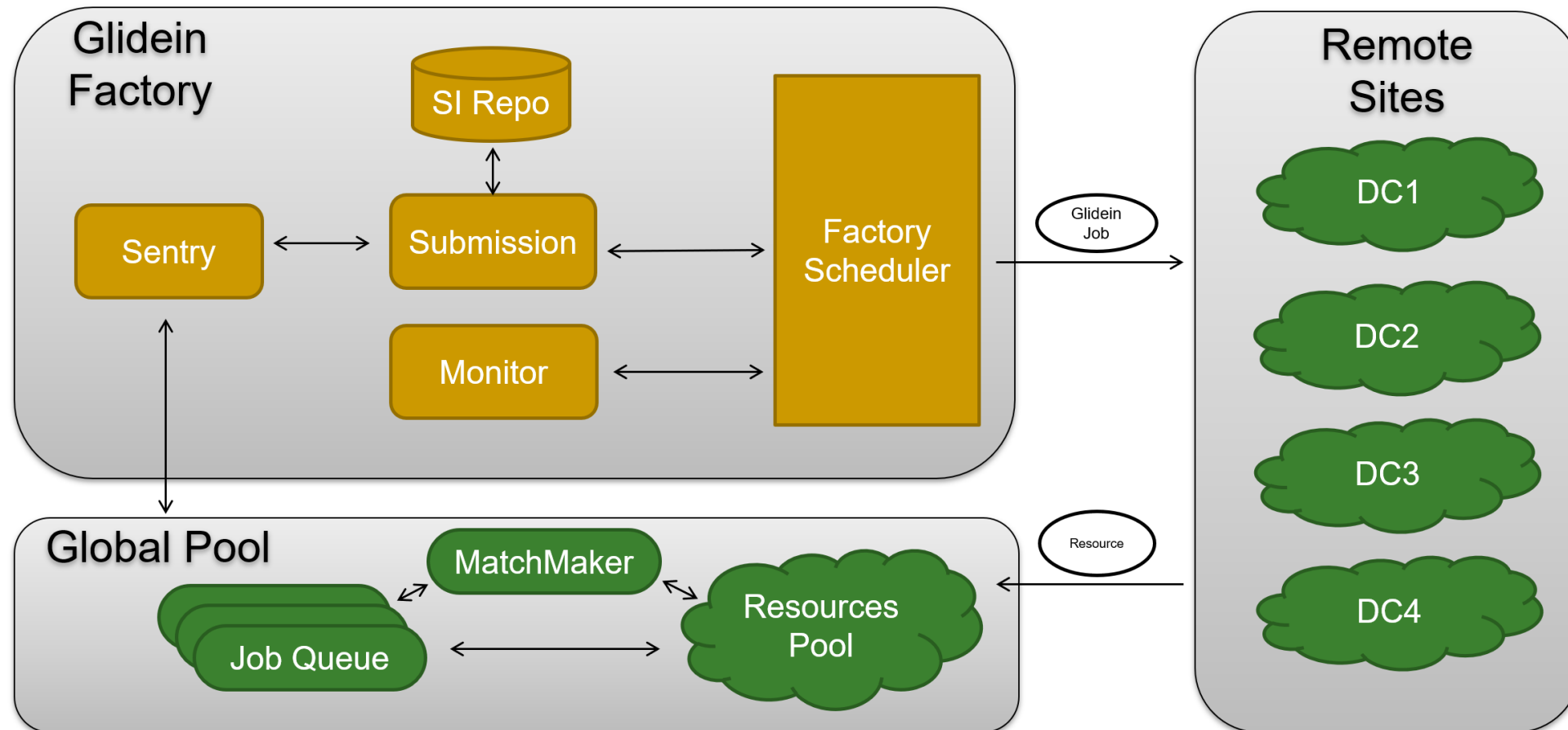
- IDTokens 和 Kerberos (Kerberos 已经在本地集群中服务多年)

### • User Interface (Workflow & HepJob)

- 作业和数据 workflow (协助实验改进数据据和计算任务 workflow)
- 通过扩展作业提交工具 HepJob 实现统一的作业入口

# Factory架构

- ❖ 采用集中式的Glidein Factory设计方案，可以更方便的接入更多的站点和集群
- ❖ 当Sentry发现有新的用户作业提交时，触发Glidein Job的向远程站点发送





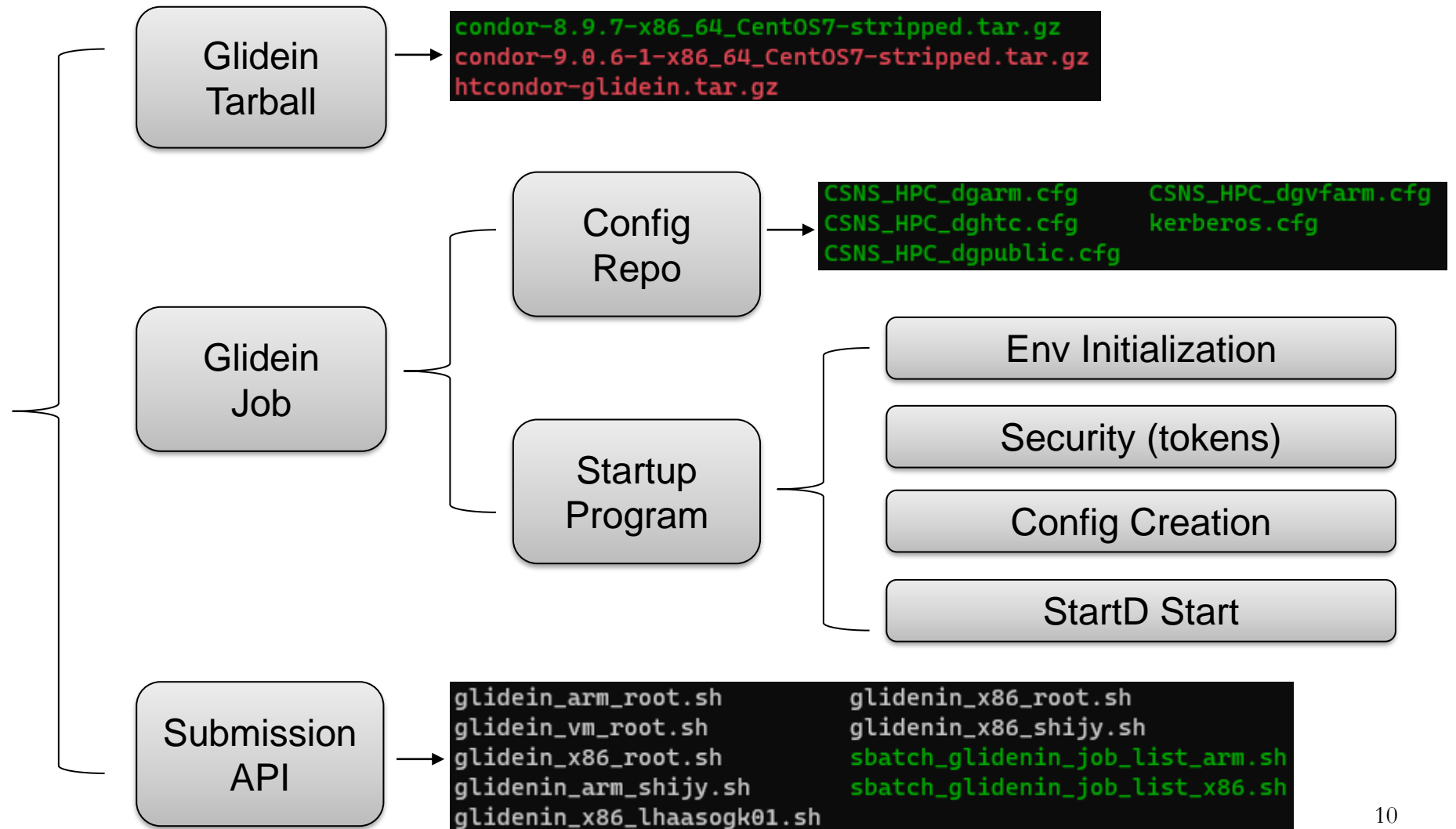
# Factory – 调度策略

- ❖ 采用混合的调度策略，以满足更多场景
- ❖ 静态资源策略（自主或协助管理的站点）
  - 站点提供固定数量的稳定资源到资源池
  - 所有资源就绪，Glidein Job可以随时提交至站点
- ❖ 动态资源策略（合作站点或机会站点）
  - 站点依据内部负载情况，提供特定数量的短期资源
  - 站点决定何时提供何种资源到资源池
    - 一个Glidein Factory客户端在站点控制Glidein Job的提交和删除
- ❖ 作业分发策略基于作业的需求
  - 所有作业均可以分发至本地HPC或HTC集群（本地集群具备相同的节点环境）
  - 数据传输大的作业分发至区域中心和大型站点（具备稳定存储空间和网络带宽）
  - 数据传输小的作业分发至大部分机会站点、边缘站点（通常不具备稳定的存储空间和网络带宽）<sup>9</sup>

# Glidein Job

## ❖ Glidein Job

- Glidein Tarball
- Config Repo
- Startup Program



# Cert&Auth with Tokens

## ❖ 基于Tokens认证方案能够增加整个系统安全性

- 同时可以解决多站点环境下的访问权限（特别是对数据系统的远程访问）

## ❖ IDTokens: 计算系统服务间认证

- 由CLAIMTOBE方式迁移至IDTokens

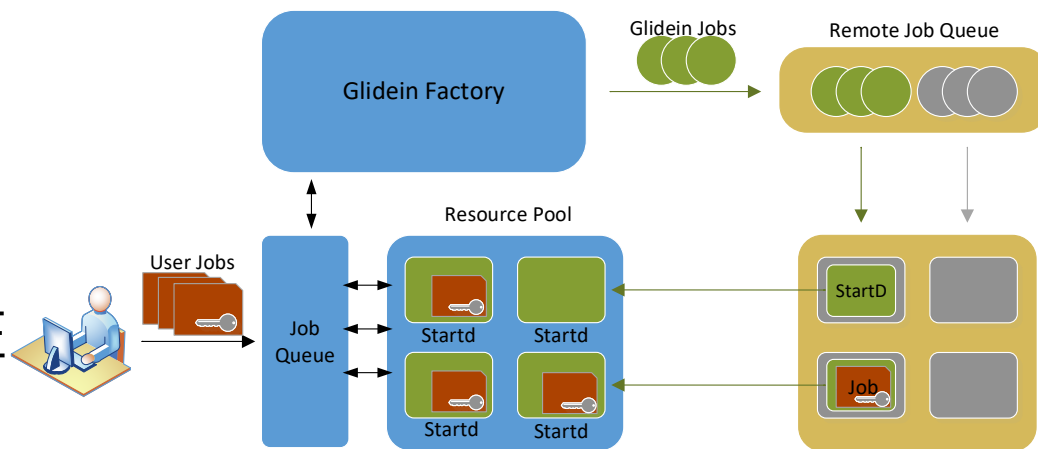
## ❖ Kerberos Tokens: 作业与服务的认证

- 尝试过直接使用HTCondor软件套件自带Kerberos认证

- HTCondor没有解决提交端和运行端的用户命名空间一致性问题

- 自研一套Kerberos Tokens分布式环境下的自动传输、初始化、更新、销毁流程

- Token票单的初始化和生效在Job Wrapper中实现，同时周期性刷新Tokens有效期



Submit Side

```
transfer_input_files = /tmp/krb5cc_10634  
+HepJob_KRB5CCNAME = "krb5cc_10634"
```



Execute Side

```
$ ls /var/lib/condor/execute/dir_6412/  
condor_exec.exe _condor_stderr _condor_stdout krb5cc_10634 tmp var
```

# 站点接入 – 本地HPC集群

❖ The system software of local HPC WN is same as that of HTC WN

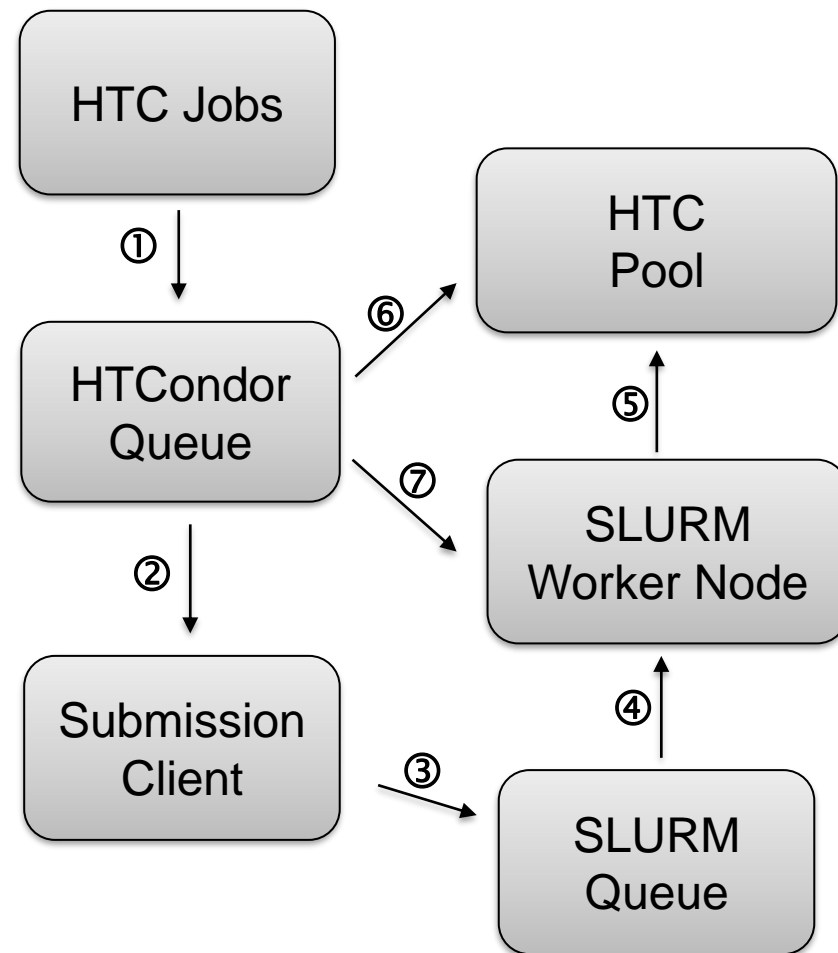
- With shared file system and same user namespace

❖ The workflow is:

- ① submit htc jobs to htcondor queue
- ② submission API receives the job requirements
- ③ submission API submits glidein jobs to slurm queue
- ④ slurm schedules glidein jobs to slurm worker nodes
- ⑤ glidein startups startd and report to htc pool
- ⑥ matchmaker
- ⑦ htc jobs are sent to slurm worker node

❖ The policy in submission API is quite easy

- Specific idle user jobs & empty worker nodes in slurm



# 站点接入 – 东莞大科学数据中心（区域中心）

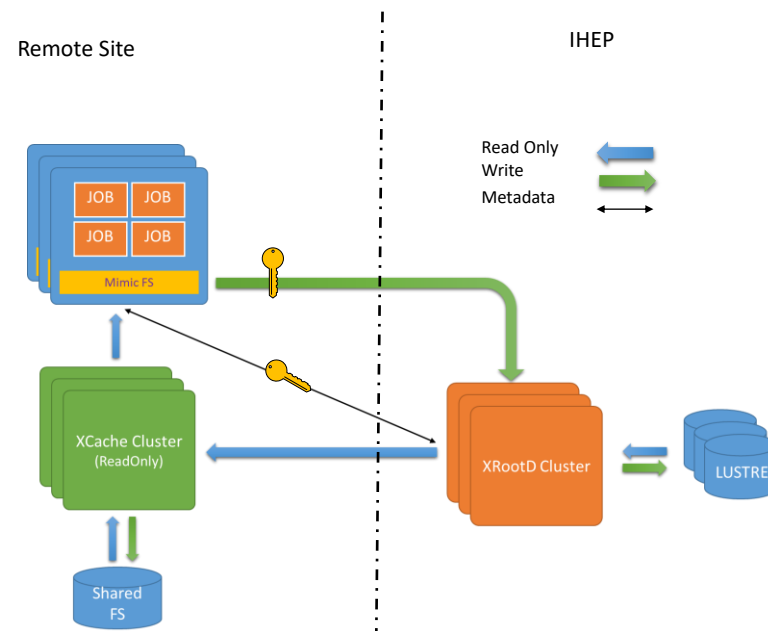
## ❖ 东莞大科学数据中心

- 基于SLURM构建的高性能计算集群
- 北京与东莞间没有统一的共享文件系统
- 北京与东莞间没有统一的用户命名空间

Could CPU: 9600 cores  
X86 CPU: 10000 cores  
ARM CPU: 9600 cores  
GPU: 80 Tesla V100  
Storage: 6PB

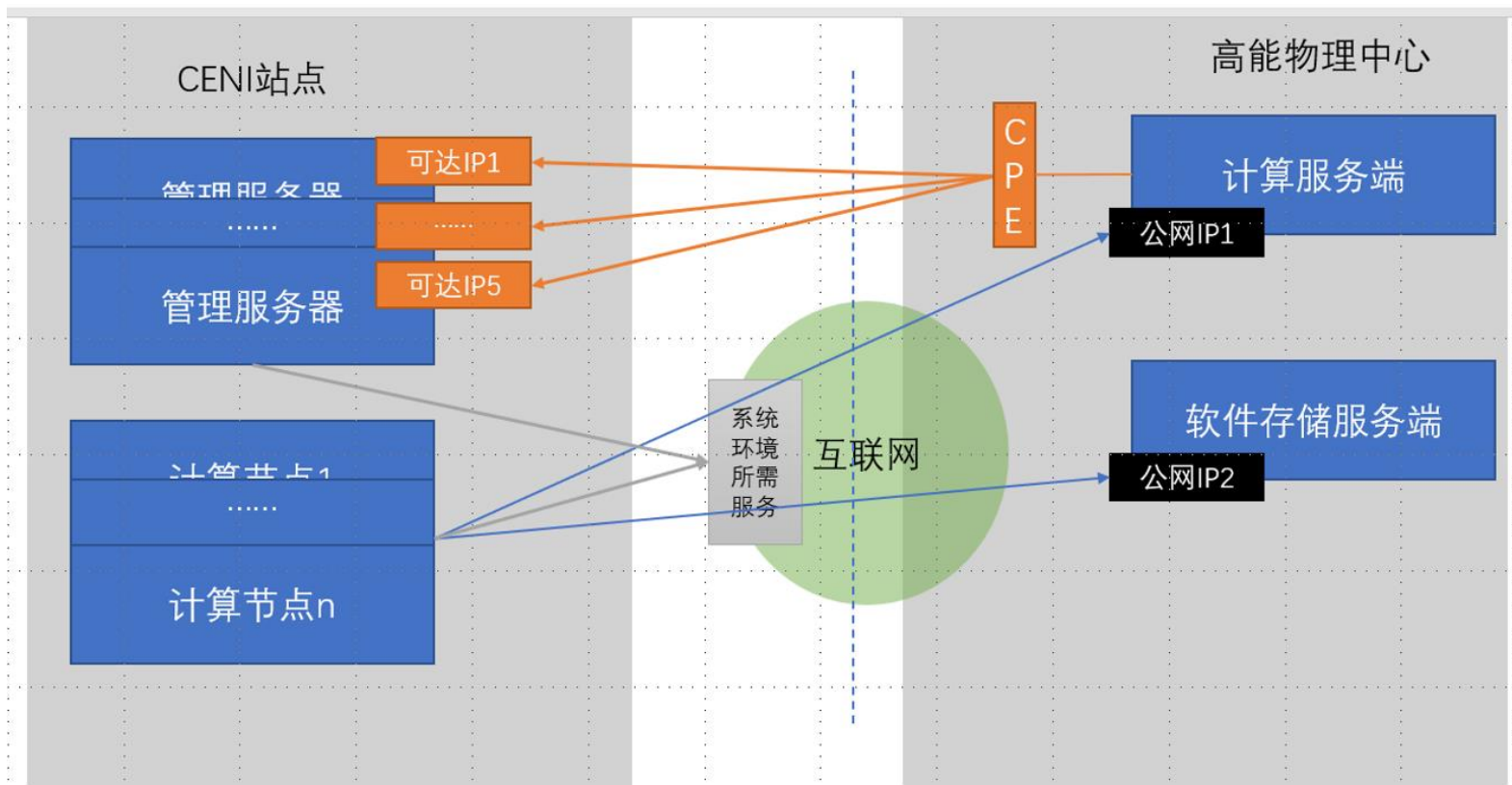
## ❖ 东莞数据中心与北京数据中心之间有稳定的网络专线，并且处于同一张内网

- 网络带宽：20Gpbs
- 可以满足大规模的数据访问和传输
  - EOS：直接内置xrootd远程访问方式
  - Lustre：借助xrootd软件，以proxy方式提供远程访问与传输



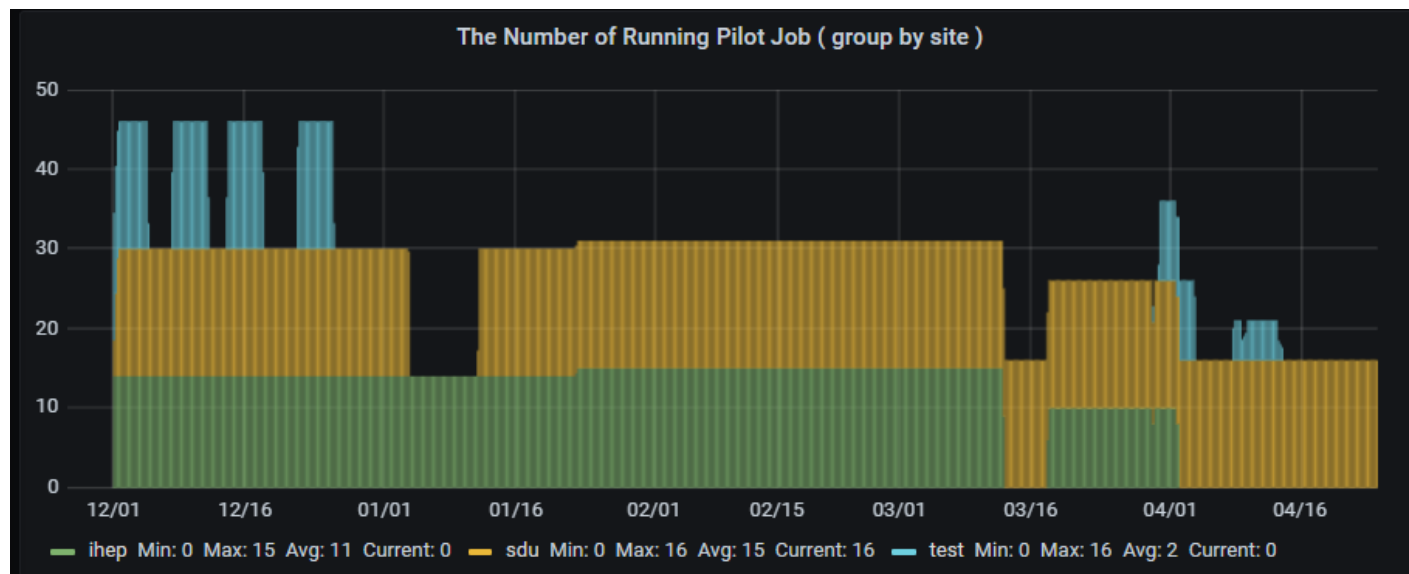
# 站点接入 – CENI算力网络&超算

- ❖ CENI算力网络目前提供5个站点资源：北京怀柔、天津、武汉、广州、南京
  - 主要需要解决网络访问受限问题：Global端无法到达CENI站点内部
- ❖ Global端提供双网连接
  - CPE (VPN)接入站点内部
  - 公网IP提供日常服务
- ❖ 超算中心采用类似方式
  - 目前正在接入成都超算
  - 双网连接方式还待验证



# 站点接入 - 合作组成员单位站点

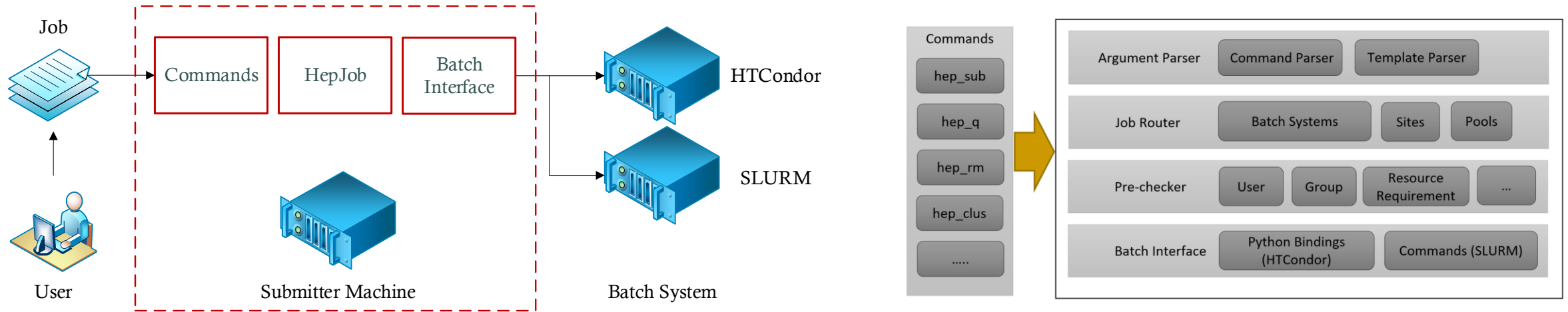
- ❖ 合作组成员单位站点（目前接入资源数量不稳定）
  - 山东大学物理学院站点, 兰州大学核物理学院站点
- ❖ 目前接入资源数量、网络带宽和存储空间受限
  - 未开展的大规模生产应用



# 作业接口

## ❖ 用户接口遵循原有使用方式（用户长期习惯本地集群方式）

- 在远程站点运行作业，用户不需要做过多的使用方式上的改动
- 远程站点相关的功能增加并打包在HepJob软件工具中



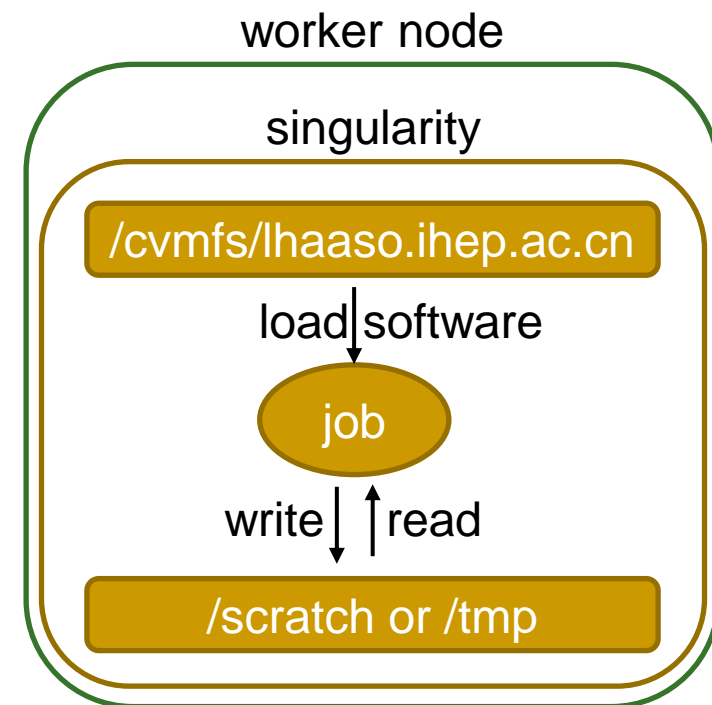
## ❖ 协助实验完善作业和数据 workflow（有助于向实验推广分布式计算资源）

- 完善和实现作业 workflow 和数据存储结构
- 如 LHAASO 实验的 mini-workflow, BESIII 实验的 boss.condor



# 作业运行环境 – Singularity容器

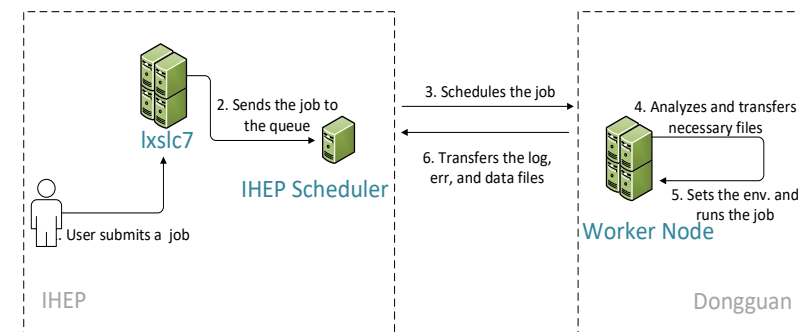
- ❖ 作业运行环境主要基于singularity容器方案实现
- ❖ 操作系统
  - Singularity镜像发布至/cvmfs/container软件存储中
  - 在用户作业到达节点时，Glidein Job依据特定操作系统镜像启动singularity容器
- ❖ 实验软件与个人软件
  - 实验软件由/cvmfs管理和提供访问
  - 个人软件由系统负责传输到远程计算节点
- ❖ 运行时存储空间
  - 计算节点的本地存储空间，如/scrach或/tmp



# 分布式计算应用 – BESIII实验

## ❖ 主要支持BESIII实验的模拟和重建数据处理（传统计算模式拓展至分布式场景）

- 设计实现远程作业的封装：基于boss.condor的功能延伸
- 实现Lustre共享文件系统远程读写：基于xrootd的proxy方案
- 高频热点数据访问：基于cvmfs的random trigger数据远程共享
- 数据库架构升级：远程分布式数据库系统

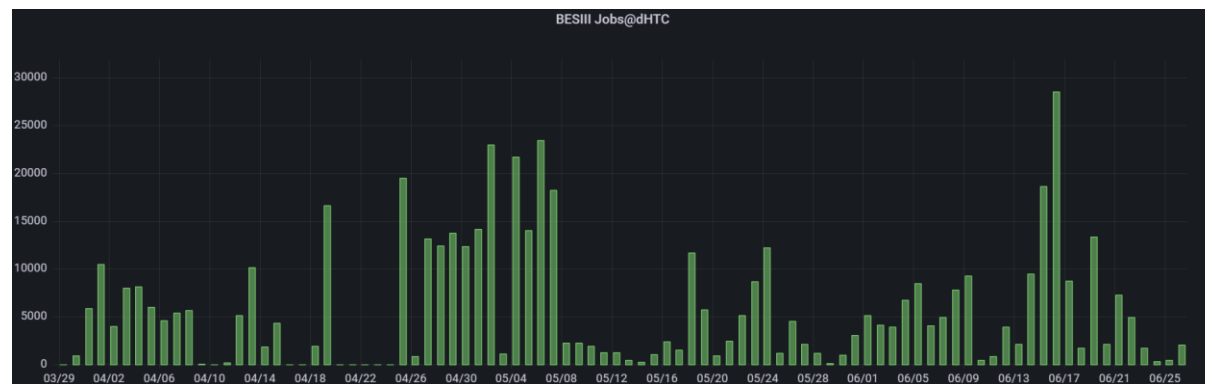
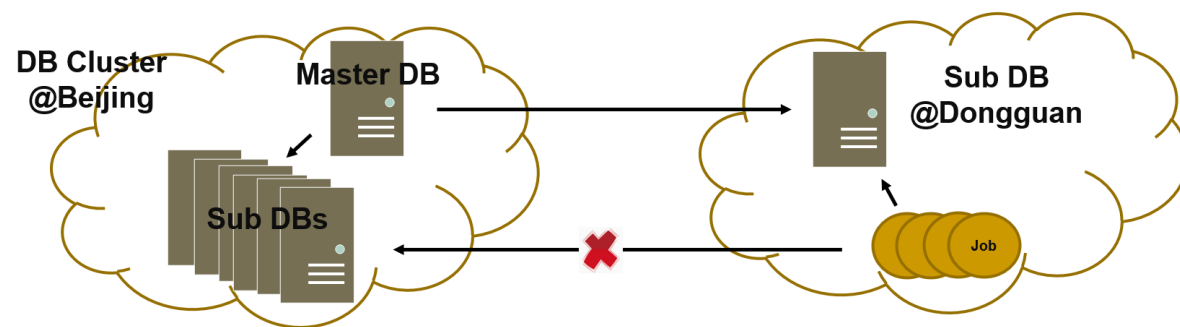


## ❖ BESIII实验可额外获得30%计算资源

- 忙时可额外获得56%计算资源

## ❖ 作业运行情况(近3月)

- 完成作业: ~52万作业量
- 完成机时: ~361万小时

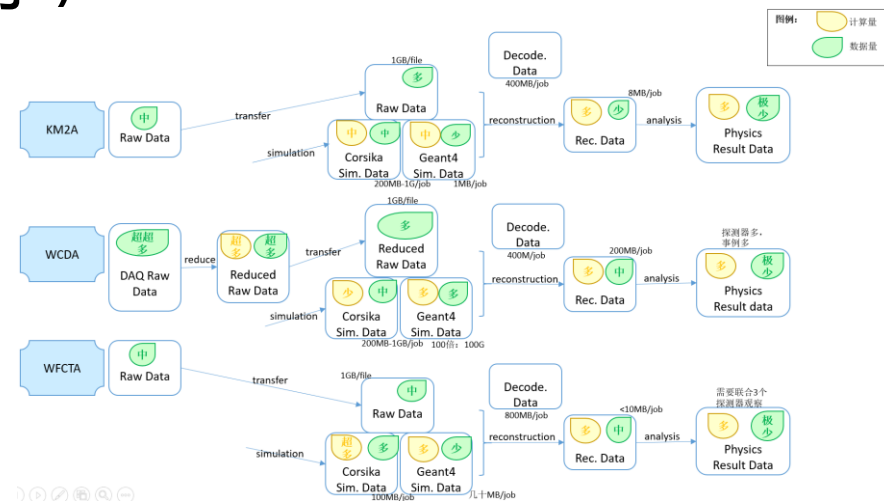


# 分布式计算应用 – LHAASO 实验

## ❖ 主要支持LHAASO实验的模拟任务（数据传输相对小）

- 已支持WCDA、KM2A、WFCTA等探测器模拟数据处理

	Corsika	Geant4(step1)	Geant4(step2)	Reconstruction	Analysis
input	tiny	middle	large	small	various
output	middle	large	small	small	small
persistent store?	No	<b>Yes</b>	No	<b>Yes</b>	No
CPU hours	long	Long	Short	Short	Various
users	Dedicate user	Dedicate user	Dedicate user	Dedicate user	Various user



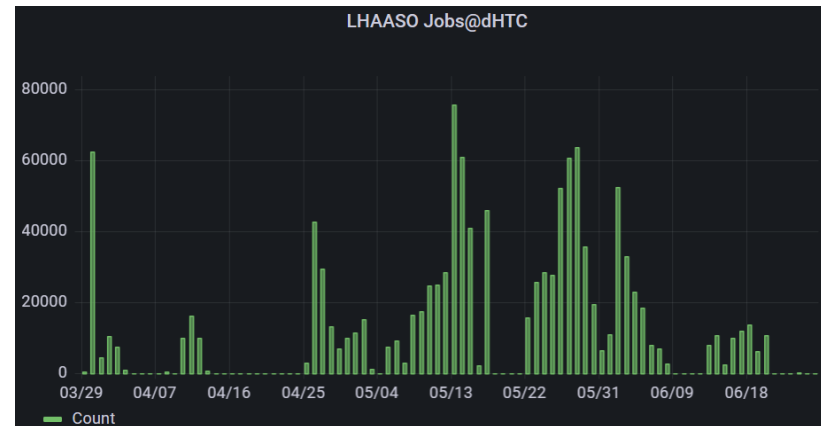
## ❖ 协助实验实现Job Workflow

- 自动组织数据：能量、粒子、事例等多个维度
- 批量作业提交：x86&ARM

```

corsika_job_example.sh      hepjob_wfcta_corsika_arm_single_job.sh  wfcta_corsika_arm.sh
hepjob_wfcta_corsika.sh    hepjob_wfcta_corsika_single_job.sh     wfcta_corsika_arm_test.sh
hepjob_wfcta_corsika_arm.sh wfcta_corsika.sh                       wfcta_corsika_test.sh
    
```

## ❖ 作业运行情况(近3月): 120万作业



# 总结

## ❖ 基于dHTC方式的分布式计算方案在高能物理国内资源整合上取得一定进展

- 软件基础设施已基本完成，并投入生产使用：HepJob/HTC Pool/Glidein Job
- 基于Kerberos Tokens实现安全认证和远程服务授权
- 基于xrootd proxy的远程数据读写
- 适应分布式计算的作业接口和作业运行环境
- 已接入若干类型数据中心和站点，并支持BESIII、LHAASO、HERD等实验

## ❖ 下一步计划

- 构建和实现完整的Factory系统，已保证整套方案的全自动运转（近期）
- 实现超算中心等具有特殊网络环境的站点接入（近期）
- 继续协助实验接入分布式计算平台（持续ing）
- 提供敏捷的站点服务部署方式（容器方式），以方便更多边缘站点顺利接入
- 完善站点接入方式，以适应更多场景



Thanks  
Q&A

