



Back-n装置的科学数据管理与服务平台

报告人：胡鹏, 李斌斌, 李亚康, 樊瑞睿, 齐法制等

部门：中国散裂中子源计算机组

日期：2025. 7. 23

1 Back-n数据管理系统介绍

2 Back-n数据服务平台介绍

3 进展、计划与展望

Back-n实验数据管理挑战

Back-n等多台实验设施

实验数据特征

实验数据量大、复杂度高

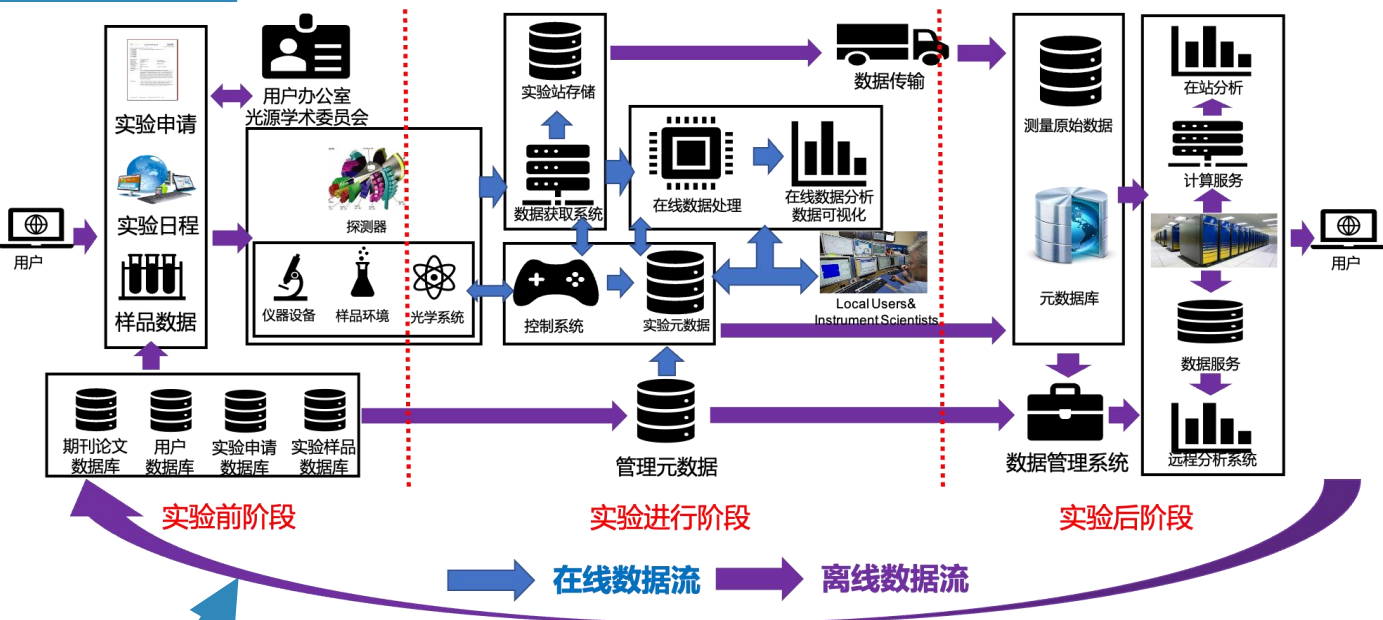
实验元数据类目众多

在线数据服务、应用

...

数据挑战

采集与分析
传输与存储
高效管理、挖掘、应用 → 科学发现



解决方案

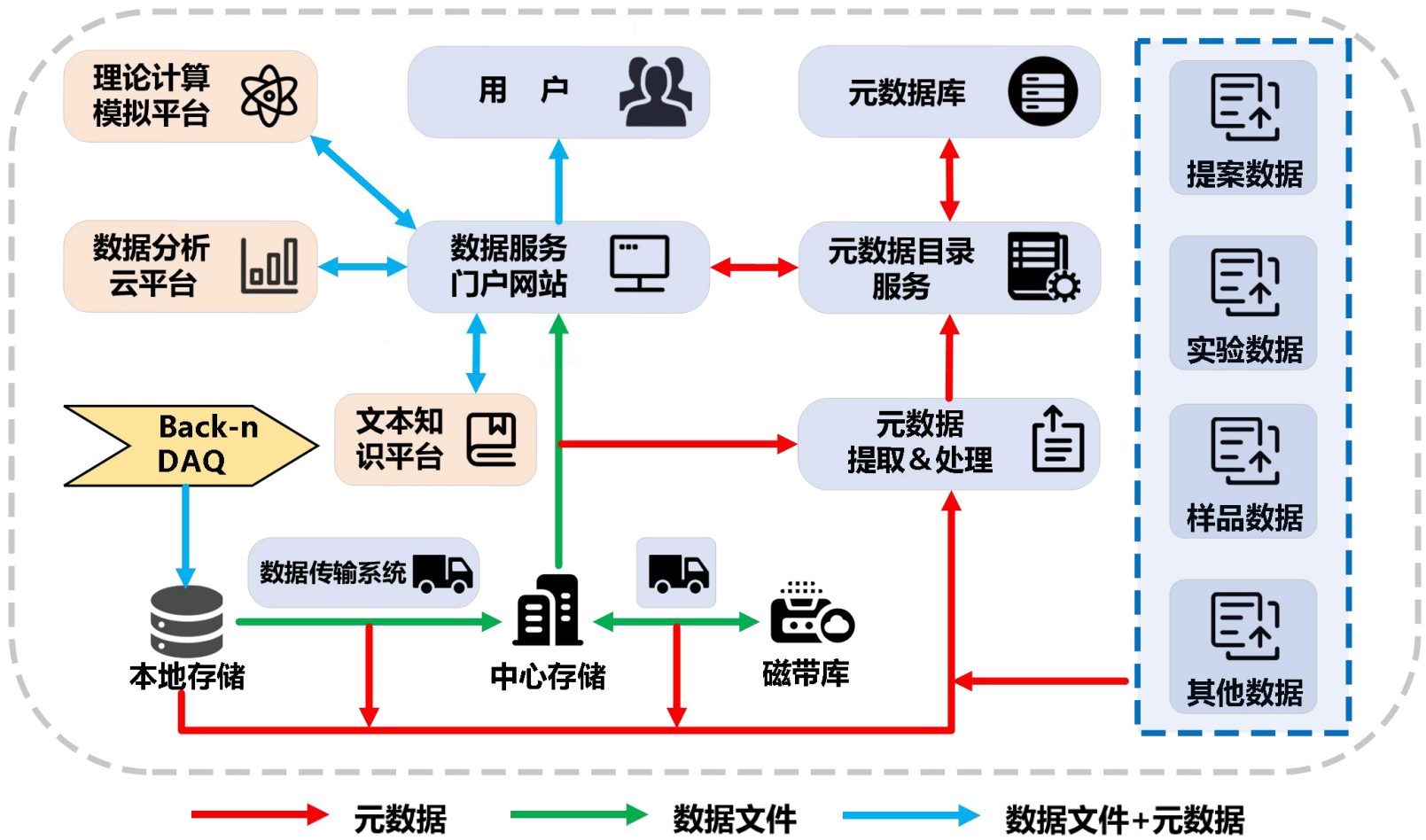


实验前→实验中→实验后全生命周期各个阶段数据、元数据的统一管理。

建设面向Back-n装置的数据管理和服务平台

科学数据管理系统——架构设计

基于数据流的系统架构设计



基于DOMAS框架的子系统设计构成:

元数据获取系统:

from DAQ, file, database...

元数据目录服务:

create, store, access, API...

数据传输系统:

multiple storage media, disk, tape...

数据服务平台:

separation of front-end and backend...

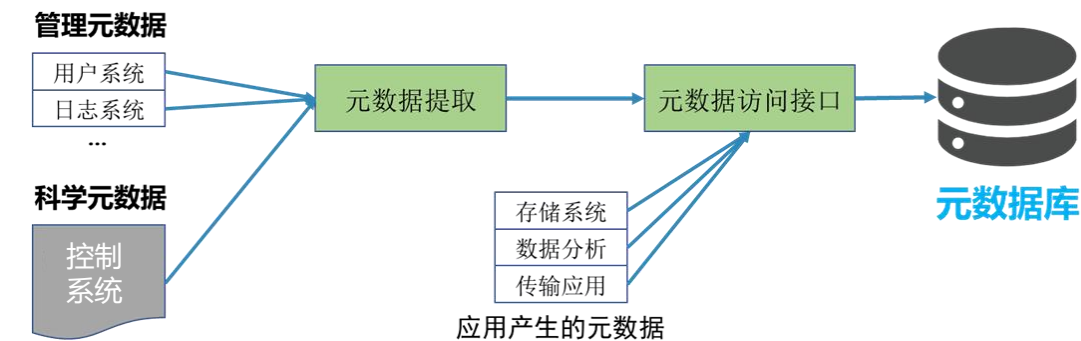
DOMAS框架已应用于BSRF、HEPS、HERD、CSNS、SHINE、HLS等装置

理论计算模拟平台、文本知识平台 on the way...

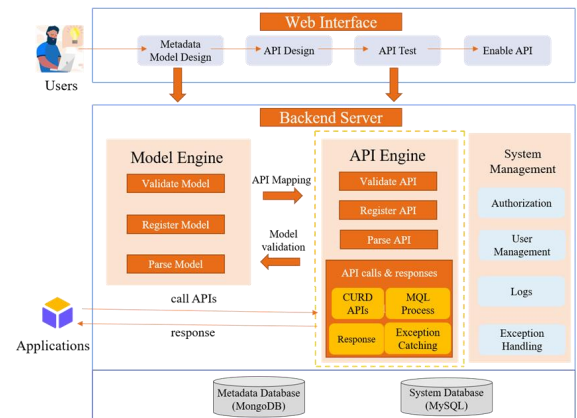
面向Back-n等设施的科学数据管理系统架构设计

科学数据管理系统——子系统

元数据采集系统



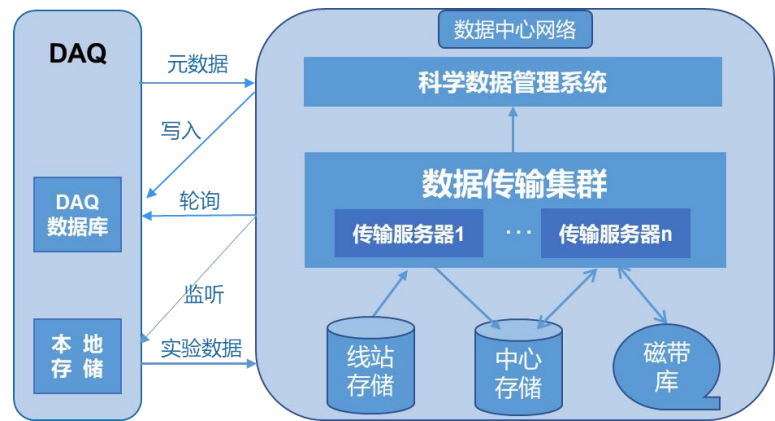
元数据目录服务



功能介绍:

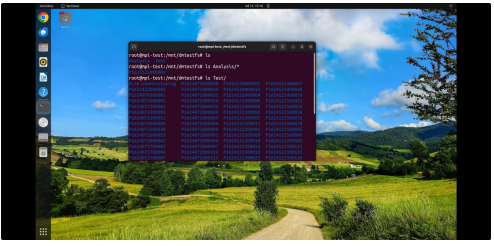
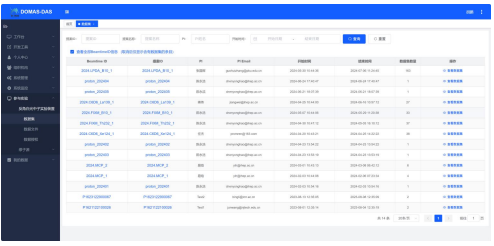
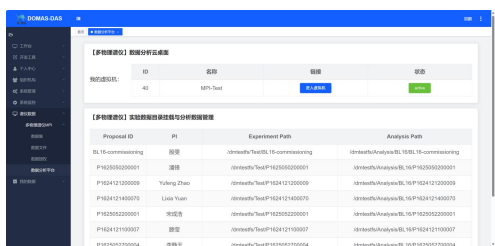
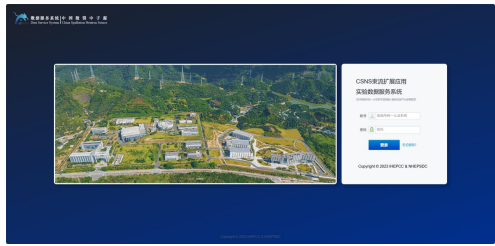
- 核心功能:** 元数据模型的可视化设计, 模型接口的自动化创建与应用
- Web可视化:** 模型设计、创建接口、接口测试、启用接口
- 核心模块:**
 - 模型引擎:** 校验、注册、解析
 - 接口引擎:** 校验、注册、解析、调用
 - 系统管理:** 用户、权限、异常、日志等
- 性能:** 灵活、动态修改和调用, 扩展性好

数据传输系统



已在**BRSF、HEPS、JUNO、LHAASO、ALICPT**部署验证, 稳定运行、效果良好。

数据服务系统

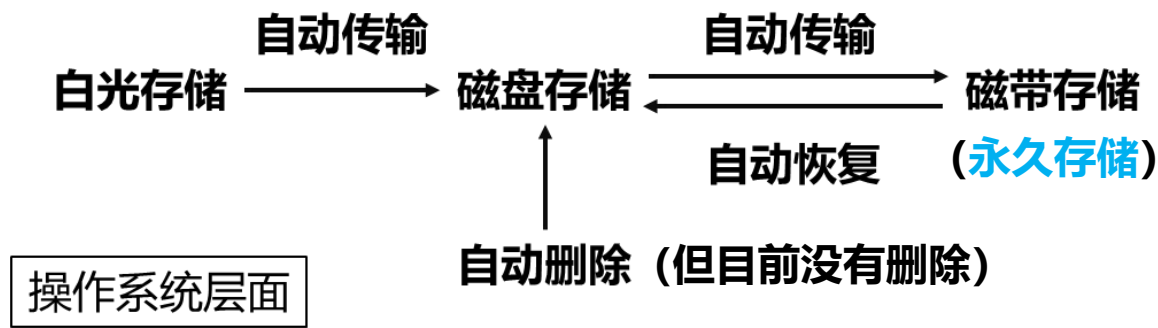


科学数据管理系统——传输与存储目录设计

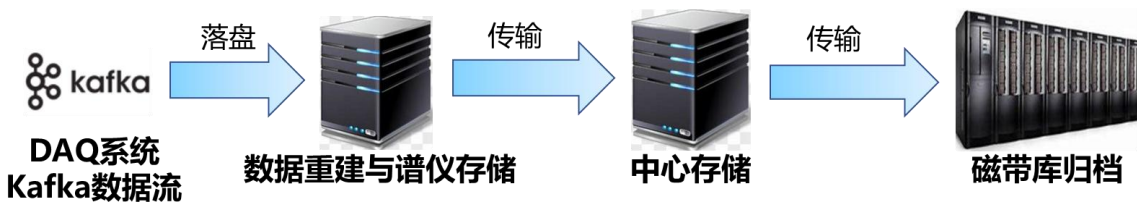
数据目录设计:

- 实验数据存储目录:
`.../raw/<instrument>/<proposalID>/<datasetID>`
- 分析数据存储目录:
`.../user/<instrument>/<proposalID>/<ana...>`
- 挂载实验数据、用户分析数据目录到云平台、计算平台

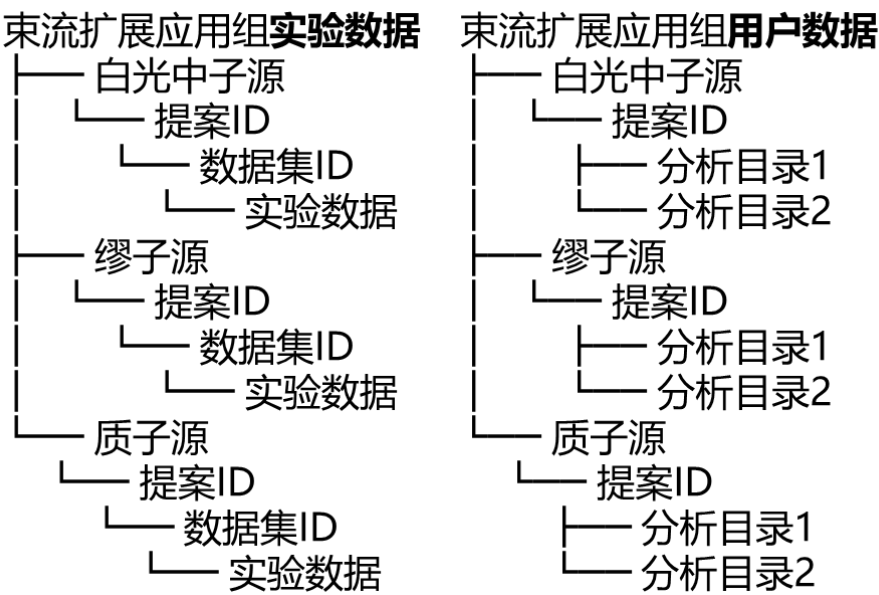
数据传输:



数据存储系统 (陈娟)



数据目录示例



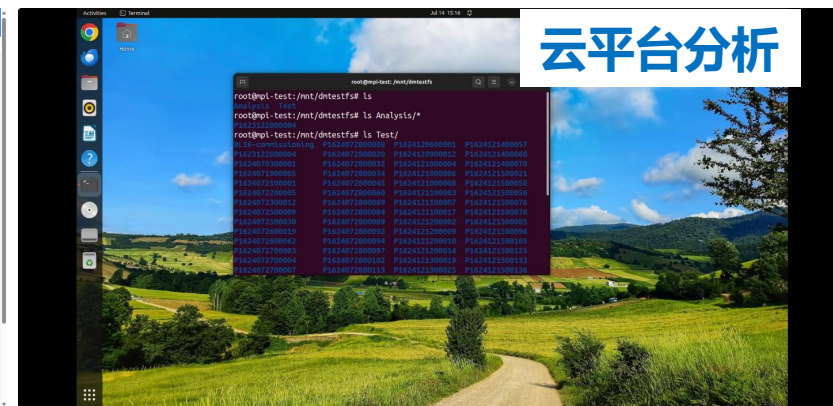
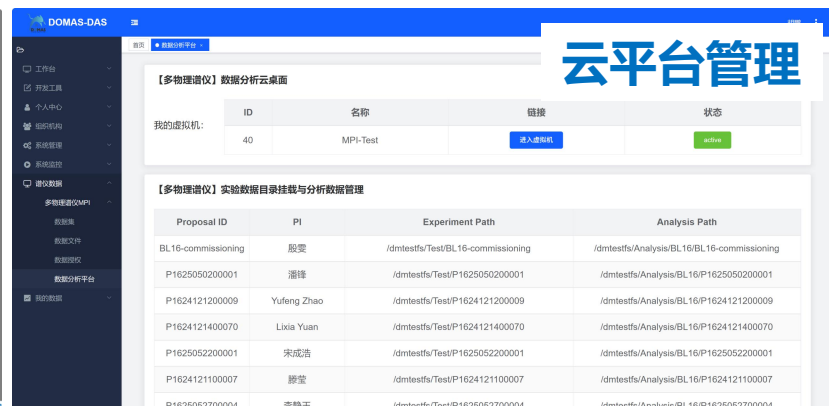
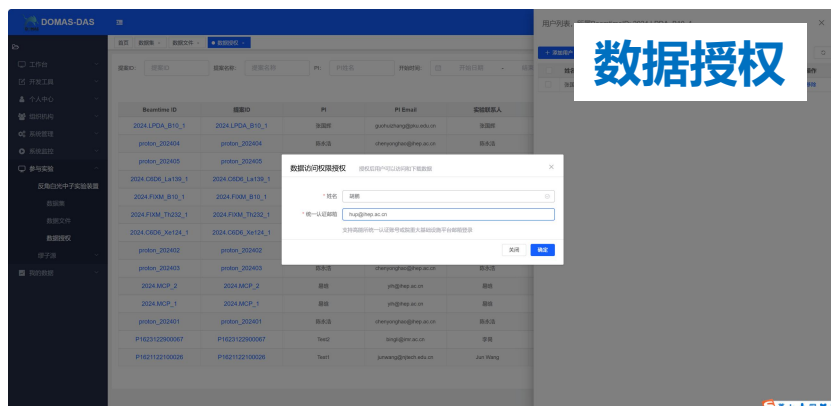
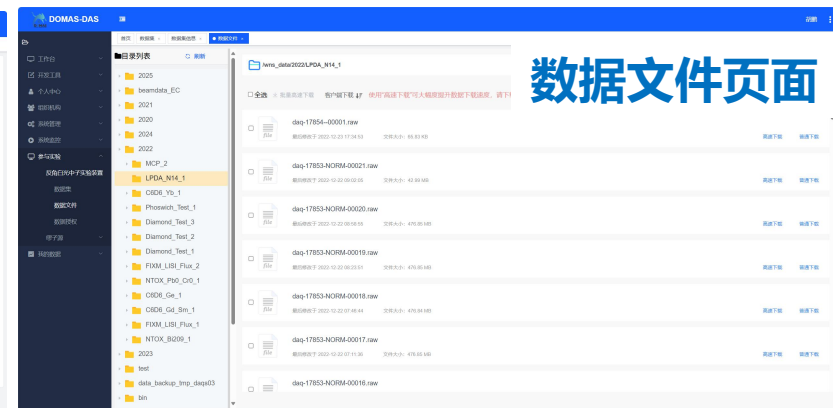
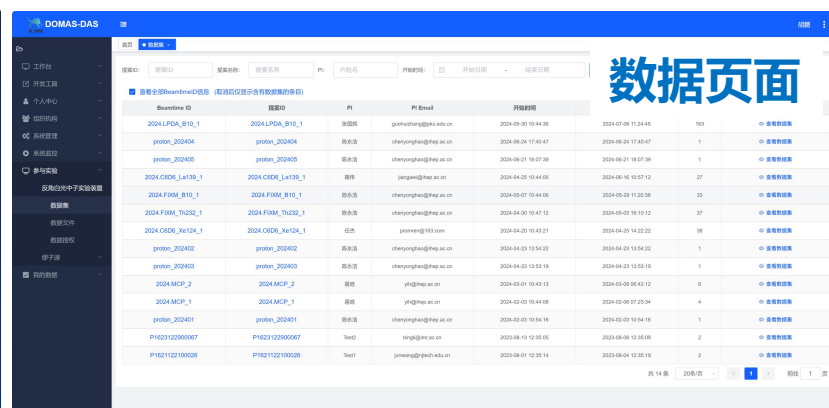
以数据存储目录设计为基础，通过ACL实现数据权限控制、保障数据安全。

1 Back-n数据管理系统介绍

2 Back-n数据服务平台介绍

3 进展、计划与展望

数据服务平台 <https://wnsdata.csns.ihep.ac.cn>



- 已接入高能所统一认证系统
- 不同角色的页面设计与开发
- 云平台分析管理与分析数据管理
- 接入存储系统、控制系统、用户系统、镭速客户端下载系统
- 其他定制化功能...

数据服务平台支持角色管理、数据检索、查看、下载、授权、可视化、在线分析等服务

数据服务平台——实验提案页面

查看实验信息

查看 BeamtimeID

仪器名称: Multi-Physics Instrument

提案名称: P1623073100040

提案描述: 通过高通量中子辐照研究材料在极端条件下的物理和化学行为

实验类型: CommissioningCalibrationDead AccessRepeat AccessXpress Access

开始日期: 2024-03-13 16:49:40

结束日期: 2024-03-13 20:32:27

负责人: 张立华

负责人邮箱: zhanglihua@pku.edu.cn

BeamtimeID: P1623073100040

备注:

查询

查看提案信息

查看提案

仪器名称: 中国散裂中子源

提案名称: Multi-Physics Instrument

提案编号: P1623073100002

提案描述: 高通量中子辐照材料在极端条件下的物理和化学行为

实验类型: CommissioningCalibrationDead AccessRepeat AccessXpress Access

开始日期: 2024-03-13 16:49:40

结束日期: 2024-03-13 20:32:27

负责人: 张立华

负责人邮箱: zhanglihua@pku.edu.cn

BeamtimeID: P1623073100040

备注:

查询

DOMAS-DAS

工作台

开发工具

个人中心

组织机构

系统管理

系统监控

参与实验

反角白光中子实验装置

数据集

数据文件

数据授权

缪子源

我的数据

首页

菜单管理

数据集

提案ID: 提案ID

提案名称: 提案名称

PI: PI姓名

开始时间: 开始日期 - 结束日期

查询

重置

检索

查看全部BeamtimeID信息 (取消后仅显示含有数据集的条目)

Beamtime ID	提案ID	PI	PI Email	开始时间	结束时间	数据集数量	实验数据	分析数据
2024.LPDA_B10_1	2024.LPDA_B10_1	张国辉	guohuizhang@pku.edu.cn	2024-05-30 10:44:36	2024-07-06 11:24:45	163	查看数据集	查看数据集
proton_202404	proton_202404	陈永浩	chenyonghao@ihep.ac.cn	2024-06-24 17:40:47	2024-06-24 17:40:47	1	查看数据集	查看数据集
proton_202405	proton_202405	陈永浩	chenyonghao@ihep.ac.cn	2024-06-21 18:07:39	2024-06-21 18:07:39	1	查看数据集	查看数据集
2024.C6D6_La139_1	2024.C6D6_La139_1	蒋伟	jiangwei@ihep.ac.cn	2024-04-25 10:44:00	2024-06-16 10:57:12	27	查看数据集	查看数据集
2024.FIXM_B10_1	2024.FIXM_B10_1	陈永浩	chenyonghao@ihep.ac.cn	2024-05-07 10:44:06	2024-05-29 11:20:38	33	查看数据集	查看数据集
2024.FIXM_Th232_1	2024.FIXM_Th232_1	陈永浩	chenyonghao@ihep.ac.cn	2024-04-30 10:47:12	2024-05-03 16:10:12	37	查看数据集	查看数据集
2024.C6D6_Xe124_1	2024.C6D6_Xe124_1	任杰	promren@163.com	2024-04-20 10:43:21	2024-04-25 14:22:22	38	查看数据集	查看数据集
proton_202402	proton_202402	陈永浩	chenyonghao@ihep.ac.cn	2024-04-23 13:54:22	2024-04-23 13:54:22	1	查看数据集	查看数据集
proton_202403	proton_202403	陈永浩	chenyonghao@ihep.ac.cn	2024-04-23 13:53:19	2024-04-23 13:53:19	1	查看数据集	查看数据集
2024.MCP_2	2024.MCP_2	易晗	yih@ihep.ac.cn	2024-03-01 10:43:13	2024-03-08 06:42:12	6	查看数据集	查看数据集
2024.MCP_1	2024.MCP_1	易晗	yih@ihep.ac.cn	2024-02-03 10:44:08	2024-02-06 07:23:34	4	查看数据集	查看数据集
proton_202401	proton_202401	陈永浩	chenyonghao@ihep.ac.cn	2024-02-03 10:54:16	2024-02-03 10:54:16	1	查看数据集	查看数据集
P1623122900067	P1623122900067	Test2	bingli@imr.ac.cn	2023-08-13 12:35:05	2023-08-08 12:35:09	2	查看数据集	查看数据集
P1621122100026	P1621122100026	Test1	junwang@njtech.edu.cn	2023-08-01 12:35:14	2023-08-04 12:35:19	2	查看数据集	查看数据集

共 14 条

20条/页

1

前往 1 页

查看实验数据集

查看分析数据集

数据服务平台——数据集页面

DOMAS-DAS

胡鹏

工作台

开发工具

个人中心

组织机构

系统管理

系统监控

参与实验

我的数据

数据集名称: 数据集名称 获取时间: 开始日期 - 结束日期 查询 重置

检索

批量高速下载 批量普通下载 客户端下载

批量数据集下载

	数据集名称	获取时间	联系人	联系人邮箱	文件数
☐	19353	2024-06-16 18:57:11	蒋伟	jiangwei@ihep.ac.cn	3
☐	19352	2024-06-16 18:55:52	蒋伟	jiangwei@ihep.ac.cn	7
☐	19344	2024-06-15 23:55:26	蒋伟	jiangwei@ihep.ac.cn	446

文件列表

批量高速下载 批量普通下载 输入关键字搜索文件 重置

文件名	数据量	保存时间	数据类型	操作
daq-19344-NORM-00001.raw	0.47GB	2024-06-15 18:56:55	wns_data	普通下载 高速下载
daq-19344-NORM-00002.raw	0.47GB	2024-06-15 18:59:55	wns_data	普通下载 高速下载
daq-19344-NORM-00003.raw	0.47GB	2024-06-15 18:50:45	wns_data	普通下载 高速下载
daq-19344-NORM-00004.raw	0.47GB	2024-06-15 19:01:15	wns_data	普通下载 高速下载
daq-19344-NORM-00005.raw	0.47GB	2024-06-15 19:01:53	wns_data	普通下载 高速下载
daq-19344-NORM-00006.raw	0.47GB	2024-06-15 19:02:32	wns_data	普通下载 高速下载
daq-19344-NORM-00007.raw	0.47GB	2024-06-15 19:03:11	wns_data	普通下载 高速下载
daq-19344-NORM-00008.raw	0.47GB	2024-06-15 19:03:51	wns_data	普通下载 高速下载
daq-19344-NORM-00009.raw	0.47GB	2024-06-15 19:04:30	wns_data	普通下载 高速下载
daq-19344-NORM-00010.raw	0.47GB	2024-06-15 19:05:10	wns_data	普通下载 高速下载

单个数据集下载

操作

[普通下载](#) [高速下载](#)

[普通下载](#) [高速下载](#)

[普通下载](#) [高速下载](#)

网页下载

客户端下载

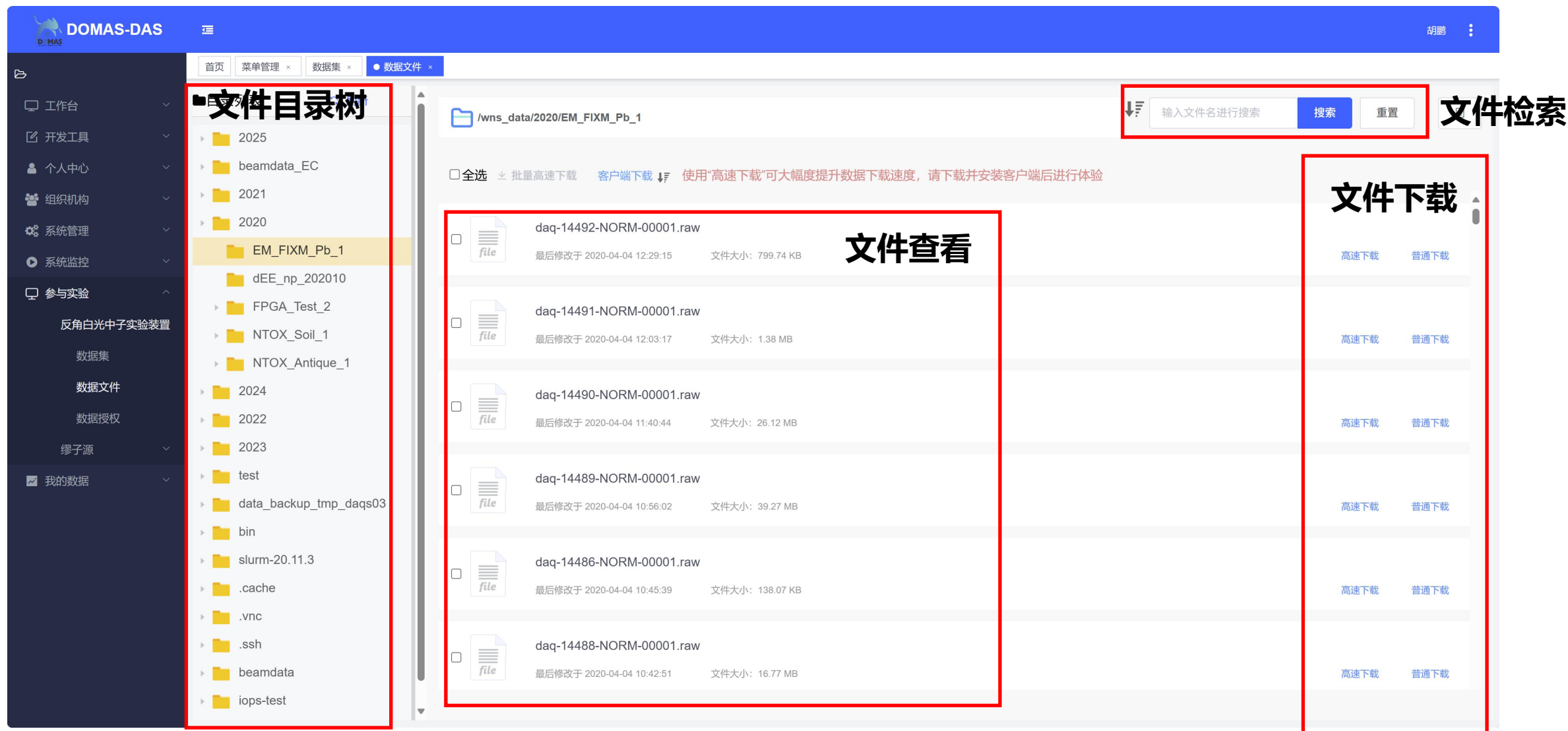
单个文件下载

共 446 条 10条/页 1 2 3 4 5 6 ... 45 前往 1 页

☐	19342	2024-06-15 18:56:20	蒋伟	jiangwei@ihep.ac.cn	42	普通下载 高速下载
☐	19339	2024-06-14 23:35:39	蒋伟	jiangwei@ihep.ac.cn	271	普通下载 高速下载
☐	19336	2024-06-14 20:34:21	蒋伟	jiangwei@ihep.ac.cn	211	普通下载 高速下载

接入镭速客户端下载后，下载速度将达到百Mb/s。

数据服务平台——数据文件页面



数据服务平台——数据授权页面

Beamtime ID: 提案ID: 提案名称: PI: PI Email: 开始时间: -

Beamtime ID	提案ID	PI	PI Email	实验联系人	实验联系人邮箱	开始时间	结束时间	操作
P0121122400087	P0121122400087	潘锋	panfeng@pkusz.edu.cn	潘锋	panfeng@pkusz.edu.cn	2021-12-29 10:52:27		数据授权
BL01-commissioning	BL01-commissioning	Ke Yubin	keyb@ihep.ac.cn	Ke Yubin	keyb@ihep.ac.cn	2020-10-06 10:55:01		数据授权
P0119121600004	P0119121600004	Xinhui Lu	xinhui.lu@cuhk.edu.hk	Xinhui Lu	xinhui.lu@cuhk.edu.hk	2020-07-13 17:53:55		数据授权
P0120062100001	P0120062100001	郑众	zhengz@iccas.ac.cn	郑众	zhengz@iccas.ac.cn	2020-07-11 16:48:28		数据授权
P0120052600002	P0120052600002	李准	lizhun@ihep.ac.cn	李准	lizhun@ihep.ac.cn	2020-07-10 19:52:05		数据授权
P0120060900001	P0120060900001	汪勇	wangyongac@gmail.com	汪勇	wangyongac@gmail.com	2020-07-09 12:55:52		数据授权
P0119121300052	P0119121300052	肖荫果	xiaoyg@pkusz.edu.cn	肖荫果	xiaoyg@pkusz.edu.cn	2020-07-09 12:23:10		数据授权
P0119121700003	P0119121700003	赵环宇	zhaohuanyu@jlu.edu.cn	赵环宇	zhaohuanyu@jlu.edu.cn	2020-07-08 10:26:24		数据授权
P0119121200001	P0119121200001	沈瑞	496661103@qq.com	沈瑞	496661103@qq.com	2020-07-06 10:23:17		数据授权

数据授权

用户列表, 所属BeamtimeID: P0121122400087

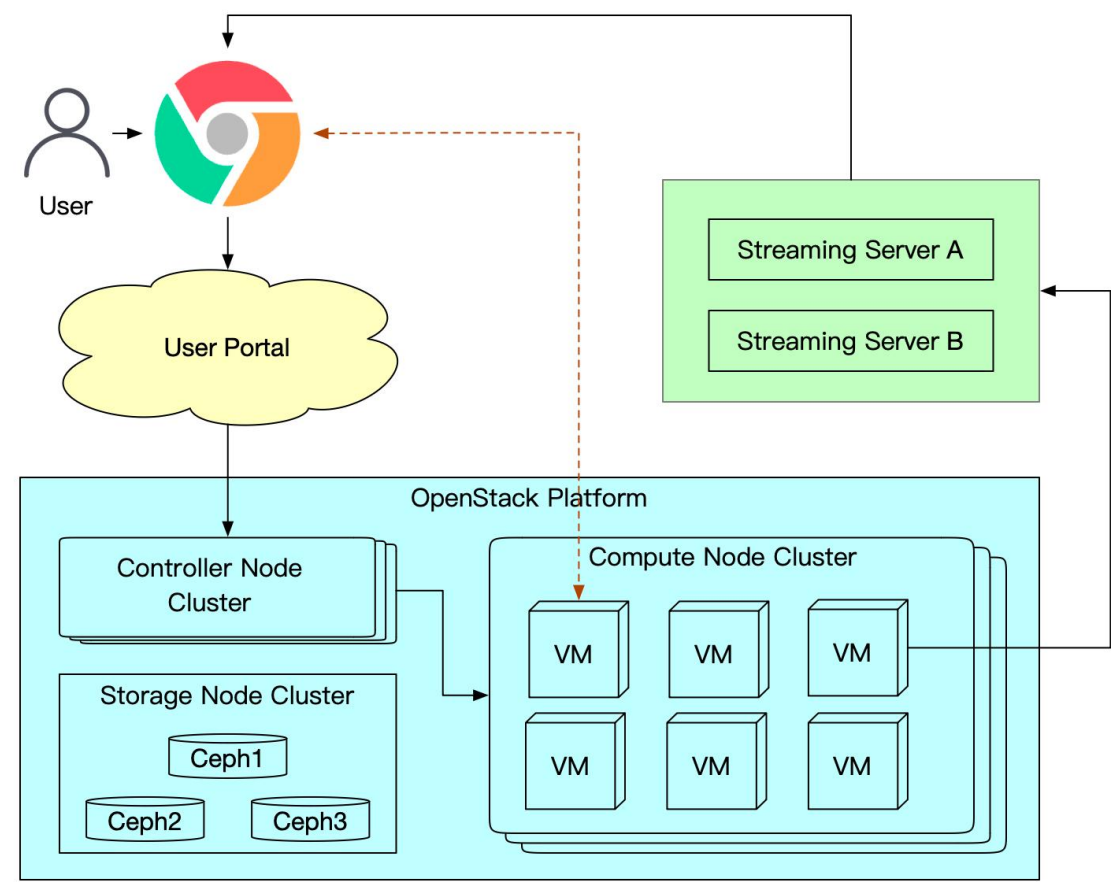
仅需要输入待授权用户的高能所
统一认证邮箱和姓名即可完成数
据授权

☐	姓名	邮箱	Web访问	操作
☐	潘锋	panfeng@pkusz.edu.cn	是	移除
☐	张明建	zhangmingjian@pku.edu.cn	是	移除

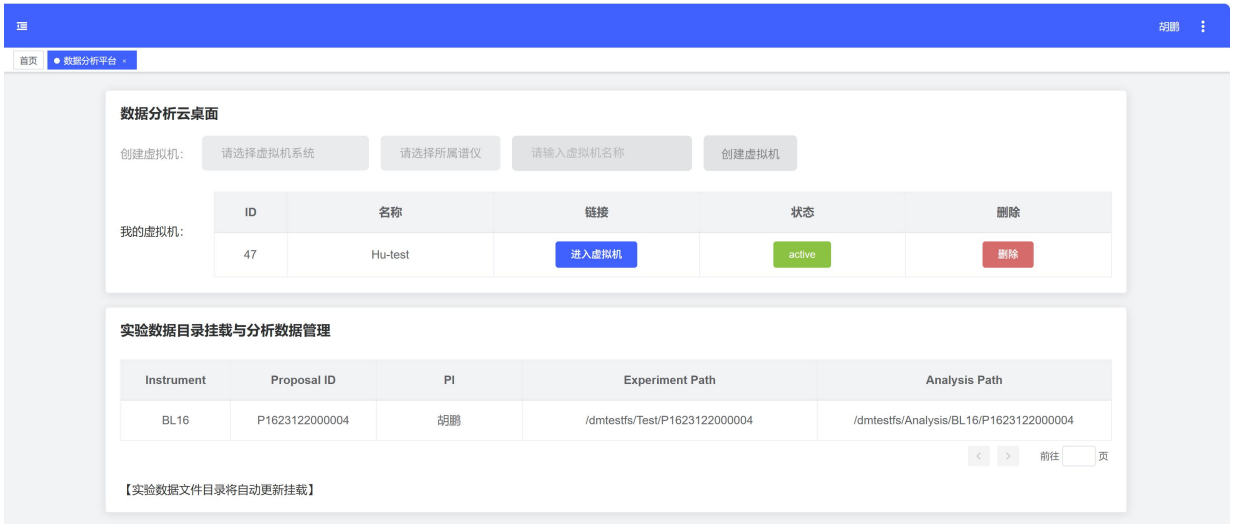
添加或移除用户

数据服务平台——数据分析

基于OpenStack的云桌面（李斌斌）

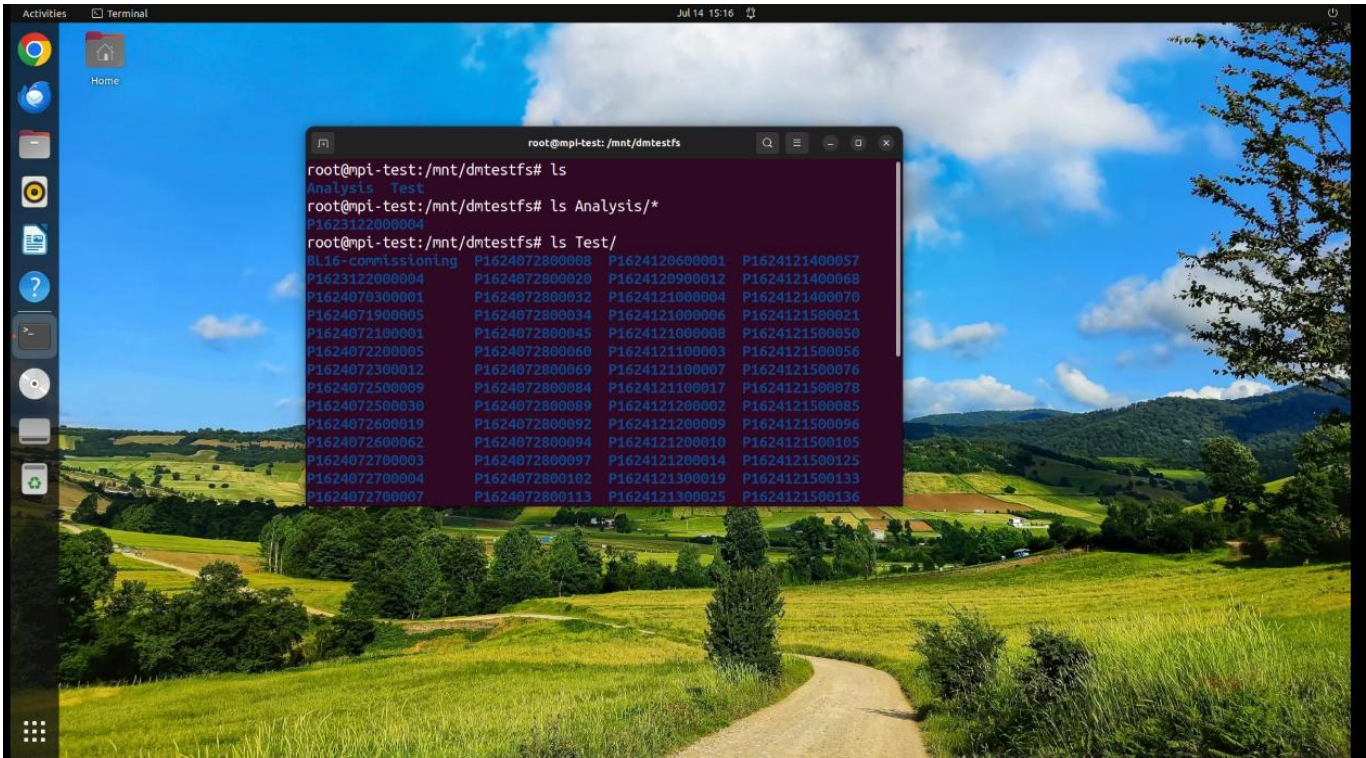


数据分析云桌面页面



- 云桌面的创建、进入、删除
- 实验数据目录、分析数据目录的自动挂载
- 系统环境、分析软件工具预先安装...

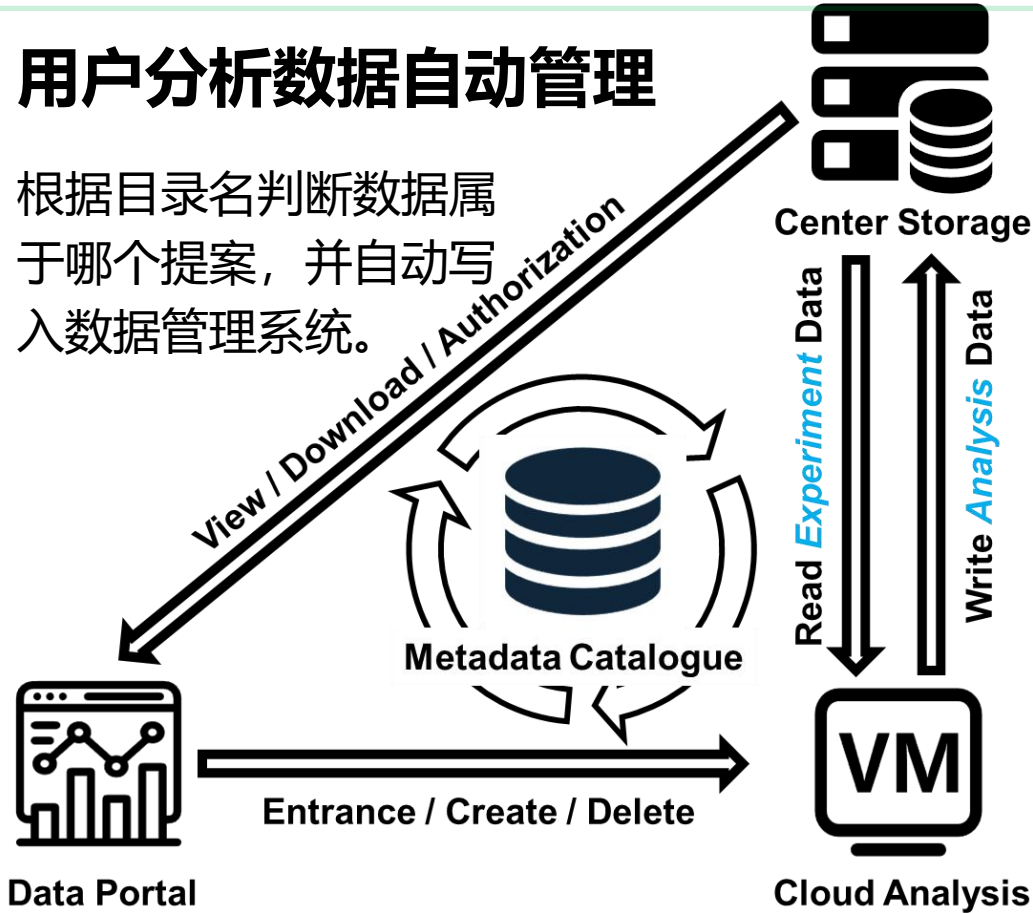
数据分析云桌面



- 操作系统：Ubuntu 桌面版 • 分辨率：2K • 低延迟
- 配置：16 Core, 64 G, NVIDIA-RTX-A4000

用户分析数据自动管理

根据目录名判断数据属于哪个提案，并自动写入数据管理系统。



将分析后数据放入指定用户分析数据目录，将自动写入到用户的提案分析数据集中。支持数据服务页面的查看、下载、授权等。

数据服务子系统——个性化功能

DOMAS-DAS

胡鹏

工作台

首页

管理员首页

参与谱仪

我的数据

首页

管理员首页

欢迎回来！ 胡鹏

欢迎使用中国散裂中子源CSNS数据服务系统

如遇到任何问题请联系计算机网络组：胡鹏（电话：15600290501 邮箱：hup@ihep.ac.cn）

多物理谱仪近期实验数据

本装置近期实验数据

Beamtime ID	提案ID	PI	更新时间	操作
P1624072800138	P1624072800138	Ji-jung Kai	2024-12-28 16:10:30	查看数据集
P1624072700025	P1624072700025	张景峰	2024-12-28 07:49:06	查看数据集
P1624072500009	P1624072500009	Jie Yan	2024-12-27 08:31:03	查看数据集
P1624072800092	P1624072800092	张显程	2024-12-26 08:46:26	查看数据集
P1624072200005	P1624072200005	马东	2024-12-25 08:33:10	查看数据集

查看更多

多物理谱仪实验数据统计

本装置实验数据统计

提案数目

989

数据集数目

7,592

数据文件数目

70,908

近3日新增数据集数目

135

公告

通知 11月添加主页新功能。

通知 10月在MPI开展联调测试。

通知 9月在MPI开展传输系统测试。

通知 8月在MPI开展数据获取系统测试。

通知 1月在MPI开展数据服务测试。

帮助文档

使用指南 CSNS介绍

快速访问

CSNS用户系统 中科院高能所

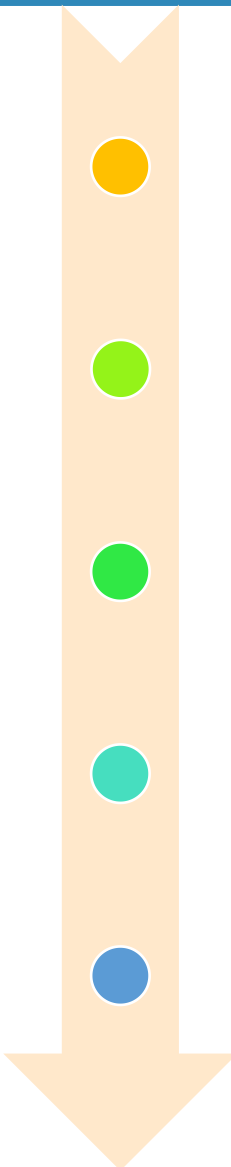
可根据装置需求快速个性化性能定制

1 Back-n数据管理系统介绍

2 Back-n数据服务平台介绍

3 进展、计划与展望

进展与计划

- 
- 2024. 10  数据服务系统测试
 - 2025. 7  数据服务系统开放公网访问, <https://wnsdata.csns.ihep.ac.cn>
 - 2025. 8  部署测试完整的数据管理系统, 并对接装置DAQ、云分析平台、存储系统等
 - 2025. 10  开放用户使用
 - Future  其他个性化功能定制

展望——理论计算模拟平台（在研）

GUI页面计算模拟作业提交（以中子散射模拟为例）

编辑并提交Neutron Cinema模拟作业

Test Model

提交作业

查看 workflow

输入代码

模型代码: Test

选择一个节点...

object {3}

Input {6}

Name: test

Seed: 123456

WlMean: 3.39

WlRange: 0.3

Slit [3]

0: 50

1: 50

2: -1000

Source [3]

0: 50

1: 50

2: -2000

Simulation {1}

HPC {4}

host: login1.csns.ihep.ac.cn

partition: NEU

node: 1

ntask_per_node: 12

Analysis {0}

(清空 object)

```
# Simulation Name: "test"
# Time: (time)
# ID: (simu_id)

from model import MySim
from Cinema.Prompt.gun import UniModeratorGun

sim = MySim(seed=123456)
sim.makeWorld()

gun = UniModeratorGun()
gun.setWlMean(3.39)
gun.setWlRange(0.3)
gun.setSlit([50,50,-1000])
gun.setSource([50,50,-2000])

# vis or production
if False:
    sim.show(gun, 1)
else:
    sim.simulate(gun, 1e6)

destination = 0
psd1 = sim.gatherHistData('psd1', dst=destination)
psd2 = sim.gatherHistData('psd2', dst=destination)
wlspec = sim.gatherHistData('wlspec', dst=destination)
if sim.rank == destination:
    psd1.plot(show=False, log=True)
    psd2.plot(show=False, log=True)
    wlspec.plot(show=False, log=True)
psd1.save('psd1.h5')
psd2.save('psd2.h5')
```

实验数据

模拟数据

联合管理、分析

形成数据库

模拟 workflow 自动调度、实时监控

Newest Workflows

NC0000000255

COMPLETED

ID: 1614

Initialize_nc_model

write_input_files_to_file_server

input_metadata_to_database

exec_simulation_on_hpc

collect_output

simulation_metadata_to_database

NC0000000254

COMPLETED

ID: 1608

Initialize_nc_model

write_input_files_to_file_server

input_metadata_to_database

exec_simulation_on_hpc

collect_output

simulation_metadata_to_database

NC0000000253

COMPLETED

ID: 1602

Initialize_nc_model

write_input_files_to_file_server

input_metadata_to_database

exec_simulation_on_hpc

collect_output

simulation_metadata_to_database

NC0000000252

COMPLETED

ID: 1600

Initialize_nc_model

write_input_files_to_file_server

input_metadata_to_database

exec_simulation_on_hpc

collect_output

simulation_metadata_to_database

number of fireworks

hours ago

number of fireworks

days ago

Database snapshot

	Fireworks	Workflows
ARCHIVED	0	0
FIZZLED	56	53
DEFUSED	0	0
PAUSED	0	0
WAITING	207	0
READY	0	3
RESERVED	3	0
RUNNING	34	30
COMPLETED	1,319	209
TOTAL	1,619	295

Summary Reports

See a visual dashboard.

Get a report of all jobs for the past:

30 minutes

24 hours

提供非常低成本的模拟作业执行及其与实验数据的融合管理

展望——数据分析 workflows、平台、智能体（在研）

分析数据

标准化数据分析

- 规约、去噪、滤波
- 谱学解析、图像解析
-

标准化:

1. 数据格式标准化

2. 数据分析软件标准化

3. 数据分析流程标准化

UI

通用用户界面/IDE

学科专用 interface

Workflow编排

Web 分析平台

快视分析反馈

Scripts

CLI

SDK

jupyter

Qt

数据可视化组件

实验数据流

元数据

原始数据

数据引擎

实验数据

控制数据

样品数据

环境数据

数据仓库

子工作流

算法

工作流

工作流引擎

Daisy kernel

jupyter

openstack

kubernetes

spark

slurm

异构计算资源

前沿创新性算法与应用软件

+

人工智能

① 知识库更新

实验提案

实验过程

科学目标

etc

知识库

RAG

Knowledge Graph

③ 检查实验数据

查询指令生成

指令执行

感知器

④ 智能标注实验数据

代码生成和执行

执行器

数据处理算法软件

机器学习、垂直专用模型

Micro-Xray CT

Segmentation

Phase-field Damage, $\phi=1$

⑤ 校验和保存

人在环路

AI

数据智能体

API接口

LLM大脑

② 识别AI任务

重建、寻峰

降噪、压缩

分割、etc

input object

→

process

→

output object

➢ input: 数据、参数、上一级的输出。...

➢ process: 加减乘除、积分、拟合、重建。...

➢ output: 数据、图像、模型。...

run_this_first

→

branching

→

branch_a

→

follow_branch_a

→

join

←

branch_b

DAISY工作流编排

数据分析智能体

自动任务模式(TODO)

DOMAS 数据管理系统

其他来源数据

数据融合工作流

感知器Agent

知识问答Agent

工作流Agent

执行器Agent

查询器Agent

DAISY数据分析工作流

DFT/MD等模拟计算程序

数据融合、对齐、补全

提供标准化、自动化、智能化的数据分析

展望——核领域数据库建设

基于元素周期表的数据库模板:

Exactly Match Elements ▾ E.g. Te O ✕ Search

1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	57-71 La-Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	89-103 Ac-Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112 Cn						

57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb	71 Lu
89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No	103 Lr

面向材料、物理、化学、核科学等领域。

基于数据库技术自定义工具、模版：



提供个性化的数据库/集服务工具、平台



感谢各位聆听!

请批评指正!

