



清华大学

基于CERNET2的互联网超算：构建与应用

清华大学/国家超算无锡中心 黄震春

2025/12/18

超算互联网

超算互联网原型构建

应用情况

挑战与展望

超算与目前存在的问题

超级计算机

(超算)

- 具备巨大算力的计算系统
- 通常由大量计算单元及相关资源组成
- 用于解决具有挑战性的科学与工程问题
- 智算也是一种定制程度比较深的专用超算



算力孤岛

- 资源分散，不同超算中心之间的算力需要跨越地理上的距离
- 归属不同，管理策略多样，算力被管理策略分割在多个孤岛之中
- 难以统一调度，无法协同使用



使用门槛较高

- 命令行+脚本提交，非计算机领域用户上手困难
- 异构带来的代码编写与优化非常困难，难以发挥超算的性能
- 管理手续繁琐，使用受限

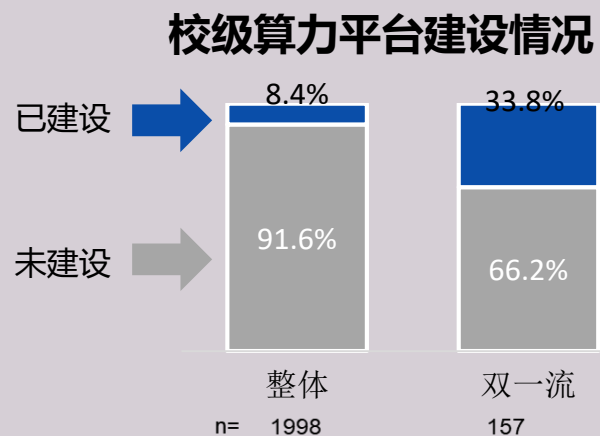


应用服务不足

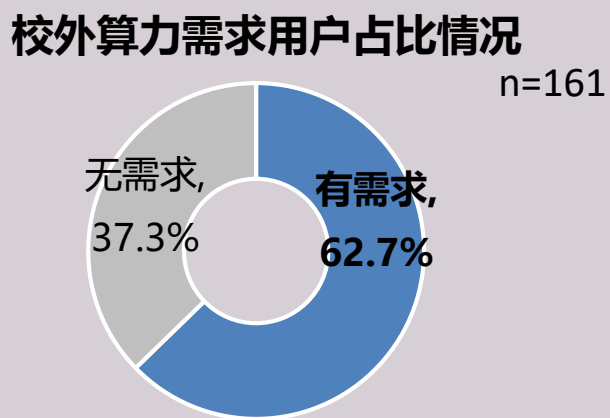
- 行业/领域支持上的不足，算力难以变成解决实际问题的能力
- 可视化、数据治理、合规审计等应用服务的缺失，使得算力难以为业务服务
- 难以实现算力与专业知识的有效融合，影响超算的应用价值

为了更好地掌握高校对算力服务的需求情况，自2022年起联合赛尔网络开展了一系列与算力需求有关的调研，调研内容包括高校对算力的建设、使用及服务需求情况

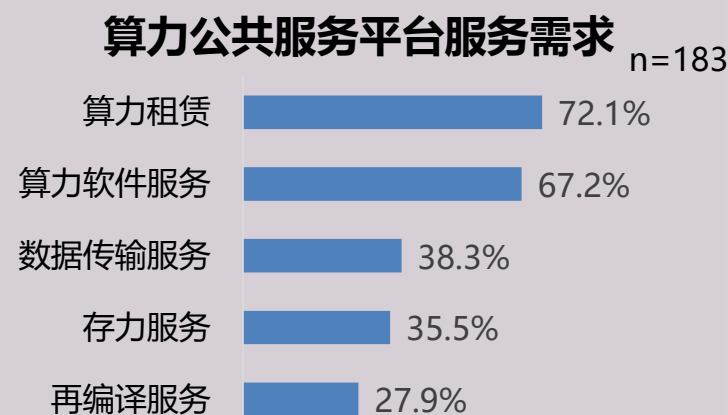
- 高校自建校级算力平台的数量较少，占比为8.4%。双一流高校群体的建设比例相对较高。
(来源：2022年算力需求调研)



- 62.7%的高校对校外算力资源有使用需求。
(来源：2024年算力需求调研)



- 高校对算力公共服务平台提供算力租赁、算力软件服务、数据传输服务等功能需求较高。
(来源：2024年算力需求调研)



超算互联网

- 以超算和高速互联网为基础的先进计算基础设施



以高速网络互联地理上分布的算力资源



构建具备算力资源共享与交易能力的公益性资源平台



支持算力、数据、软件、应用等资源的共享与交易



面向各领域与产业用户构建具有专业服务能力的领域应用平台



打破算力孤岛

- 高速互联网支持，跨越地理局限
- 统一接口与管理策略，进一步打破算力孤岛藩篱
- 构建统一的二级调度体系，支持跨地域跨中心的协同



降低使用门槛

- 标准化的调用接口，提供从API到WebUI的多种使用方式，降低上手难度
- 即时部署或者预部署的并行程序，经过充分优化，更好地发挥硬件的算力
- 统一用户体系、统一资源管理与交易，屏蔽繁琐管理



支撑应用服务

- 面向应用领域的并行软件服务，用算力解决行业和应用领域的实际问题
- 汇聚通算、超算、智算等多种算力资源及相关应用服务，支持应用平台的构建，支撑实际业务
- 聚合应用开发者、领域专家等智慧资源，实现对科学与工程的有力支撑

- 扩大超算的应用群体
- 探索超算中心、超算软件开发者和超算用户三方共赢的新型超算商业模式
- 实现超算中心从提供算力到提供服务的重大转变
- 为应用领域的实际业务提供应用平台支撑

超算互联网

超算互联网原型构建

应用情况

挑战与展望

原型建设

- 基于CERNET2/IPv6的高速互联
- 超算互联平台的研制与算力资源的高效接入
- 资源与服务的发布、交易与评价
- 面向应用领域和行业的超算互联网平台搭建

面向重点领域的
应用平台



教学科研

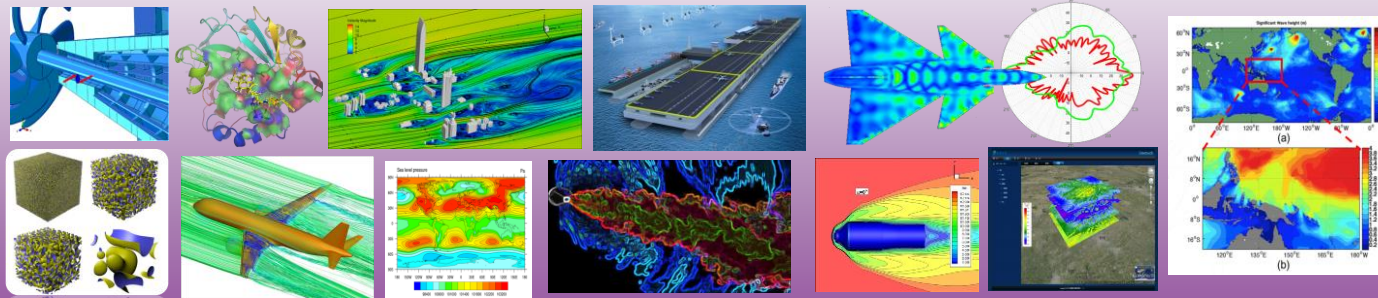


工程计算

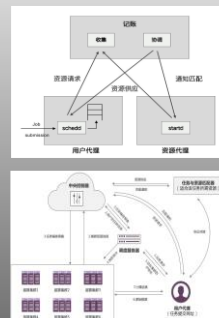


AI/大数据

超算应用
软件开发
与超算资源
市场化



超算资源
共享与算
力集成调
度管理



资源
协同
共享

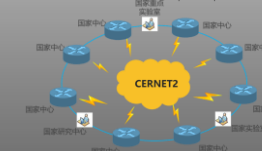
任务
统一
调度



应用
开发
支持



市场
化运
营



高速
数据
传输



算力
池化
整合

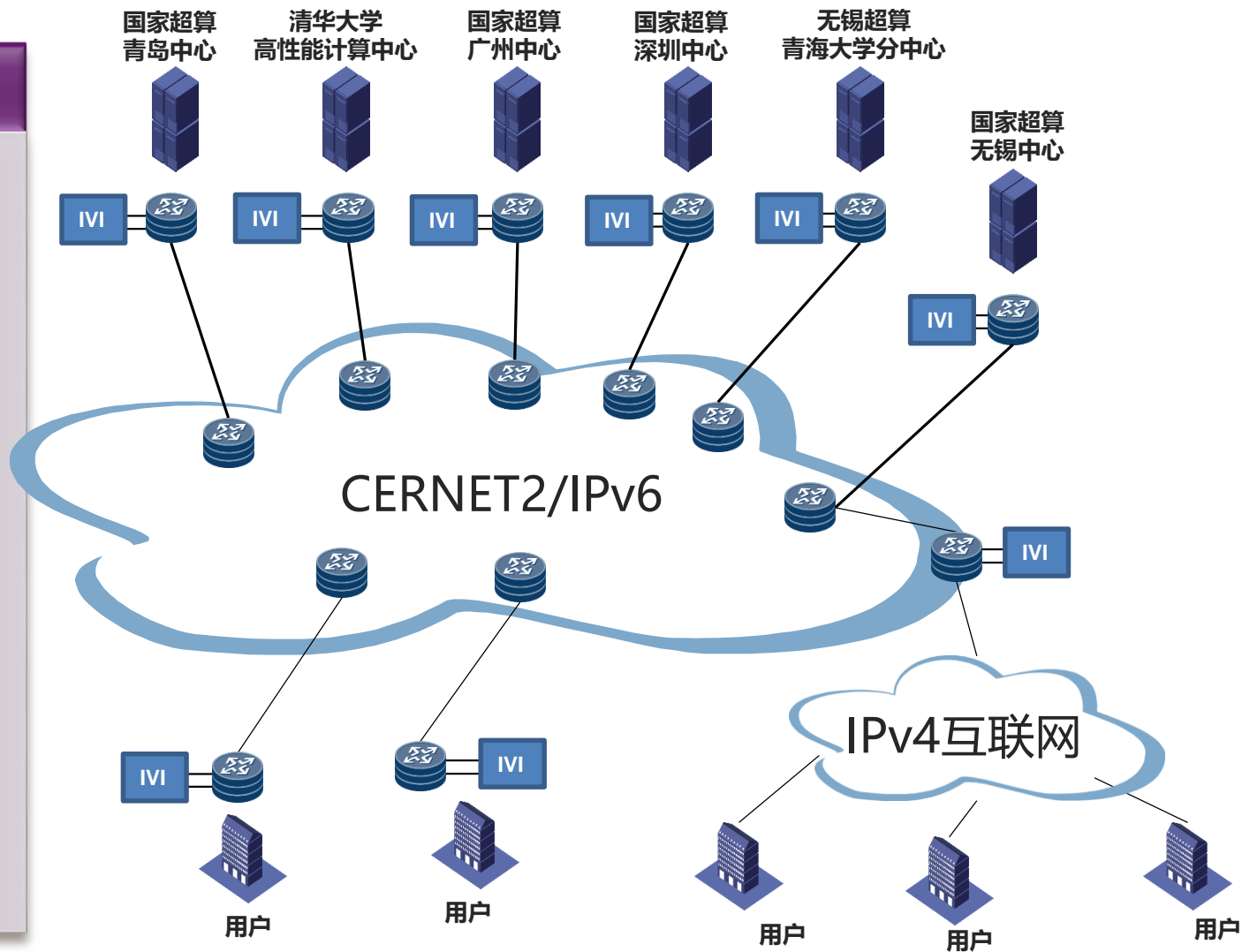
基于高速互联网络
与超级计算资源的
超算互联网硬件环境

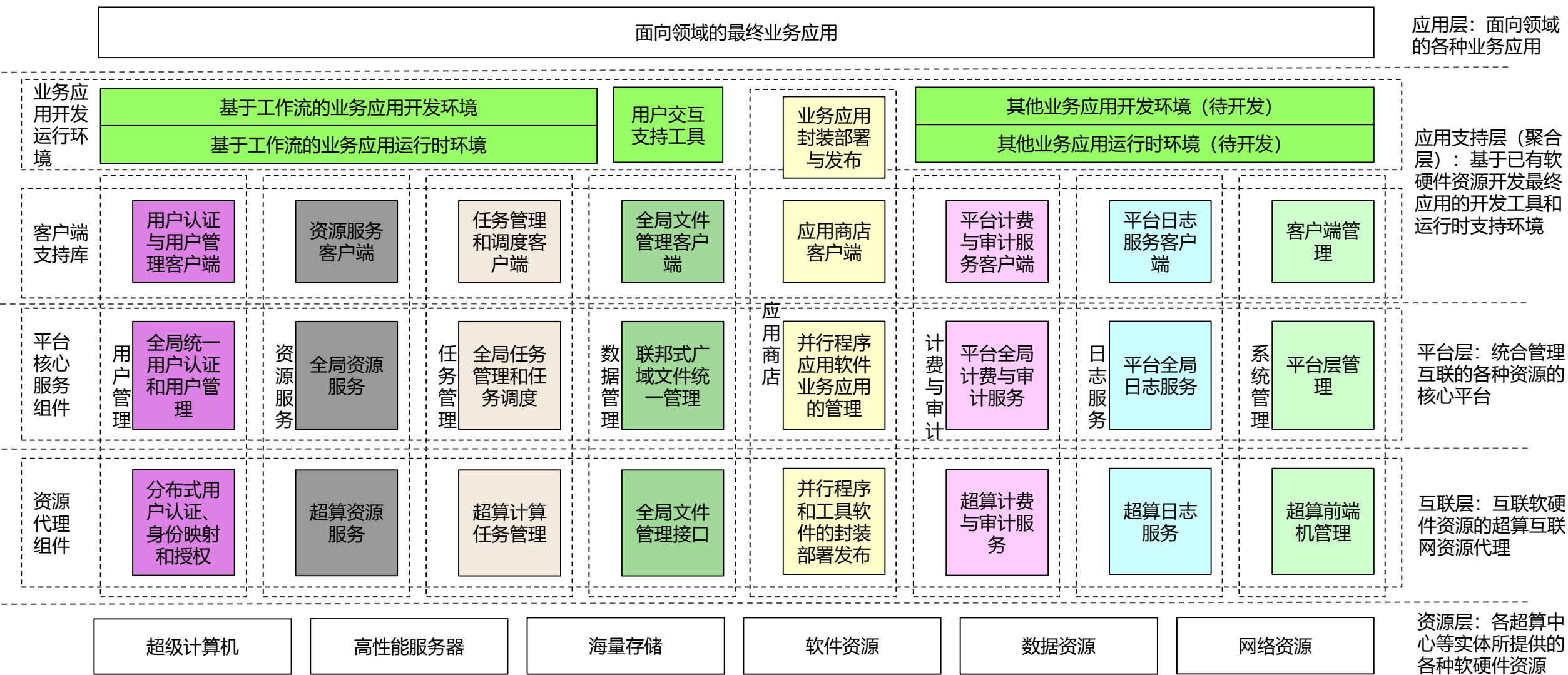


图例	带宽
	100G
	10G
	≤ 10G

基于CERNET2构建核心资源网，跨越地理局限

- 各超算中心使用CERNET2专用（或者虚拟专用）高速网络互联，提供10Gbps-100Gbps的接入带宽
- 专用网络不影响超算中心原有互联网络运行结构
- 使用IVI设备提供IPv6-IPv4翻译，实现跨超算中心的虚拟专用网络并提供超算更习惯使用的IPv4互联
- 目前：
 - 通过CERNET2-IVI设备互联超算中心6家，其他方式互联超算中心3家
 - 互联计算能力（包括即将接入）超过3.5Eflops
 - 累计接入超算网络带宽690Gbps





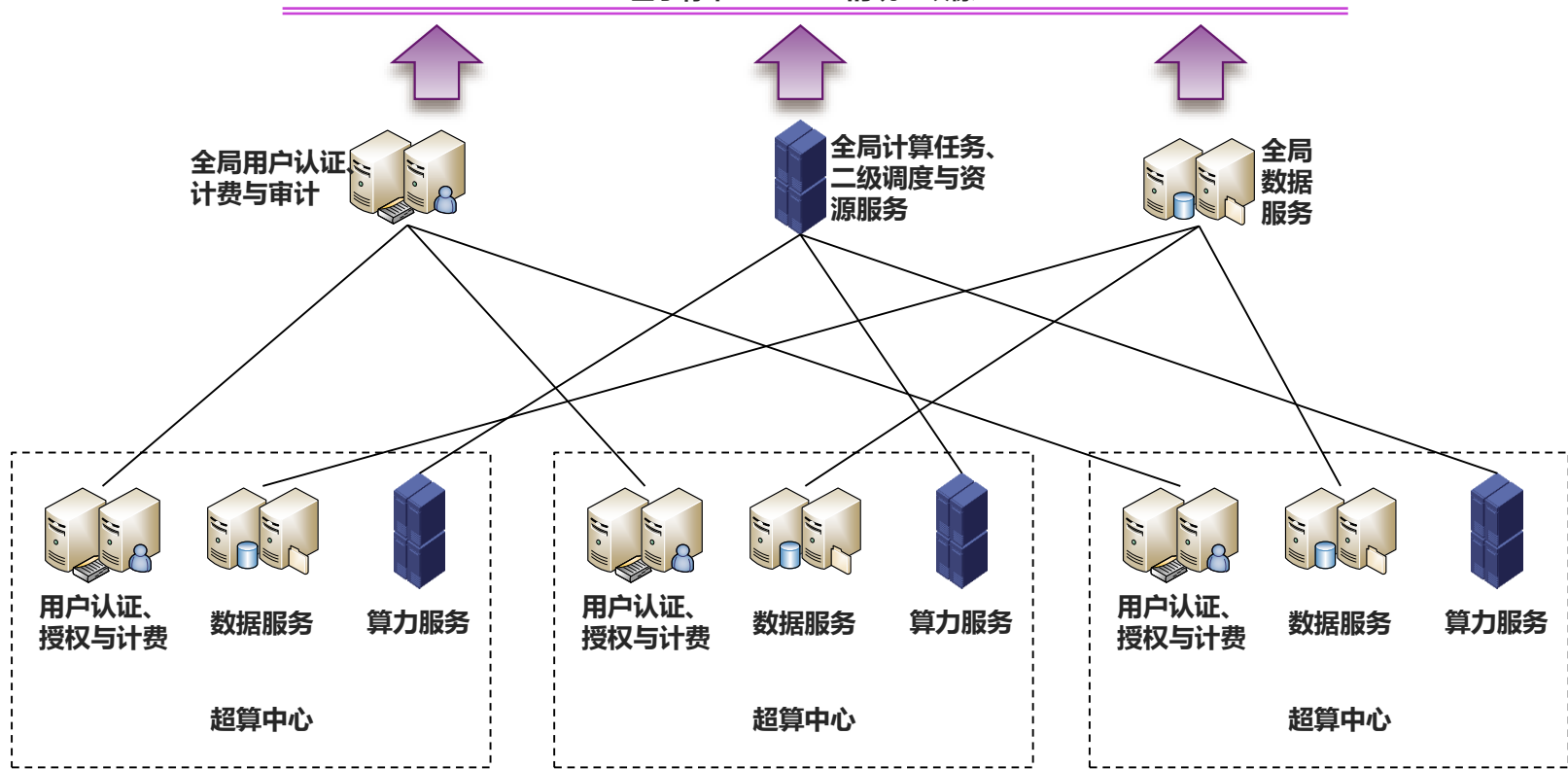
基于标准Web API,
面向用户提供统一映像

- 设计标准化的Web API调用接口
- 适配不同超算中心, 提供统一接口与管理策略, 打破算力孤岛藩篱
- 对接多个超算中心的算力, 提供基于计算任务机制的标准资源服务与算力服务
- 构建统一的二级调度体系, 支持跨地域跨中心的协同

打破算力孤岛, 降低使用门槛



基于标准Web API的统一映像



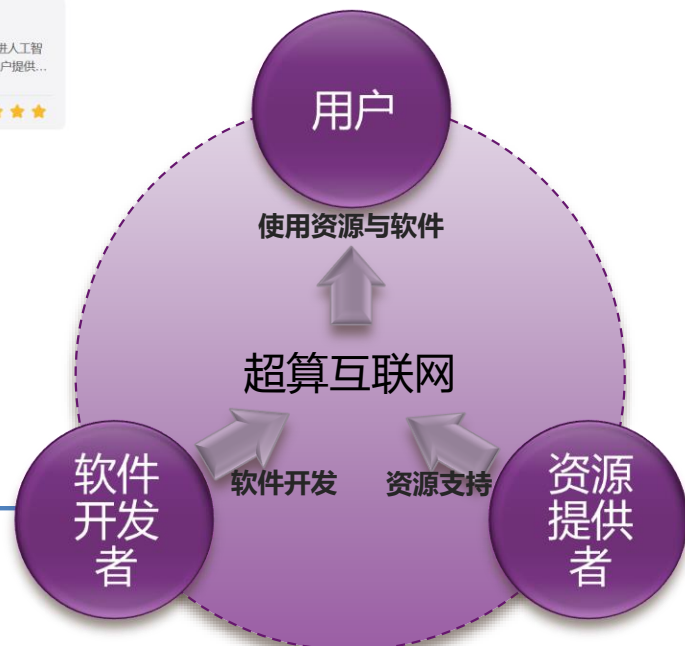
应用商店:

连接超算中心、软件开发者和用户，实现三方共赢

- 超算中心提供算力、存储等资源，支持软件的开发与运行
- 软件开发者使用超算中心的资源，开发应用软件，并在应用商店上架
- 用户从应用商店采购软件，使用超算中心的资源运行软件，支持实际业务

降低应用门槛，支撑应用服务

- 提供面向应用领域的经过充分优化的专业并行软件服务
- 更好地发挥硬件的算力，解决行业和应用领域的实际问题
- 聚合应用开发者的智慧资源，实现对科学与工程的有力支撑
- 支持超算中心从提供算力到提供服务的转变



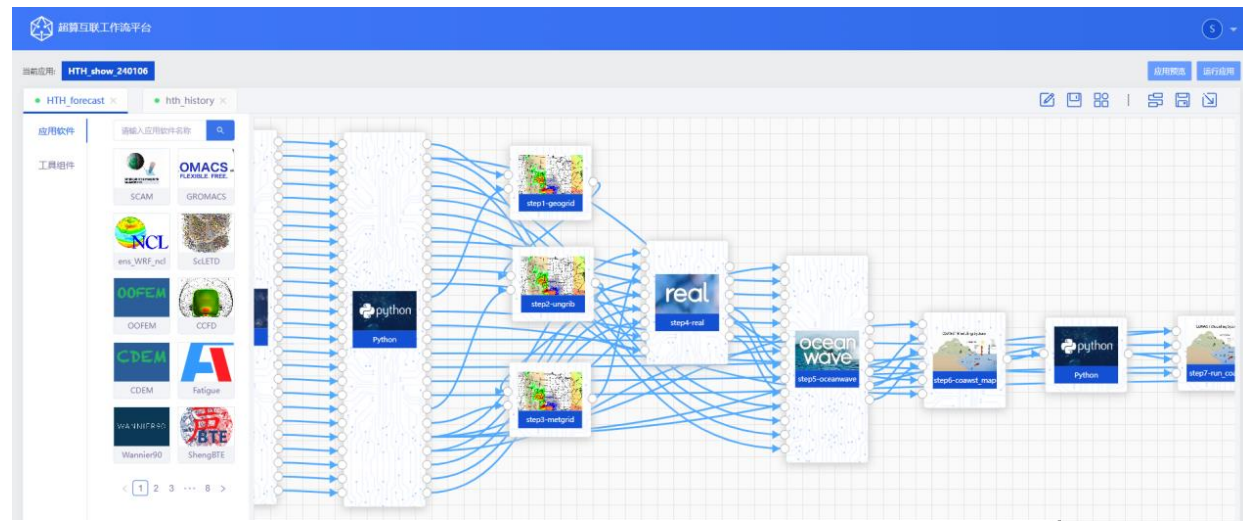
面向领域的应用平台

面向领域的应用平台构建

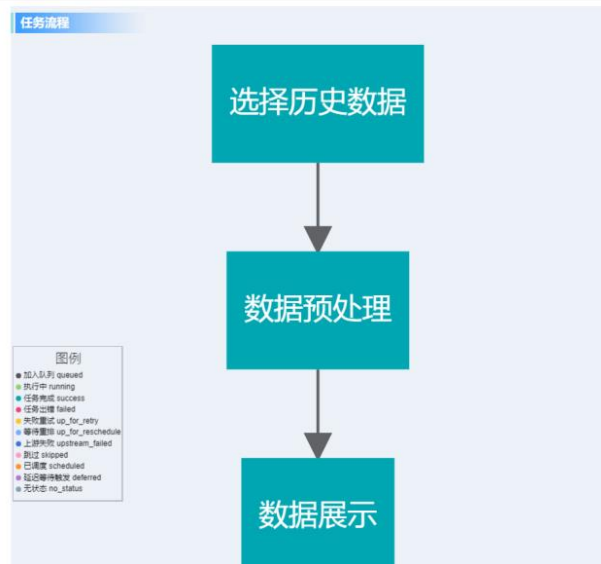
- 面向领域提供一系列并行程序和工作流组件，构建和发布应用系统
- 支持自定义交互界面，实现复杂交互
- 面向领域提供科学数据可视化服务

支撑应用服务，降低使用门槛

- 汇聚通算、超算、智算等多种算力资源及相关应用服务，支持应用平台的构建，支撑实际业务
- 聚合领域专家智慧，支持科学与工程计算需求



任务详情



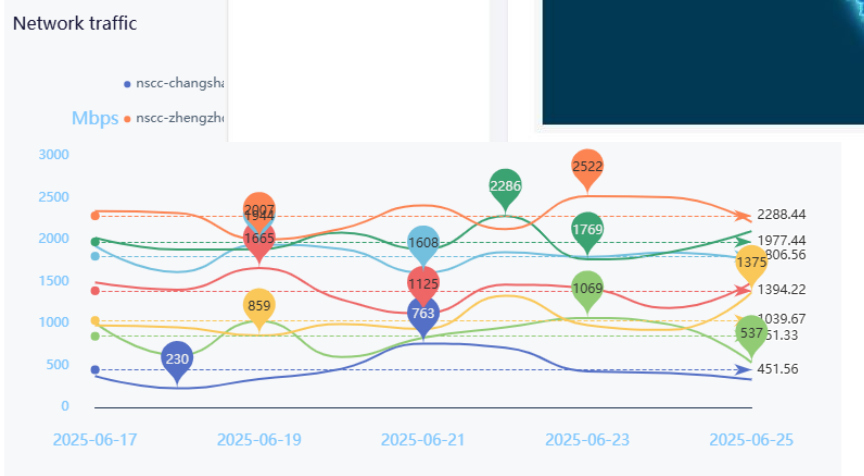
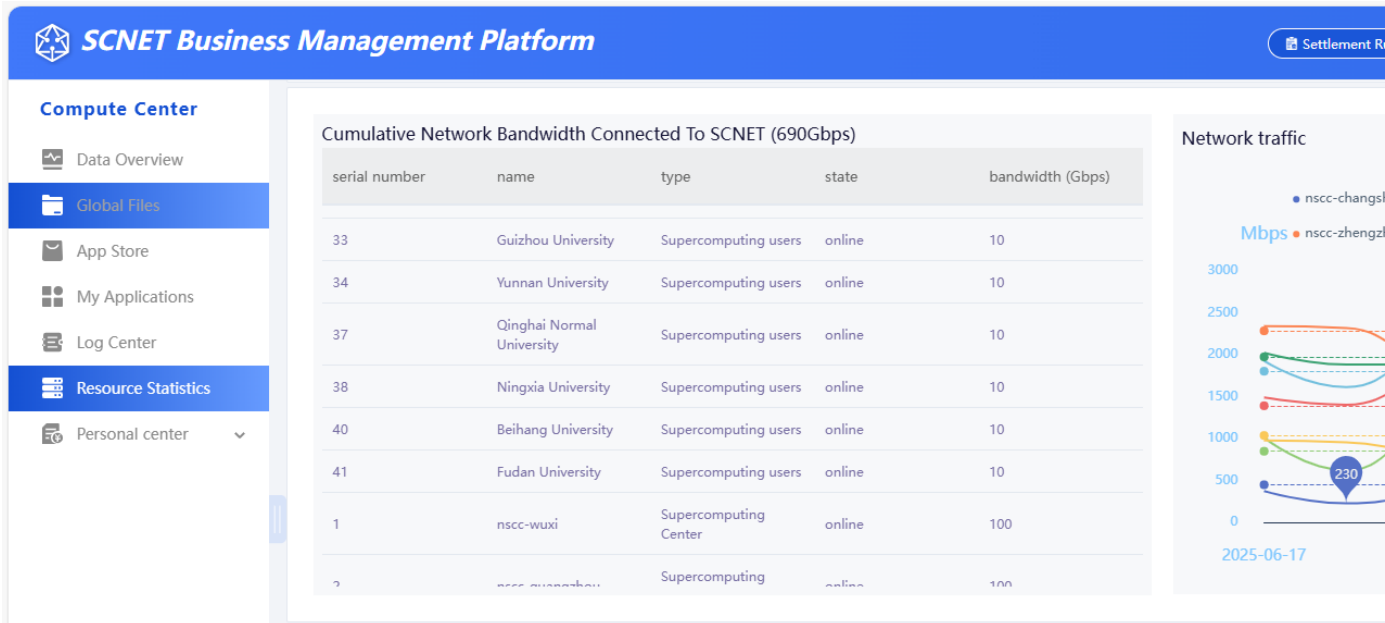
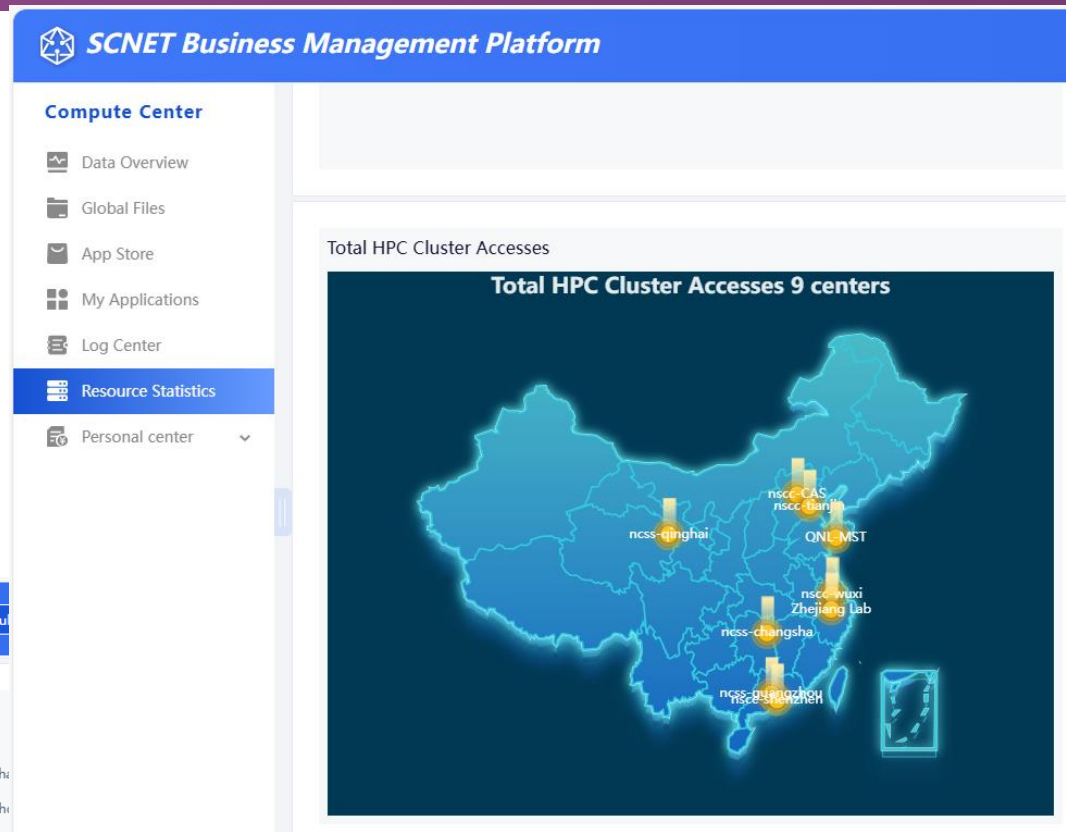
超算互联网

超算互联网原型构建

应用情况

挑战与展望

- 通过CERNET2-IVI设备互联超算中心6家，其他方式互联超算中心3家，互联带宽最高达100Gbps
- 累计接入超算网络带宽690Gbps



- 部署软件超过120款
- 面向天气气候、生物医药等应用领域
- 多种超算并行软件和智算大模型

SCNET Business Management Platform

Compute Center

- Data Overview
- Global Files
- App Store

Software Statistics

Multiscale HPC for Biology & Materials: 17

Biomedical/Materials Hybrid Compu...

Multiphysics Systems Simulation...

Computational Biomedicine & Materials: 14

Environmental Governance & Disaster ...

应用商店

应用名称: 请输入名称 | 类型: 请选择类型 | 分类: 请选择分类 | 查询

应用系统

- gromacs_test_lcy1202** (99人用过)
- gromacs_lcy_1119test** (99人用过)
- BWA** (99人用过)
- htseq** (99人用过)
- TrimGalore** (99人用过)
- FastQC** (99人用过)
- GCE** (99人用过)
- Minimap2** (99人用过)
- inference** (99人用过)
- visualization** (99人用过)

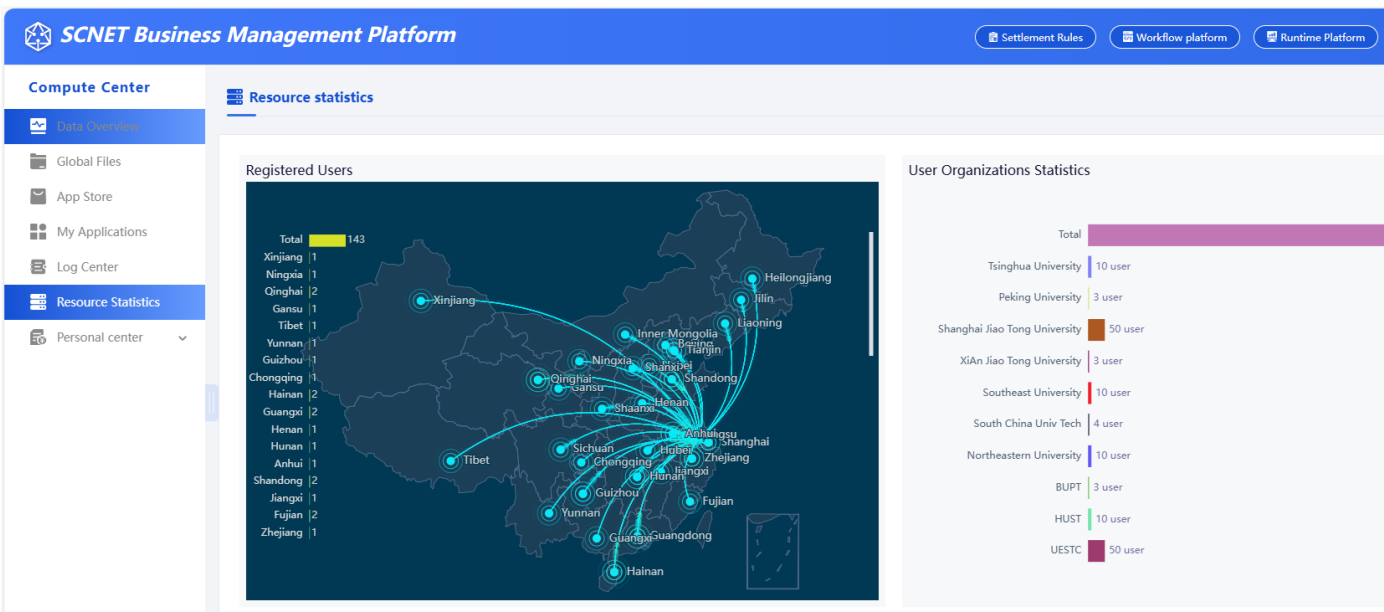
应用系统

- helili_ceshi** (99人用过)
- SWBind** (99人用过)
- 面向多尺度生物和材料的...** (99人用过)
- 人工智能算力和数据服务...** (99人用过)
- Deepseek** (99人用过)

1 2 3 4 5 6 7 >

用户应用情况

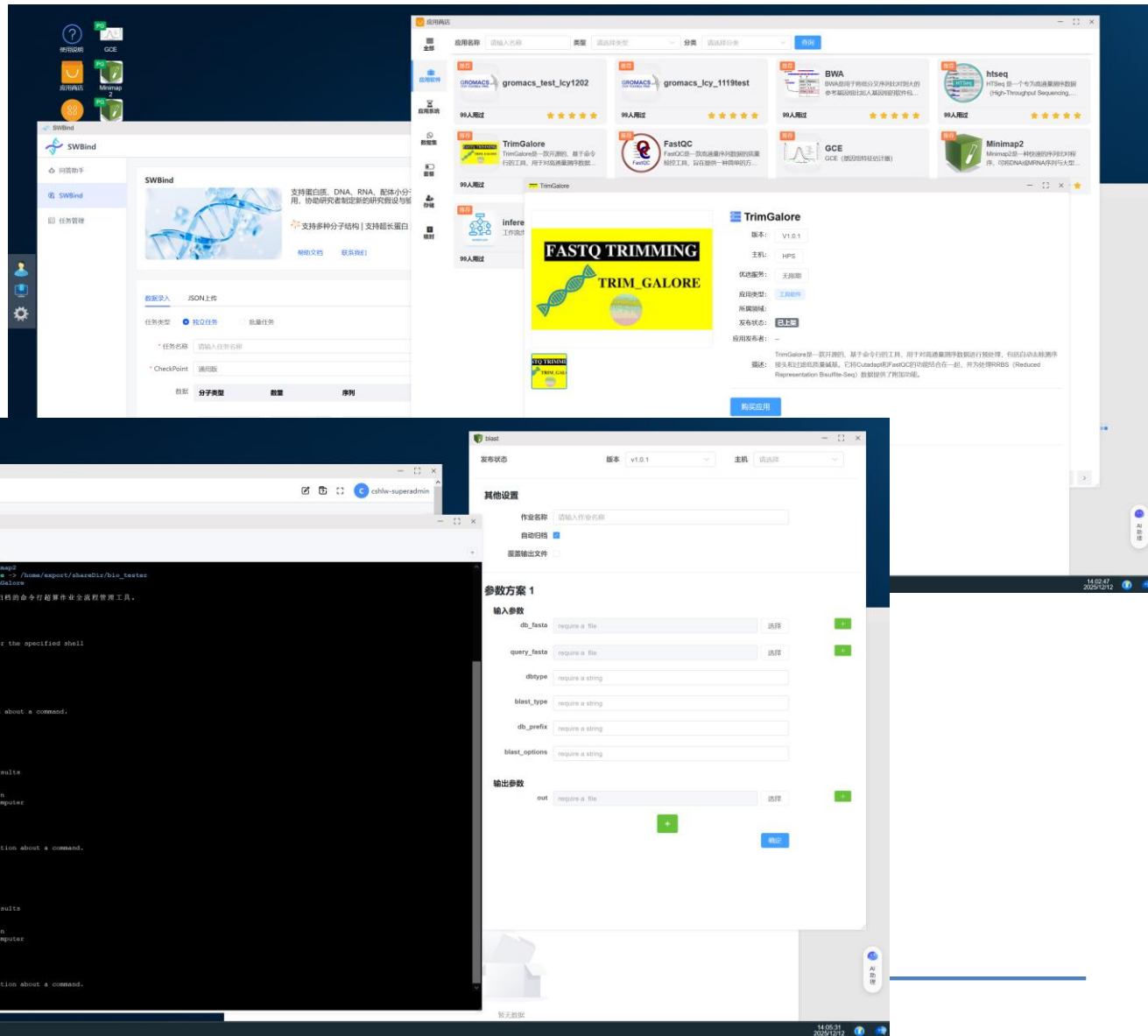
- 支撑 143 所高校完成大量计算任务
- 平台累计注册用户数1161
- 平均日用户访问量 ≥ 10000



- 与相关领域专家合作，面向天气气候、生物医药等领域，构建多个典型应用与应用平台
- 欢迎试用

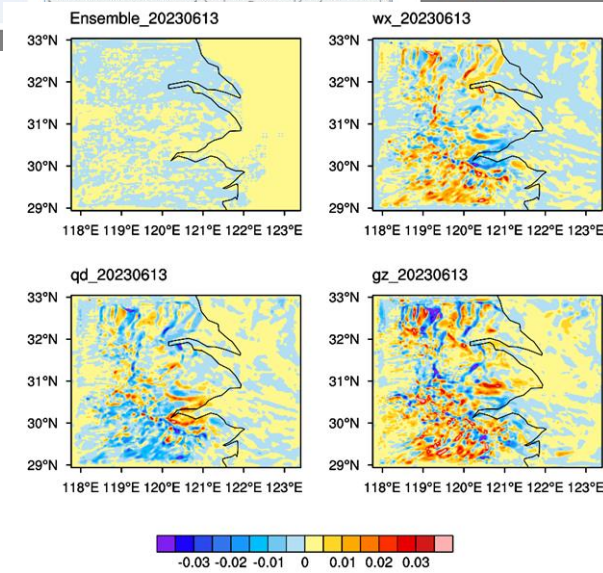
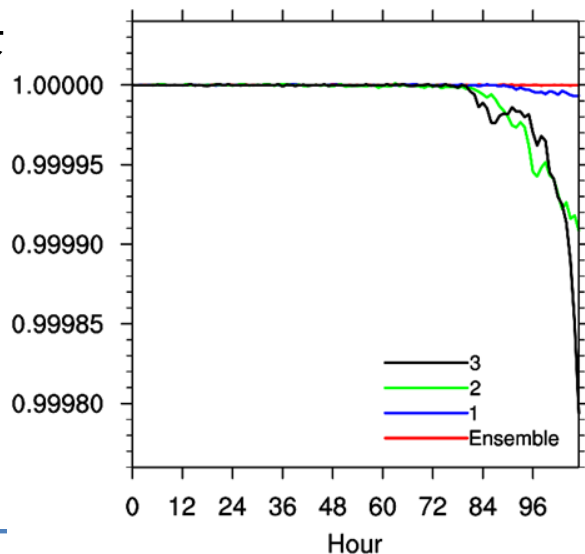
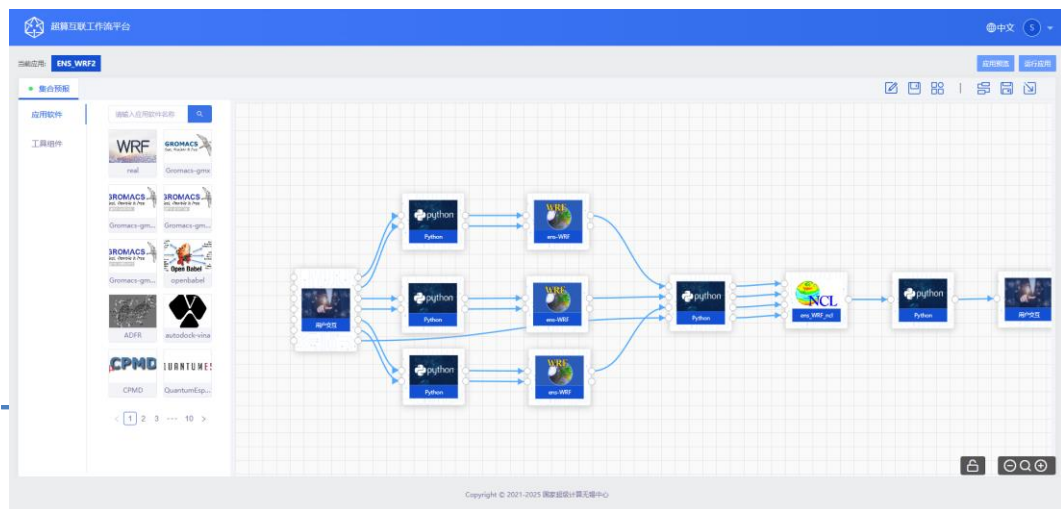
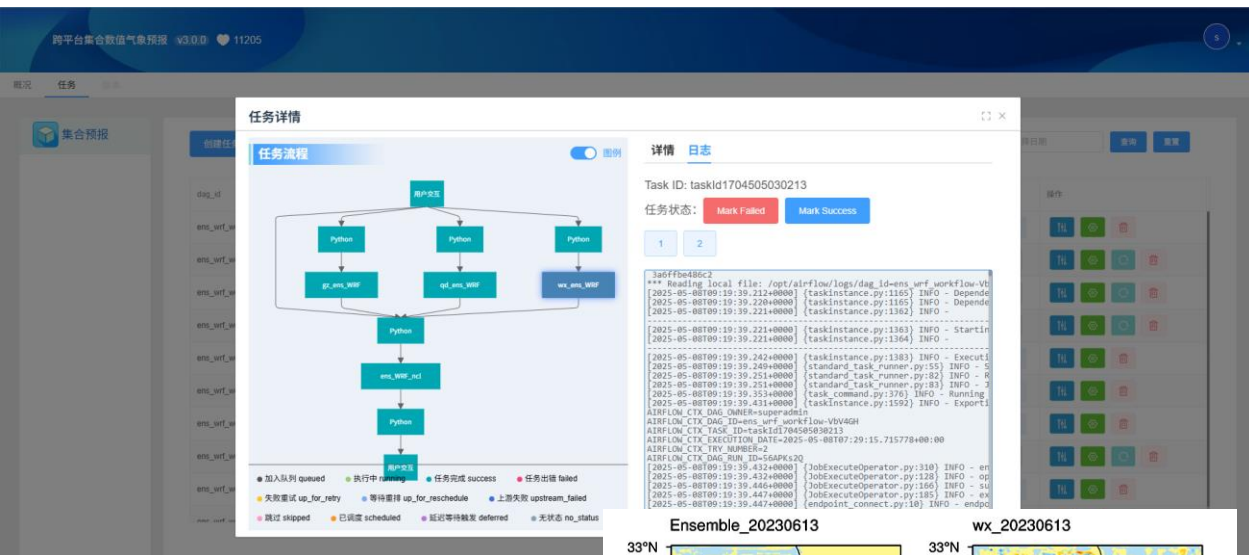
■ 面向生命科学的科研支持平台

- 集成式云桌面支持，方便使用
- 集成GROMACS、AutoDock、BWA、htseq、FastQC、GCE、Minimap2、blast等生命科学科研常用工具
- 集成对标AlphaFold的自主知识产权结构预测工具SWBind
- 支持命令行、GUI、API等多种软件调用方式



■ 集合天气预报实验平台

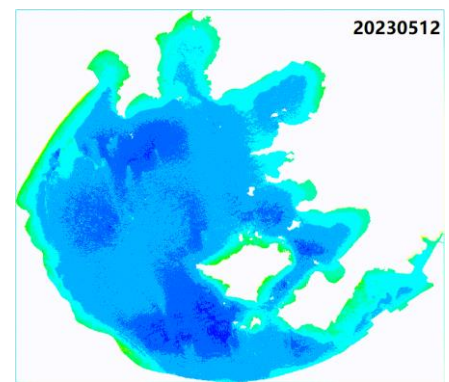
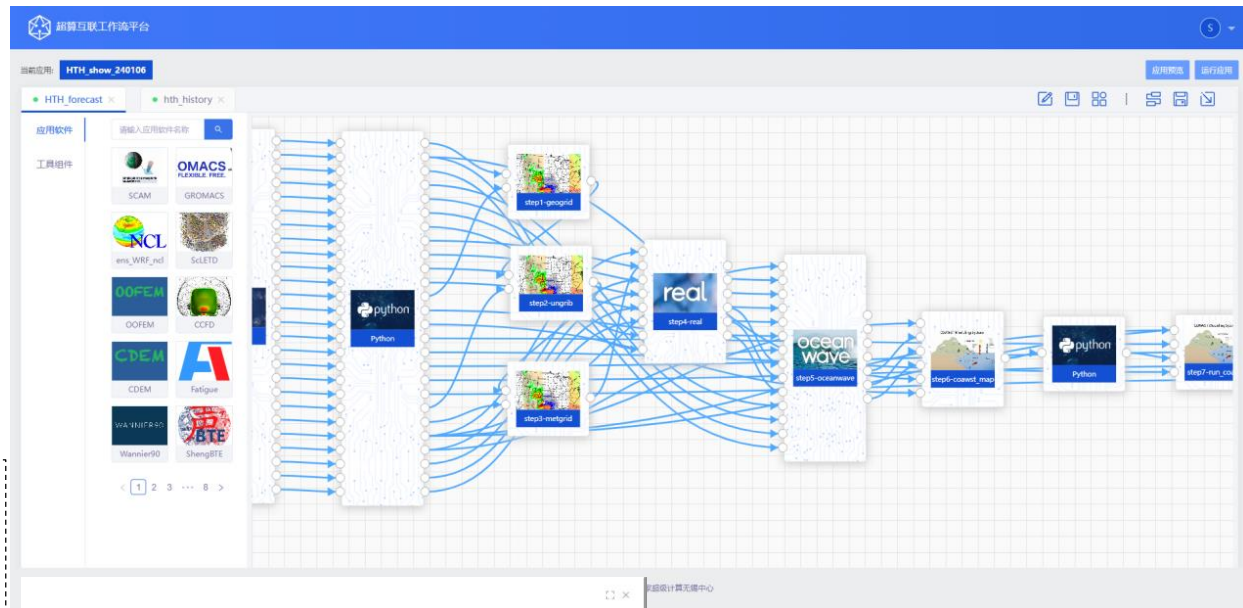
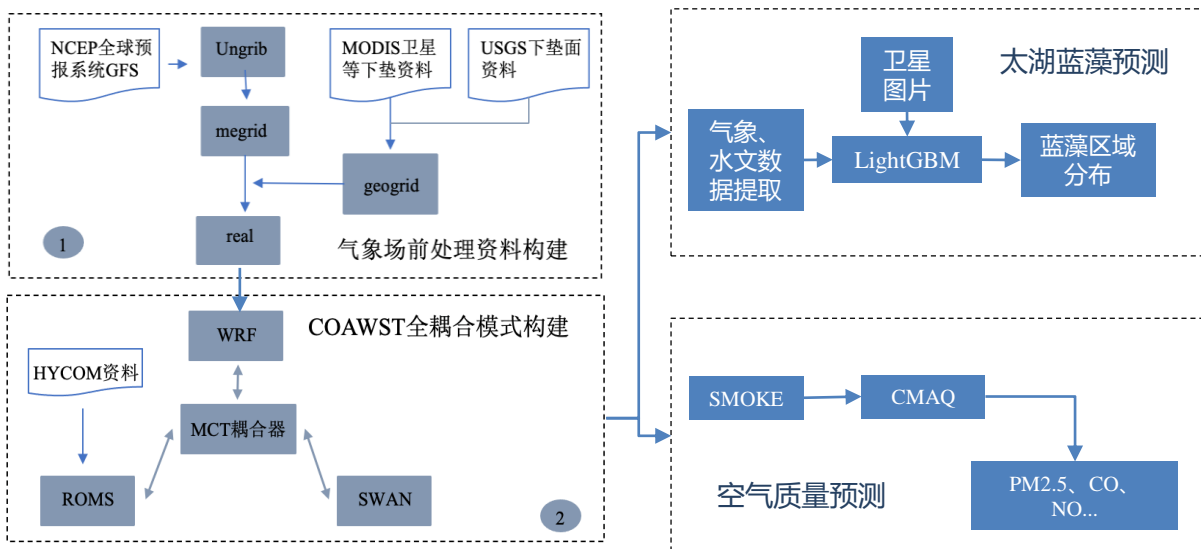
- 发展一种多架构、跨平台集合数值天气预报模拟方法，采用多参数化方案，充分提高模式物理过程的概率分布
- 采用集约化处理算法降低跨平台舍入误差所引起的结果偏差，从而提高预报准确性和稳定性
- 融合多平台的算力资源，降低单个超算平台上进行多初值集合预报的工程量，解决单个超算平台算力不足的问题



三种参数化方案集合预报和单平台预报结果与ERA5再分析资料的相关性对比图 (红线表示集合预报, 其余为单个实例预报)

■ 环太湖环境预测平台

- 天气模式+空气质量+蓝藻
- 国内首个蓝藻预测系统
- 基于WRF 流程化定制系统构建
wrf+roms+swan 全耦合模式



蓝藻预测准确率与当地卫星影像及官方数据对比, 准确率接近95%

■ 三江源地区流域多变量预测平台

- 研制面向青藏高原水圈系统的深度学习框架
- 引入河网拓扑、上下游关系和汇流特性等水文先验，将 Granger 因果构建的水文-气象因果链映射到图神经网络结构中，实现物理过程与数据驱动的协同建模
- 在拓扑显式建模、物理可解释性、非马尔可夫/非高斯/非稳态过程表征与变化环境适应性上具备显著优势
- 为青藏高原区域水圈变化趋势分析提供科学支撑

任务详情

任务流程

用户交互

Python

inference

visualization

Python

用户交互

流程信息

基本信息

名称: 三江源大模型

版本: v0.0.1

描述: 通过空-天-地多模态数据融合技术, 对来自卫星、地面观测站和气象雷达的异构数据进行有效融合, 预测模型精度和实时性。接着, 为了处理多维度气候与生态动态系统, 该研...

应用图片:

加入队列 queued

执行中 running

任务完成 success

任务出错 failed

失败重试 up_for_retry

等待重排 up_for_reschedule

上游失败 upstream_failed

跳过 skipped

已调度 scheduled

延迟等待触发 deferred

无状态 no_status

Copyright © 2021-2025

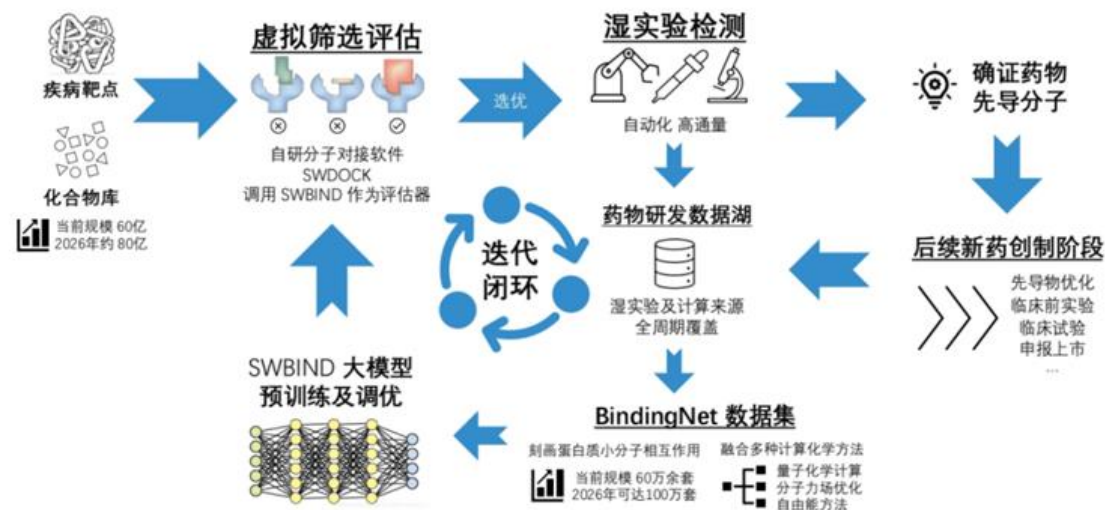
9:58:31 2025/11/21

■ 新型新药研发平台

- 整合高性能药物设计软件（自主开源可商用）、全面的化学小分子数据库（数十亿量级）与用户友好可靠的研发流程
- 部署数量规模国内领先的多种小分子化合物数据库（60亿余种化合物，存储量约1.5 PB）和具备自主知识产权的蛋白-小分子结构活性数据集BindingNet 3.0
- 部署SWDock 2.0，对接后重打分软件、不断优化的SWBind大模型和用于构建虚拟化合物库数据的完整工具链
- 支持“蛋白质组对接”研究课题，已完成人类蛋白质组（约2万个蛋白质靶点）与上市药物及自建30万虚拟化合物库的大规模分子对接任务，已在十余个蛋白靶点上完成亿级虚拟化合物库对接任务



■ 量子化学与实验来源数据集驱动，构建面向新药创制大模型基础设施



超算互联网

超算互联网原型构建

应用情况

挑战与展望

■ 技术

- 跨越地理和管理上的距离，提供标准的服务
- 不同架构的算力资源如何保持效率并统一访问
- 实现算力、数据、网络甚至电力的联合调度，提高效率降低成本
- 在已有框架下支持超算软件的开发与优化，支持更加灵活的资源使用形式
- 提供专业领域应用平台解决方案，拓展应用领域

■ 商务

- 资源提供者、软件开发者和用户三方都满意并有所收获的商业模式，吸引更多力量的加入
- 引入应用领域的专家和智慧，构建专业化的领域应用平台并推广应用

■ 基础设施

- 通过高速网络互联各种算力，打破算力藩篱，汇聚资源与智慧，构建专业化的算力应用平台，为应用领域提供专业的算力服务



清华大学

谢谢!