

中国科学院高能物理研究所  
*Institute of High Energy Physics*  
*Chinese Academy of Sciences*



高能所计算中心  
IHEP Computing Center

# Daisy : HEPs数据处理 软件框架

胡誉

**ON BEHALF OF IHEP-CC/HEPS-CC**

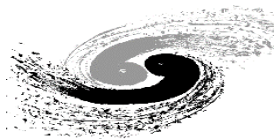
2022年08月18日

---

第三届 (2022年) 高能物理计算暑期学校



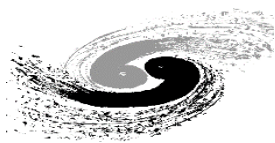
- **科学软件:使命和任务**
- **Daisy设计和开发现状**
- **学科软件团队合作开发与集成 (生态发展)**



- **科学软件:使命和任务**

- Daisy设计和开发现状

- 学科软件团队合作开发与集成 (生态发展)



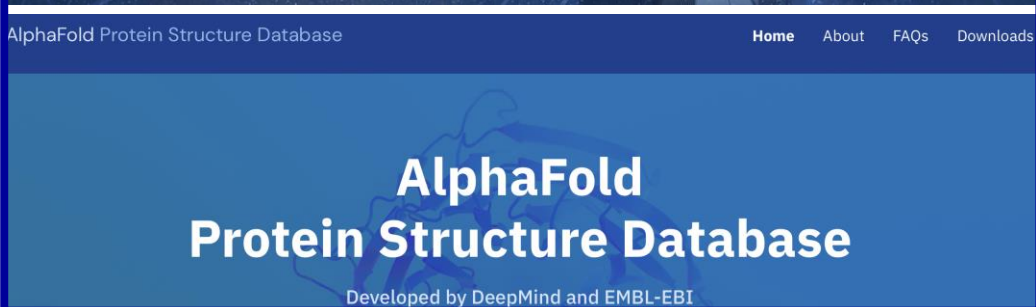
# 用户大科学装置时代的大数据科学

高能所計算中心  
IHEP Computing Center

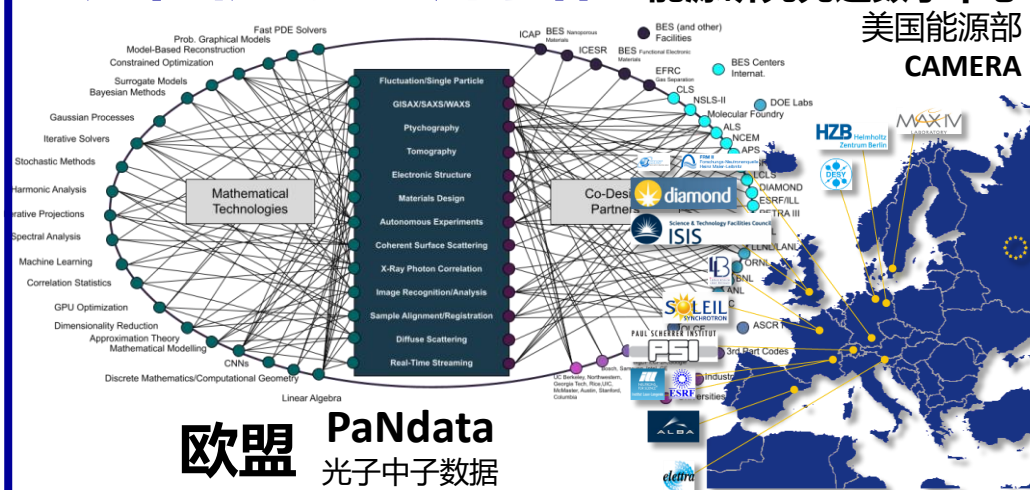
## 用户类大科学装置



## 数据/数据库、算法、软件、算力

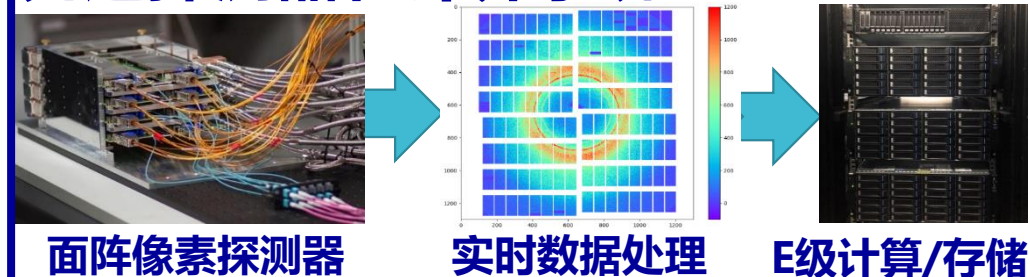


## 全方位协同创新网络



欧盟 PaNdata  
光子中子数据

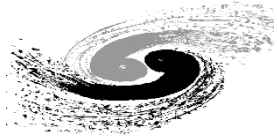
## 先进探测器和计算系统



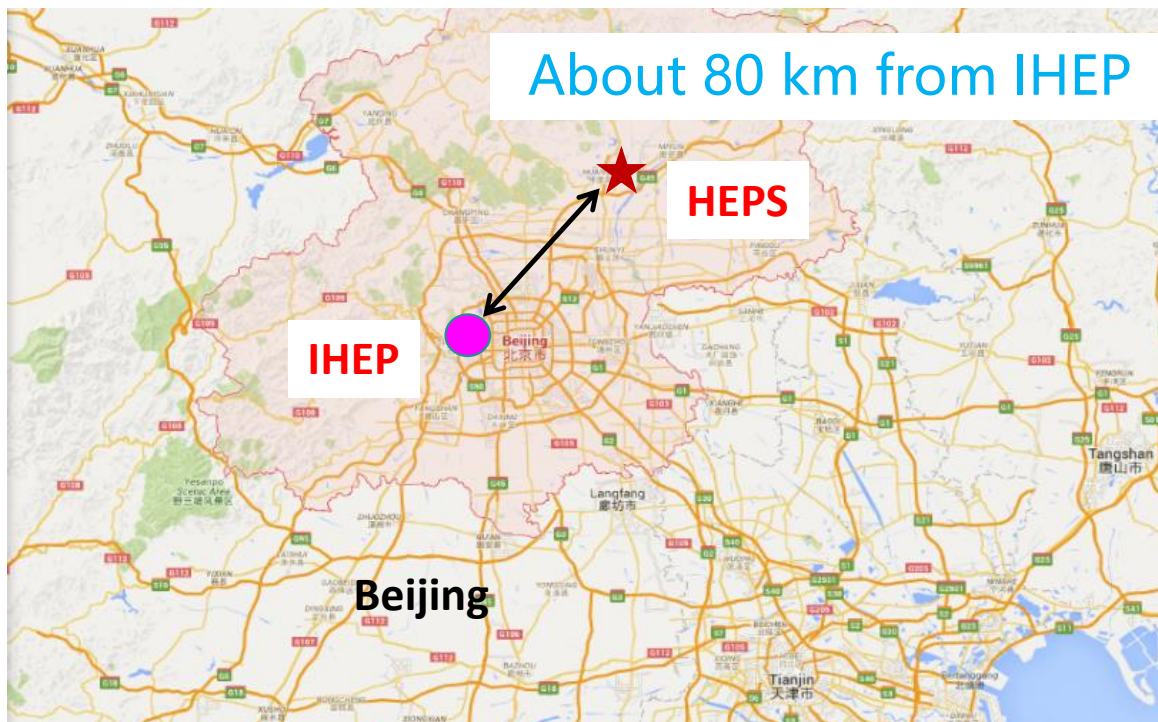
APS Scientific Computing Strategy

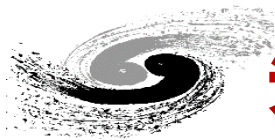
September 24, 2021

复杂科学研究需要新科学范式



- 第四代同步辐射光源，束流能量达 6GeV，亮度达  $10^{22}$  phs/s/mm<sup>2</sup>/mrad<sup>2</sup>/1% B.W.，发射度优于 0.06 nm·rad
- 首期建设14条用户线站和1条测试线，建设周期2019年中 – 2025年底。将来HEPS将容纳~90多条光束线站。





# 先进光源及其科学软件挑战

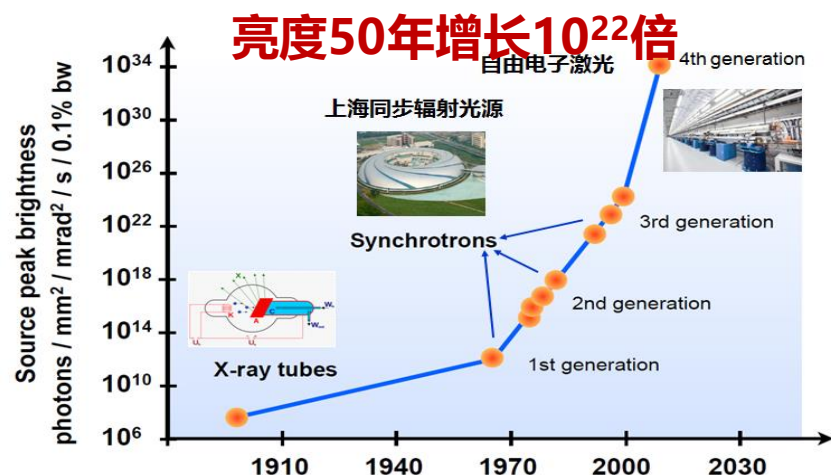


高能所计算中心  
IHEP Computing Center

基础和前沿交叉研究的重要平台，解决多领域重大科学问题的国之重器

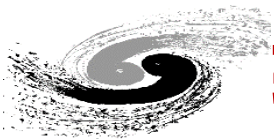
- 实验技术 → 高通量、多模态、超快频率、原位及动态加载
- 实验控制 → 复杂、高精度、多尺度、快反馈
- 实验数据 → 超大数据量（单线站数百PB/年）
- 数据处理和分析算法、软件多样化
- 用户广泛

实验过程的复杂度、实验数据高通量，对科学软件提出了巨大挑战  
需要各阶段互相协作的整体软件框架支持，实现实验过程的自动化



HEPS一期数据统计表（天）

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



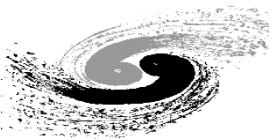
## ■非常必要，否则先进光源实验开展会面临巨大问题！

- 当前软件完全依赖欧美国家，面临卡脖子的风险
- 老一代光源采用的分离式、手动模式无法满足需求
- 当前国内外均无完全满足先进光源的全生命周期软件框架
- 复杂的控制、超大数据量处理需要高性能、自动化、联动的全生命周期软件系统支持

## ■尽快开展，时间窗口短！

- 我国光源类设施发展迅速，部分已有设施已经凸显此问题，未来几年刚性需求将更加明显
- 目前国内外处于几乎同一起跑线
- 但国外（主要是欧美国家）已经启动了相关项目
- 中国先进光源需要自己的软件框架和系统支持！

|                | 一代/二代/三代                         | 先进光源                                             |
|----------------|----------------------------------|--------------------------------------------------|
| 实验控制           | 简单                               | 复杂：多模态、动态加载                                      |
| 数据采集           | 简单                               | 高通量、多数据源                                         |
| 数据通量           | Kb/s-Mb/s                        | Gb/s-Tb/s                                        |
| 在线分析           | 线站单机（CPU）                        | 实时反馈，异构+集群处理（CPU/GPU）                            |
| 数据管理           | 无                                | 元数据+实验数据                                         |
| 数据共享           | 无                                | 需要，国家科学数据管理办法                                    |
| 数据访问           | U盘/移动硬盘拷贝                        | 在线网络访问（单次实验达数百TB）                                |
| 数据分析           | 办公室电脑                            | 大规模异构计算/并行计算，AI                                  |
| 学科软件           | 国外/开源/商业/自研                      | 国外/开源/商业/自研，AI                                   |
| 全生命周期科学软件框架及应用 | 需要，但不是必须的，实验控制复杂度、数据处理需求等可以无关联实施 | 必须，多模态控制、超高通量、实时反馈、数据处理需要软件集成，支持线站自动化高效运行和数据分析处理 |



# 国外趋势：软件框架和系统研发以应对挑战

国际上积累了多种领域软件，但只能解决个别专向问题。先进光源**已经在布局发展集成式的软件系统。**

## ■ 美国：



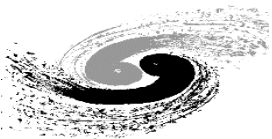
能源部**五大光子科学装置**2020.12联合启动**先导项目**，成立“Data Solution Task Force”

开展：Bluesky + Xi-Cam + Analysis Tools联合研制  
实验控制 数据采集 数据分析 数据管理 分析工具 集成

## ■ 欧洲：



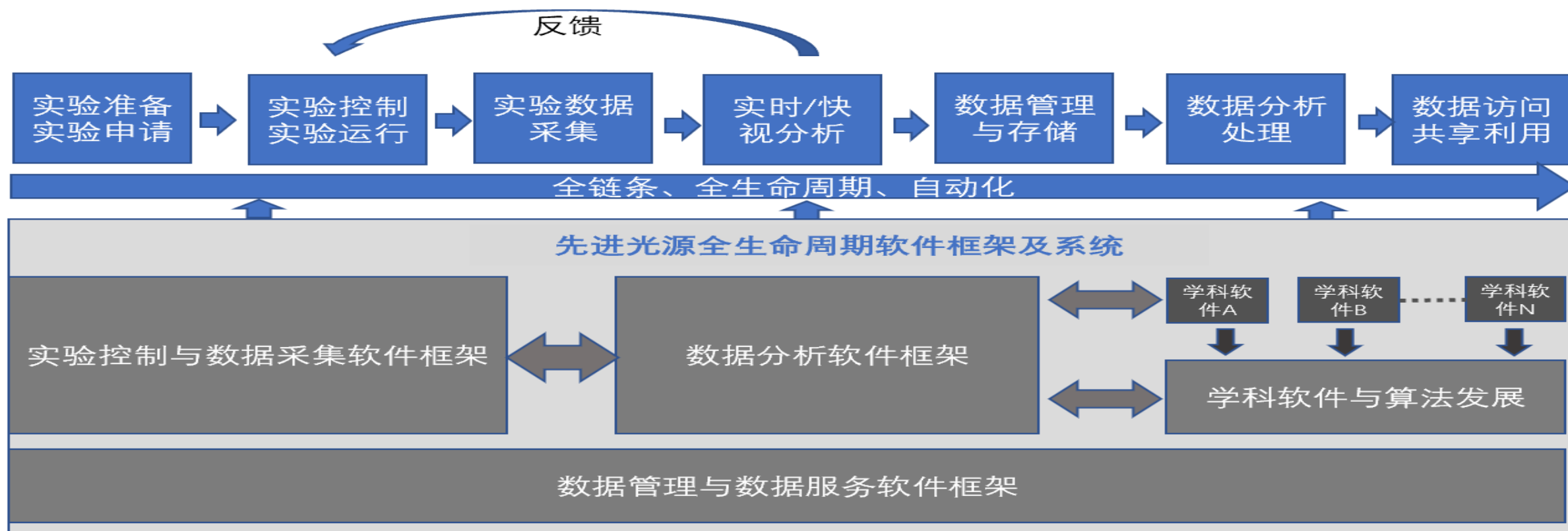
欧盟第七框架计划资助的PaNdata（“光子和中子数据基础设施”项目），**汇集了14个欧洲研究设施**，以合作解决共性问题，为数据交互提供标准和工具，并促进开展多装置联合实验



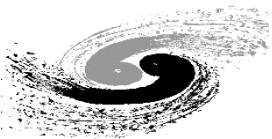
# 覆盖实验过程和实验数据全生命周期



高能所計算中心  
IHEP Computing Center



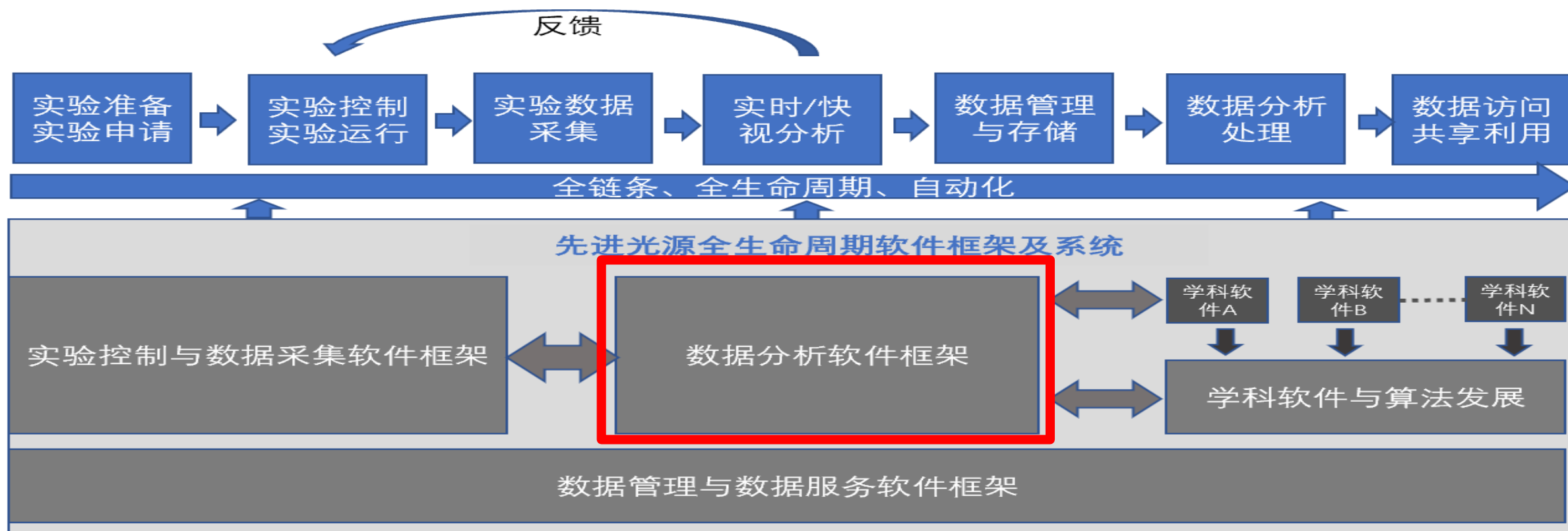
- 围绕实验和数据全生命周期过程需求及挑战
- 多组织多系统联合，发展先进光源实验全生命周期软件框架及系统，实现先进光源实验和数据处理自动化
- 支持发展线站数据分析新方法、新软件，支持已有方法学和软件的自动化框架集成



# 覆盖实验过程和实验数据全生命周期



高能所计算中心  
IHEP Computing Center



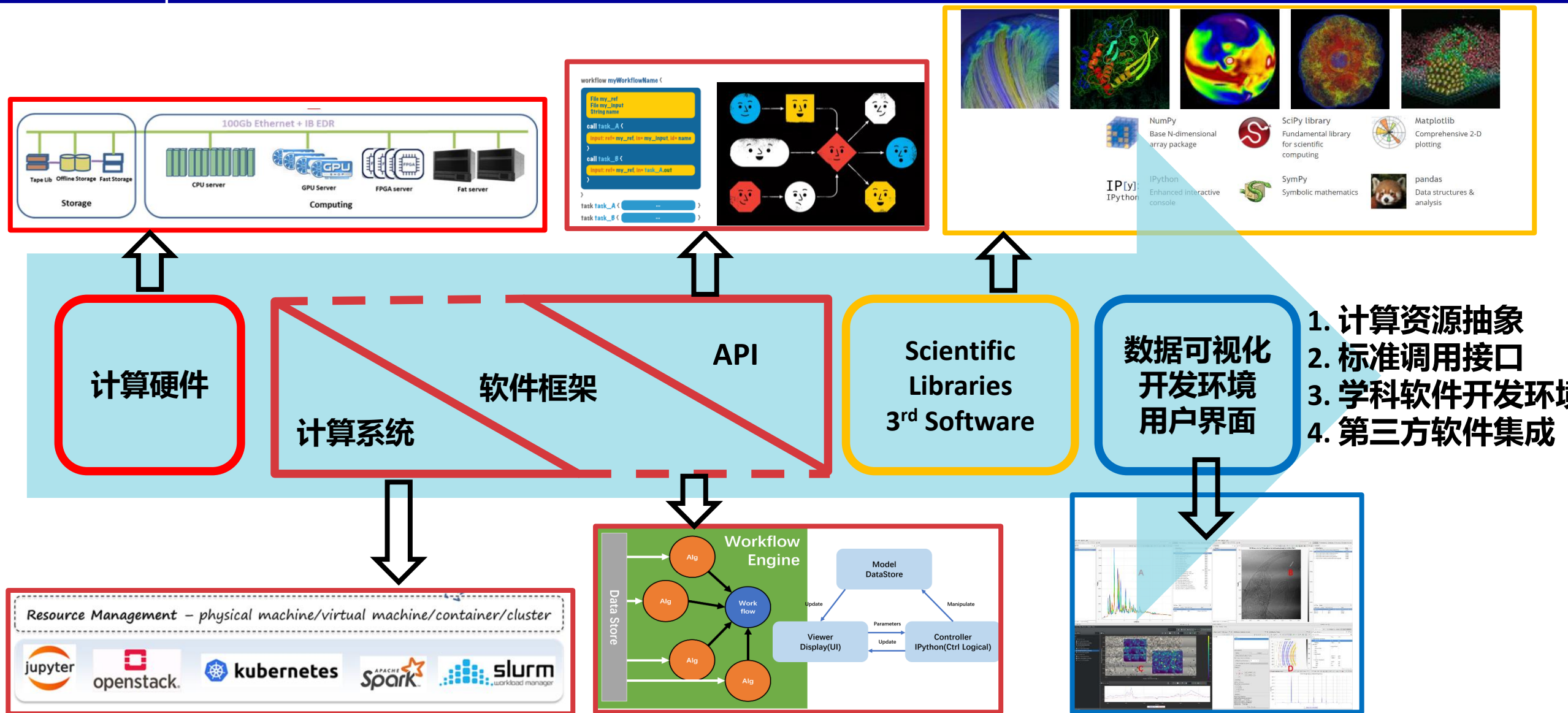
- 围绕实验和数据全生命周期过程需求及挑战
- 多组织多系统联合，发展先进光源实验全生命周期软件框架及系统，实现先进光源实验和数据处理自动化
- 支持发展线站数据分析新方法、新软件，支持已有方法学和软件的自动化框架集成

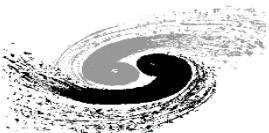


# 科学软件框架的任务



高能所计算中心  
IHEP Computing Center



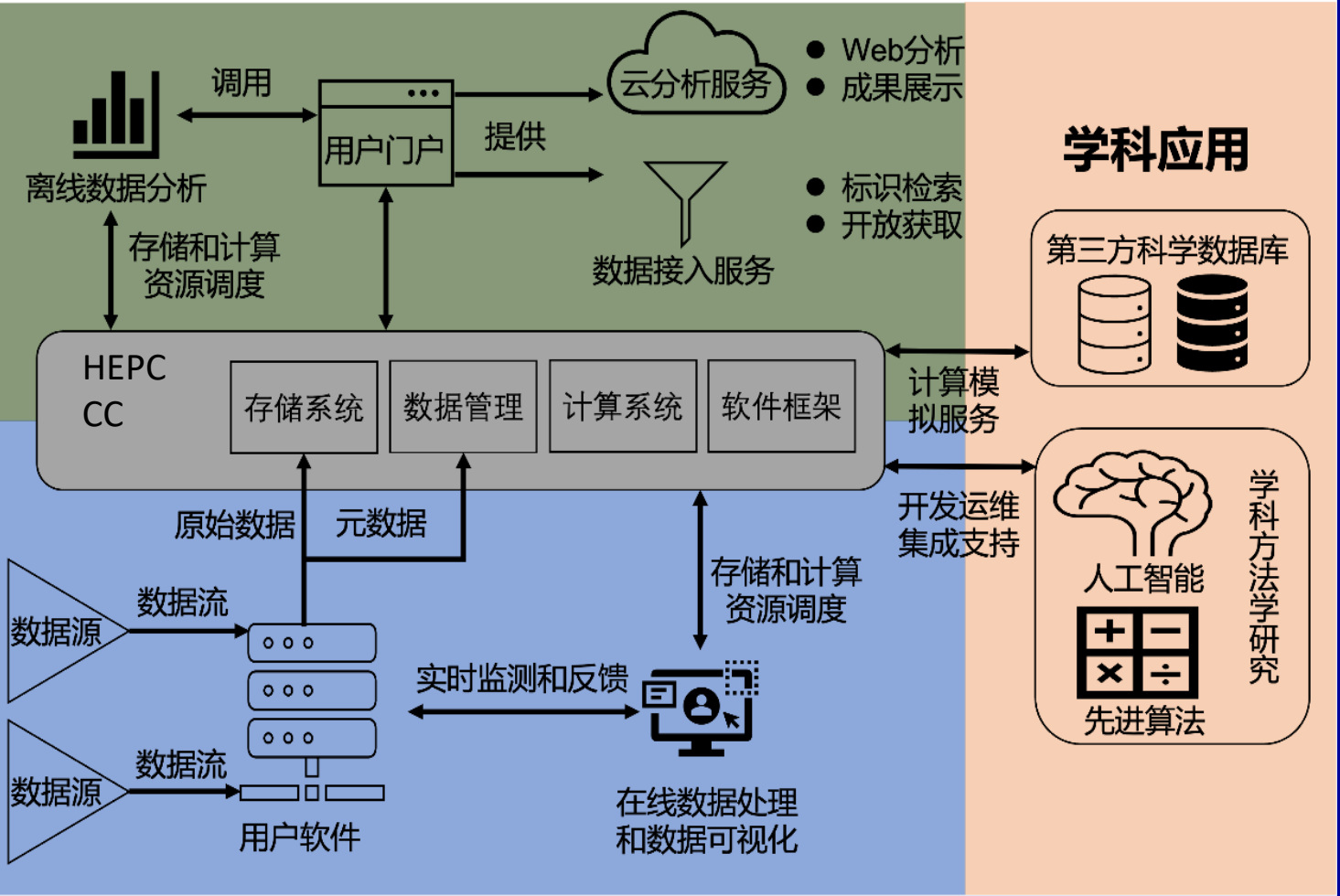


# 支持在线分析/离线处理/方法学开发和集成

高能所计算中心  
IHEP Computing Center

离线计算服务

实时计算服务



## • 在线数据处理和分析

- 指导实验 (可视化/估算/表征)
- 数据产品生成
- 文件/流 (全部/部分)

## • 离线和远程数据分析

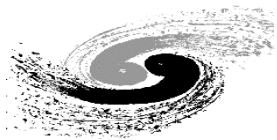
- 规约, 重建, 建模和模拟
- 庞大的数据容量 (数据托管, 算力)

## • 计算基础设施

- HPC 集群 (Spark/Slurm)
- 单个工作站 (虚拟机)
- 接入方式: JupyterHub (notebook app), 远程桌面 (traditional app), .....

## • 软件部署和容器化

- 利用容器封装软件环境,
- 方便移植、复用软件环境。复现分析结果



- 科学软件:使命和任务
- **Daisy设计和开发现状**
- 学科软件团队合作开发与集成 (生态发展)



# Inspired by Many Projects



高能所計算中心  
IHEP Computing Center

## DAWN SCIENCE

DAWN, the Data Analysis Workbench, is an Eclipse based application for scientific data analysis



Visualisation

Comprehensive visualisation of data in 1D, 2D and 3D



Development Environments

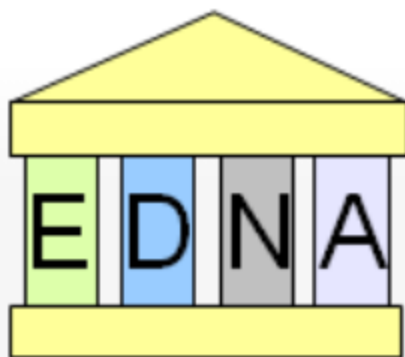
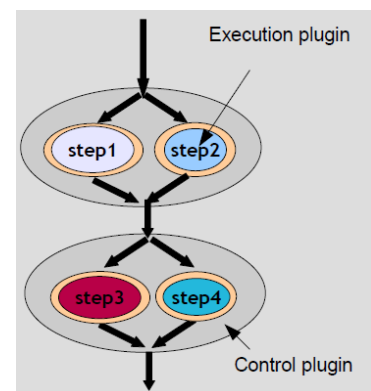
Integrated Python, Jython and Eclipse plug-in development, debugging and execution



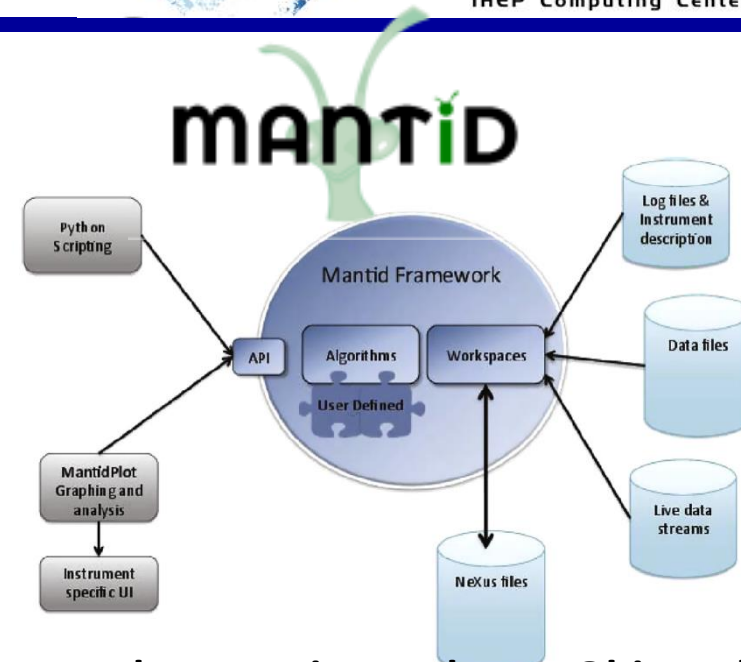
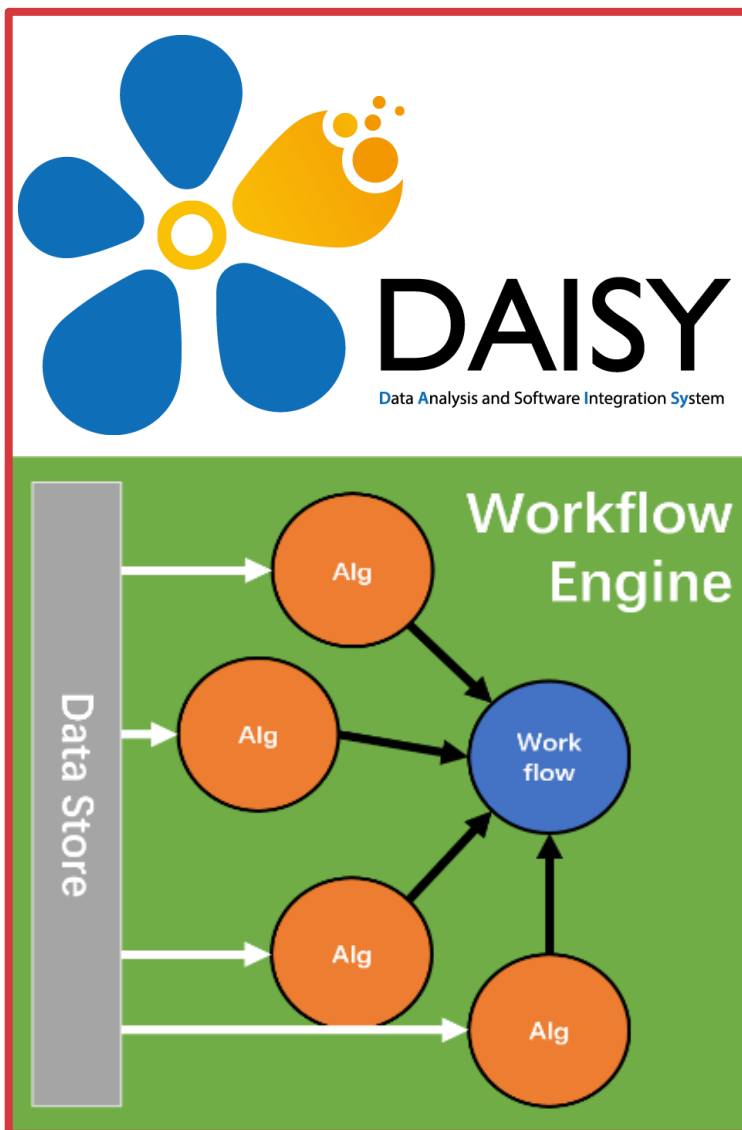
Processing and Workflows

for visual algorithms analysing scientific data

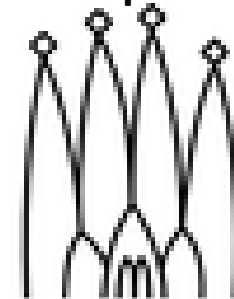
## Data Visualization and Processing & IDE



## Decouple Control and Execution Module

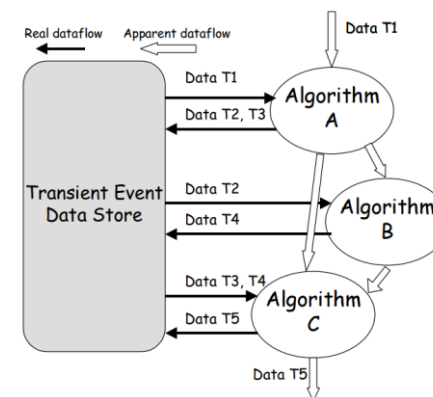


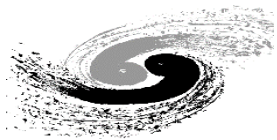
## Decouple Execution and Data Obj Module



## Gaudi

Integrate with existing code via Algorithm



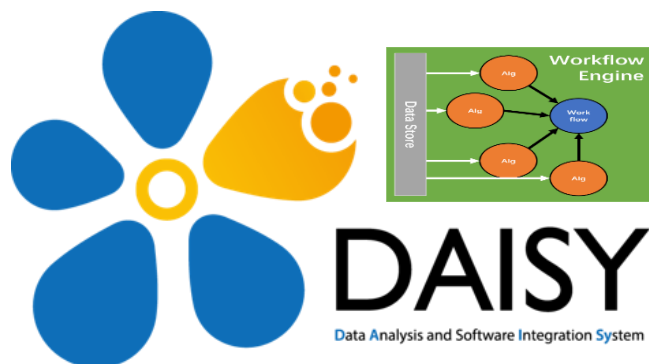


# Daisy Overview

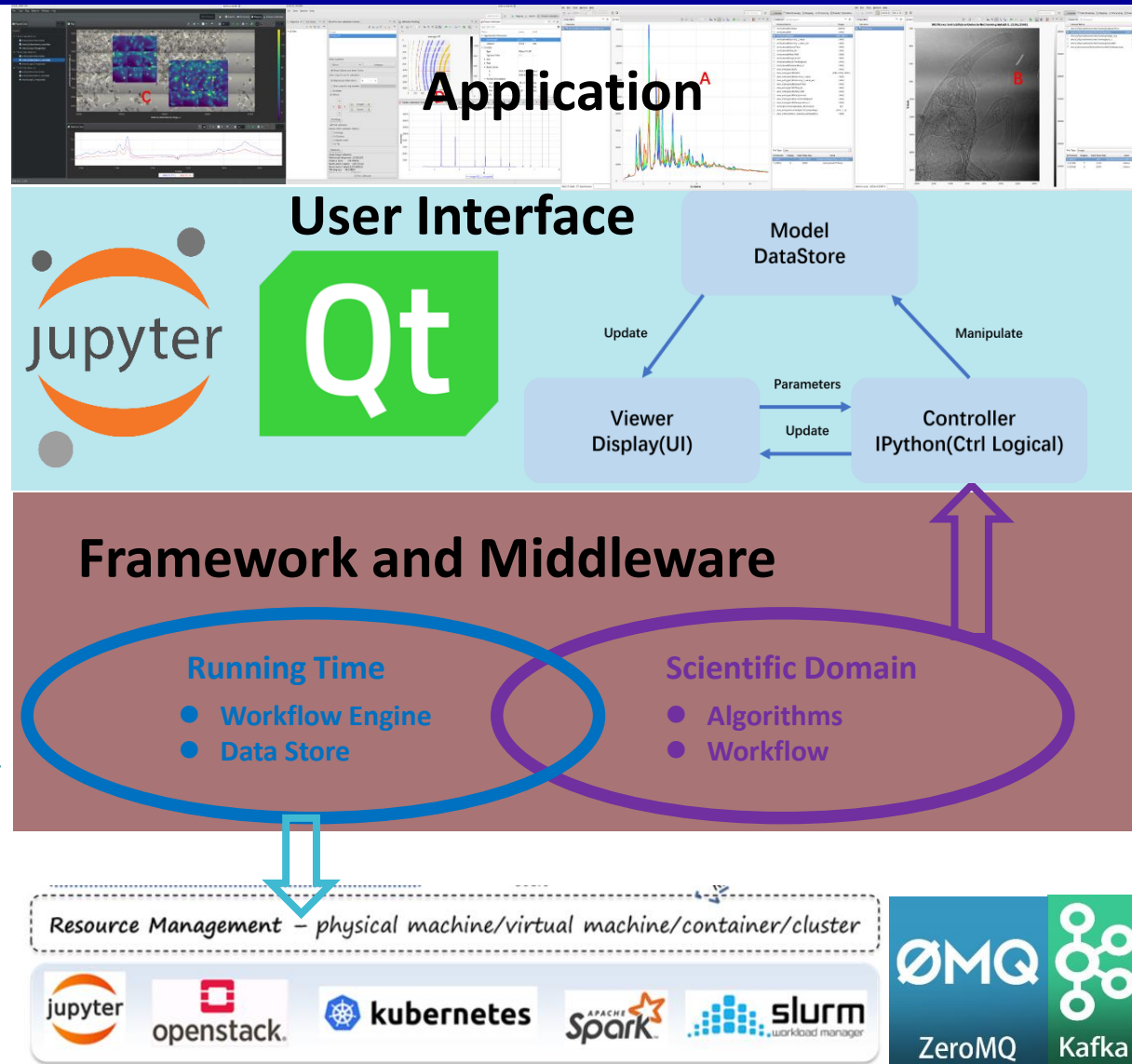
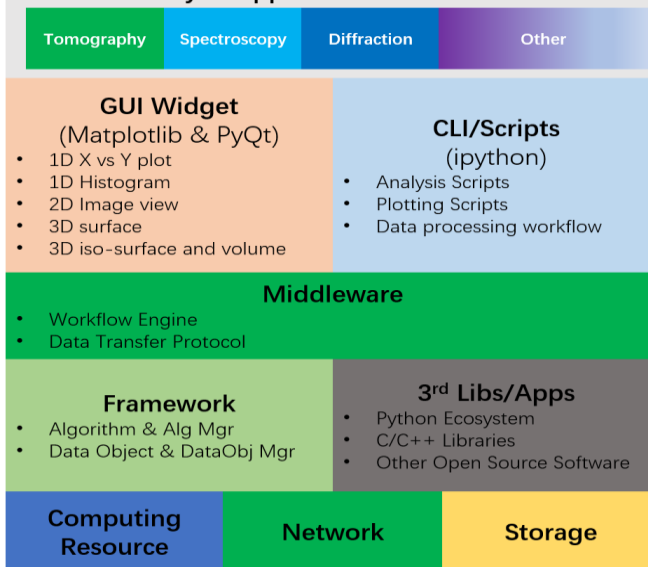


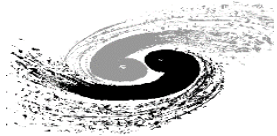
高能所計算中心  
IHEP Computing Center

Data analysis integrated software system for X-ray experiments



## Scientific Analysis Application





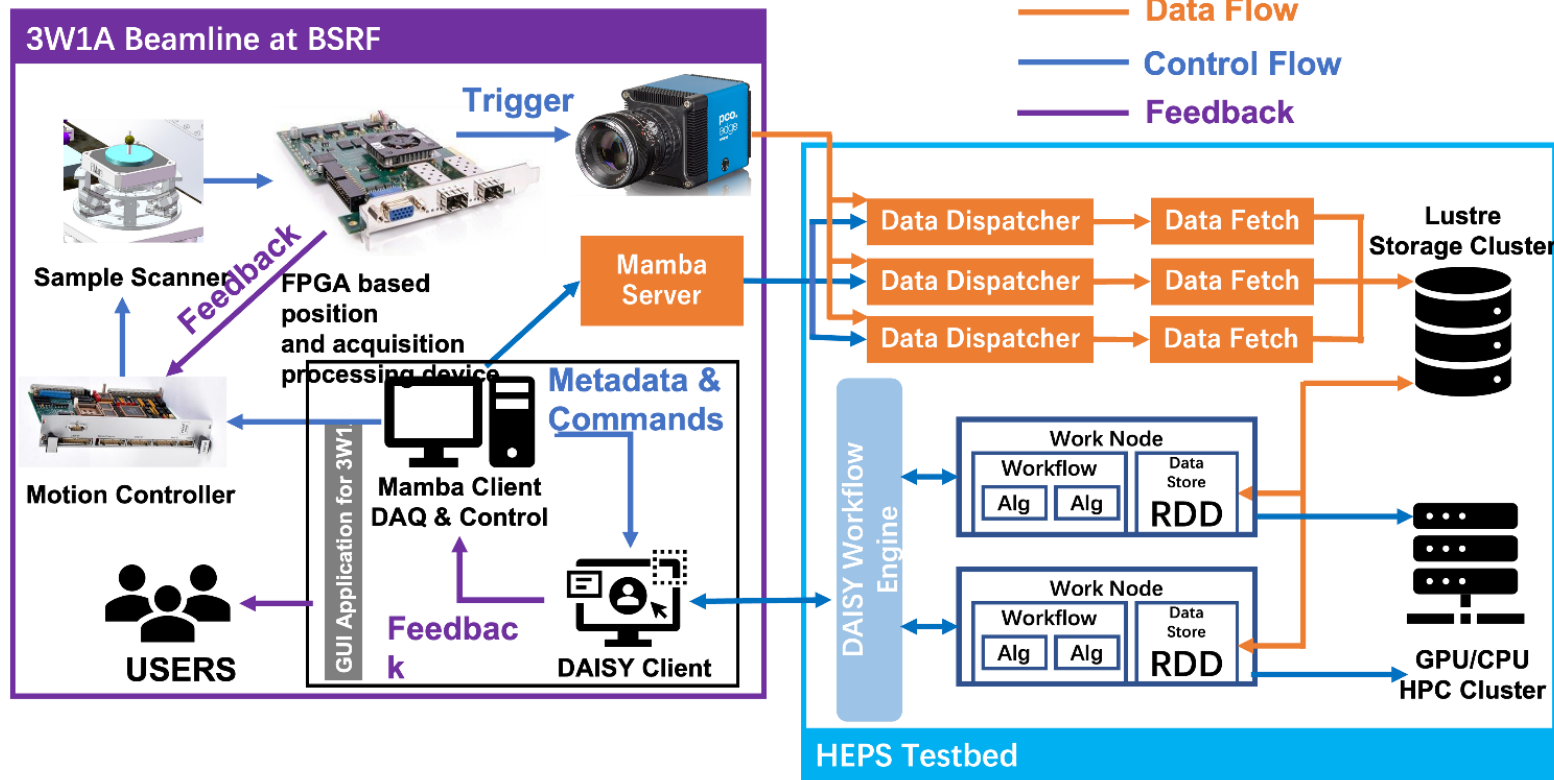
## Scientific Domain

- Algorithms
- Workflow

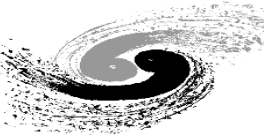
- 通过算法集成方法学
- 工作流的编排
  - ❖ 多模态方法学编排
  - ❖ 数据处理/控制反馈
  - ❖ 数据分析/人工智能

## Running Time

- Workflow Engine
- Data Store



- workflow引擎管理数据仓库以及工作流运行时环境
- 数据仓库管理数据对象



A

```
import numpy as np
import tomopy
from Daisy import DaisyAlg
class AlgTomopyRecon(DaisyAlg):
    def __init__(self, name):
        super().__init__(name)

    def initialize(self):
        self.data = self.get("Datastore").data()
        self.LogInfo("initialized, Tomopy Reconstruction")
        return True

    def execute(self, input_dataobj, theta, center, alg_type, output_dataobj):
        projs = self.data[input_dataobj]
        thetas = self.data[theta]
        dataobj = tomopy.recon(projs, thetas, center=center, algorithm=alg_type)
        self.data[output_dataobj] = dataobj
        return True

    def finalize(self):
        self.LogInfo("finalized")
        return True
```

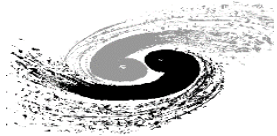
- 继承算法基类
- 每个算法必须实现initialize、execute和 finalize 三个方法，对应算法初始化、算法执行和算法销毁三个阶段

一个算法一般包括 一组输入数据对象，一组输出数据对象，一组计算参数。

从 datastore 拿到输入数据

数据处理的主要工作在execute函数完成

将输出数据注册回datastore，注册后的数据可被其它模块访问，或保存到输出文件



# Workflow 实现 (Python)



```
@Daisy.Singleton
class WorkflowCTReconstruct(Daisy.PyWorkflow):
    def execute(self):
        self.engine['loadhdf5'].execute(input_path='/entry/tomo', output_dataobj='tomodata')
        self.engine['loadhdf5'].execute(input_path='/entry/dark', output_dataobj='darkdata')
        self.engine['loadhdf5'].execute(input_path='/entry/flat', output_dataobj='flatdata')
        self.engine['normalize'].execute(projs_dataobj='tomodata', darks_dataobj='darkdata', \
                                         flats_dataobj='flatdata', output_dataobj='normdata')
        self.engine['angles'].execute(input_dataobj='normdata', output_dataobj='thetas')
        self.engine['minuslog'].execute(input_dataobj='normdata', output_dataobj='mlogdata')
        self.engine['reconstruct'].execute(input_dataobj='mlogdata', theta='thetas', \
                                           center=1030, alg_type='fbp', output_dataobj='recodata')
        self.engine['savehdf5'].execute(input_dataobj='recodata', output_path='/entry/reco')

wf = WorkflowCTReconstruct('WorkflowCTReconstruct')
wf.initialize(workflow_engine='PyWorkflowEngine', \
              workflow_environment = init_dict, algorithms_cfg = cfg_dict)
wf.execute()

data = wf.data_keys()
algs = wf.algorithm_keys()

wf.finalize()
```

B

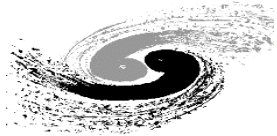
- 工作流和算法一样，也必须实现initialize、execute 和 finalize 三个方法

业务逻辑的实现，按序列调用算法或者子工作流

- self.engine 和 self.engine.datastore 为初始化时导入的工作流引擎以及对应的数据仓库
- 在工作流执行过程中，不同的工作流引擎和数据仓库对应不同的计算环境
- self.engine 和 self.engine.datastore 分别维护一组 key-value字典，可以通过算法名和数据对象名，访问算法对象和数据对象

工作流初始化时需要指定工作流引擎

算法对象初始化参数列表和算法运行参数可以由 JSON 对象或者 Python 字典对象表示



# Daisy用户接口: Daisy Workbench



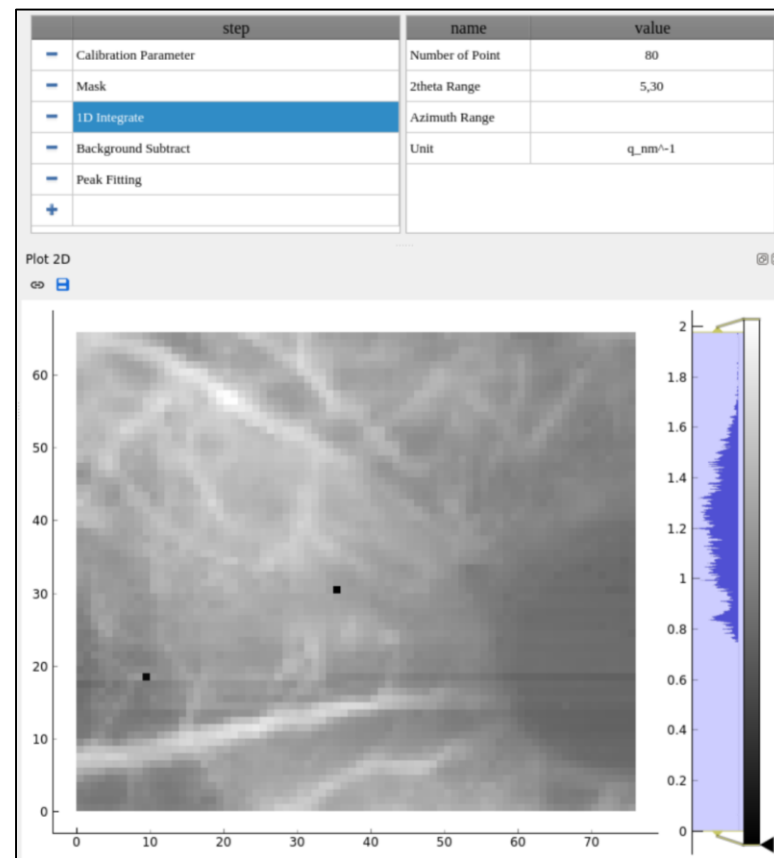
高能所計算中心  
IHEP Computing Center

- 基于PyQt5 的通用 GUI, 包含数据对象列表, 算法列表, 数据展示/可视化, 以及提供给开发者的 IDE模块。
- 领域方法学专用 interface, 包括用户自研以及第三方软件现有的GUI。

## Interface for XRD

The screenshot displays the Daisy Workbench interface with several components labeled in red:

- Workspace/数据对象**: A panel on the left showing a list of data objects, including 'tooth'.
- Python editor**: A central window for editing Python code.
- IPython**: A window for running IPython code.
- System Memory Usage**: A panel showing memory usage (1.72/62.76 GB (2%)).
- Plots**: A panel showing a 2D plot of data.
- Message**: A panel showing system messages.
- 数据展示**: A window showing a table of data.
- 算法列表**: A panel showing a list of algorithms.



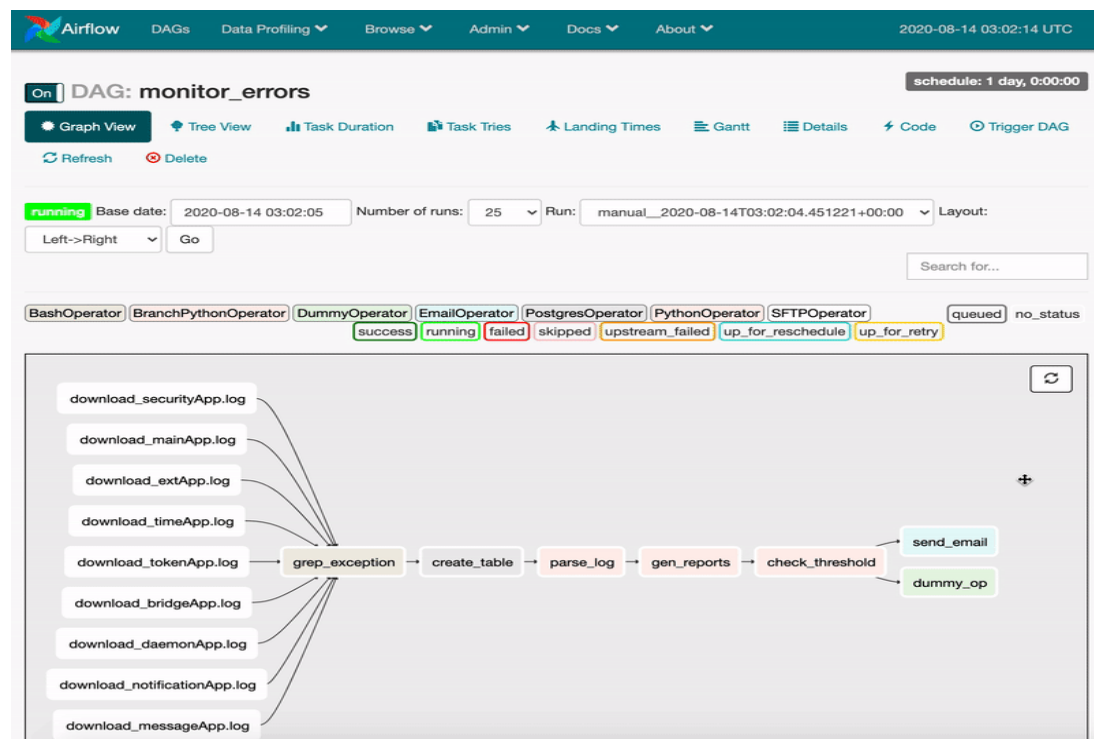
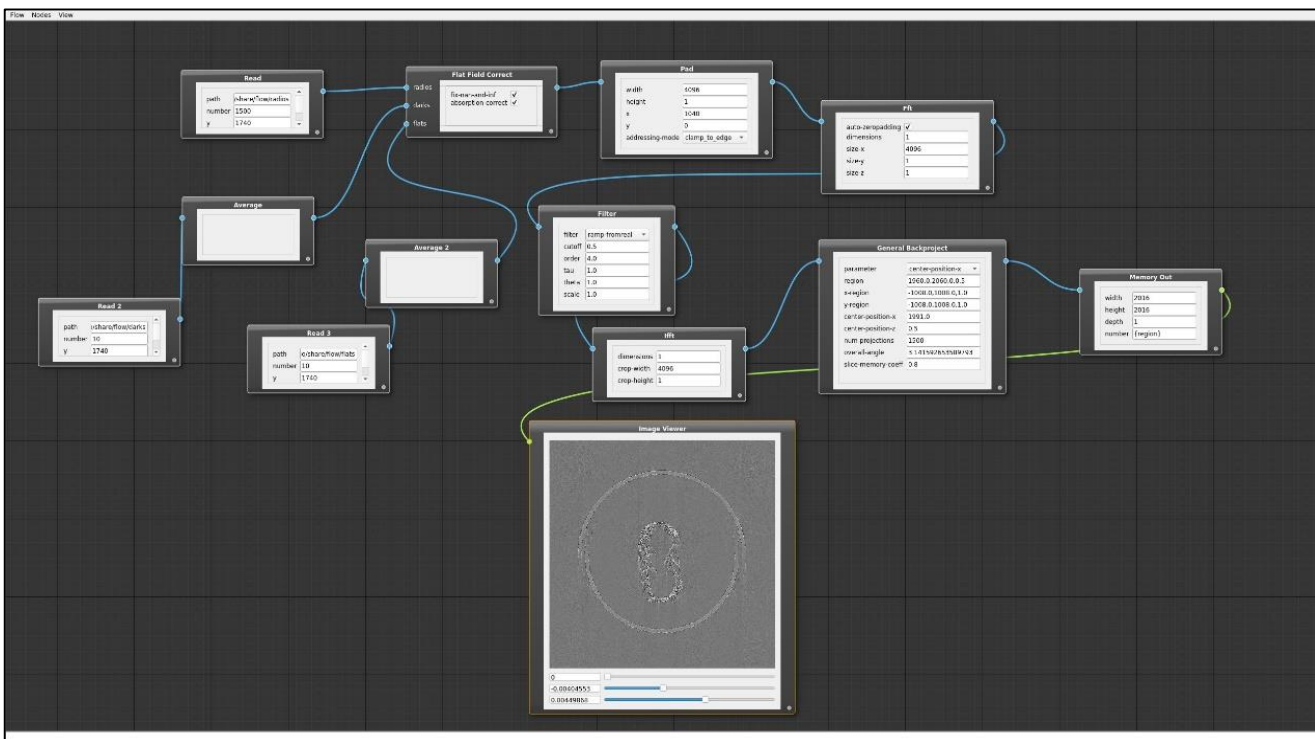


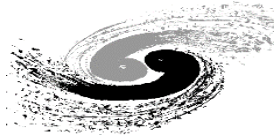
# Daisy用户接口——工作流编辑工具



高能所計算中心  
IHEP Computing Center

- 以有向无环图的方式可视化编排 workflow，支持 Daisy 以及 Daisy 外部算法，支持多种类型的 workflow（单机， 分布式）。
- 提供 Daisy workflow 运行状态状态监控以及中间数据可视化。
- workflow 可以转换为标准工作流语言 CWL 供其他 workflow 引擎调用（如 Airflow），支持 CWL 语言的解析和导出。
- 亦可作为通用的 workflow 编辑工具独立部署运行。

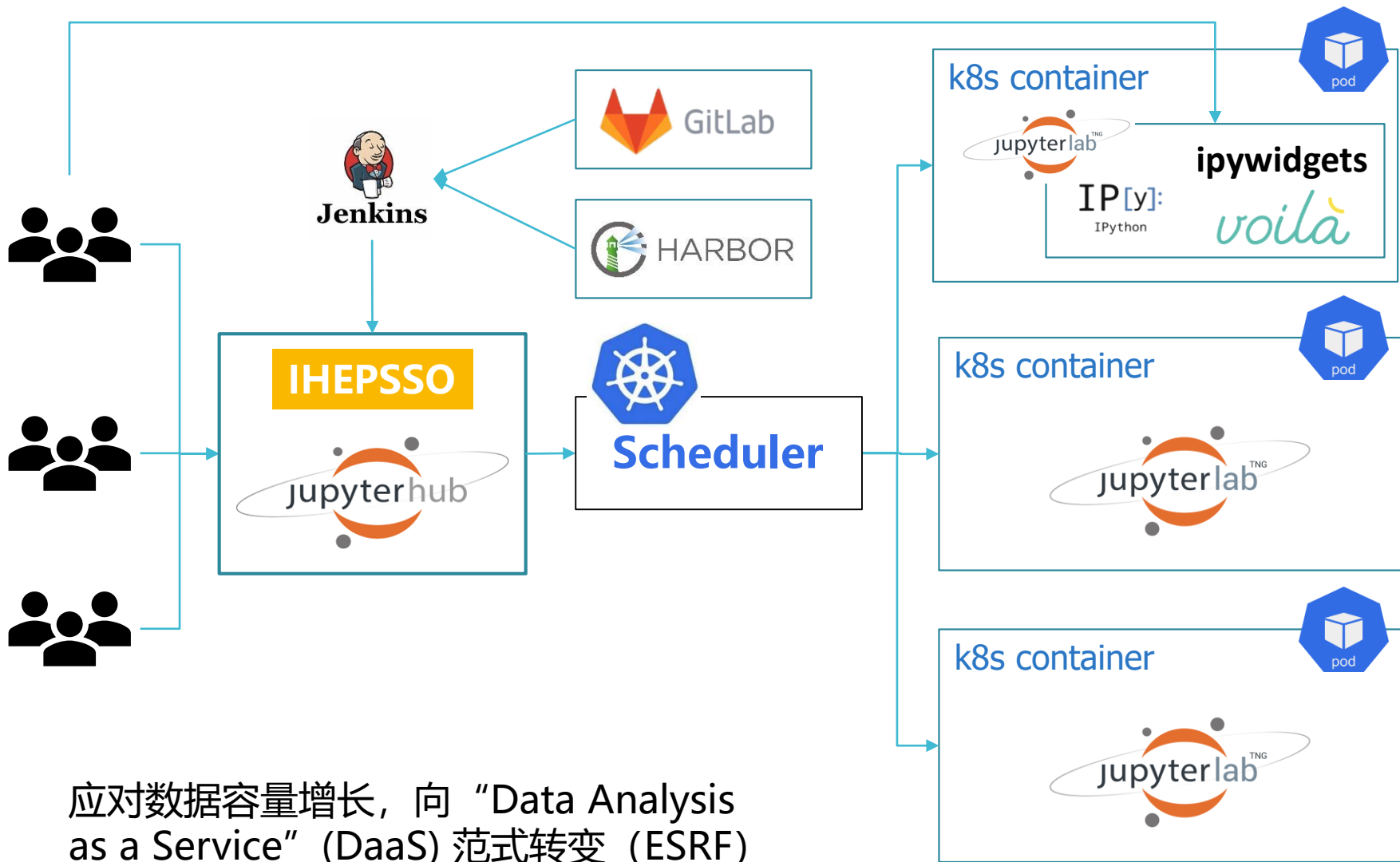




# Daisy用户接口——web 数据分析平台

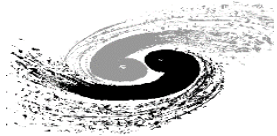


高能所计算中心  
IHEP Computing Center



应对数据容量增长，向 “Data Analysis as a Service” (DaaS) 范式转变 (ESRF)

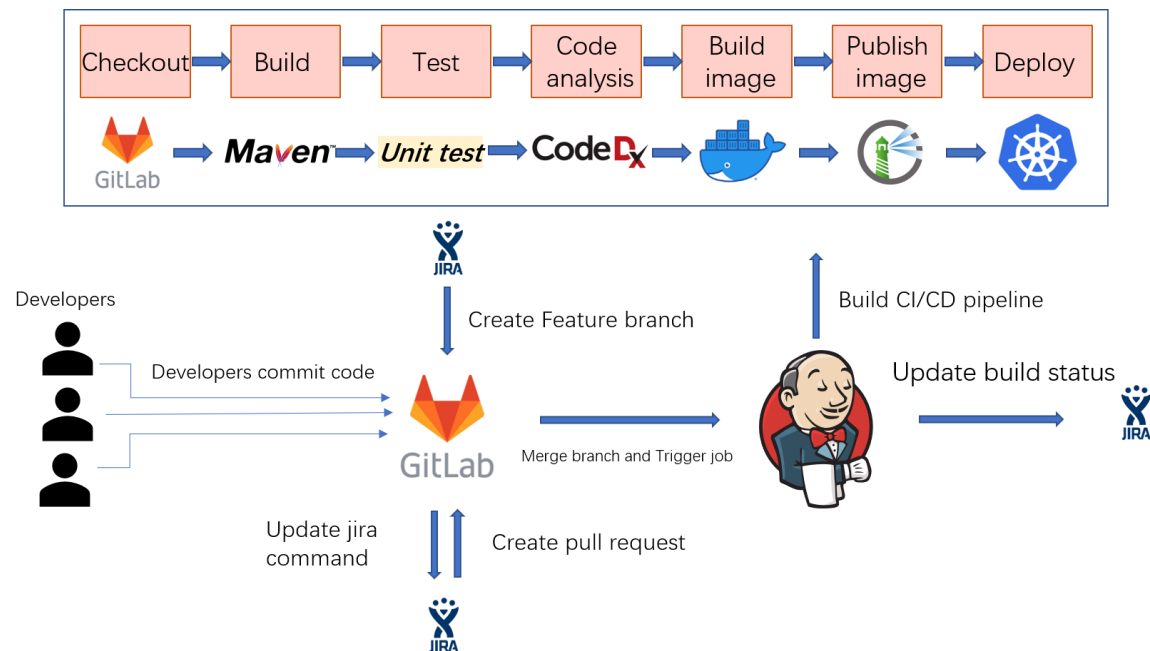
- 前端基于 jupyterlab 生态
- 用户通过web浏览器访问 web portal，经过认证后由分析平台为其分配相应的计算资源，并启动相应的web交互式界面
- 用户不用自己下载、安装软件，配置计算环境，随时随地都可以通过浏览器访问服务
- 用户不用自己更新服务，服务始终保持最新版本
- 用户友好界面，适合不同专业程度的用户
- 用户不受操作系统限制



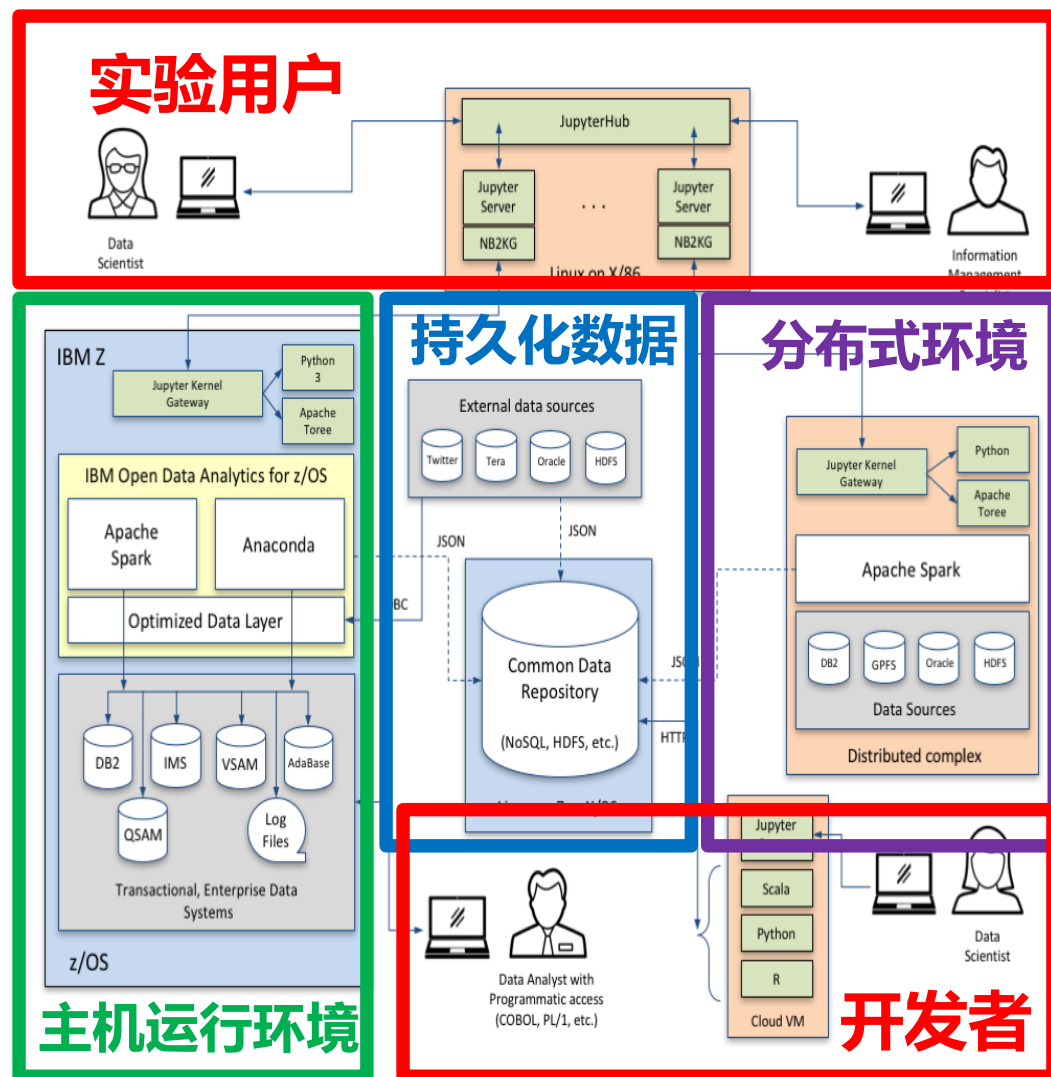
# Daisy 开发者和用户支持



高能所計算中心  
IHEP Computing Center



- 提供基于 JupyterLab 和 PyQt5 的用户开发界面
- 提供软硬件一体化开发环境，并计划提供软件测试框架，支持 CI/CD
- 提供软件主页及丰富的用户文档
- 不同学科应用以 Plugin/Package 形式插入到框架中，支持独立发布





## 代码仓库:

Daisy: <https://code.ihep.ac.cn/hepscc/Daisy>

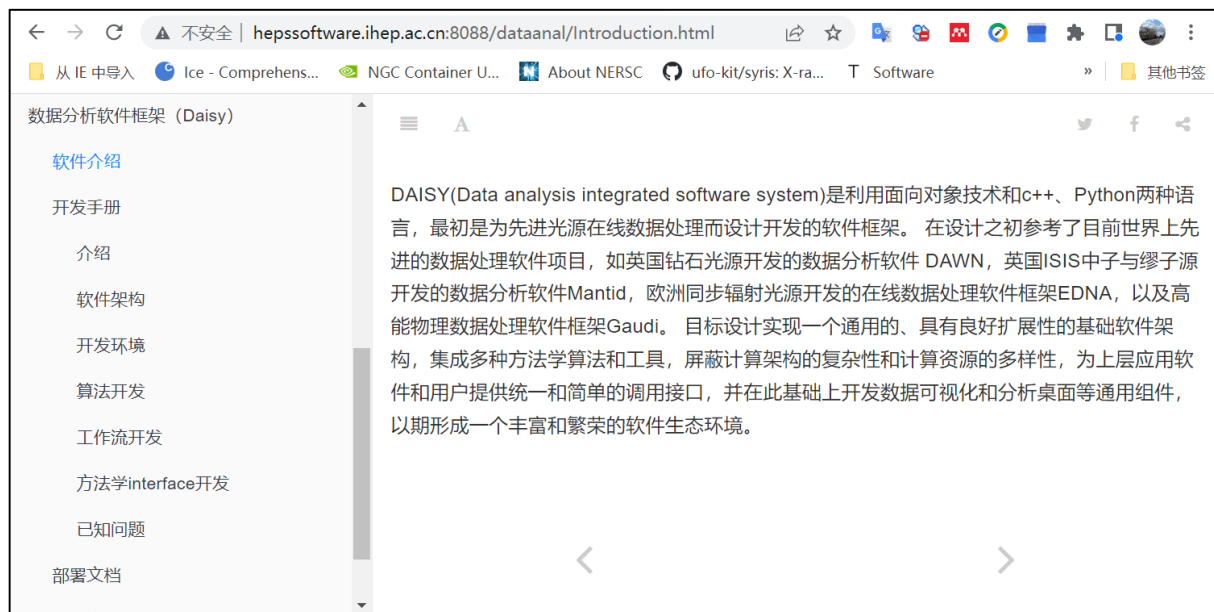
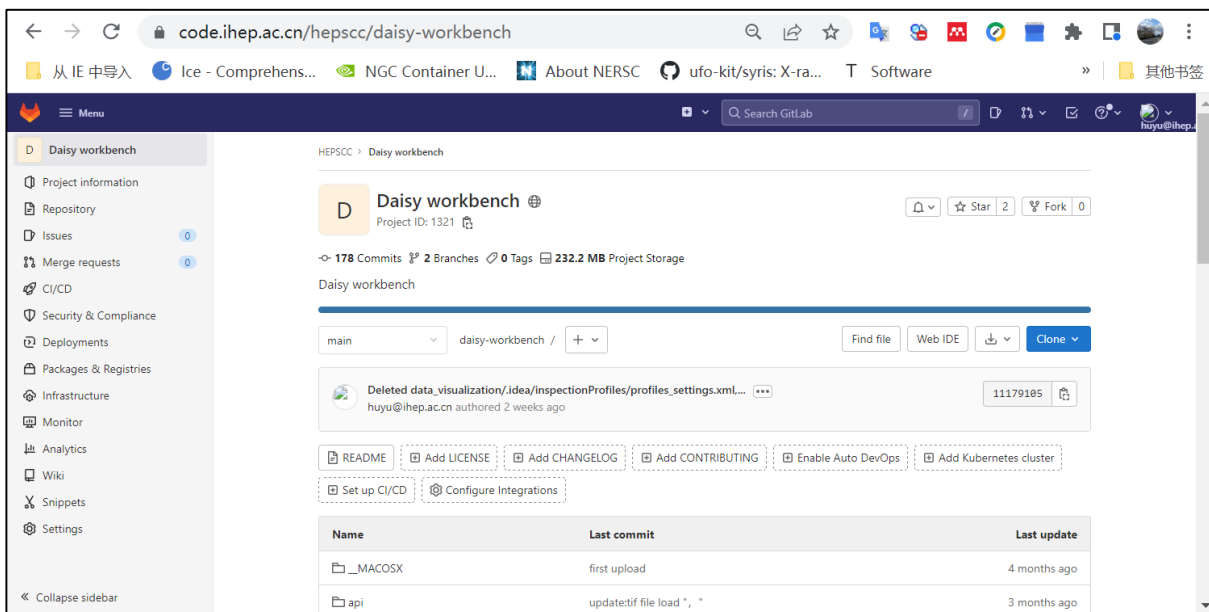
Daisy workbench: <https://code.ihep.ac.cn/hepscc/daisy-workbench>

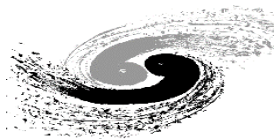
Docker 容器: <https://code.ihep.ac.cn/hepscc/opendata-daisydev-image>

## 用户文档:

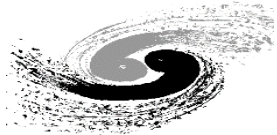
<http://hepssoftware.ihep.ac.cn:8088/dataanal/>

欢迎贡献代码!

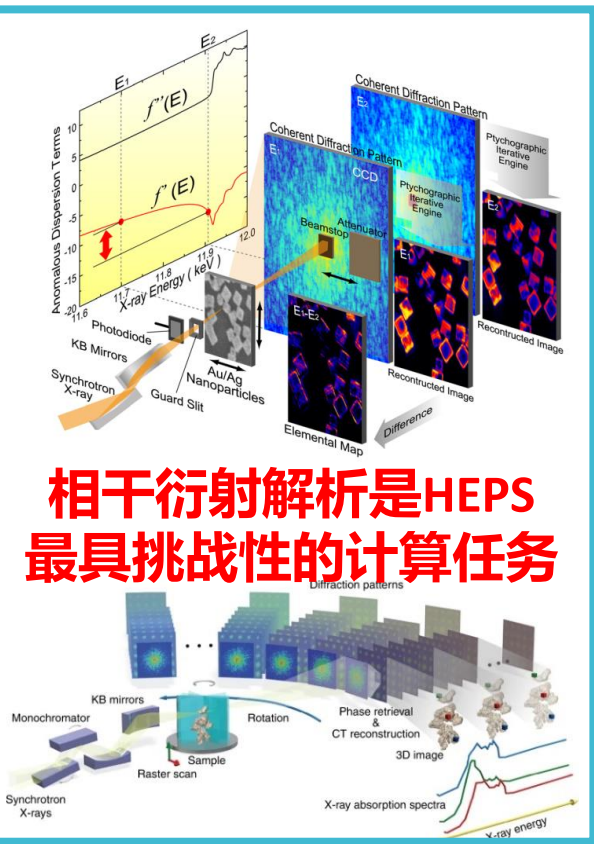




- 科学软件:使命和任务
- Daisy设计和开发现状
- **学科软件团队合作开发与集成 (生态发展)**



# HEPS相干衍射算法在Daisy中的开发与集成

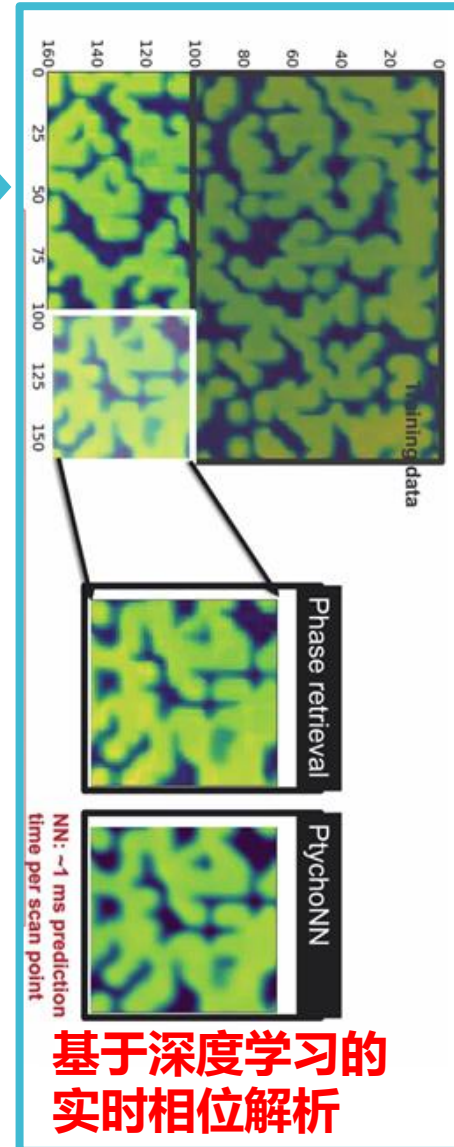
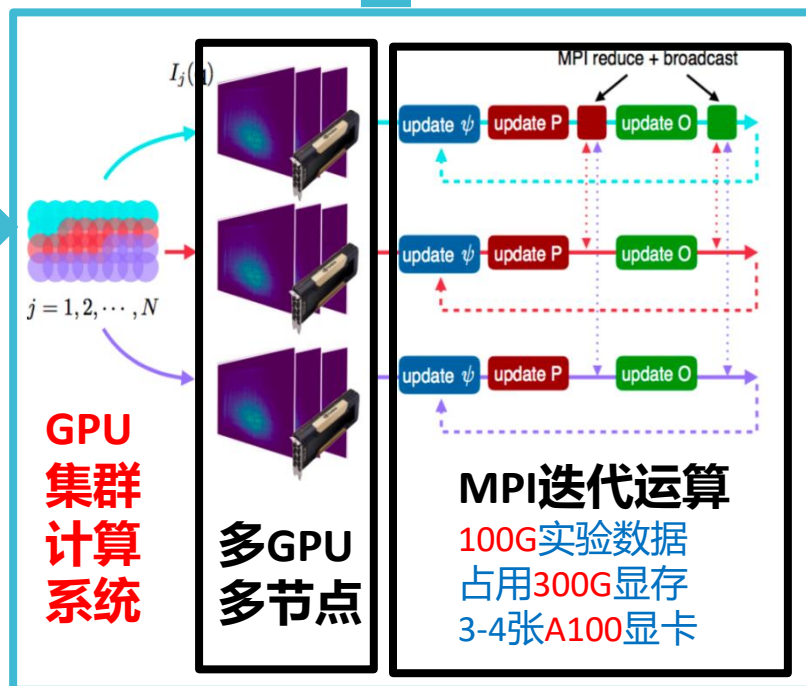
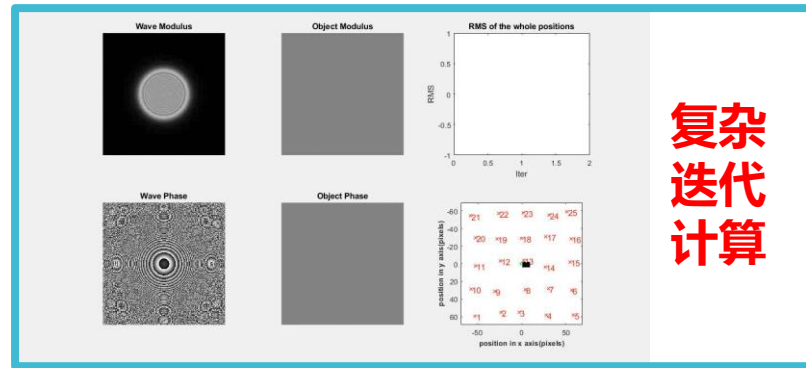
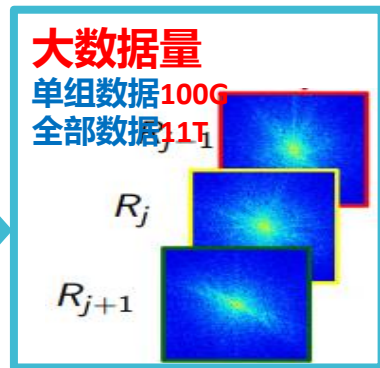


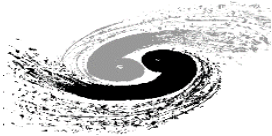
相干衍射解析是HEPS  
最具挑战性的计算任务

## 合作线站

- HEPS-B2 硬X射线纳米探针

大量的数据帧，且图像重建依赖迭代算法  
需要借助异构资源，分布式进行加速





# HEPS断层扫描重建及AI应用在Daisy中的开发与集成

高能所计算中心  
IHEP Computing Center



## 合作线站

- HEPS-B7
- HEPS-BE
- BSRF-3W1

数据导入

数据预处理

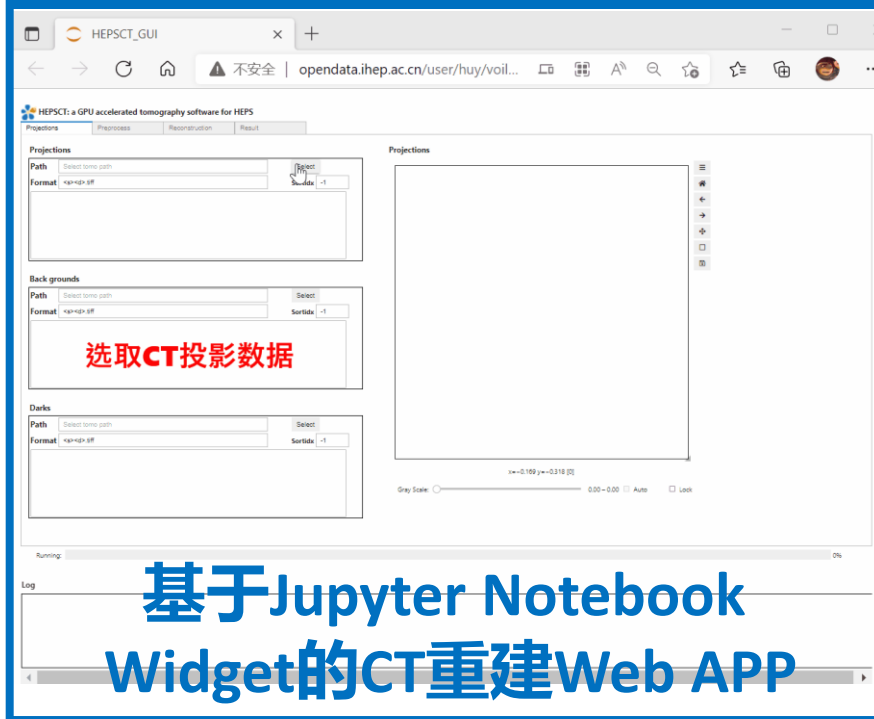
微米CT校正

纳米CT校正

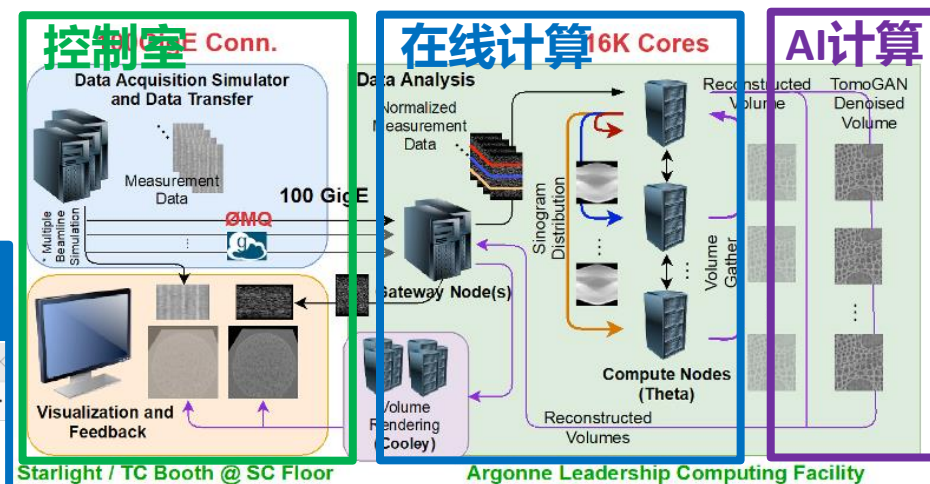
重建、后处理

显示

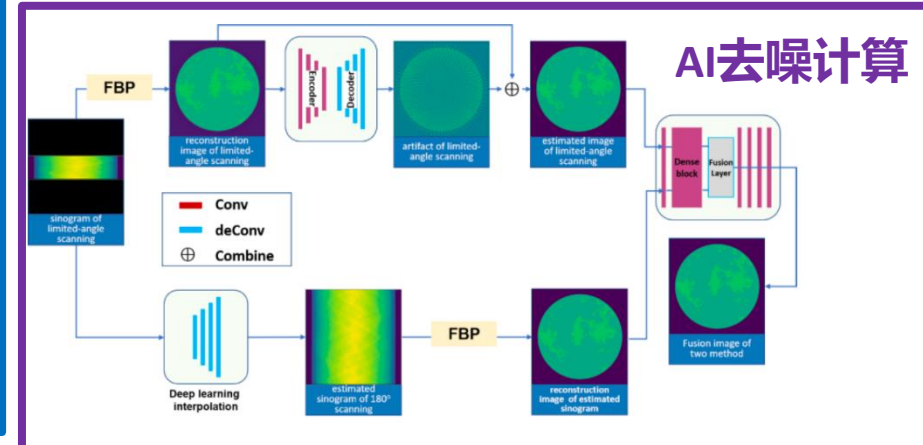
## 在线重建计算

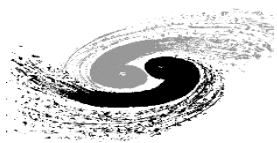


## 传统算法+AI去噪



中国科学院“十四五”网信专项  
面向国产先进计算资源的人工智能应用支撑平台



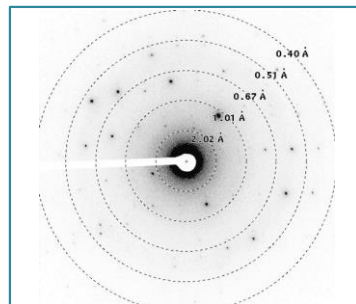


# 物理所软物质实验室 生物大分子结构解析软件集成以及AI应用



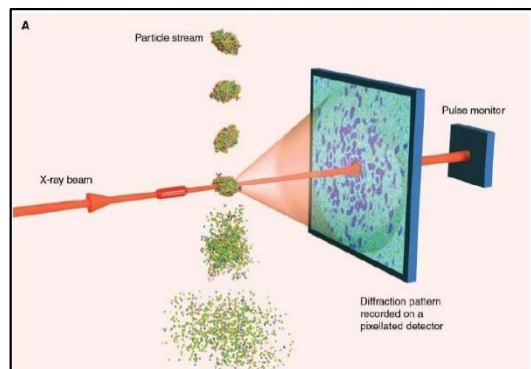
高能所计算中心  
IHEP Computing Center

## HEPS-BA线站

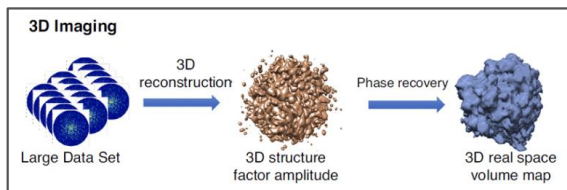


多个  
装置  
实验  
数据

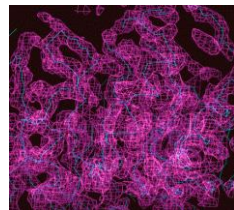
## XFEL高通量数据解析



## 高通量数据三维重建



相位恢复 模型构建



人工智能结构预测



基于  
软件  
框架  
实现  
工作  
管线

## 合作线站

- HEPS-BA

基于 AI 的  
结构精修

中国科学院物理研究所  
Institute of Physics, Chinese Academy of Sciences

北京凝聚态物理国家研究中心  
BEIJING NATIONAL LABORATORY FOR CONDENSED MATTER PHYSIC

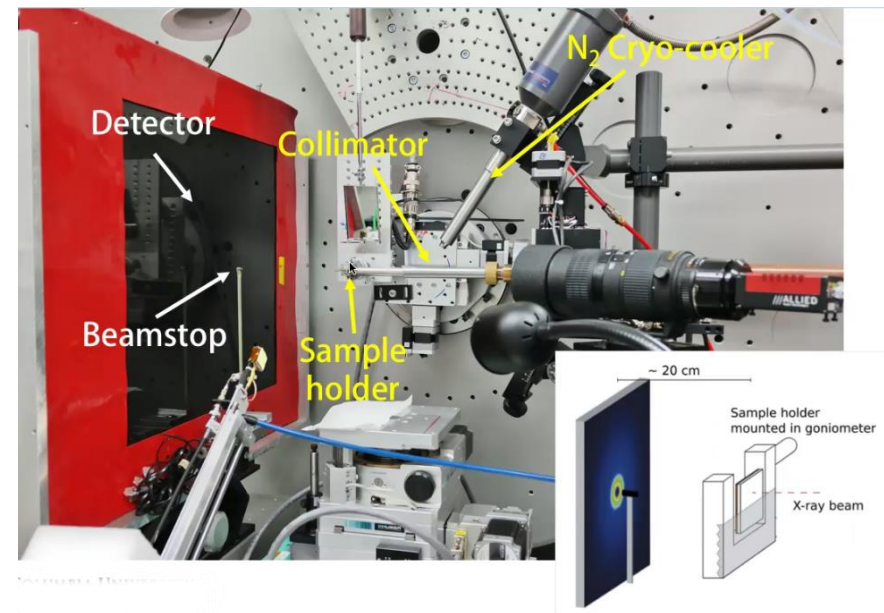
范海福 院士

丁玮 副主任工程

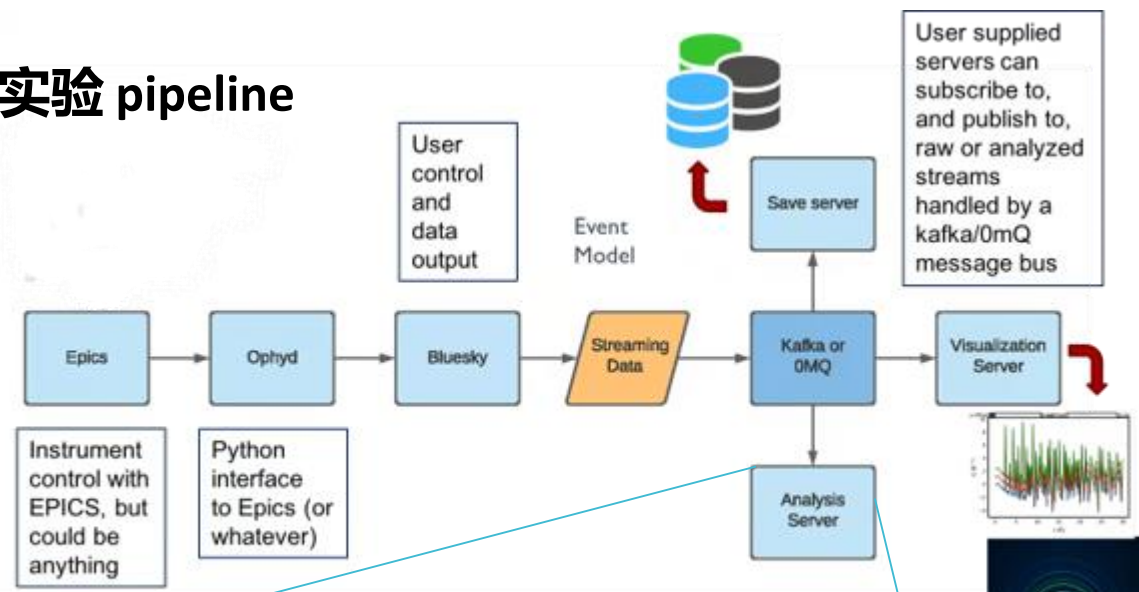
郝权教授 (HKU)

软物质物理实验室

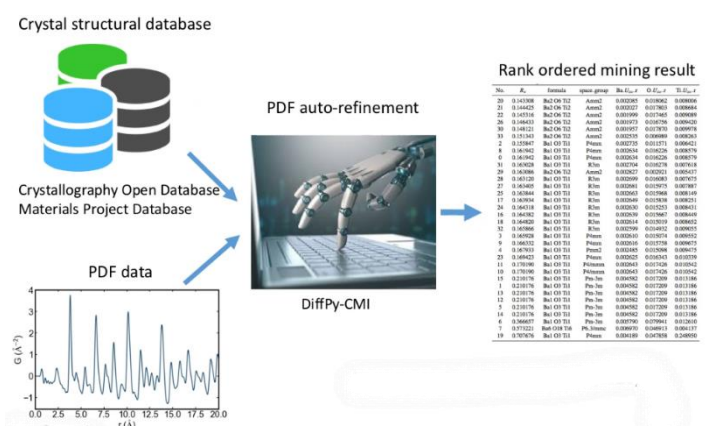
望石智慧  
STONEWISE



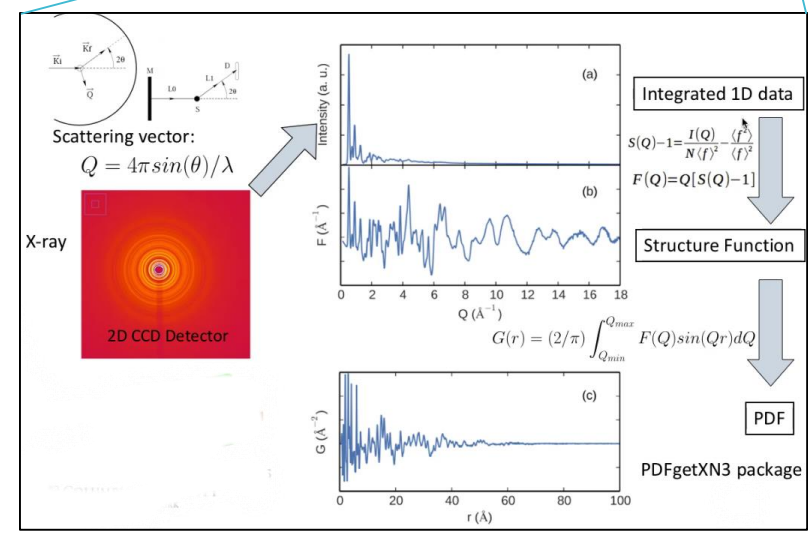
### PDF线站实验 pipeline



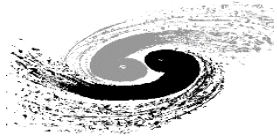
合作线站  
BSRF-3W1A  
HEPS-B1, B3, B6



利用机器学习进行结构精修



数据处理  
工作流



# 空间天文数据产品工具开发与集成



高能所计算中心  
IHEP Computing Center



## 数据分析 可视化

HMXT

Proposal data select Work flow

Target: Her\_X\_1 Begin date: 年/月/日 End date: 年/月/日

Obs. ID: P0101308001 Exp. ID: P01013080101

More search condition Condition clear Search

| ind. x | obsid       | tsstart   | tsstop    | obsdate                | obsend                 | duration | targetid | pt   | proposaltype | ind. y | target  | ra      | dec     |
|--------|-------------|-----------|-----------|------------------------|------------------------|----------|----------|------|--------------|--------|---------|---------|---------|
| 1      | P0101308001 | 173603274 | 173664655 | 2017-07-07 01:00:10.92 | 2017-07-07 01:00:10.92 | 161381   | 7047     | HMXT | CST          | C      | Her_X_1 | 254.458 | 35.3423 |
| 22     | P0101308002 | 178539881 | 178528427 | 2017-08-08 05:00:51.18 | 2017-08-08 05:00:51.18 | 88548    | 7047     | HMXT | CST          | C      | Her_X_1 | 254.458 | 35.3423 |
| 25     | P0101308003 | 176648731 | 176689486 | 2017-08-08 07:00:24.43 | 2017-08-08 07:00:24.43 | 49755    | 7047     | HMXT | CST          | C      | Her_X_1 | 254.458 | 35.3423 |
| 336    | P0101308004 | 191100795 | 191184850 | 2018-01-20 20:22:03.12 | 2018-01-20 20:22:03.12 | 75055    | 7047     | HMXT | CST          | C      | Her_X_1 | 254.458 | 35.3423 |
| 539    | P0101308005 | 191218581 | 191385464 | 2018-01-20 20:22:03.12 | 2018-01-20 20:22:03.12 | 166883   | 7047     | HMXT | CST          | C      | Her_X_1 | 254.458 | 35.3423 |
| 416    | P0101308006 | 194265558 | 194300989 | 2018-02-26 20:20:24.56 | 2018-02-26 20:20:24.56 | 39041    | 7047     | HMXT | CST          | C      | Her_X_1 | 254.458 | 35.3423 |
| 518    | P0101308007 | 200103187 | 200137966 | 2018-05-06 07:00:33.04 | 2018-05-06 07:00:33.04 | 34809    | 7047     | HMXT | CST          | C      | Her_X_1 | 254.458 | 35.3423 |
| 593    | P0101308008 | 200246411 | 200283275 | 2018-05-06 07:00:33.04 | 2018-05-06 07:00:33.04 | 34864    | 7047     | HMXT | CST          | C      | Her_X_1 | 254.458 | 35.3423 |
| 596    | P0101308009 | 200382343 | 200367267 | 2018-05-06 07:00:33.04 | 2018-05-06 07:00:33.04 | 34864    | 7047     | HMXT | CST          | C      | Her_X_1 | 254.458 | 35.3423 |
| 599    | P0101308010 | 200435457 | 200464589 | 2018-05-06 07:00:33.04 | 2018-05-06 07:00:33.04 | 29132    | 7047     | HMXT | CST          | C      | Her_X_1 | 254.458 | 35.3423 |
| 602    | P0101308011 | 200815631 | 200800933 | 2018-05-06 07:00:33.04 | 2018-05-06 07:00:33.04 | 34808    | 7047     | HMXT | CST          | C      | Her_X_1 | 254.458 | 35.3423 |

Log

```
2021-06-04 09:35:40.428 - [WARNING] The env value of TERM will be overwr
2021-06-04 09:35:40.428 - [WARNING] The env value of PATH will be overwr
2021-06-04 09:35:40.429 - [WARNING] The env value of SHLV will be overwr
2021-06-04 09:35:40.429 - [WARNING] The env value of PYTHONPATH will be o
2021-06-04 09:36:44.930 - [INFO] Start execution!
2021-06-04 09:36:45.137 - [INFO] hepical : #####
2021-06-04 09:36:45.141 - [INFO] hepical : HMXT HE task, hepical is runn
2021-06-04 09:36:45.142 - [INFO] hepical : PILParSet Warning: parameter
2021-06-04 09:36:45.144 - [INFO] hepical : PILParSet Warning: the file /s
308001/P010130800101 HE pl.fits will be overwritten!
#####
2021-06-04 09:38:11.023 - [INFO] hepical : HMXT HE task, all events from
2021-06-04 09:38:11.024 - [INFO] hepical : HMXT HE task, glitch removal e
2021-06-04 09:38:11.026 - [INFO] hepical : HMXT HE task, hepical is runn
2021-06-04 09:38:11.028 - [INFO] hepical : #####
2021-06-04 09:38:11.747 - [INFO] hepical : #####
2021-06-04 09:38:12.389 - [INFO] hepical : #####
2021-06-04 09:38:12.392 - [INFO] hepical : HMXT HE task, hepical is runn
2021-06-04 09:38:12.393 - [INFO] hepical : PILParSet Warning: parameter
```

HMXT

Proposal data select Work flow

Low level

ANG\_DIST: 0.04 ELV: 10 COR: 8

T\_SAA: 300 TN\_SAA: 300 High level selection

High level

INPUT\_RA: 0 INPUT\_DEC: 0 LE\_DYE\_E... 30

HE\_LC\_Mi... 24 HE\_LC\_M... 120 HE\_LC\_Bi... 1

ME\_LC\_Mi... 120 ME\_LC\_M... 546 ME\_LC\_Bi... 1

LE\_LC\_Mi... 100 LE\_LC\_MA... 1170 LE\_LC\_Bi... 1

HE\_EVEN... N/A

Run

Start run Go show Go confirm

ME Gti LE Gti

Result

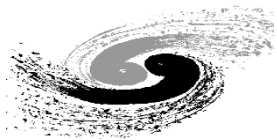
Show result Download result Fit

HE lc ME lc LE lc

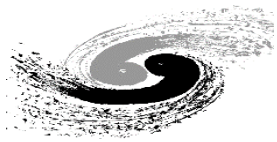
## 数据访问 数据管理

## 目标

- 组织规范各天文项目数据处理软件
- 构建数据产品生产 pipeline, 产生高级数据产品



- 围绕 HEPS 需求，设计了先进光源全生命周期的科学软件框架
- 软件原型已经形成，并进行了测试床验证
- 围绕软件框架，学科方法学开发和应用集成工作正在开展中
- 适用于光源实验，同样在空间项目具有应用前景
- 软件持续开发和用户社区是科学软件生态发展关键



Thank you!

