

第二十届全国科学计算与信息化会议



散裂中子源科学计算平台

中国科学院高能物理研究所/中国散裂中子源

侯丰尧

2023年7月10日

报告内容

- 当前物理学发展的技术挑战
- 散裂中子源介绍
- 科学计算平台建设及实践

当今物理学研究面临的挑战



“一方面，现代物理学的实验越来越复杂，费用越来越高，其中每一项实验都需要好几年去准备和执行。现在高能物理实验需要周期很长，而不幸的是，以后会变得越来越长。另一方面，高能物理理论也越来越复杂，理论物理学家之间、理论物理学家和实验物理学家之间，越来越充满隔阂”。

---杨振宁《物理学的未来（1961年）》后记，2018年

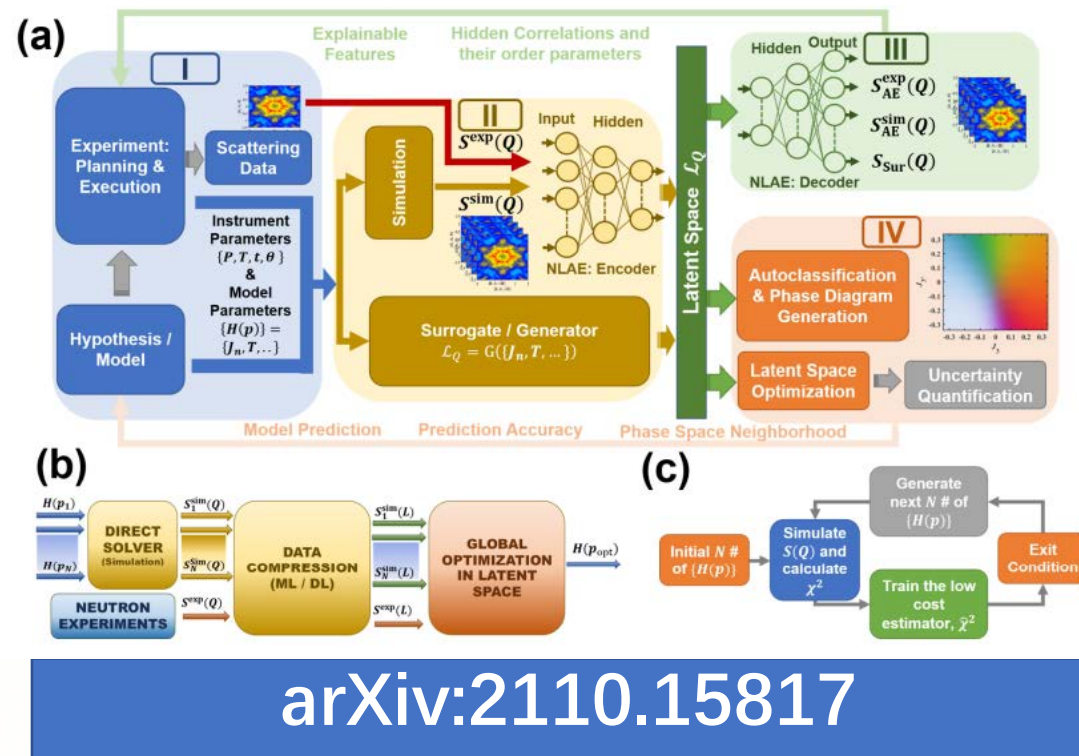
AI + 数据 + 科学计算的融合成为当前研究热点

■ 领域大模型

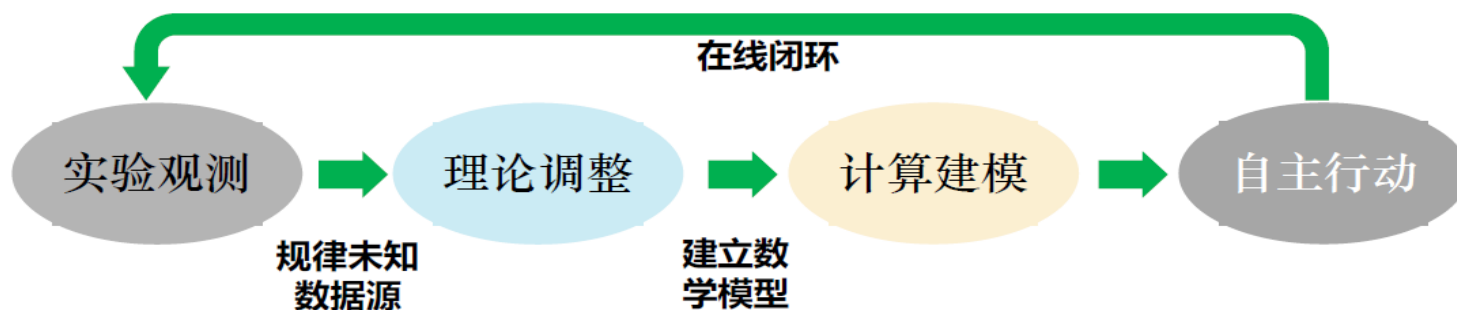
- 基于已有的理论知识和海量科学数据，构建领域知识大模型

■ 科研活动智能化

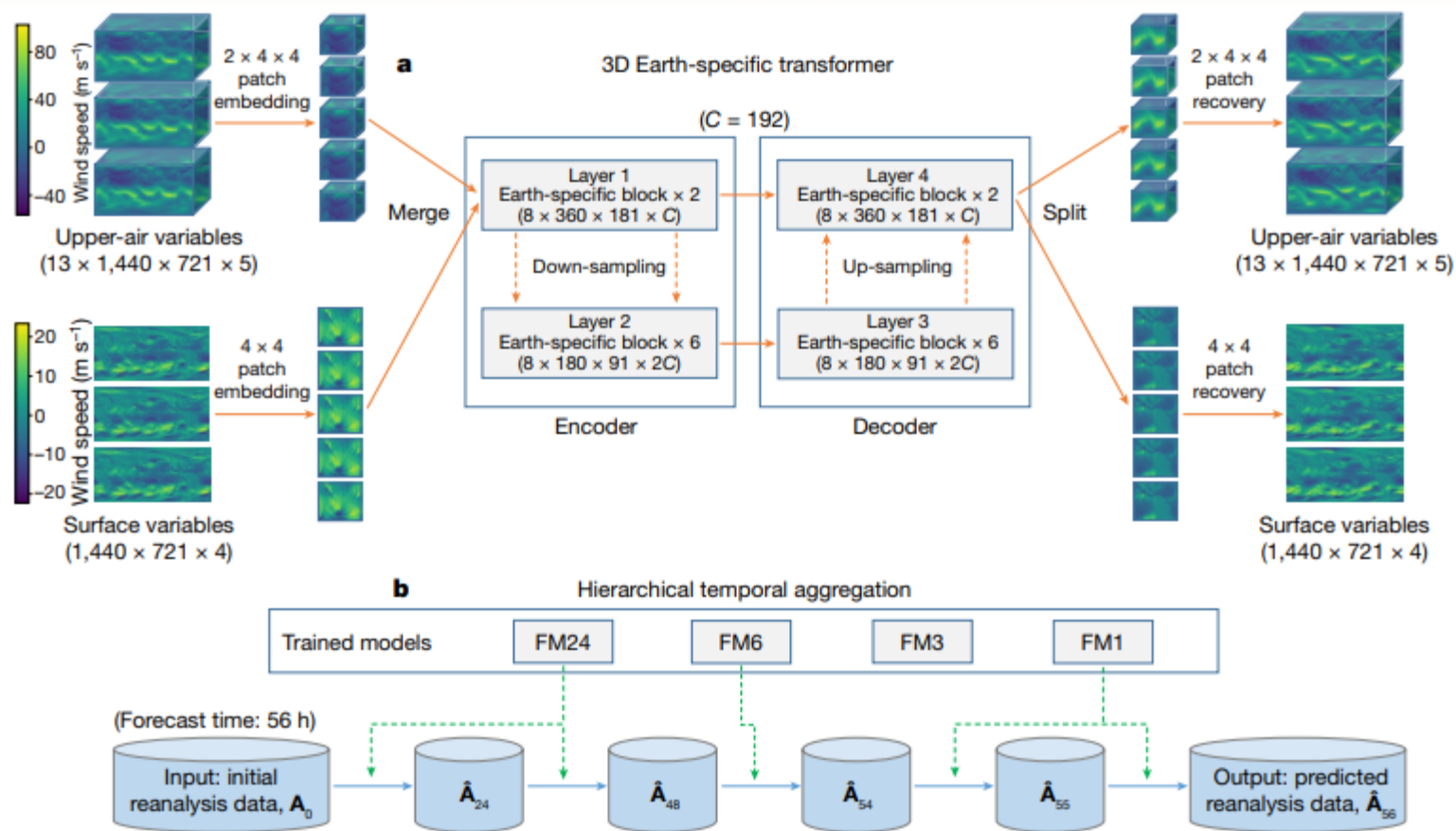
- 实现科研活动从计划、推理、实验、分析等过程的智能化



通过多范式快速循环迭代，互相激荡，产生顿悟和知识涌现

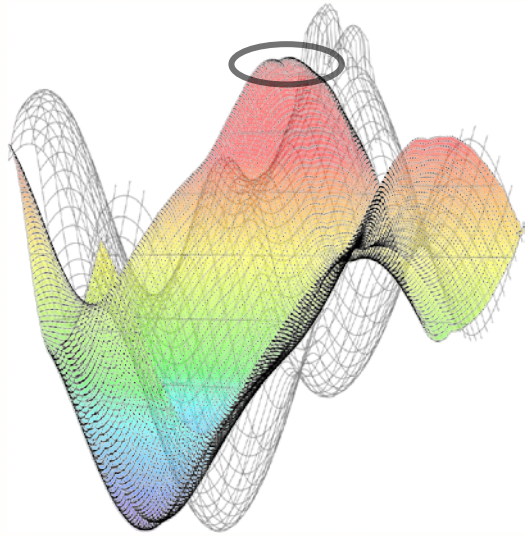


基于大模型的相关研究及应用成为热点

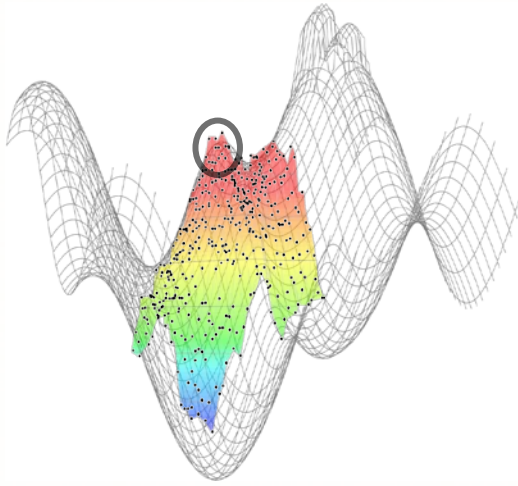


Bi, K., Xie, L., Zhang, H. et al. Nature (2023).
<https://doi.org/10.1038/s41586-023-06185-3>

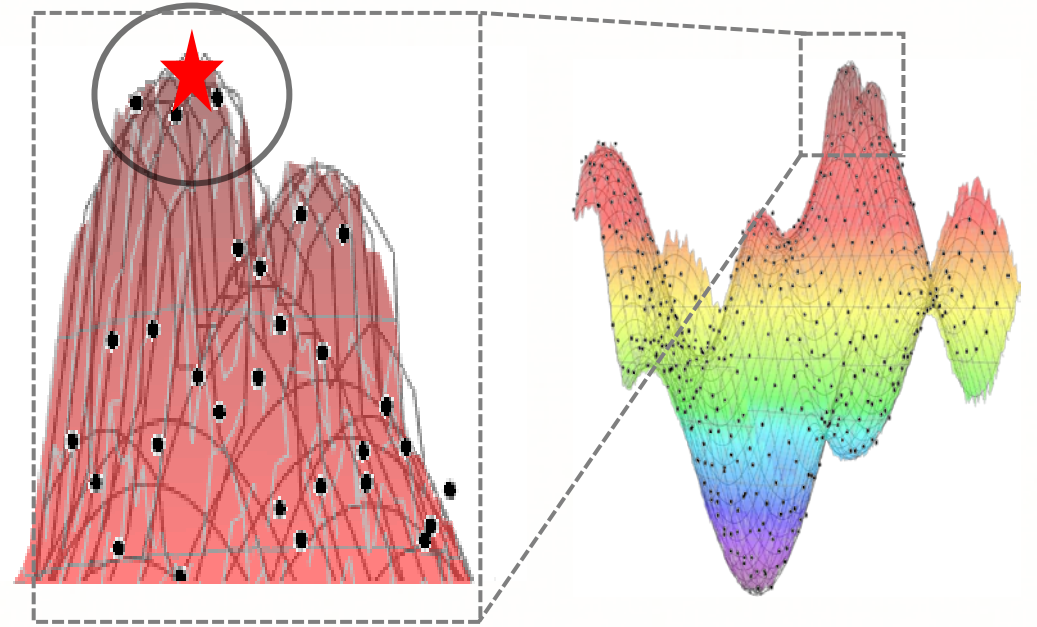
基础物理中大模型与实验数据的融合与挑战？



+



?



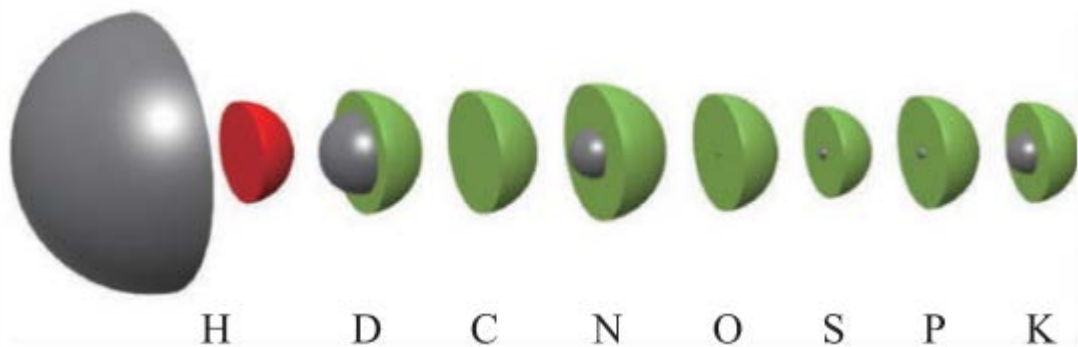
1. 基于海量计算数据
构建大模型

2. 少量实验数据, 微调
模型参数

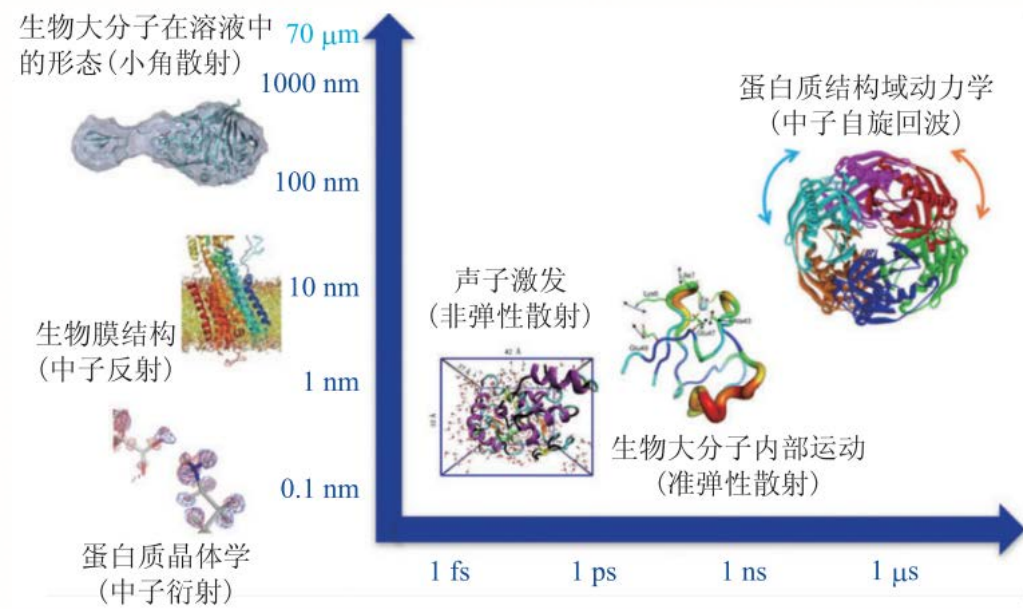
3. 理实交融获得高精度的大模型

中子散射是探索物质微观结构的有力手段

- 不带电、不存在库仑势垒，**穿透力强**
- 具有磁矩，可**与原子磁矩发生电磁作用**
- 对轻元素敏感、**能分辨同位素**
- 波长与原子分子尺寸相当，可**研究物质结构和运动状态**



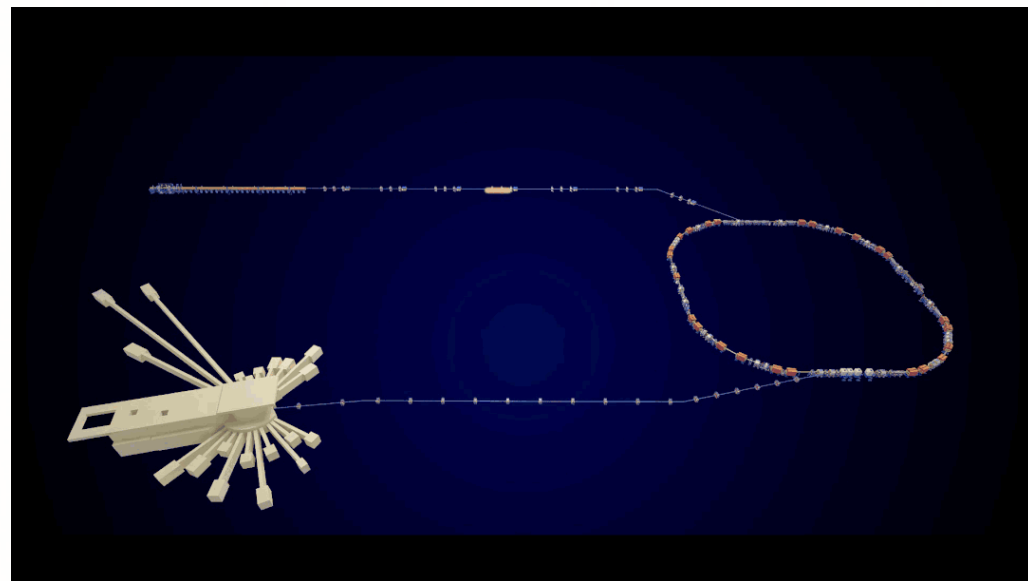
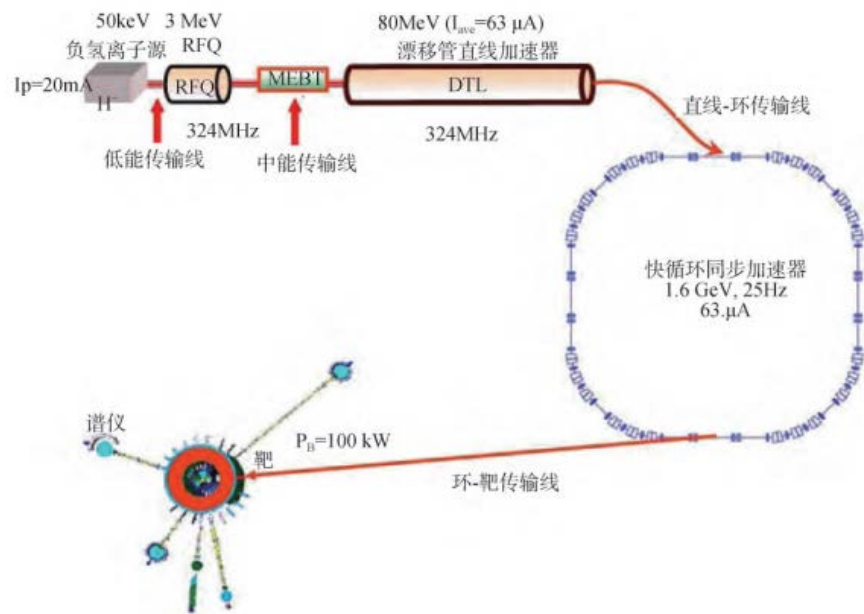
不同元素对中子的非相干散射截面和相干散射长度。非相干散射截面用左半球表示，相干散射长度用右半球表示。红色的半球表示氢原子的相干散射长度为负，绿色的半球表示其他原子的相干散射长度为正



不同中子散射技术在探索生物体系结构和动力学方面的应用

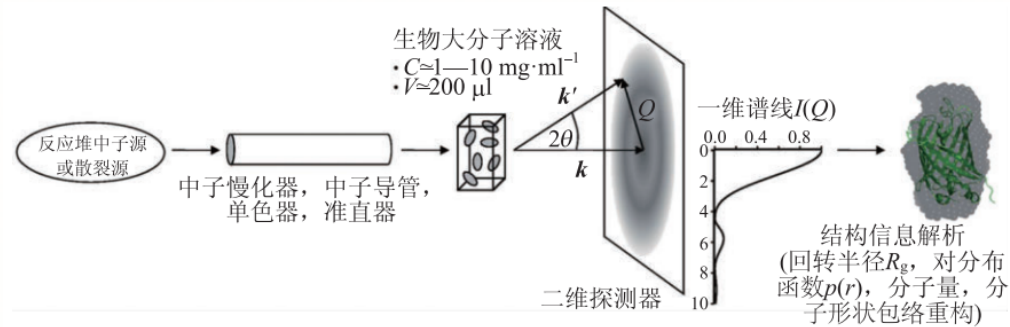
散裂中子源是探索微观世界的超级显微镜

- 依托高能物理技术建设的多学科交叉科研平台
- 新一代科学研究与材料研发的大科学装置，全球第四台
- 不使用核燃料，热功率低，只产生极少量活化产物，环境友好
- 高脉冲通量、丰富的高能短波中子、优越的脉冲时间结构、低本底

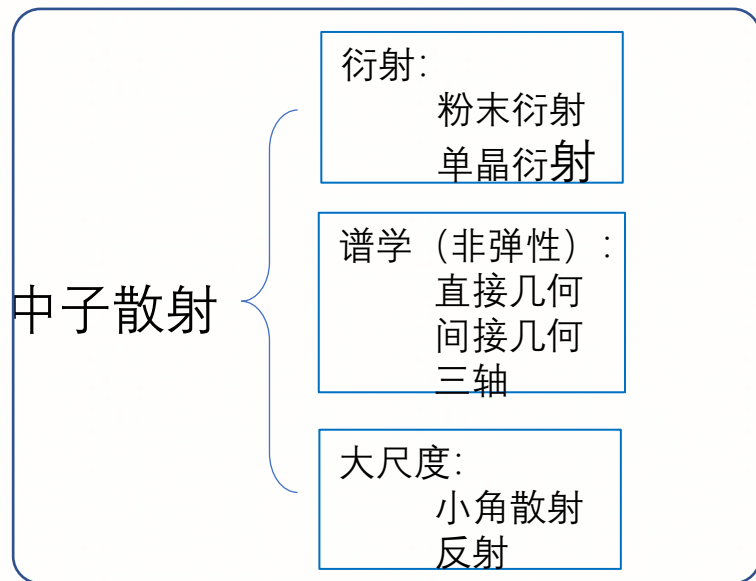


中子散射过程及其应用

■通过测量散射后中子能动量变化，研究样品微观结构和运动规律。



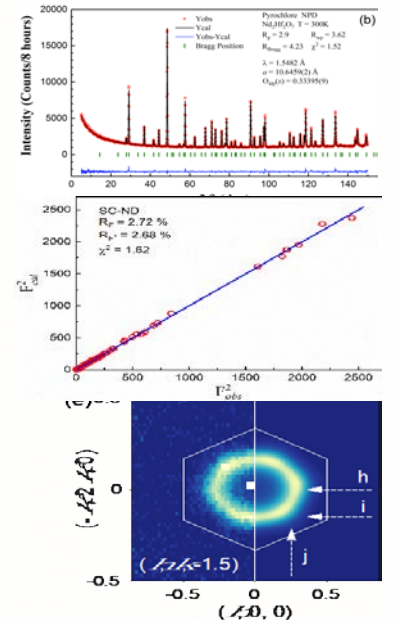
■不同中子散射技术在生物体系结构和动力学方面的应用



数据处理: Mantid
 时间零点校正
 机械零点校正
 探测器效率校正
 样品吸收校正
 强度归一化
 单位转换
 数据格式转换



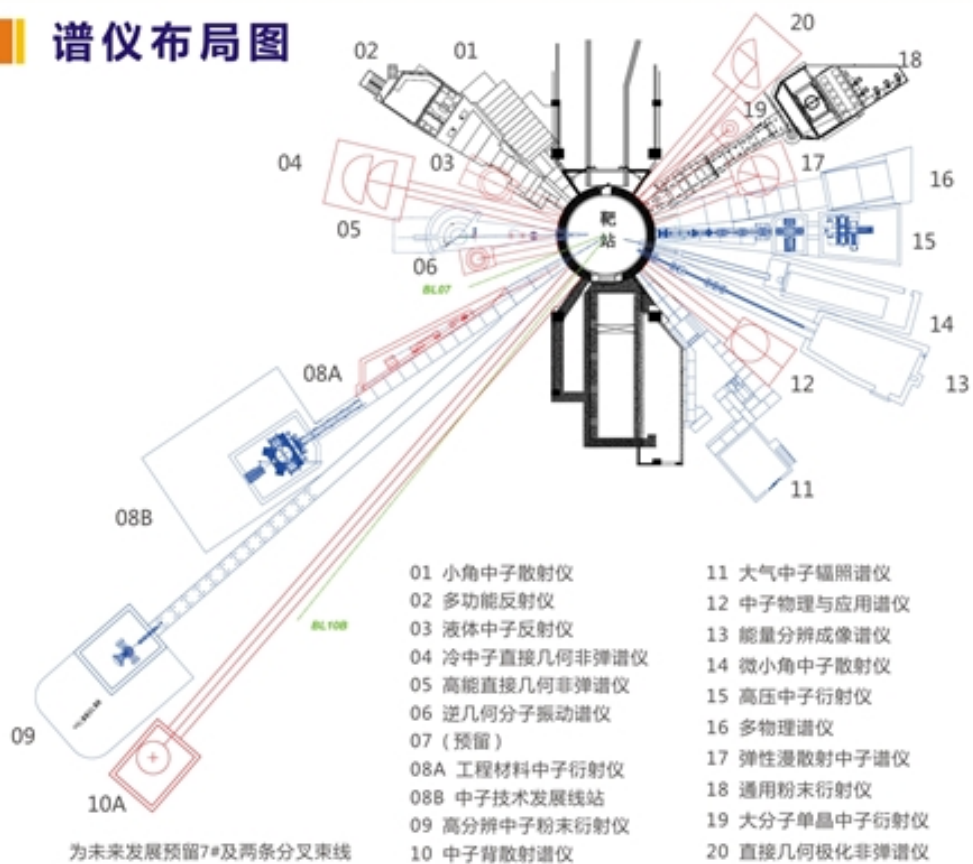
数据分析:
 Fullprof
 Jana
 Shelx
 RMC-discord
 Olex



数据需求：CSNS运行产生海量的实验数据

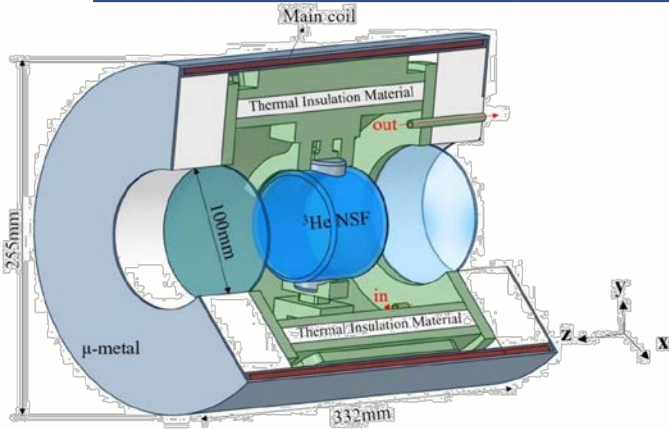
■ 实验过程产生大量的实验数据需要及时处理、分析

■ 谱仪布局图



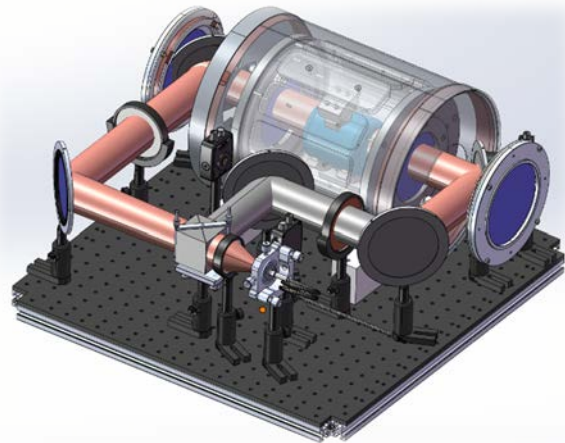
谱仪名称	CSNS每天产生数据量 (Byte/Day)	CSNSII每天产生数据量(Byte/Day)
能量分辨中子成像谱仪	2048G	8192G
高压中子衍射谱仪	150G	600G
工程材料中子衍射谱仪	150G	600G
高分辨中子粉末谱仪	50G	200G
微小角中子散射谱仪	60G	240G
多物理谱仪	512G	2048T
大气中子辐照谱仪	21G	84G
高能直接几何非弹性中子散射谱仪	150G	600G
合计	3.14T	12.56T

仿真需求：CSNS II 立项提出了新的建设任务



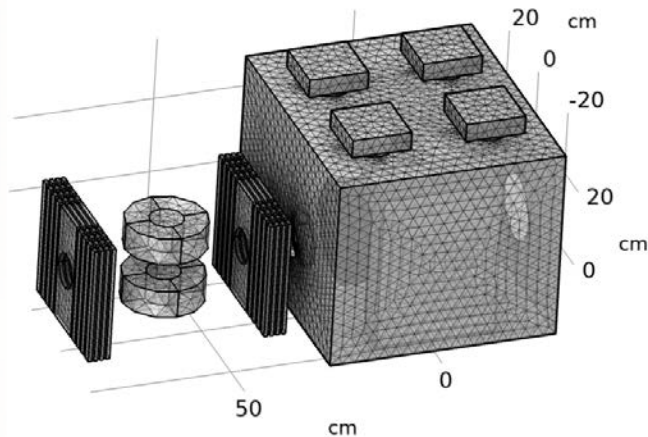
极化氦三系统温度场模拟

- 模拟设计参数：网格划分（5000万）、热源设置、边界条件（流固耦合）。
- 计算时间：约30个小时（16核并行计算）
- RAM占用：约80-100GB



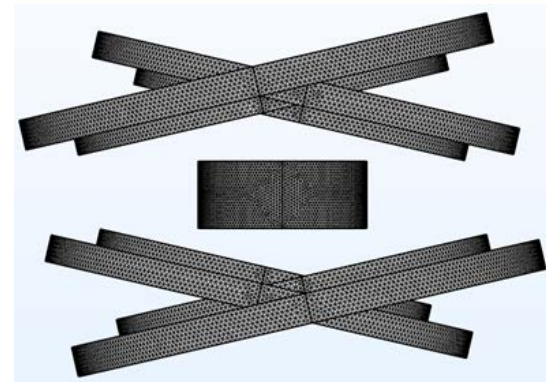
极化氦三系统激光光场模拟

- 使用模块：射线光学模块
- RAM占用：30 GB
- 网格划分尺寸：1 mm（激光覆盖区域）



极化氦三系统磁场梯度模拟

- 模拟设计参数：0.1cm-0.2cm（局部）+0.5cm-6.3cm
- 计算时间：约10小时（25核多核计算时长）
- RAM占用：约60-80 GB



广角极化氦三系统磁场梯度模拟

- 模拟设计参数：线圈大小、绕线匝数、绕线层数、线圈电流大小、线圈电流方向、线圈之间的夹角角度、气室区域大小、气室区域磁感应强度、气室区域磁场梯度

中子散射和同步辐射光源具有很强的互补性

- 中子散射和同步辐射都是多学科交叉前沿研究和应用的大科学平台，是研究物质结构和动力学的强有力工具，互为补充。



计算需求：多模态多尺度复杂系统的研究



第二代(HLS)
光子能量~ 10^3 eV
空间分辨率~30nm
能量分辨率~0.2eV



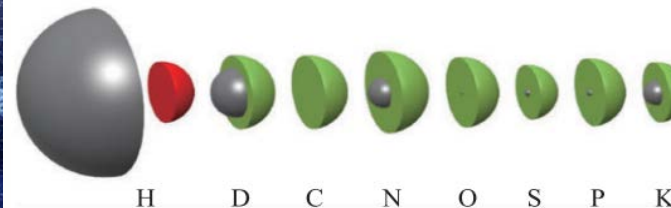
第三代(SSRFII)
光子能量~ 10^5 eV
空间分辨率~10nm
能量分辨率~100meV
时间分辨率~10ps



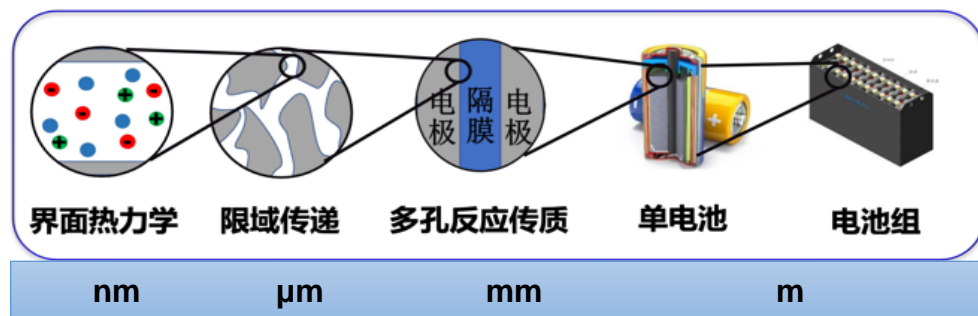
第四代(HEPS)
光子能量~ 10^5 eV
空间分辨率~10nm
能量分辨率~meV
时间分辨率~ps



低能光源适合电子结构的研究；
中高能光源适合原子结构及演变的研究；
亮度越高光源空间、时间及能量分辨率越高；



H、Li、C、O等元素对中子有较好相干性；
具有磁矩，可与原子磁矩发生电磁作用；
不带电、不存在库仑势垒，穿透力强；



量子力学 分子动力学 有限元计算

多尺度建模与计算模拟技术；
多相体系及其演化需要联合多种技术；

散裂中子源科学计算服务平台基础环境

■ 硬件系统

- 登录节点: CPU/GPU/ARM
- 计算节点: CPU/GPU/ARM
- 网络: 以太网、IB
- 存储: 共享存储

■ 管理系统

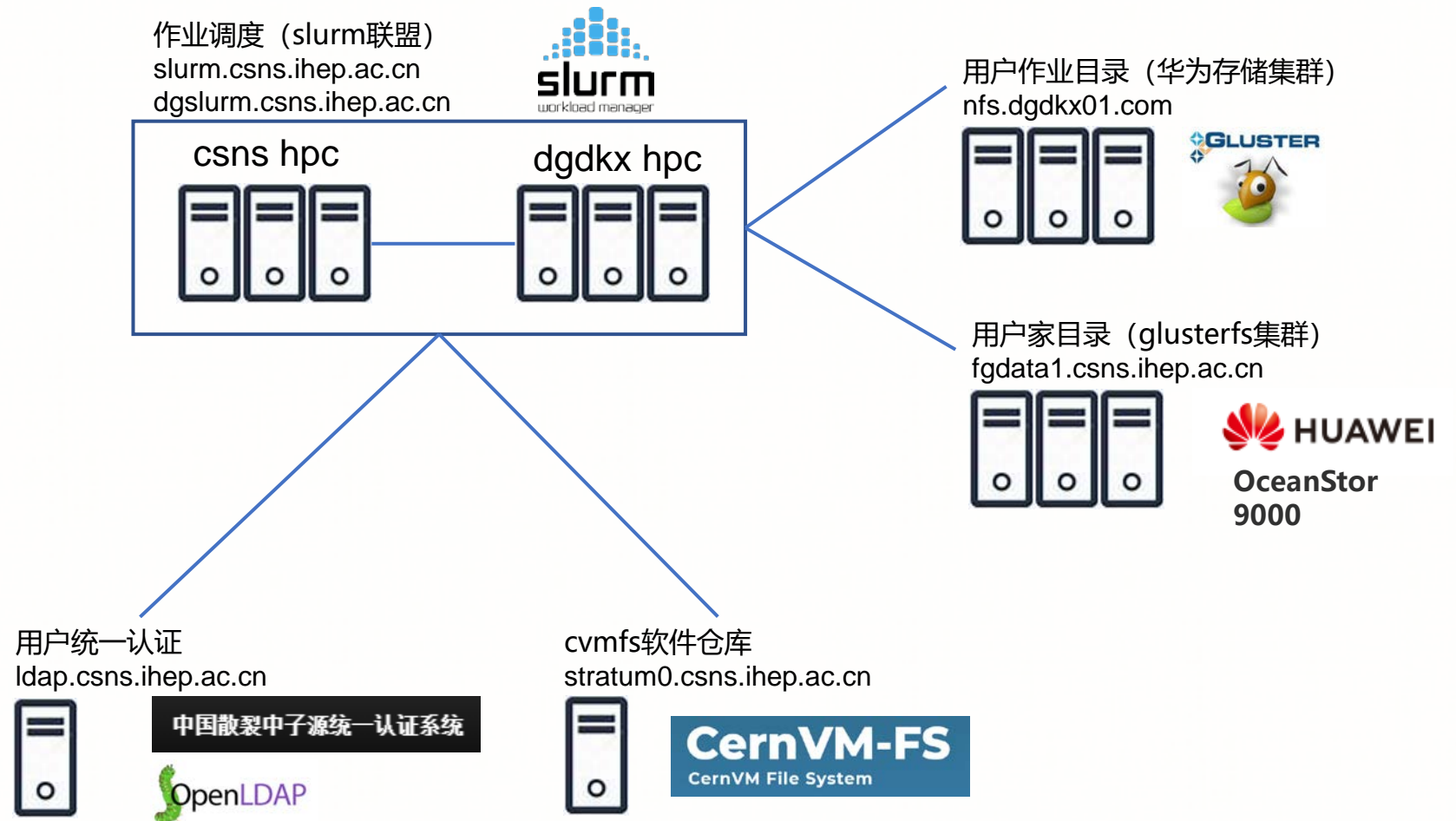
- 用户管理: OpenIldap
- 资源管理: SLURM
- 监控管理: Grafana
- 软件管理: CVMFS

■ 软件系统

- 操作系统: Linux
- 管理软件: Slurm
- 编译环境及并行库:
- 应用软件: ...

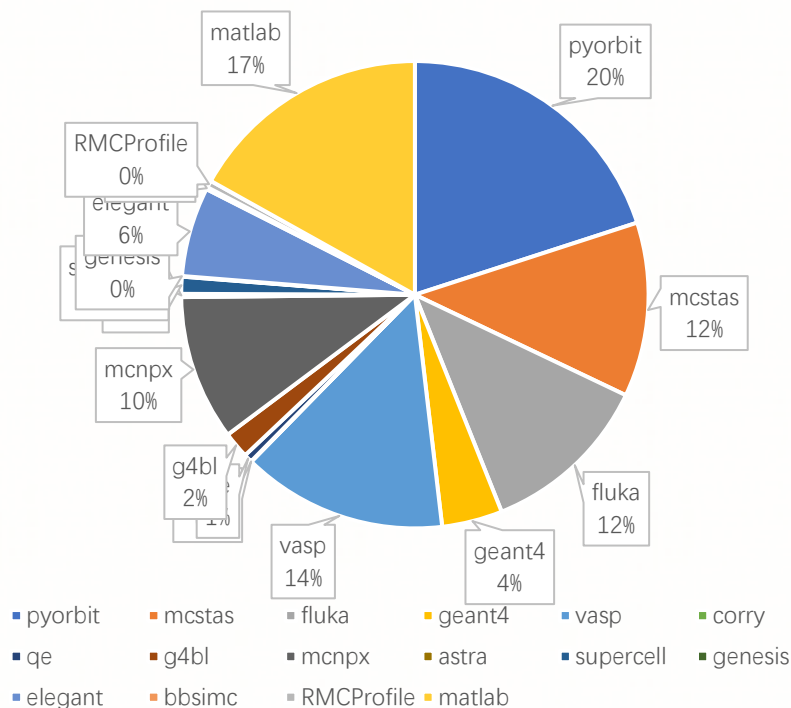
■ 配套系统

- 机房、空调、消防、电力

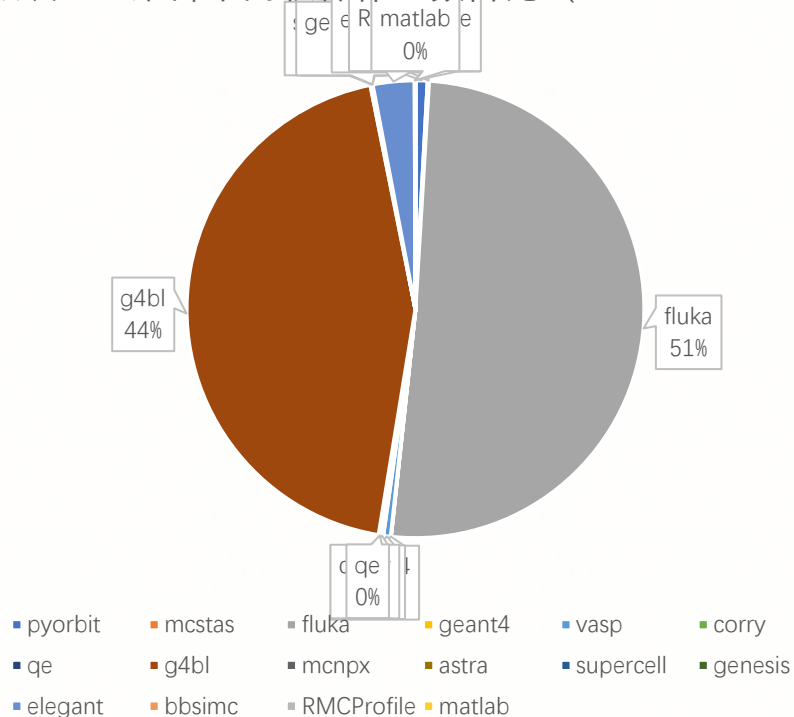


散裂计算集群不同软件使用情况

散裂HPC集群不同软件机时占比 (2022.5-2023.5)



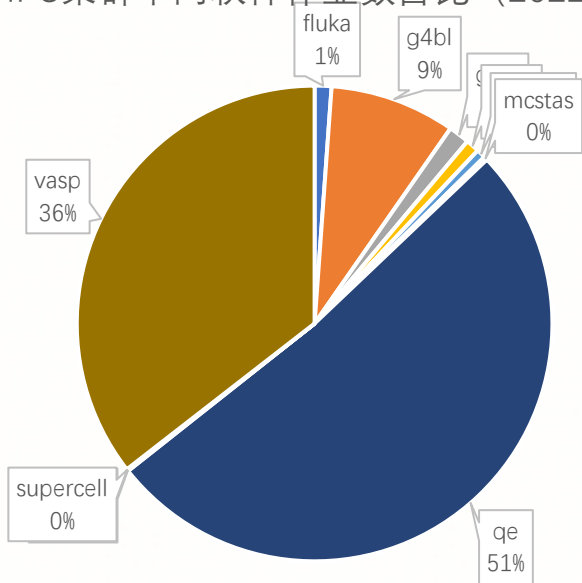
散裂HPC集群不同软件作业数占比 (2022.5-2023.4)



- 作业数最多的依次是束线设计的Fluka(50.76%)和辐射评估软件g4bl(44.27%);
- 作业机时占比最多的依次是加速器设计pyorbit(20.01%)、工程模拟Matlab(16.90%)、第一性原理计算VASP (14.20%)、束线设计mcstas(12.07%)、辐射评估Fluka(11.87%)

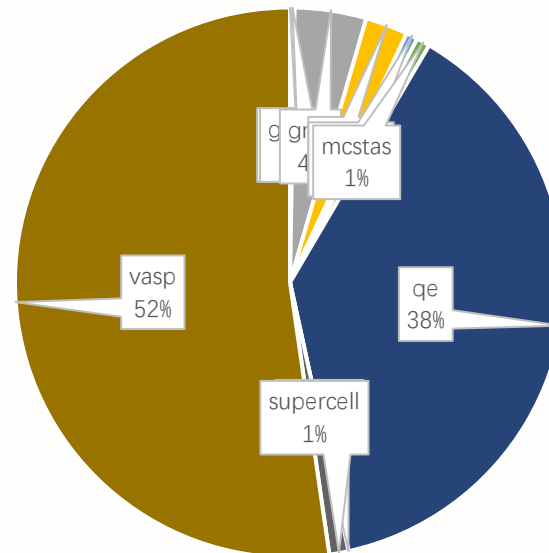
大科学数据中心不同软件使用情况

东莞大科学HPC集群不同软件作业数占比 (2022.05-2022-04)



■ fluka ■ g4bl ■ gmx ■ lqcd ■ mcnpx ■ mcstas ■ qe ■ geant4 ■ supercell ■ vasp

东莞大科学HPC集群不同软件机时占比 (2022.05-2023.04)



■ fluka ■ g4bl ■ gmx ■ lqcd ■ mcnpx ■ mcstas ■ qe ■ geant4 ■ supercell ■ vasp

- 作业数最多的依次是材料计算QE(51%)、第一性原理VASP(35.51%)、辐射评估软件g4bl(8.54%);
- 作业机时占比最多的依次是第一性原理VASP(52.29%)、材料计算QE(38.13%)、分子动力学GMX (4.21%)

散裂中子源科学计算平台 (SC@CSNS)

- 整合现有软硬件计算资源，面向用户建设的作业提交、文件管理、应用集成的在线科学计算服务。
- 2022年12月开放试用，2023年1月正式上线。

计算模拟

工程仿真

机器学习

应用集成

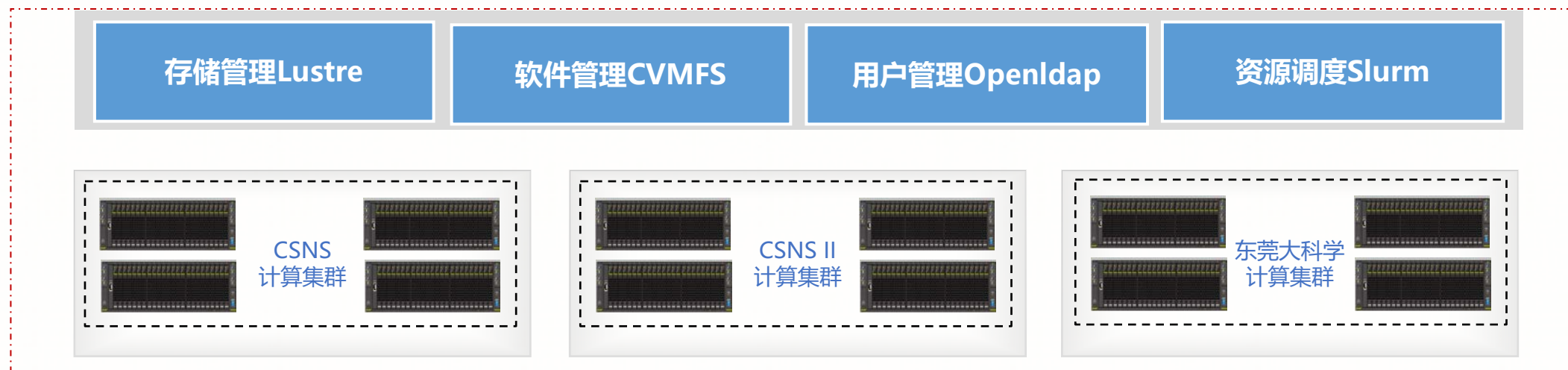
科学计算平台 (SC@CSNS)

散裂计算中心

东莞大科学

.....

散裂中子源科学计算平台系统架构



散裂中子源科学计算平台功能与应用



科学计算服务

工程计算

图形桌面

仿真计算

作业管理

数据管理

命令行 [E-Shell]

支持与服务



houfy

我的应用

更多应用

① 相关应用的版权信息，请去往 [版权声明](#) 查看

BASE

基础模板



模板 命令行 图形

COMSOL

COMSOL



Quantum
ESPRESSO



MCNPX



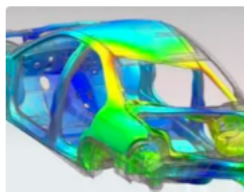
ROOT-CERN

LAMMPS

LAMMPS

GROMACS

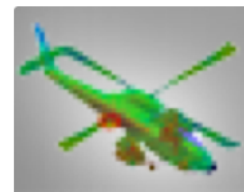
GROMACS



ANSYS Autodyn

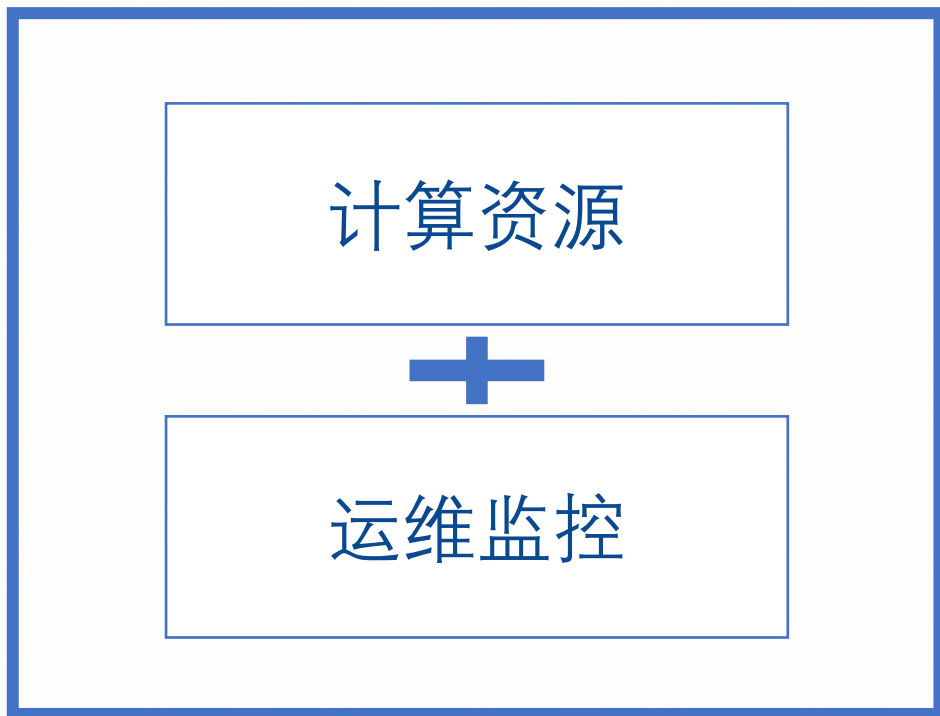
FLUENT

ANSYS Fluent

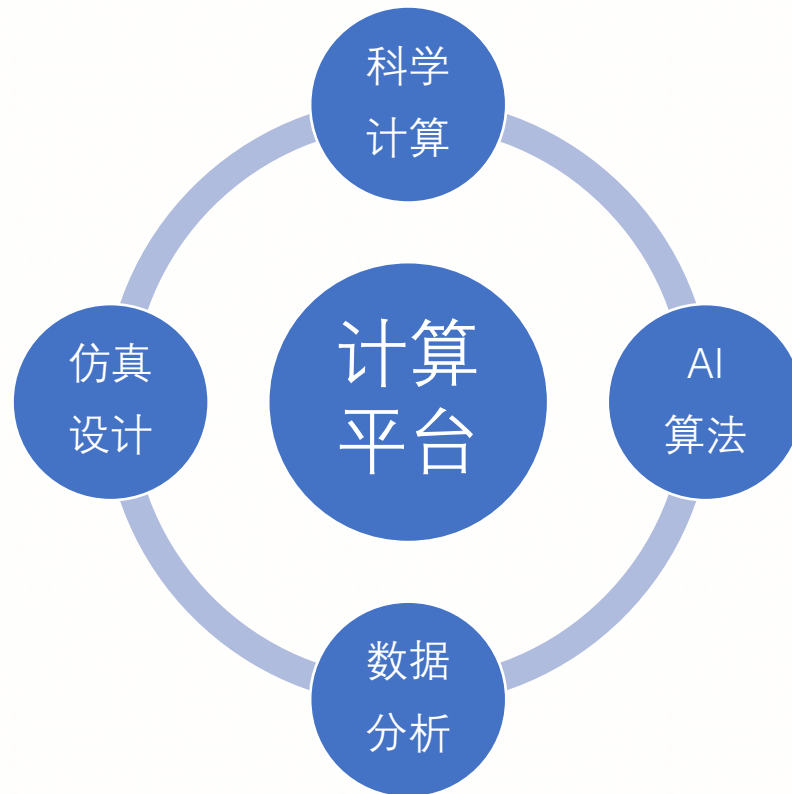
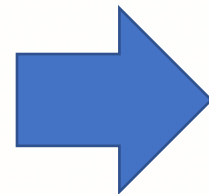


ANSYS HFSS

散裂中子源科学计算平台定及目标



资源管理

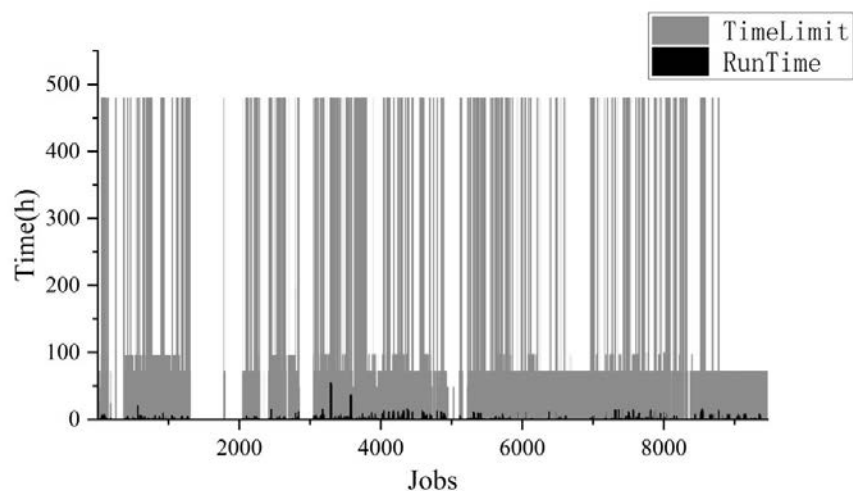


用户服务

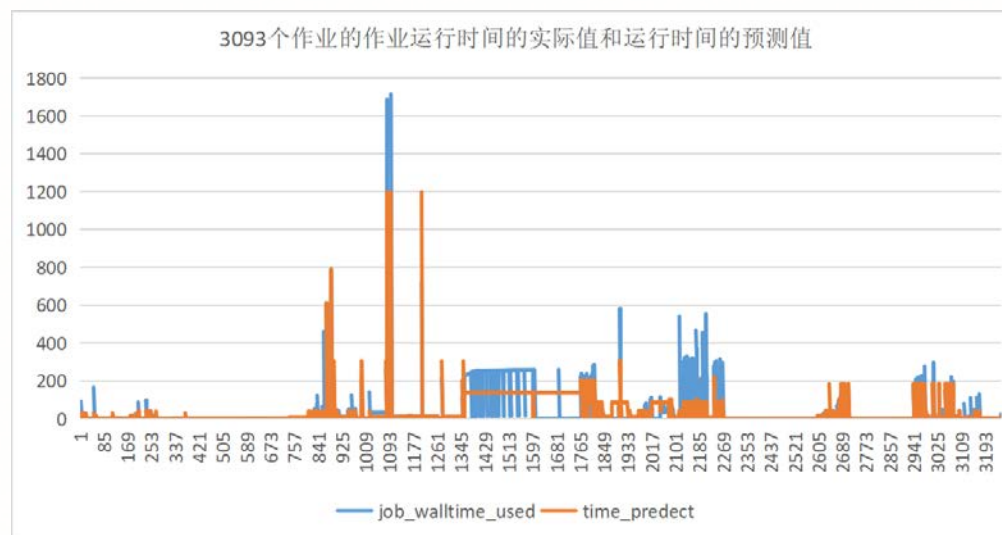
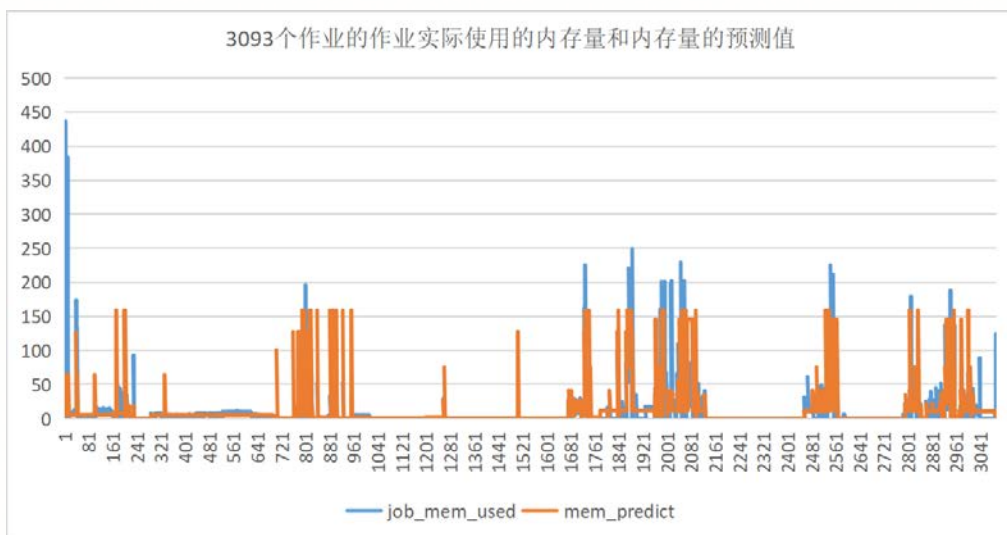
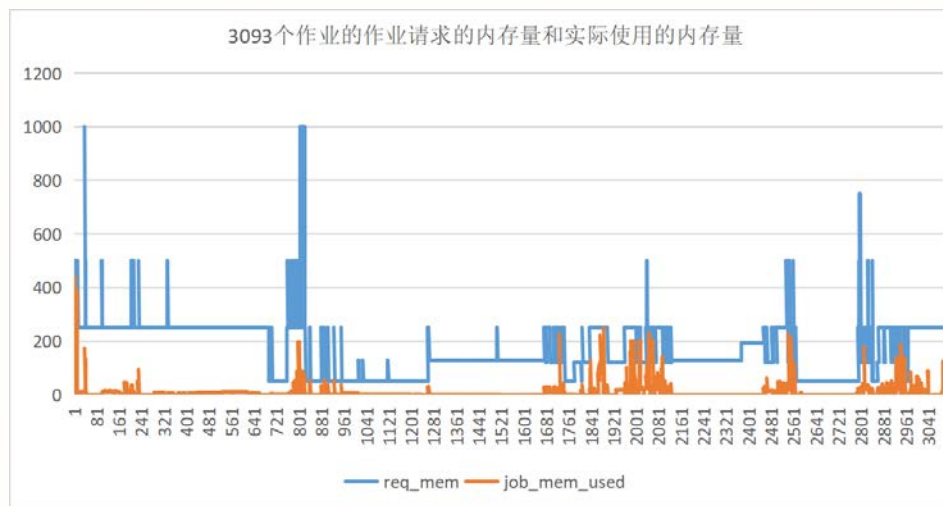
目标：计算即服务(CaaS, Computing as a Service)

CSNS高性能计算集群运行及优化

CSNS计算平台用户申请资源与实际占用资源对比



基于用户作业特征采用回归预测的资源占用情况分析



CSNS高性能计算集群运行及优化

■如何改善计算集群性能？

■系统优化：

- 优化计算节点配置，改善计算集群稳定性

■调度优化

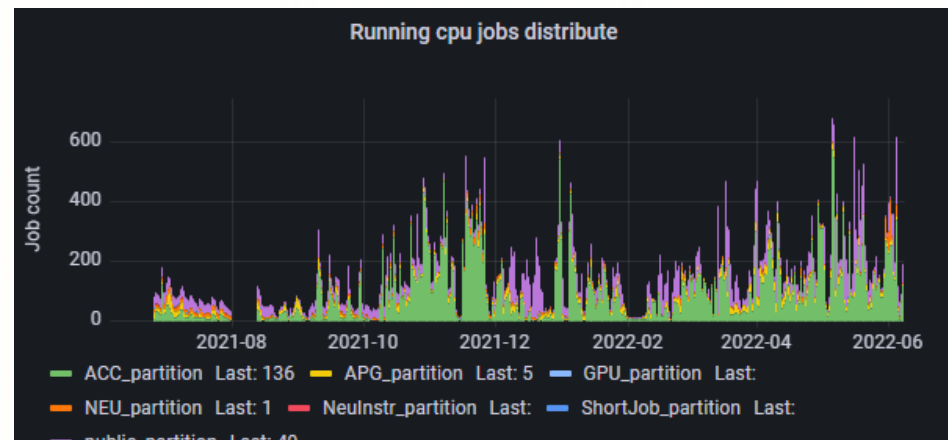
- 调整作业调度策略，缓解计算集群系统压力

■性能优化：

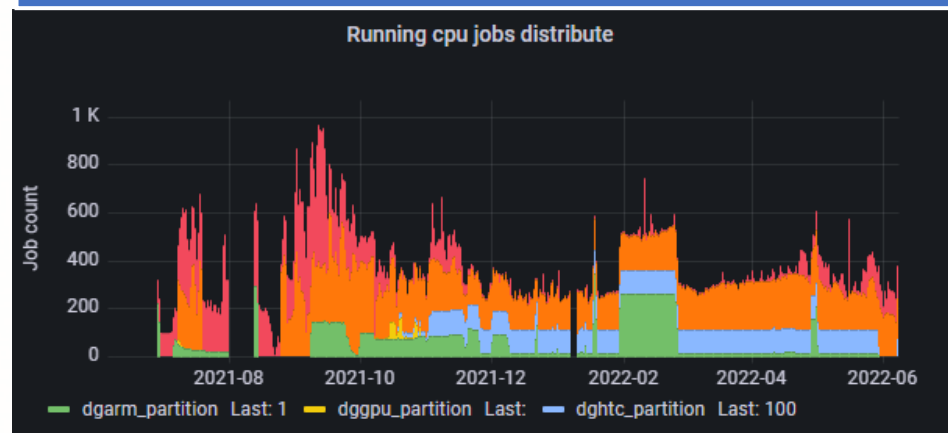
- 分析ARM节点作业特征，发现内存带宽影响较大，实测提升内存带宽性能可提升24.7%

■应用优化

- 分析作业（Quantum Espresso）运行特征，优化作业I/O方式，明显提升应用性能。



散裂中子源集群2021.6-2022.5



东莞大科学中心2021.6-2022.5

谢谢大家!

