面向光源学科数据分析场景的交互式计算平台

胡庆宝 徐吉平
huqb@ihep.ac.cn
高能物理研究所计算中心

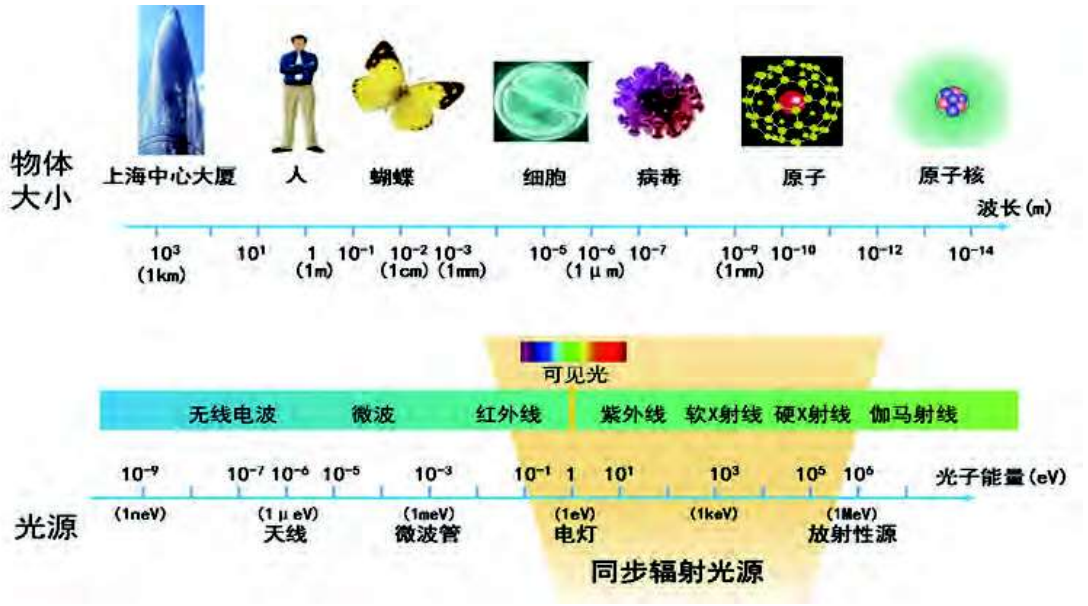
2023-07

提纲

- 背景简介
- 光源学科计算的挑战
- 光源交互式计算平台设计
- 计算平台关键技术及进展
- 总结和展望

先进光源介绍

- 基础研究离不开先进的科学装置。作为国家重大科技基础设施的重要组成部分，同步辐射光源，如同一台超级显微镜，能够让科学家在原子和分子的层次上去研究物质的内部结构，也可以做大分子结构的研究，真正做到“俯察品类之盛”，成为支撑众多学科前沿基础研究与高新技术研发不可或缺的试验手段。
- 当前拥有合肥光源、上海光源、大连先进光源，正在建设的高能同步辐射光源（HEPS)、上海硬X射线自由电子激光装置、合肥先进光源，它们为基础科学和工程科学等领域的突破性创新提供不同寻常的研究机遇。

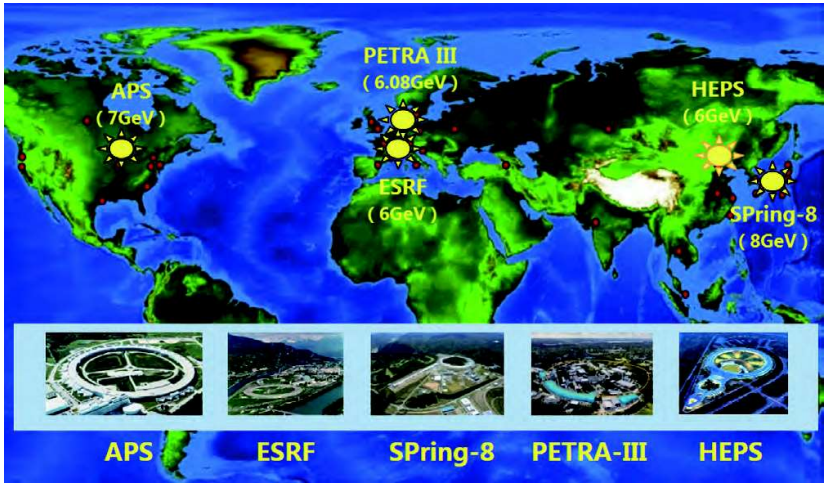


先进光源介绍

- 第一代：北京同步辐射装置——依托北京正负电子对撞机的光源
- 第二代：合肥同步辐射光源——专门为同步辐射的应用而设计
- 第三代：上海同步辐射光源——低发射度、大量采用插入件的专用光源
- 第四代：高能同步辐射光源——更亮度的光照亮微观世界

高能同步辐射光源

HEPS 建设周期为 6.5 年，2019 年年中完成所有前期准备工作、开工建设，2019-2025 年完成工程建设，2025 年底验收并投入运行。



数据特点：

线站多、数据规模大 PB量级

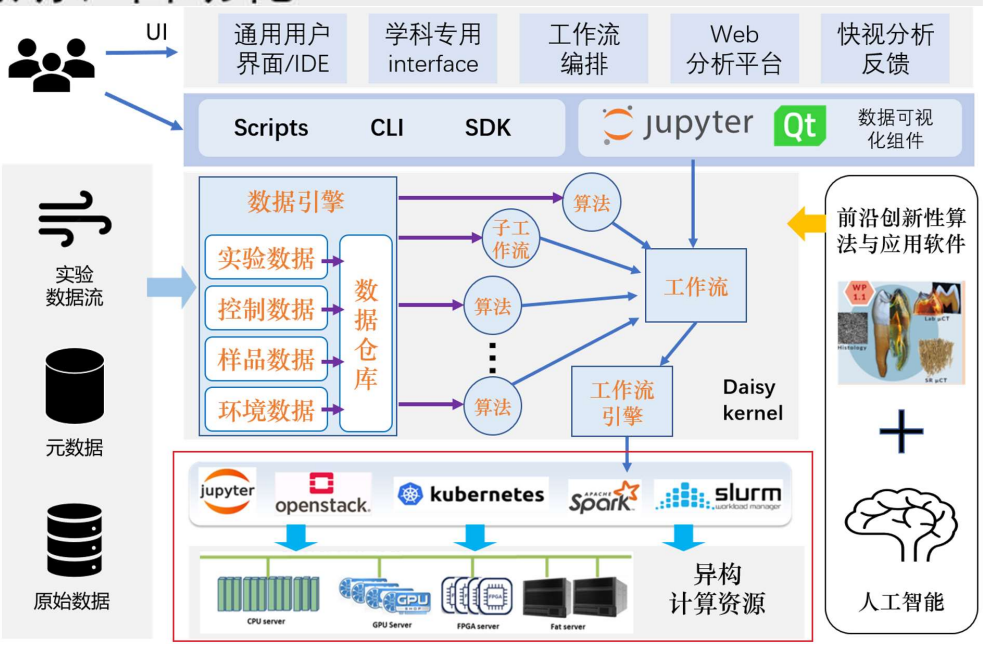
多模态、跨尺度、高帧率

面向多学科、多方法学

线站编号	每天峰值产生数据量（TB/Day）	每天平均产生数据量（TB/Day）
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
合计		805.00

面向先进光源的数据处理平台

- 计算平台提供底层核心服务：
  - 异构资源调度
  - 分析环境配置
  - 网络通信管理
  - 存储权限控制



提纲

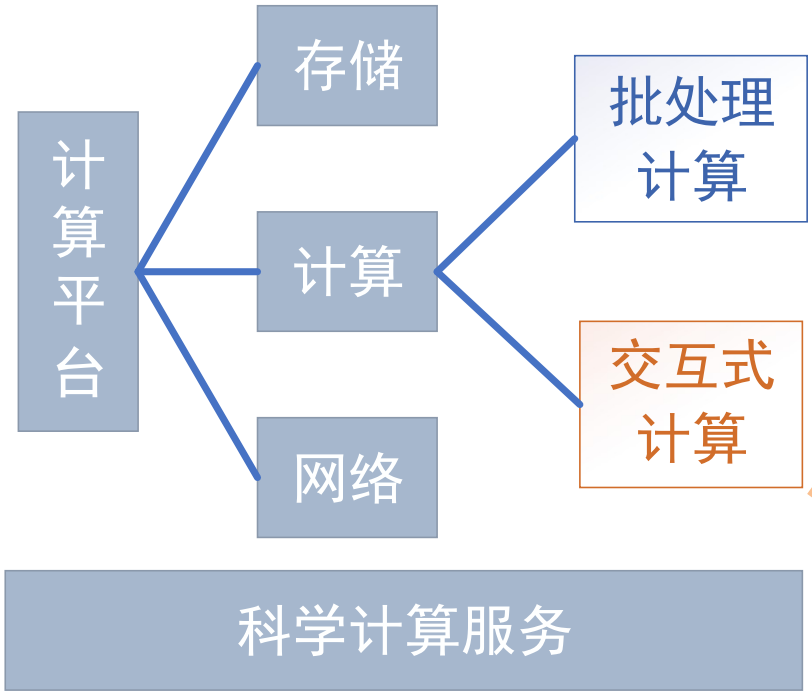
- 背景简介
- 光源学科计算的挑战
- 光源交互式计算平台设计
- 计算平台关键技术及进展
- 总结和展望

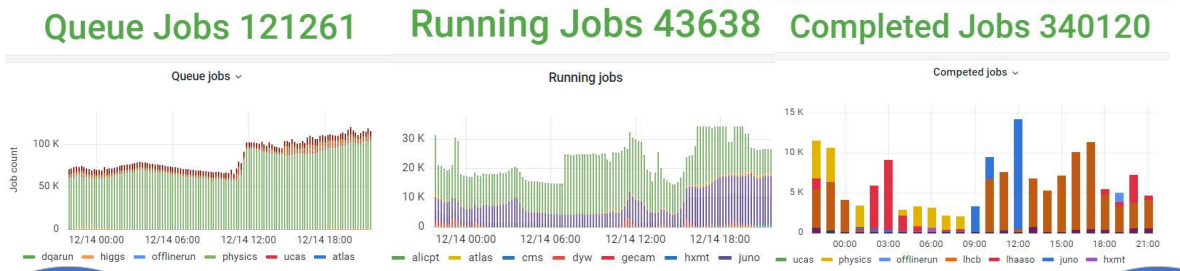
高能物理VS光源学科

	高能物理离线数据处理	HEPS光源在线/离线数据处理
需求	单一领域、 用户集中、 数据开放访问	14个线站涉及多学科领域、 多用户群体； 用户数据保密级别高
计算模式	开源物理软件 → cvmfs 计算环境统一， 离线批处理计算	开源软件和商业软件 → 灵活部署 Linux 、Windows，远程桌面 → 多OS、多环境支持 在线及离线计算→需要从小时级到秒级快速反馈

- 异构资源调度
- 分析环境配置
- 网络通信管理
- 存储权限控制

高能物理VS光源学科

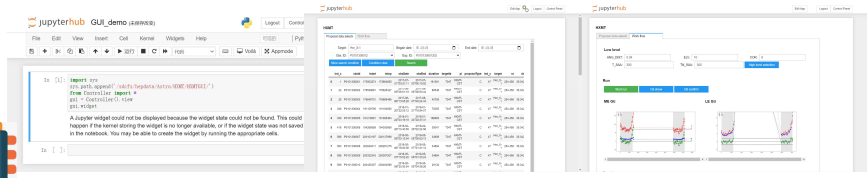




提交作业 → 队列排队 → 分配资源 → 运行作业 → 产生结果

高通量（无干预）

请求密集型，
单位时间内算的多



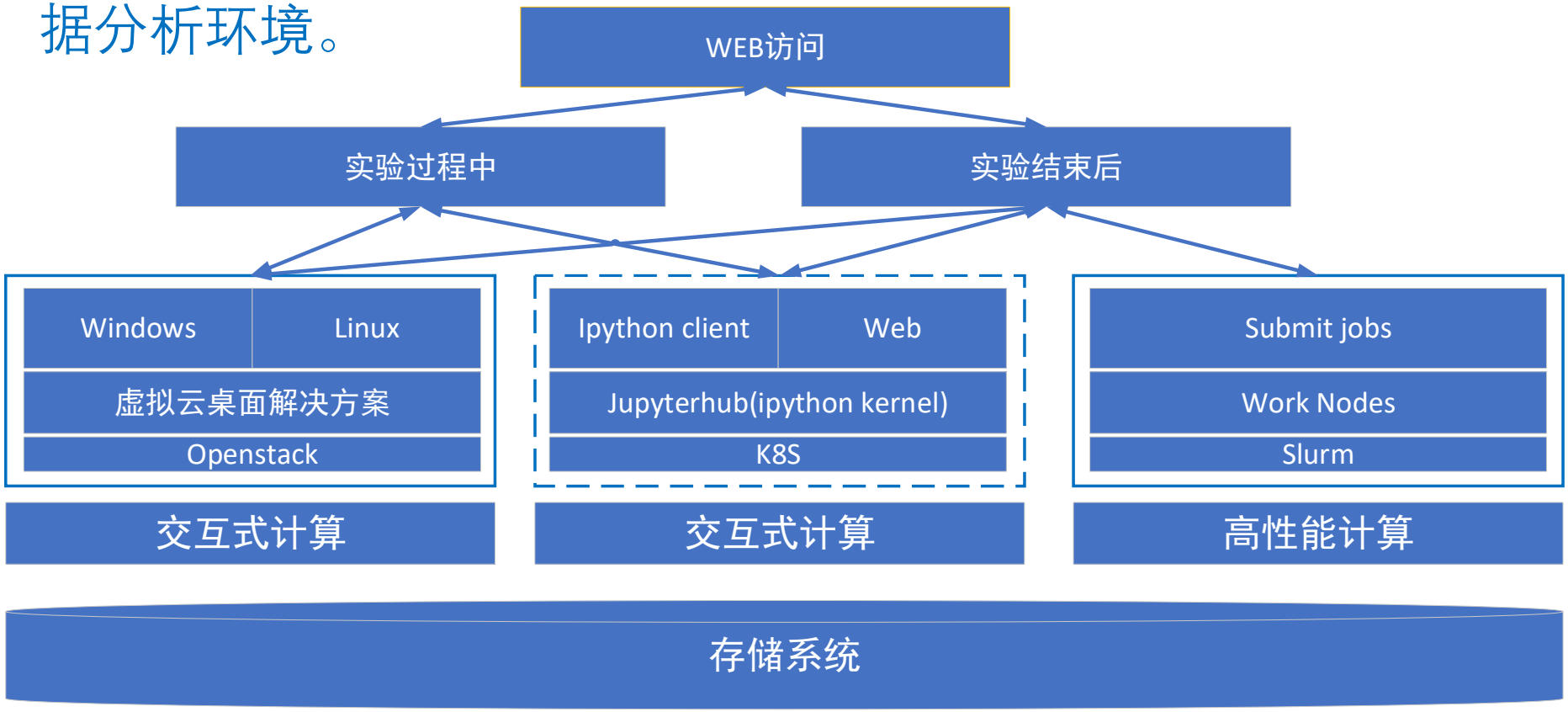
选择分析环境 → 初始化环境 → UI访问 → 交互式分析 → 呈现结果

环境特异性，
交互实时性

实时性（交互式）

光源交互式计算平台

- 以web为入口，提供随时随地可接入的，与集群环境同视图的数据分析环境。



提纲

- 背景简介
- 光源学科计算的挑战
- 光源交互式计算平台设计
- 计算平台关键技术及进展
- 总结和展望

平台关键技术

- 异构资源调度
  - 容器虚拟化
  - KVM虚拟化

  屏蔽底层基础设施的异构性支持资源隔离
- 定制化分析环境
  - 容器编排
  - 镜像管理

  实现异构资源集中管理与灵活调度

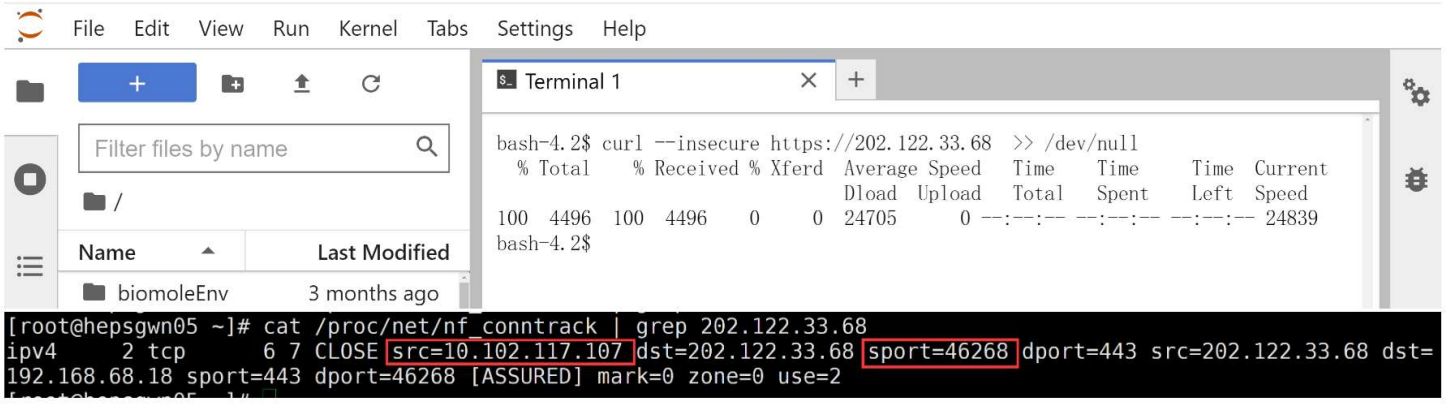
  为不同用户提供定制化计算环境

  按计算需求动态扩展算力

平台关键技术

- 网络通信管理
  - 网络策略控制
    - 黑名单策略，保护计算环境核心IP段，用户分析环境禁止访问数据中心核心IP段。
  - 用户网络行为反向溯源（基于OMAT平台实现）
    - 部署交互式计算平台物理节点路由信息采集模块。高频复制物理节点本机路由信息，记录物理节点pod IP、物理机端口映射、目的IP、目的端口信息。
    - SOC平台发现物理节点出现网络安全事件后，可以反向溯源具体的分析环境的使用用户。

	Time	objstat	h_dstip	h_srcip	h_sport	h_dport	poduser
>	Jul 5, 2023 @ 17:47:36.000	external	10.111.244.52	10.102.45.221	33610	8081	huqb
>	Jul 5, 2023 @ 17:47:36.000	external	10.111.244.52	10.102.45.221	33604	8081	huqb
>	Jul 5, 2023 @ 17:47:29.000	external	10.111.244.52	10.102.45.221	33604	8081	huqb

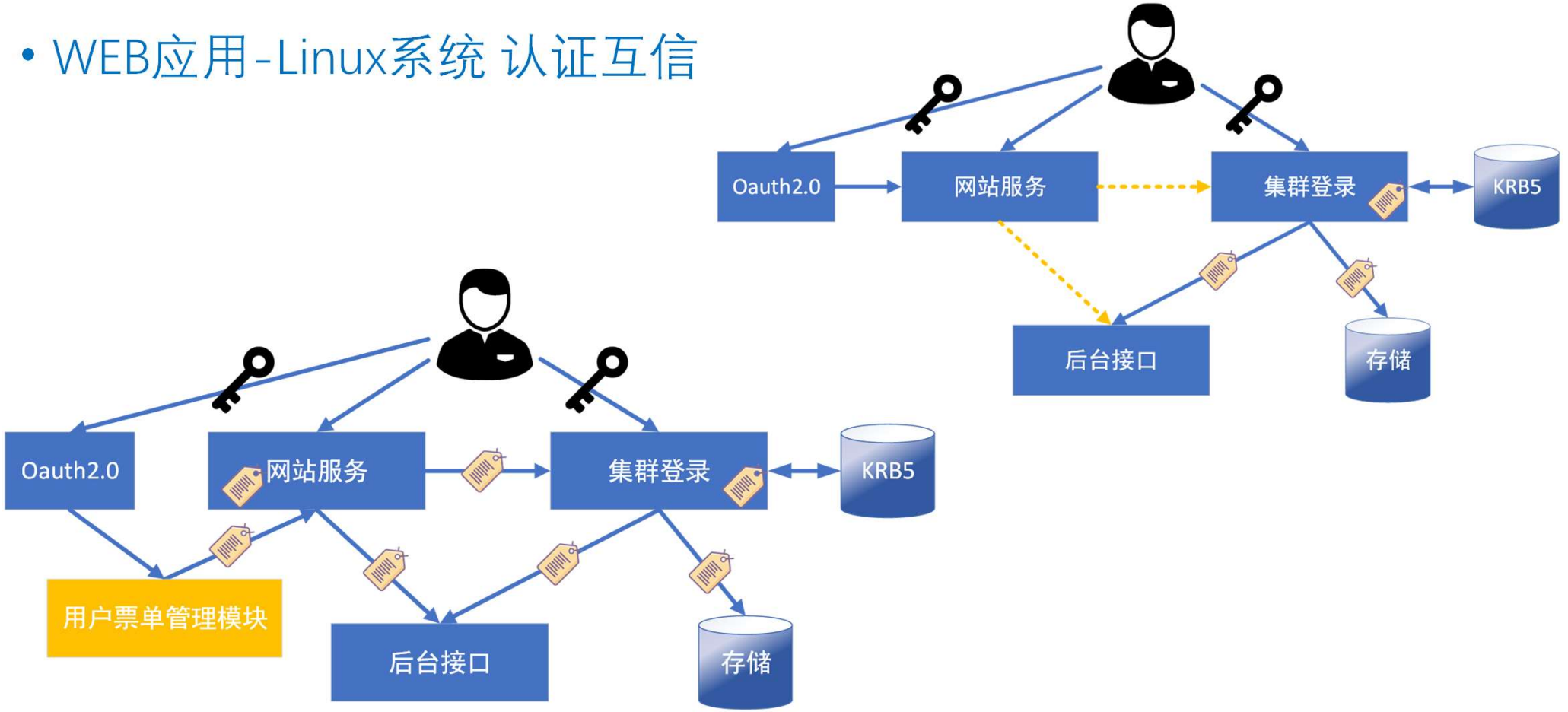


平台关键技术

- 存储权限控制
  - web访问和传统集群模式拥有相同的存储访问习惯。
  - S3 VS POSIX
    - 在S3中，ACL是通过为每个对象或存储桶定义访问策略来实现的。S3的访问策略是基于JSON格式的，可以定义不同的权限和访问控制规则。通过在策略中指定特定的用户、组或公共访问权限，可以控制对存储桶和对象的读取、写入和删除等操作。S3还提供了预定义的访问策略模板，可以简化ACL的配置。
    - 在POSIX中，ACL是通过扩展文件系统的权限模型来实现的。传统的POSIX权限模型包括所有者、所属组和其他用户的读、写和执行权限。而通过使用ACL，可以对更细粒度的用户和组设置访问权限。POSIX ACL使用一种特定的语法来定义权限规则，可以为每个文件或目录指定不同的用户和组的访问权限。
    - S3 适用性广　POSIX更符合用户习惯，性能更高。

平台关键技术

- WEB应用-Linux系统 认证互信

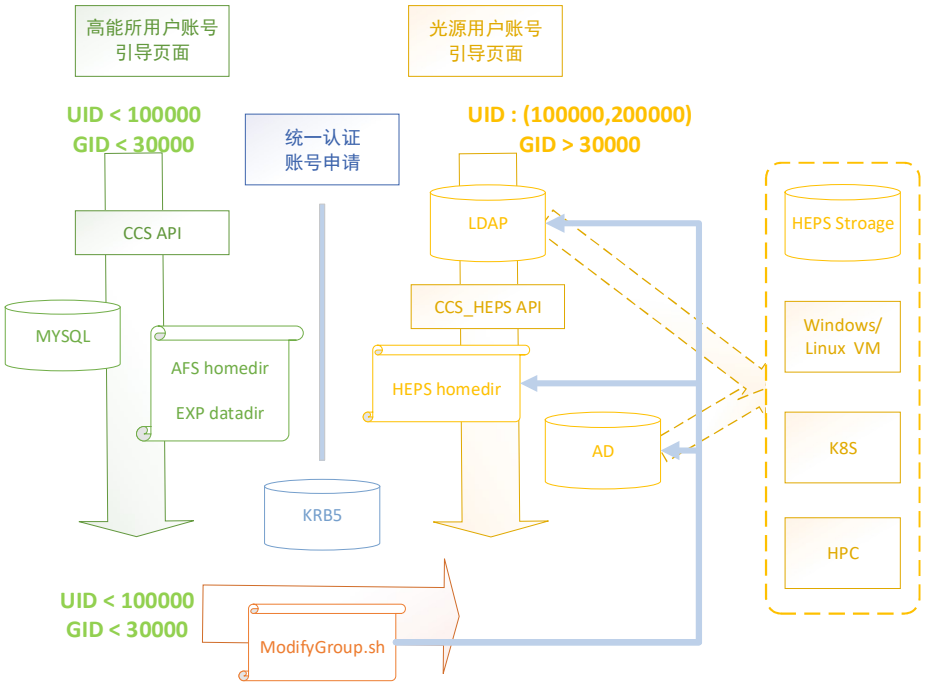


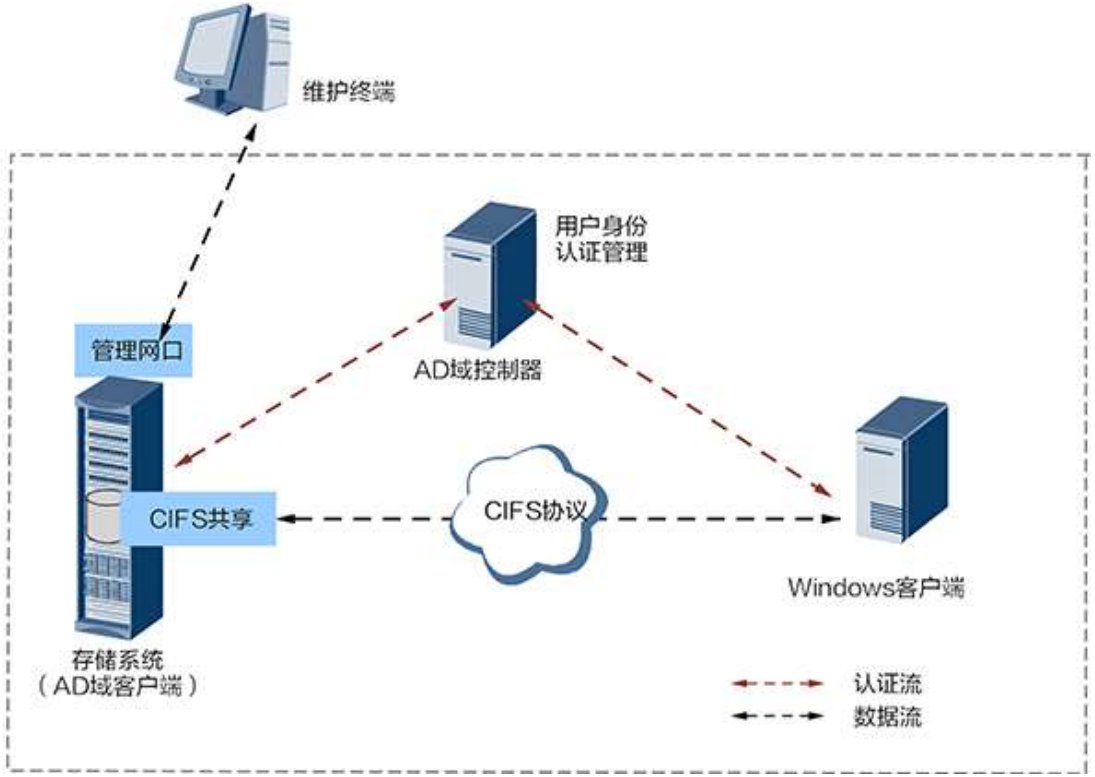
平台关键技术


登录 您正在使用高能所统一认证系统登录 HEPS交互式计算平台， 一键通行更轻松
jupyterhub
账号
huqb@ihep.ac.cn
密码
请输入高能所统一认证系统密码
登录
忘记密码?
欢迎使用HEPS交互式计算平台
Sign in with IHEPSSO / 使用高能所统一认证账号登陆
jupyterhub
Home
Token
Admin
huqb
Logout
File Edit View Run Kernel Tabs Settings Help
/
Name
Last Modified
biomoleEnv 5 months ago
huqbhomedir a year ago
pfiles a year ago
Untitled Fol... 2 years ago
Untitled Fol... 9 months ago
a 5 months ago
env_shell.json 5 months ago
Terminal 1
bash-4.2$ whoami
huqb
bash-4.2$ klist
Ticket cache: FILE:/tmp/krb5cc_10517
Default principal: huqb@IHEPKRB5
Valid starting Expires Service principal
07/11/2023 13:22:22 07/13/2023 13:22:22 krbtgt/IHEPKRB5@IHEPKRB5
renew until 07/18/2023 13:22:22
bash-4.2$ pwd
/home/huqb
bash-4.2$ id huqb
uid=10517(huqb) gid=600(u07) groups=600(u07),100(users),1088(newtest),1065(hepsbl),140(hxmt),1040(manager),1055(qc),1056(sdc)
bash-4.2$

平台关键技术

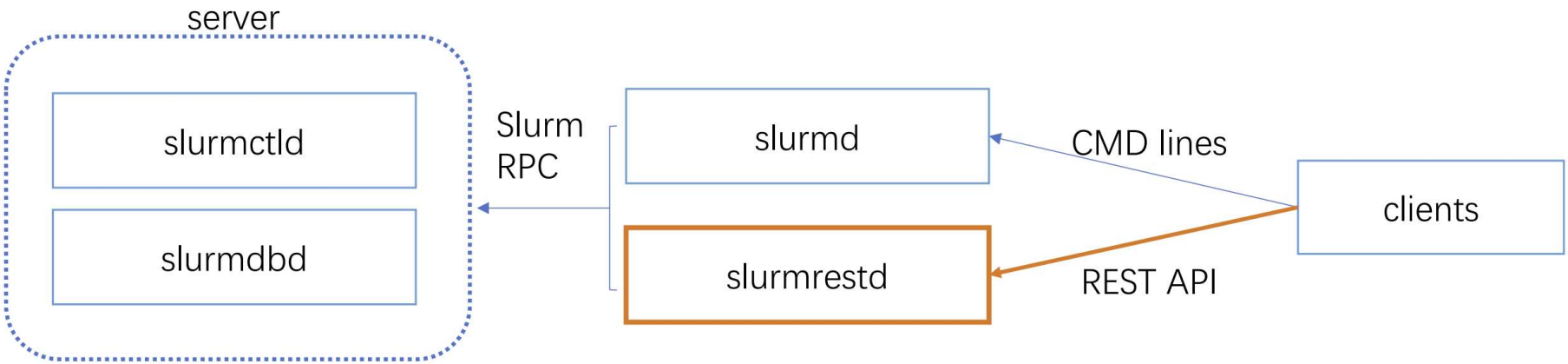
- Windows-Linux系统 认证互信





平台关键技术

- HPC作业提交 认证互信
  - 支持无状态的认证模式。



提纲

- 背景简介
- 光源学科计算的挑战
- 光源交互式计算平台设计
- 计算平台关键技术及进展
- 总结和展望

总结和展望

- 面向先进光源数据处理特点，设计建设了面向光源的交互式计算平台，整合容器、虚拟化、票单认证、网络路由等技术，保障了计算平台的安全、高效、灵活的特性。基本满足了光源的计算需求。
- 在应用接入方面，已支持windows、Linux虚拟云桌面，交互式分析支持CT成像分析、AlphaFold等应用，未来计划接入更多应用场景。

Thanks for your attentions!

谢谢！