



国家高能物理科学数据中心

National HEP Science Data Center



高能所计算中心

IHEP Computing Center

高能物理相关领域计算平台介绍

姜晓巍

高能所计算中心

2024-05-20



1 高能物理相关领域计算平台概况

2 高能所计算平台现状

3 高能所计算平台服务

4 总结及展望

计算平台支撑科研活动



高能物理相关领域 – 数据中心



上海光源(SSRF):
11000 CPU cores
28 GPU cards
23 PB storage

上海自由电子激光(SHINE):
计算平台

合肥光源(HSL-II):
计算平台

高能所东莞分部
(CSNS, LQCD, LHAASO等):
35000 CPU cores
10 PB storage

兰州大学-核物理学院
(BES, LHCb等):
8000 CPU cores
6 PB storage

华中师大
(ALICE, LQCD等):
10000 CPU cores
10 PB storage

上海交通大学
(ATLAS等):
100000 CPU cores
128 GPU Cards
40 PB storage

高能所
(BES, LHAASO, JUNO, HEPS等):
100000 CPU cores
200 GPU cards
100 PB storage

山东省高研院
(AMS等):
30000 CPU cores
2 PB storage

山东大学-物理学院
(BES, LHAASO, JUNO, BelleII等):
1500 CPU cores
1 PB storage

中国科技大学-物理学院
(BES, ATLAS等):
2200 CPU cores
32 GPU cards
1.5 PB storage

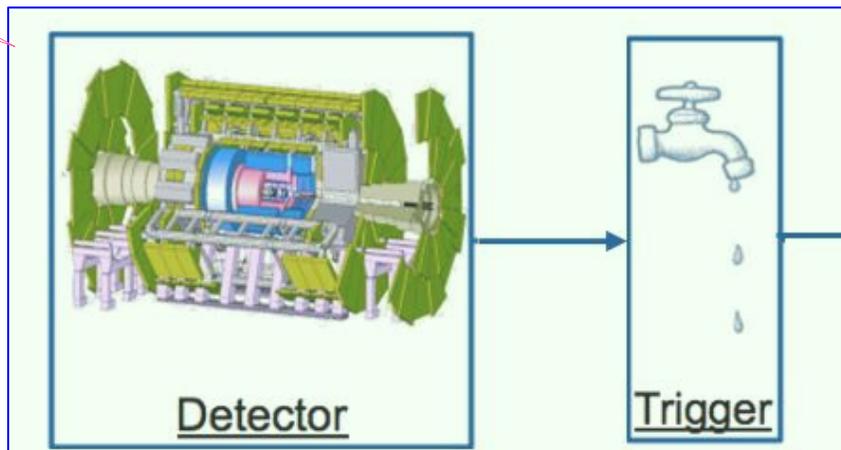
北京大学-物理学院
(BES, CMS等):
1000 CPU cores

数据主要来源为各数据中心官网或相关文献 (截止时间为2023年)

粒子物理类 – 仍大量需求CPU资源和存储资源



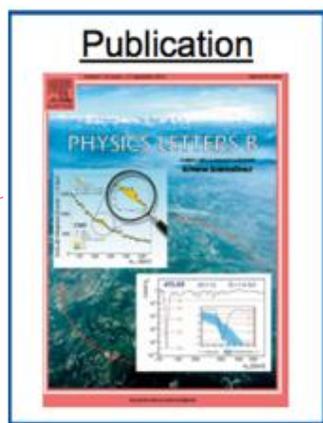
在线
取数



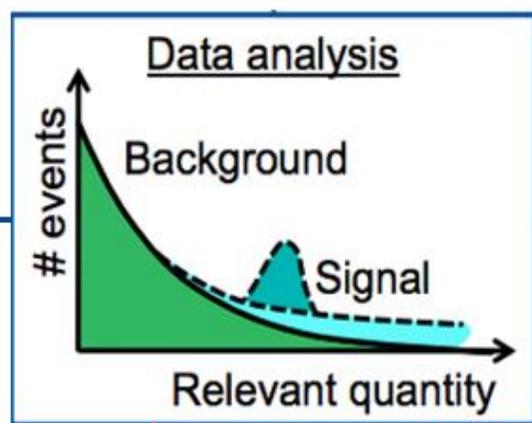
原始
数据



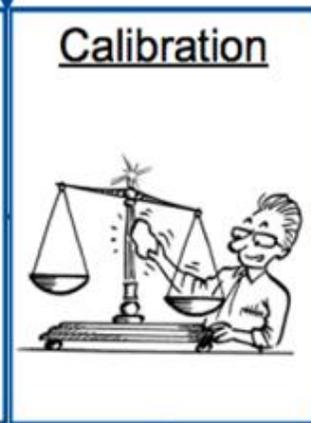
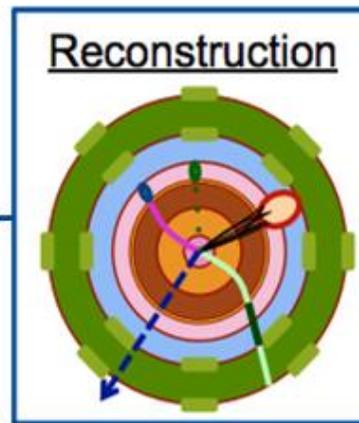
离线模拟
→ 模拟数据



物理
成果



不断反
复分析



刻度+重建
→ 重建数据

光源类 – 高通量数据IO & 交互式在线数据分析



● 交互式数据分析平台（流式计算）

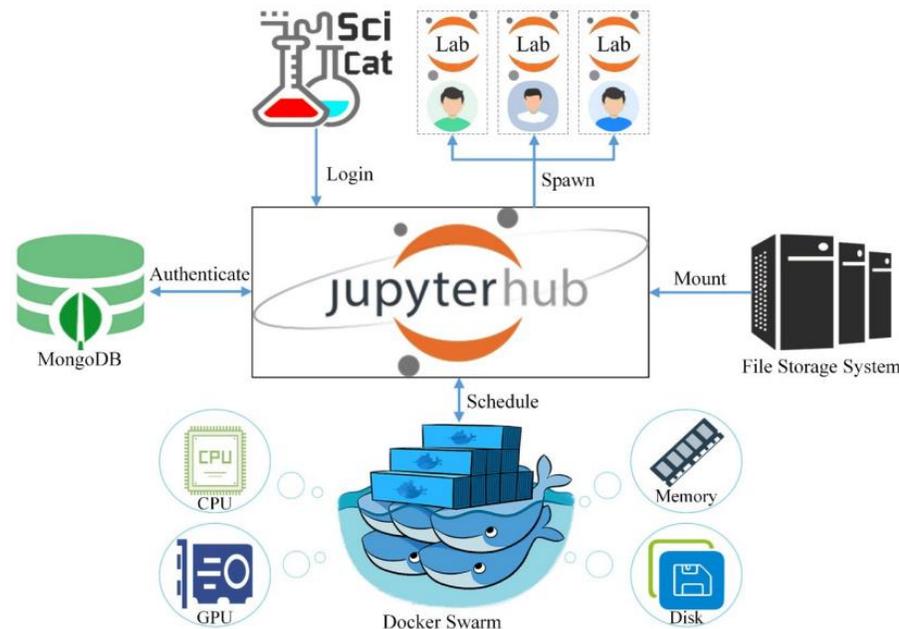
- 大多基于Jupyterhub开发
- 底层为计算平台（结合虚拟技术）

● 超大数据量

- 要求存储实时IO性能较高

● 学科交叉融合

- 软件框架、数据管理、学科方法



合肥光源交互式在线数据分析



线站名称	每天数据量峰值 (TB)	每天数据量平均值 (TB)
B1	600.00	200.00
B2	500.00	200.00
B3	8.00	3.00
B4	10.00	3.00
B5	10.00	1.00
B6	2.00	1.00
B7	1000.00	250.00
B8	80.00	10.00
B9	20.00	5.00
BA	35.00	10.00
BB	400.00	50.00
BC	1.00	0.20
BD	10.00	1.00
BE	25.00	11.20
BF	1000.00	60.00
总计		805



- 单核 -> 多核
 - 多进程、多线程
- CPU -> GPU
 - 人工智能兴起 (大模型)
- 国产设备
 - ARM架构CPU
 - 国产GPU (如曙光DCU)

高能所计算平台 - 多实验、多站点场景





● 资源规模

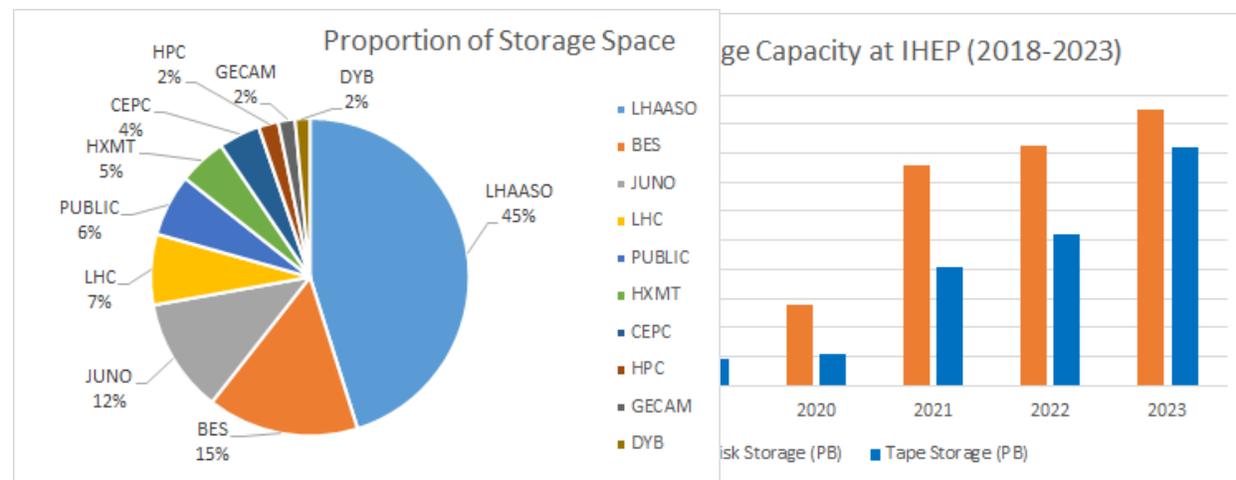
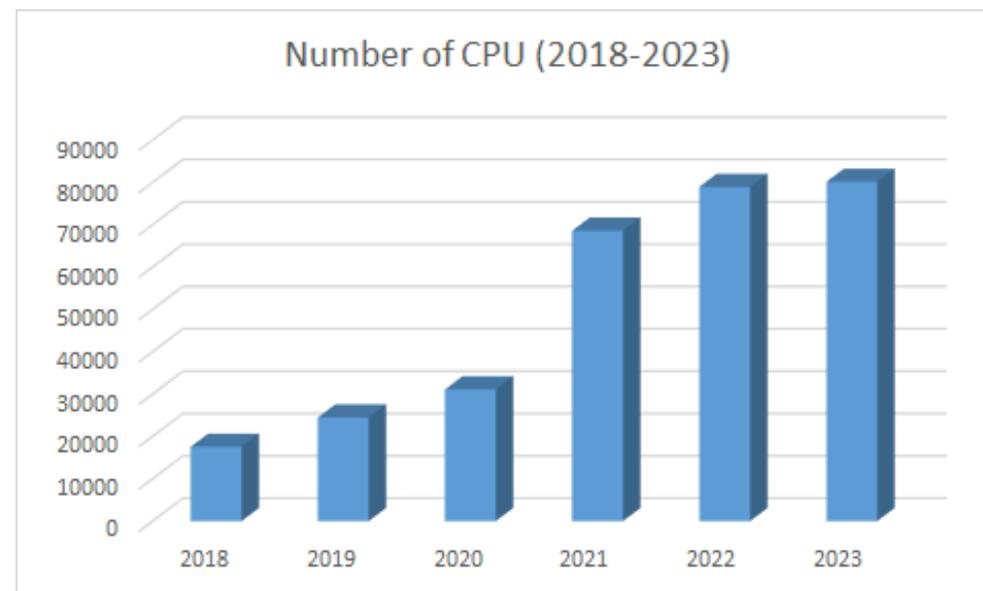
- ~100K CPU cores
- ~100PB Disk Storage
- ~80PB Tape Storage

● 支持实验或应用

- 28 experiments/applications (12 new)
- HEPS/LQCD/HERD/LHCb/Ali-CPT/GECAM/EP/...

● 多站点或数据中心

- Beijing(IHEP), Donguan(CSNS), Daocheng(LHAASO), Jiangmen(JUNO)



● 主要支持:

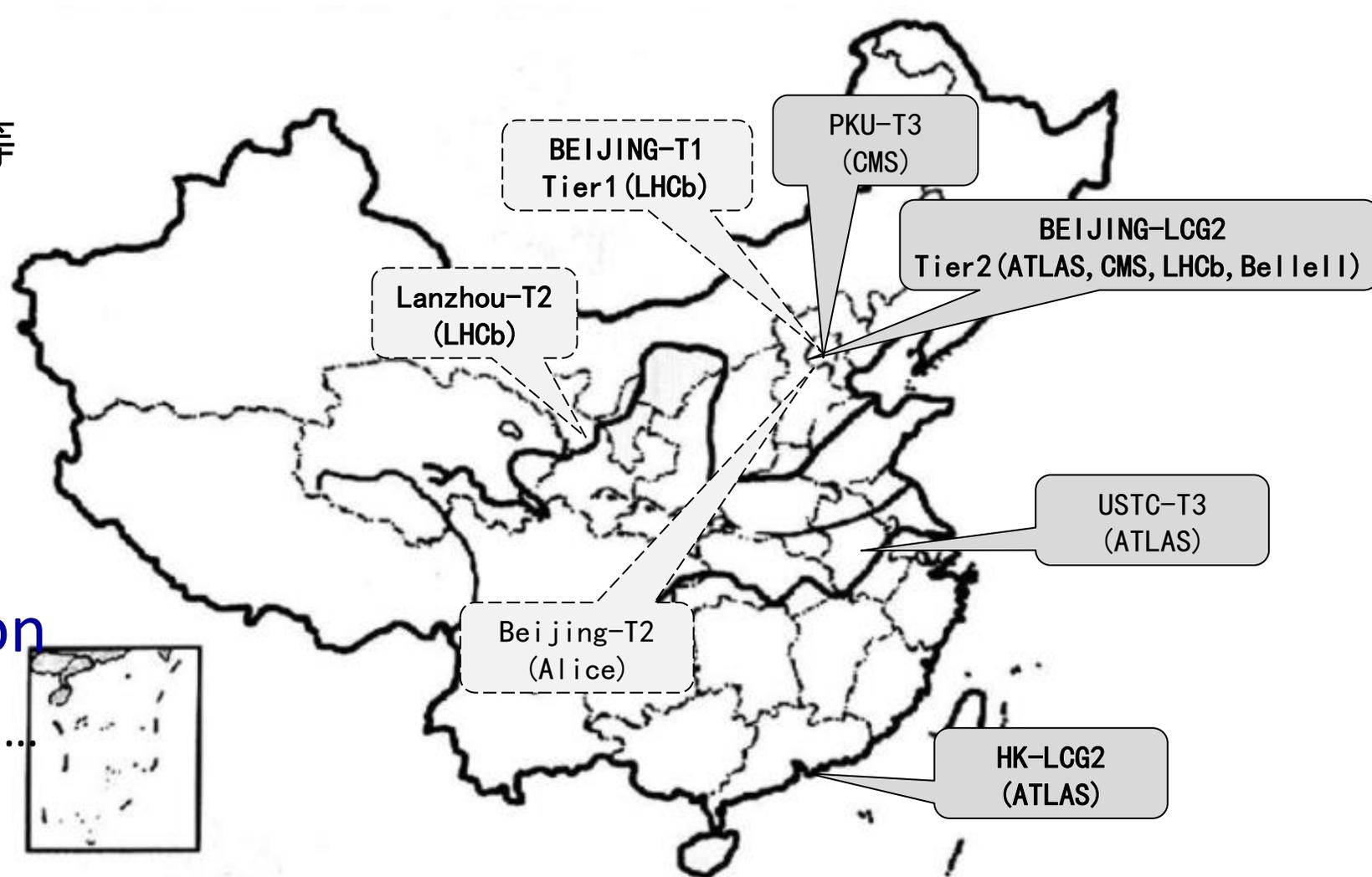
- LHC, BelleII, JUNO, CEPC等

● WLCG Beijing-T1

- LHCb T1 站点 (6月正式)
- 3200 CPU cores
- 3.2PB disk, 3PB tape

● WLCG T2 site federation

- ATLAS, CMS, LHCb, BelleII,...
- LZU, SDU, USTC, ...





● 国际网络链路升级为100Gbps（多实验共享）

■ CSTNET (CNIC), GEANT, CERN and CN-IHEP 合作共建

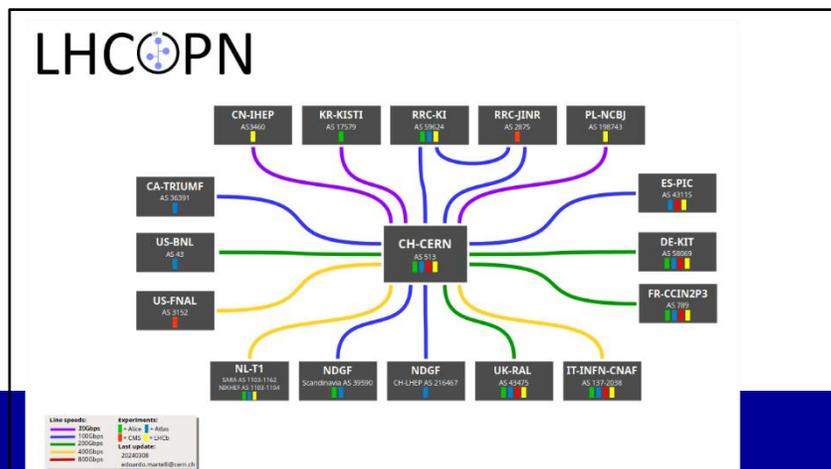
■ LHCOPN: 20Gbps for Beijing-T1

◆ CSTNET: 20Gbps

◆ GEANT: more than 20Gbps promised

■ LHCONE: 100Gbps shared to T2 sites

◆ 高能所计算平台是CSTNET最主要用户





• 所级人工智能共享计算平台

- 面向高能所人工智能应用需求（AI数据分析、大模型等）
- 提供用于人工智能应用的GPU计算资源、存储资源
- 提供人工智能用户Portal，集成通用算法和工具，简化平台使用方式

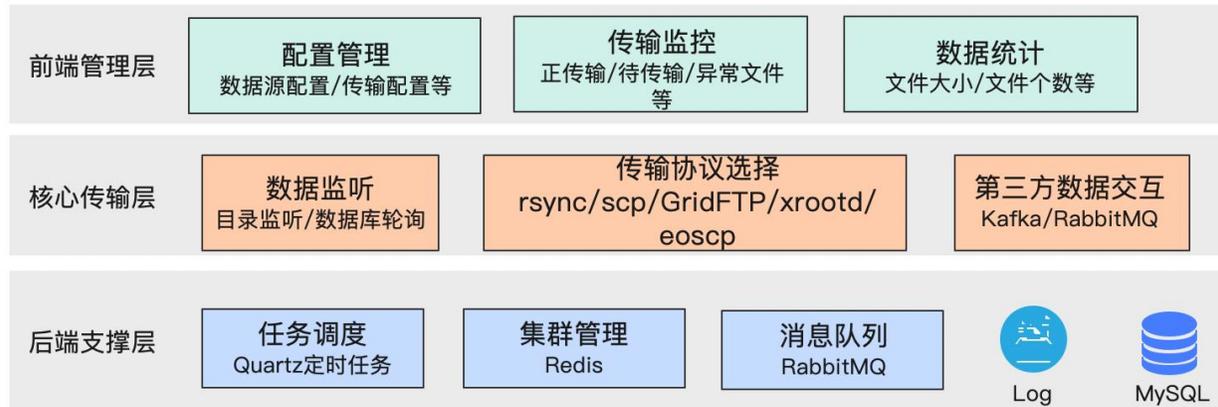


服务 - 数据传输



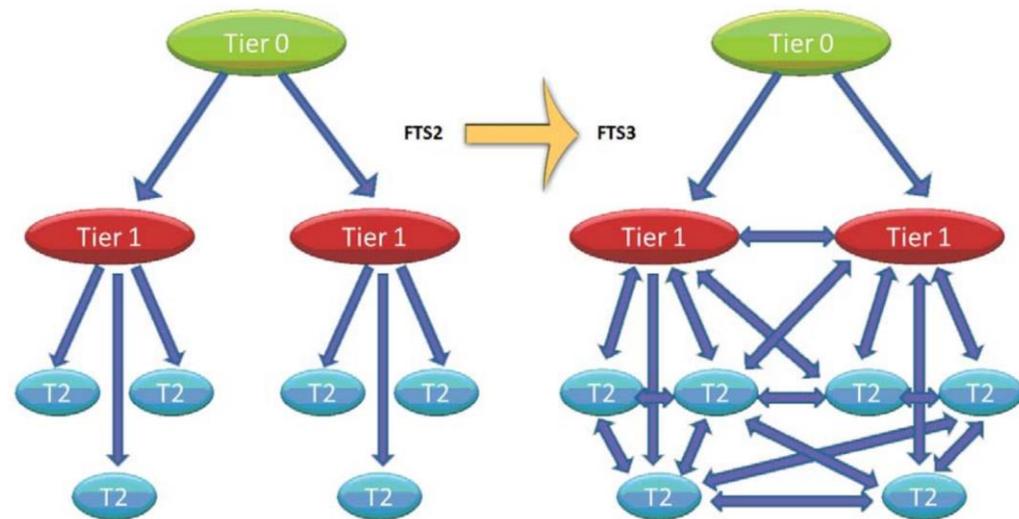
• HiDTS: High Performance Data Transfer System

- 端到端的海量数据传输系统
- 高可靠性：完整性、一致性检查，失败重传
- APIs：支持目录和数据库；监控；记账统计
- 应用：LHAASO、DYW



• FTS: File Transfer Service

- 分布式数据传输系统：多站点场景
- 支持多网格协议：http, xrootd, tpc, gridftp
- 支持多VO(多实验)
- 广泛应用于网格系统中：LHC, BelleII, JUNO等

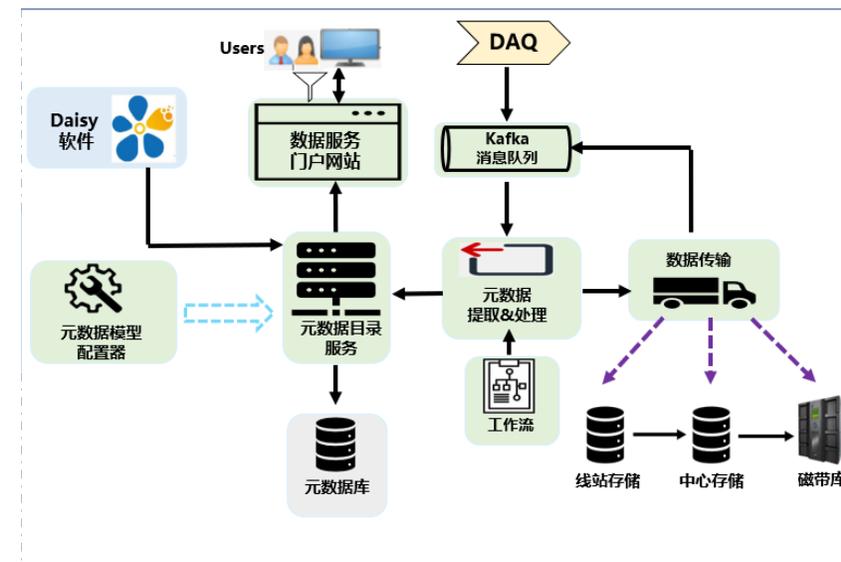




● 数据管理系统(DOMAS): 数据全生命周期管理

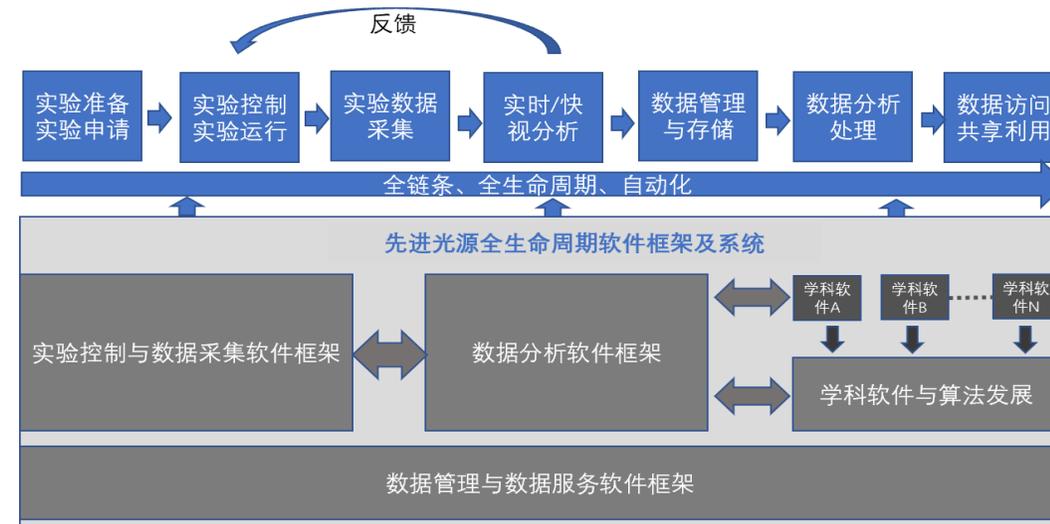


- 标准与规范: 数据策略、数据格式、元数据标准的制定
- 元数据目录: 实现对元数据的目录设计和存储, 更快速和便捷地检索数据
- 元数据获取: 从不同系统 (DAQ、存储、用户服务、数据分析) 获取元数据
- 数据存储: 数据存储目录设计、权限设计
- 数据服务: 提供数据访问、下载、分析与追踪服务
- 开放接口: 标准接口, 满足各系统/模块访问数据或元数据



● 数据管理服务

- 对海量实验数据合理存储、组织和管理
- 提供数据全流程“一站式”服务
- 数据格式多样, 提升数据的利用、共享能力及效率
- 支持不同实验站数据流程(策略)的多样与变化





• 公开共享科学研究数据促进数据的检查和再利用

- 开放科学数据的保护、管理、维护和共享

• Open Data

- 数据开放策略
- 提供数据集
- 配套软件
- 支持文档
- 示例分析

L1: Publications and additional documentation

L2: Simplified data for outreach & education

L3: Reconstructed data

L4: Raw data

• 应用：LHAASO、CSNS、卫星项目

Found 3 results.

Sort by: Best match asc

LHAASO-Science-adj2778

Paper "Very high-energy gamma-ray emission beyond 10 TeV from GRB 221009A" is published in Science Advances Vol 9, Issue 46 on 15th November 2023....

LHAASO 2023 Publication

LHAASO-LE18750-2023

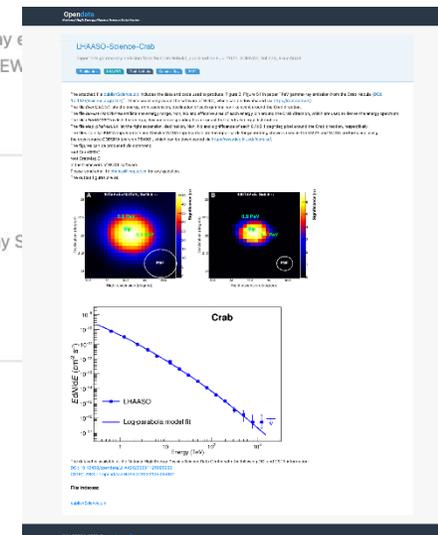
Paper "Measurement of ultra-high-energy diffuse gamma-ray emission with LHAASO-KM2A" has been published in PHYSICAL REVIEW LETTERS...

LHAASO 2023 Publication

LHAASO-ApJS - AAS47142R2

The catalog paper 'The First LHAASO Catalog of Gamma-Ray Sources' is published in Astrophysical Journal Supplement Series on September, 2023.

Publication LHAASO Gamma-Ray 2023





• 提供更丰富的运行状态

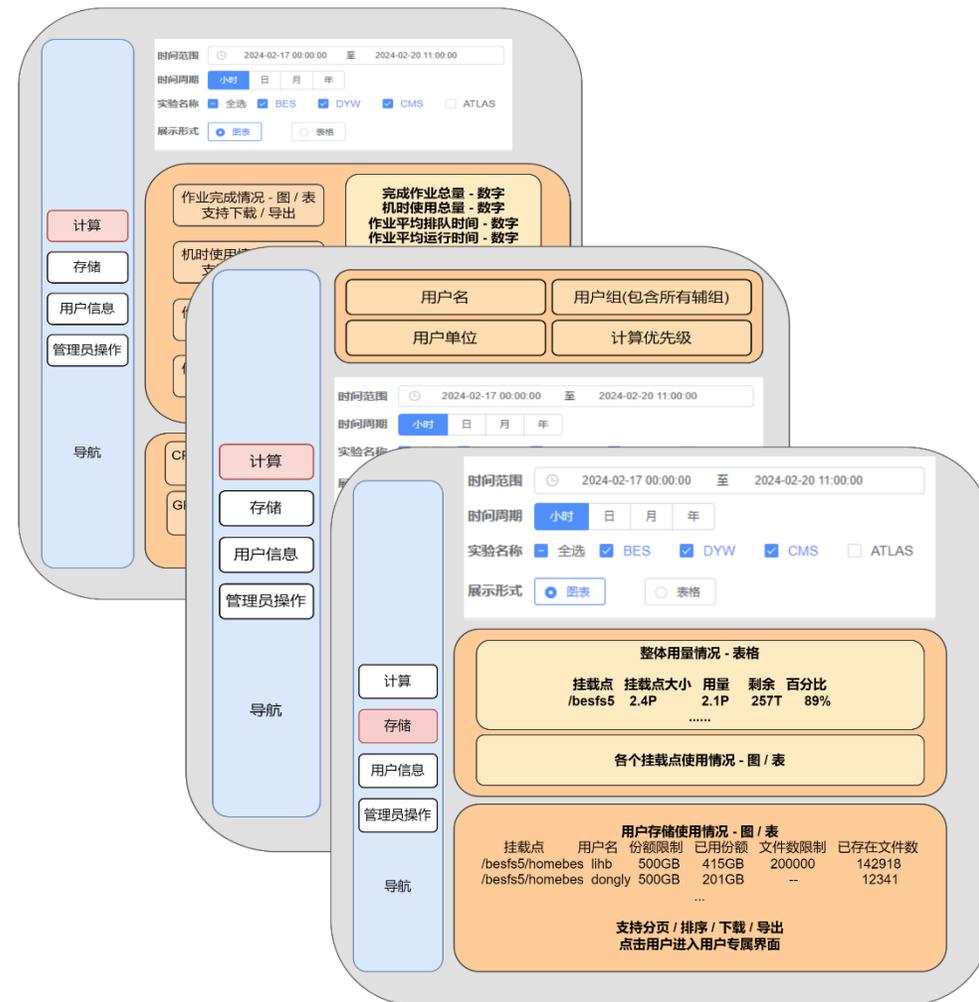
■ 计算、存储状态查询

- ◆ 计算协调人能够查看本实验用户使用的详情信息
- ◆ 账号, 用户姓名, 机时, 脚本路径, 用户单位

■ 自动形成周报/月报发送给计算协调人

■ 提供实验协调人更多自主管理权限

- ◆ 限制提交或是限制用户作业上限
- ◆ 邮件警告
- ◆ 通知上级 (学生 -> 导师 / 职工 -> 部门负责人)





- 高能物理相关领域的数据中心数量和规模不断增大，资源需求也呈现多样性
 - 粒子物理类、光源类等
- 高能所计算平台长期为各实验提供计算、存储、网络等各类资源和服务
 - AI平台、数据管理、数据传输、Opendata等
- 计算平台建议或意见，请联系：
 - 石京燕：shijy@ihep.ac.cn
 - 姜晓巍：jiangxw@ihep.ac.cn

Thanks!
Q&A