

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



国家高能物理科学数据中心
National HEP Data Center



高能所计算中心
IHEP Computing Center

面向先进光源的数据处理 workflow 管理系统

孙浩凯 sunhk@ihep.ac.cn

中科院高能所 计算中心

2023年07月10日



1 背景介绍

2 挑战

3 技术方案

4 项目进展

5 未来计划

背景 - 先进光源



- 我国目前正在运行的前三代光源：
BSRF、合肥光源、SSRF等
- 正在/计划建设的下一代先进光源：
HEPS、合肥先进光源、SHINE等

特点是

- ❖ 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



● 先进光源对数据处理提出了极高要求：

- 自动化 -- 在线实时计算、反馈
- 高效化 -- 并行、异构加速
- 智能化 -- 数据+知识驱动AI
- 易用化 -- 操作简便、用户友好

● 多方法学开发与集成

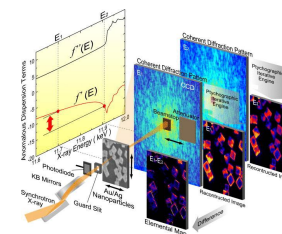
- 底层框架 -- IO、计算、调度等
- 统一平台 -- 算法/模型展示、调用等

HEPS一期数据统计表

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

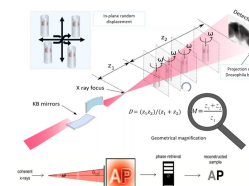
- 从光源软件开发人员和线站科学家的角度

需要一个**统一、通用的系统**来汇总和呈现其集成、开发和优化的学科（谱学、成像、衍散射等）方法学，从而易于实验用户的测试和使用；



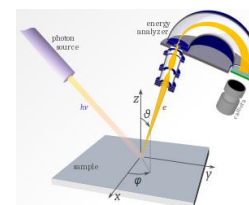
- 从软件工程学和技术方案的角度

需要一个面向多线站的高度抽象、架构清晰的**算法管理和编排系统**，向下可与服务器硬件解耦、灵活调度计算和存储资源，向上能以逻辑统一的方式便于方法学开发人员实现和优化其算法及后期维护；

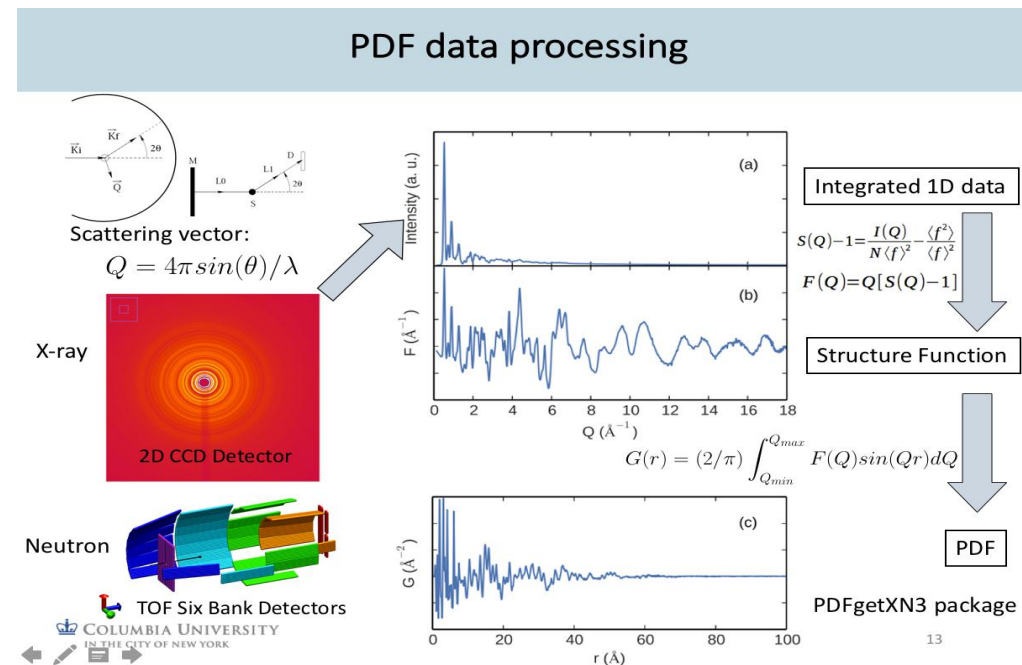


- 从实验用户的角度

需要一个可视化的、无需编程基础的**实验分析流程管理和监控系统**，快速实现、部署和调整数据处理全流程，做到可控、可复用和结果易复现。



- workflow (Workflow) 的含义非常朴素：一个任务中的各个步骤和整体规划
- workflow管理：统筹、编排、监控
- 目的是更好 (合理、高效、直观) 地完成 “任务”

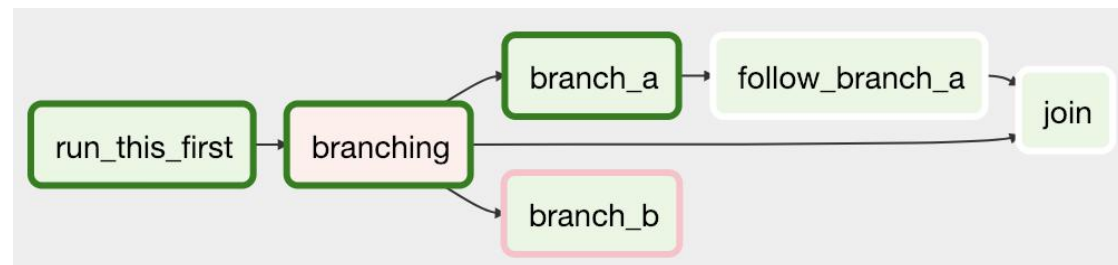


- 工作流的最小单元

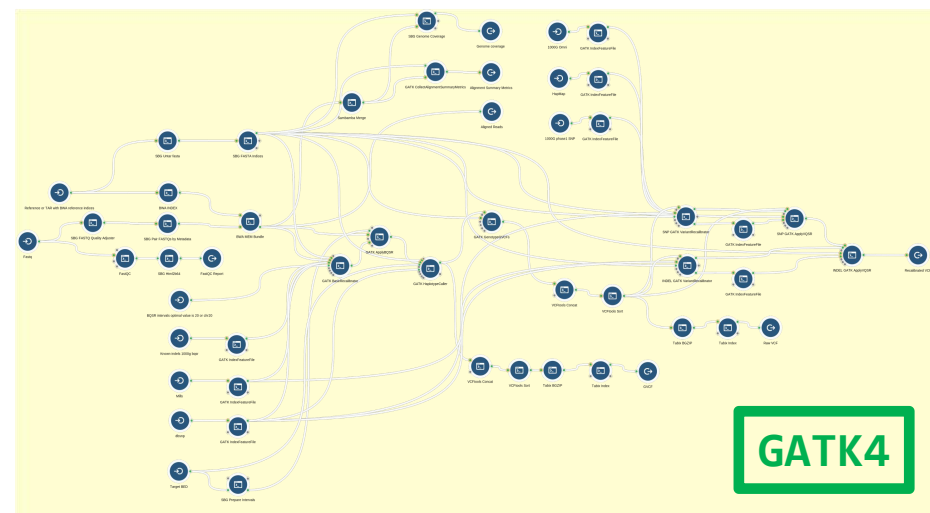


- **input** : 数据、参数、上一级的输出。。。
- **process** : 加减乘除、积分、拟合、重建。。。
- **output** : 数据、图像、模型。。。

- 有向无圈图 (Directed Acyclic Graph , DAG)



- 工作流可以非常复杂



技术方案 - 软件框架



- 数据处理软件框架 **Daisy** 核心架构
- 满足新一代光源数据处理需求的衍生技术模块
- 基于软件框架的学科专用应用软件以及针对灵活数据处理需求的通用工作流编排系统

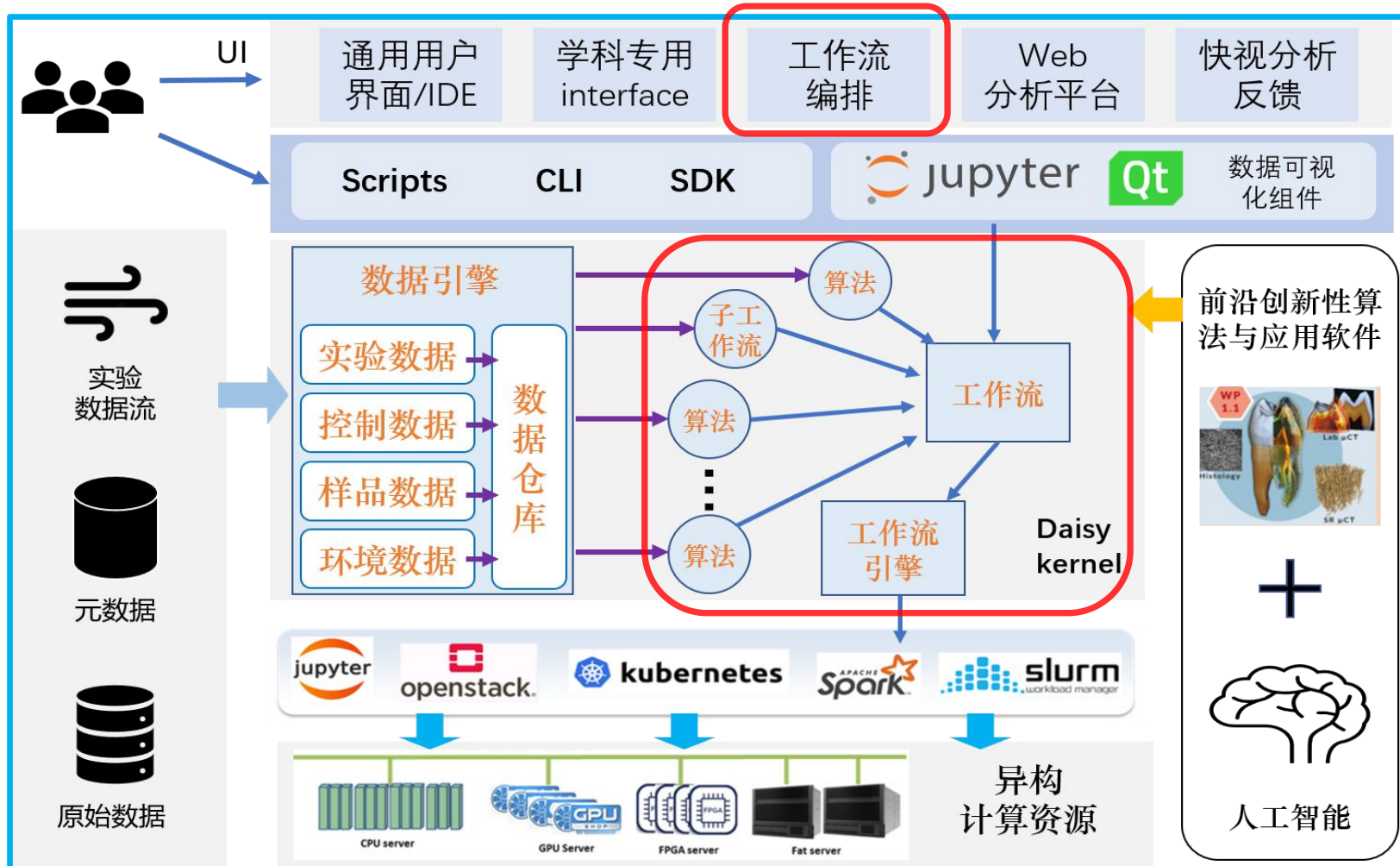
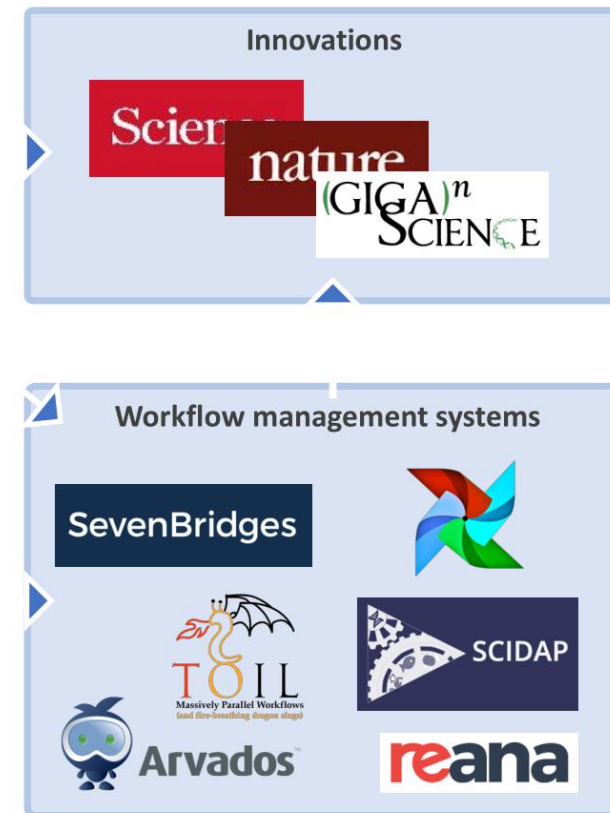
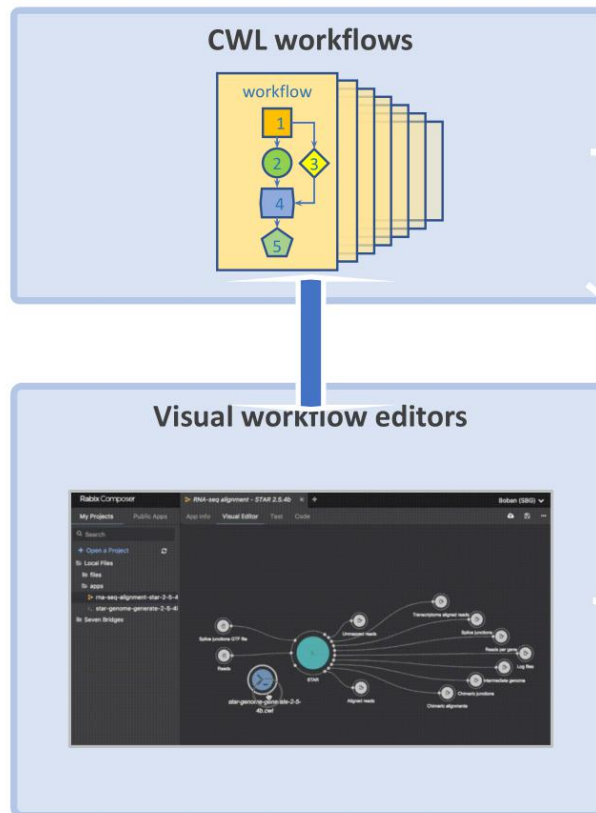


图 1. Daisy软件框架结构图及工作流系统模块的定位

技术方案 - 组成模块



- ✓ workflow描述语言
- ✓ 可视化编排界面
- ✓ 解析、执行引擎
- ✓ 日志、监控系统



SevenBridges



技术方案 - 架构图



- 前台是可视化编排和监控界面
- 中台是统一的通用 workflow 描述语言和运行日志
- 后台是 workflow 解析、调度、派发和日志记录、监控系统。

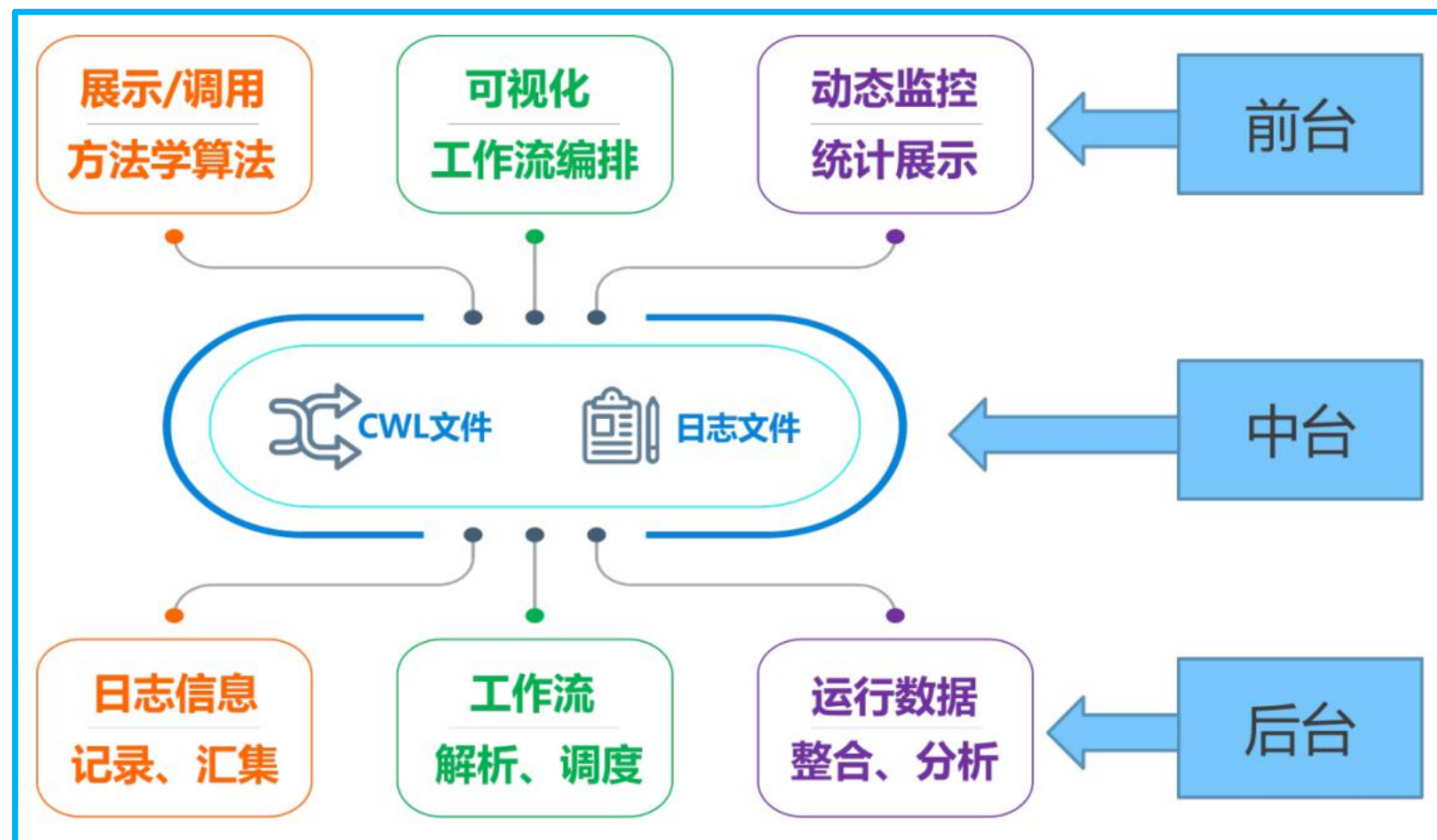


图 2. Daisy workflow 管理与编排系统架构图

前台：可视化编排和监控界面

- 已有算法**自动生成**计算节点和**集成**的方法学软件生成 workflow
- 鼠标**拖拽式**编排 workflow 和配置计算资源
- workflow **保存、读取、修改、复用**等
- workflow 执行时**动态监控和快反馈**、自动统计分析的图形化用户界面

中台：统一的通用 workflow 描述语言和运行日志

- 将 workflow 管理系统的图形化前台和负责业务逻辑处理的后台**解耦合**
- 选用了 **Common Workflow Language (CWL)** 做为通用的 workflow 描述语言
- workflow 文件和日志文件通过 Kafka (Redpanda) 做为**消息中间件**，前台和后台异步通信

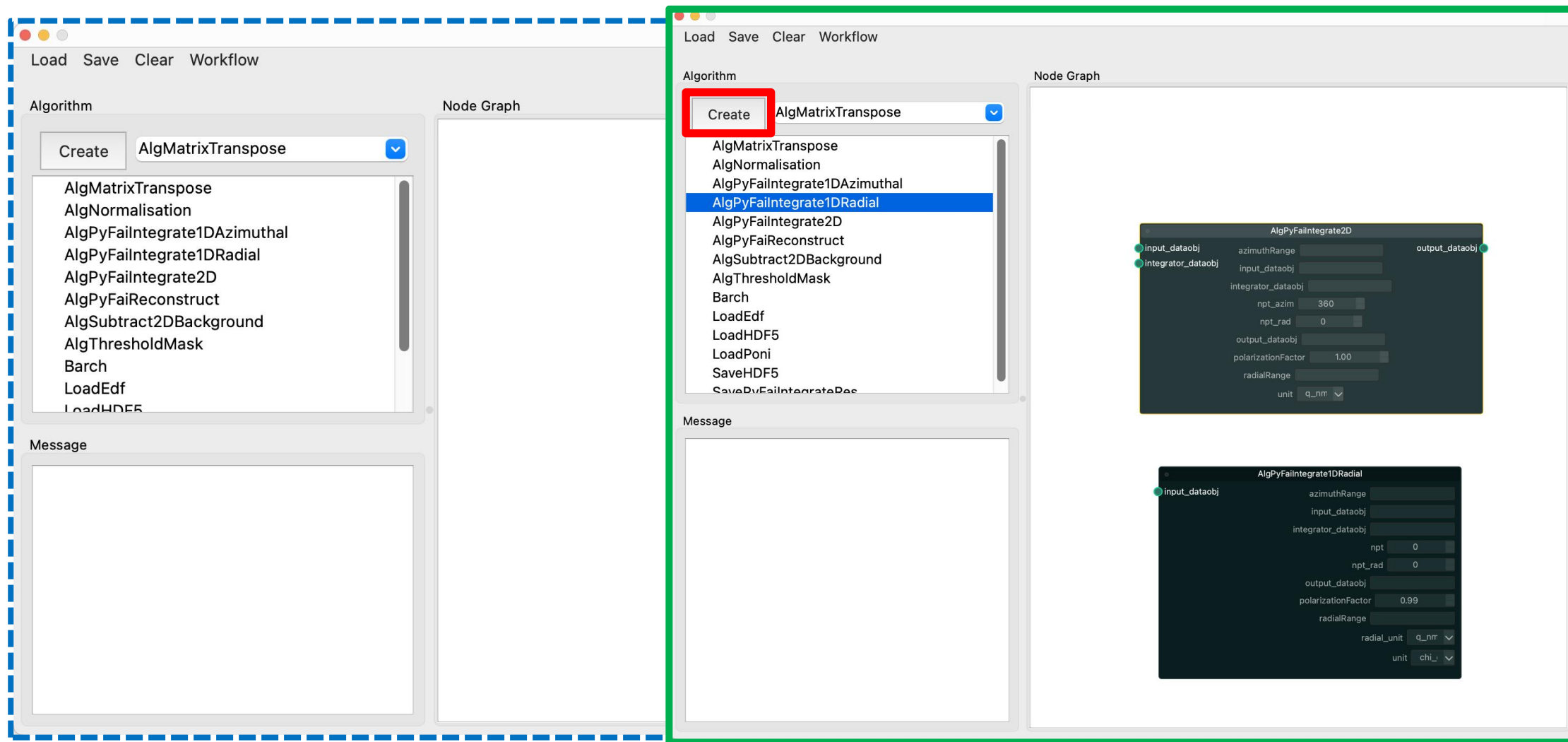
后台：统一的通用 workflow 描述语言和运行日志

- **CWL 文件解析器**，将其转换具体的 workflow 计算节点和运行逻辑
- 代码生成器引擎转换成软件框架核心层 workflow 引擎模块上可运行的代码来**调用计算引擎**
- **日志系统**的底层库是 quill (低延迟) 和 spdlog (高通量)，精简包装后释放给 Daisy 中核心层和各个子系统使用。

目前，开发了“两套”系统：

- 考虑到先进光源上数据处理中**批流一体、实时分析、离线处理**等特点；考虑到与软件框架**Daisy**各模块的对接，**定制化**开发了一套先进光源专用的 **workflow 管理系统**
- 考虑到粒子天体物理、科学数据管理等其他领域对 **workflow 的需求**，开发了一套**通用型 workflow 管理系统**
- 两套系统共享**统一的用户界面和使用逻辑**，背后都是标准的**CWL workflow 描述语言**。

通过 Daisy Workbench -> Interfaces -> Node Graph 进入：



The screenshot displays the Daisy Workbench interface, specifically the Node Graph view. The 'Create' button in the Algorithm list is highlighted with a red box. The 'AlgPyFailIntegrate1DRadial' algorithm is selected in the list. The Node Graph area shows two algorithm nodes: 'AlgPyFailIntegrate2D' and 'AlgPyFailIntegrate1DRadial'. The 'AlgPyFailIntegrate2D' node has inputs 'input_dataobj' and 'integrator_dataobj', and an output 'output_dataobj'. The 'AlgPyFailIntegrate1DRadial' node has an input 'input_dataobj' and an output 'output_dataobj'. The interface also includes a 'Message' panel at the bottom left.

编排 workflow 并运行：

The image shows a workflow editor interface with two main panels. The left panel, titled 'Algorithm', contains a list of available algorithms including AlgMatrixTranspose, AlgNormalisation, AlgPyFailIntegrate1DAzimuthal, AlgPyFailIntegrate1DRadial, AlgPyFailIntegrate2D, AlgPyFailReconstruct, AlgSubtract2DBackground, AlgThresholdMask, Barch, LoadEdf, LoadHDF5, LoadPoni, SaveHDF5, and SavePyFailIntegrateRes. The right panel, titled 'Node Graph', displays a workflow with two nodes: 'LoadHDF5' and 'SaveHDF5'. The 'LoadHDF5' node has input parameters 'input_dataobj' (zation/05-12/test.h5 Load) and 'input_path' (/data), and an output parameter 'output_dataobj' (data). The 'SaveHDF5' node has an input parameter 'input_dataobj' (data) and output parameters 'output_filename' (zation/05-12/test4.h5 Load) and 'output_path' (/data). A red box highlights the 'output_dataobj' parameter of the 'LoadHDF5' node and the 'input_dataobj' parameter of the 'SaveHDF5' node, indicating a data flow connection. Below the 'Node Graph' panel, a terminal window shows the execution log, including file sizes, counts, and elapsed time for the 'SaveHDF5' node.

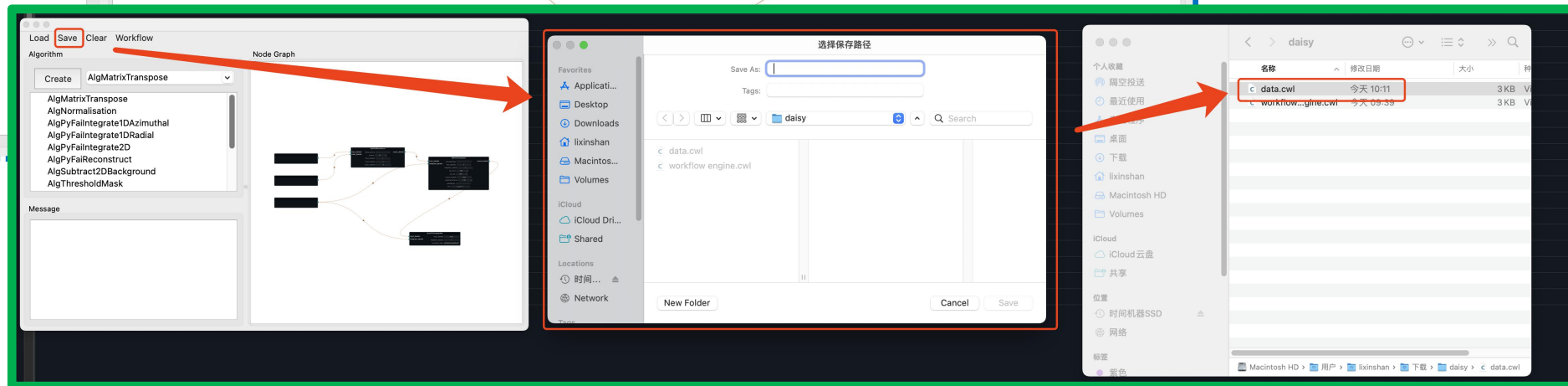
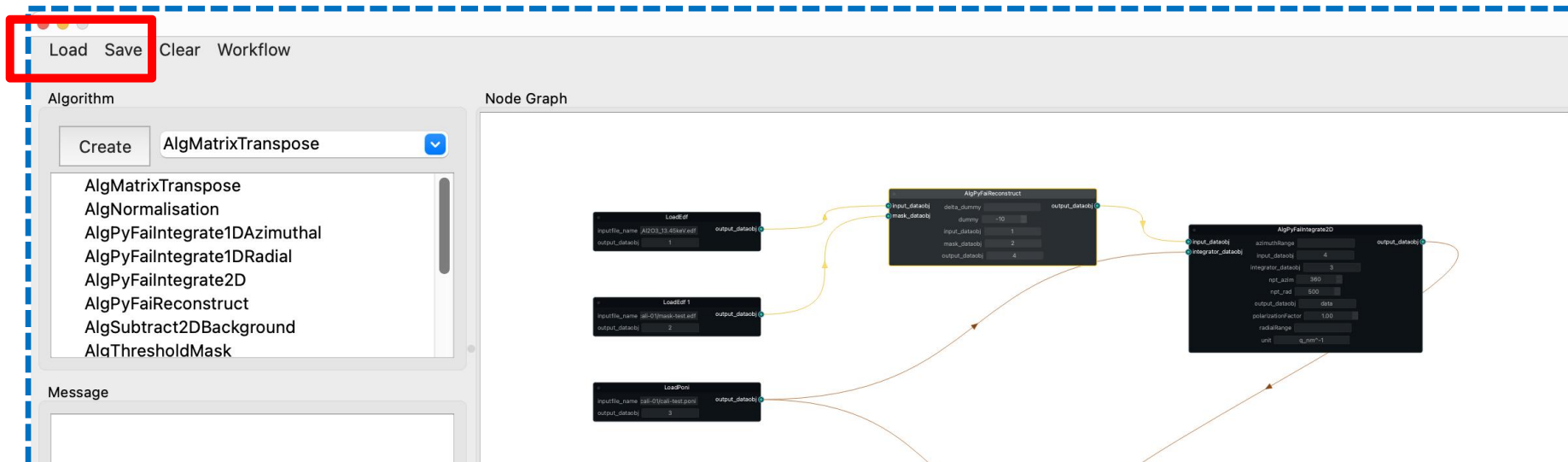
```
data_visualization/05-12/  
k/Daisy/DataHdlerAlg/  
python3.11/site-packages/  
B  
python3.11/site-packages/  
/Users/lixinshan/PycharmProjects/data_visualization/windows/widgets/message/  
Stream.py:46: size=346 B, count=2, average=173 B  
/Users/lixinshan/PycharmProjects/data_visualization/framework/Daisy/DataHdlerAlg/  
LoadHDF5.py:42: size=320 B, count=2, average=160 B  
start run SaveHDF5  
start run {"input_dataobj": "data", "output_filename": "/Users/lixinshan/PycharmProjects/  
data_visualization/05-12/test4.h5", "output_path": "/data"}  
SaveHDF5 run successfully  
Elapsed time: 0.002953 seconds  
/Users/lixinshan/PycharmProjects/data_visualization/framework/Daisy/DataHdlerAlg/  
LoadHDF5.py:50: size=905 KiB, count=3, average=302 KiB
```

算法数据的可视化：

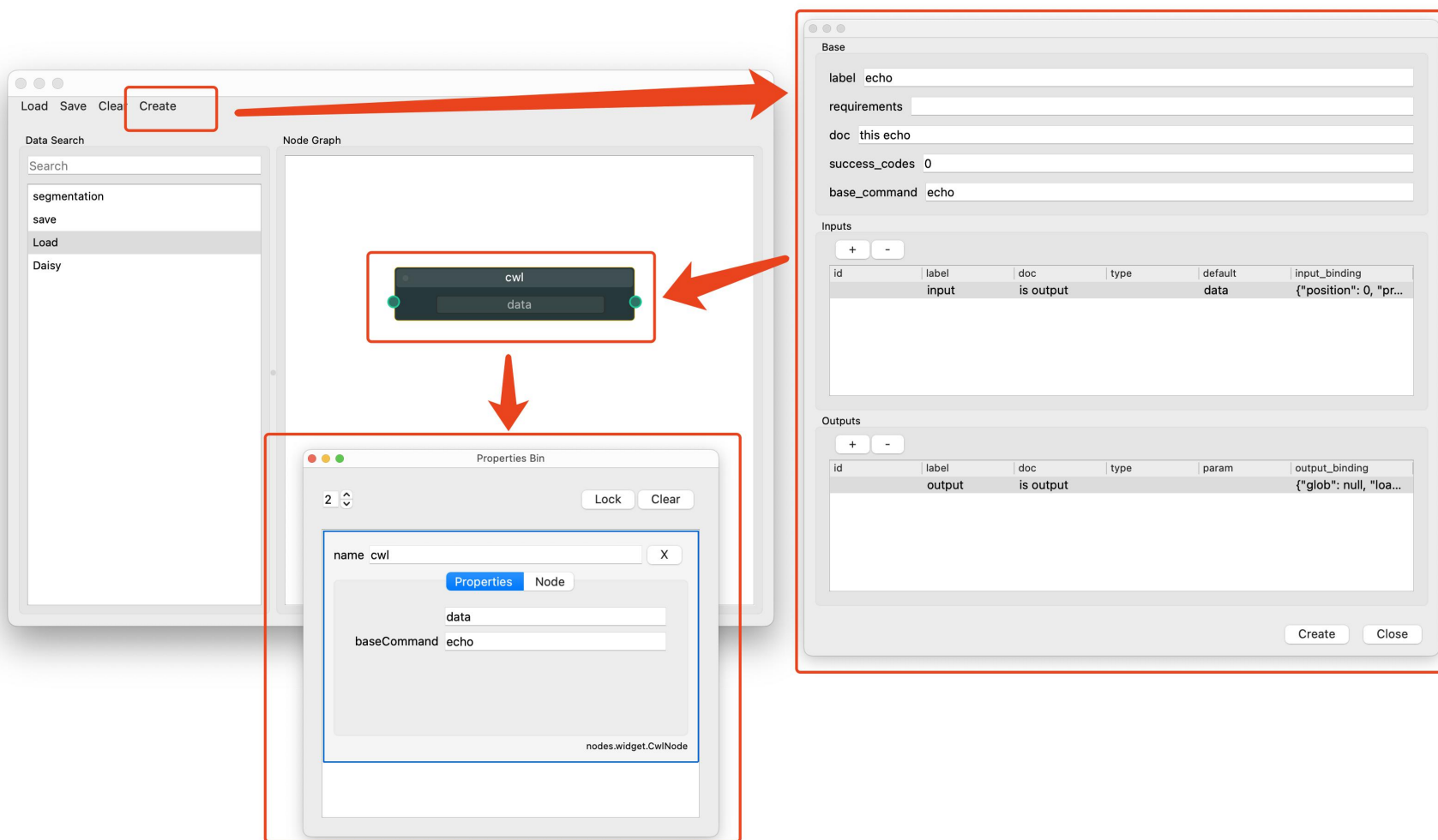
The screenshot displays a workflow visualization interface with the following components:

- Algorithm List:** A sidebar on the left lists various algorithms such as AlgPyFaiIntegrate2D, AlgPyFaiReconstruct, AlgSubtract2DBackground, AlgThresholdMask, Barch, LoadEdf, LoadHDF5, LoadPoni, SaveHDF5, and SavePyFaiIntegrateRes.
- Node Graph:** A central workspace showing a sequence of nodes: LoadHDF5 (green) and SaveHDF5 (green). A context menu is open over the SaveHDF5 node, with options: show plot, show visualizer 2D, delete, add input, and add output. A red arrow points from the 'show plot' option to a dialog box.
- Message Console:** A bottom-left window showing execution logs for 'AlgSubtract2DBackground'.
- Options Dialog:** A dialog box with 'Axis: Y' and 'Index: 0' fields, and 'Cancel' and 'OK' buttons.
- 2D Plot:** A window titled 'data' showing a 2D heatmap visualization. The x-axis ranges from 0 to 600, and the y-axis ranges from 0 to 100. The plot shows a complex pattern of colors (green, blue, purple) with vertical stripes.

读取和保存 workflow :



从零开始创建新的 workflow 节点：



The image illustrates the process of creating a new workflow node in a software interface. It consists of three main panels:

- Node Graph:** Shows a central node labeled 'cwl' with a 'data' input field. A red box highlights this node, and an arrow points from the 'Create' button in the top-left panel to it.
- Properties Bin:** A window titled 'Properties Bin' showing the configuration for the selected node. It includes fields for 'name' (set to 'cwl'), 'data', and 'baseCommand' (set to 'echo').
- Base Node Configuration:** A detailed configuration window for the 'Base' node. It includes fields for 'label' (set to 'echo'), 'requirements', 'doc' (set to 'this echo'), 'success_codes' (set to '0'), and 'base_command' (set to 'echo'). It also displays 'Inputs' and 'Outputs' tables.

id	label	doc	type	default	input_binding
	input	is output		data	{"position": 0, "pr...

id	label	doc	type	param	output_binding
	output	is output			{"glob": null, "loa...

- 日志系统

信息窗的stdout/err捕捉 → 基于日志文件/信息中间件的保存、解析和流处理等

- 监控系统

简易的回调函数嵌入 → 基于日志流分析的实时监控、错误定位、统计展示等

- 与Daisy数据IO和计算引擎对接

- 界面优化、细节功能完善。。。