



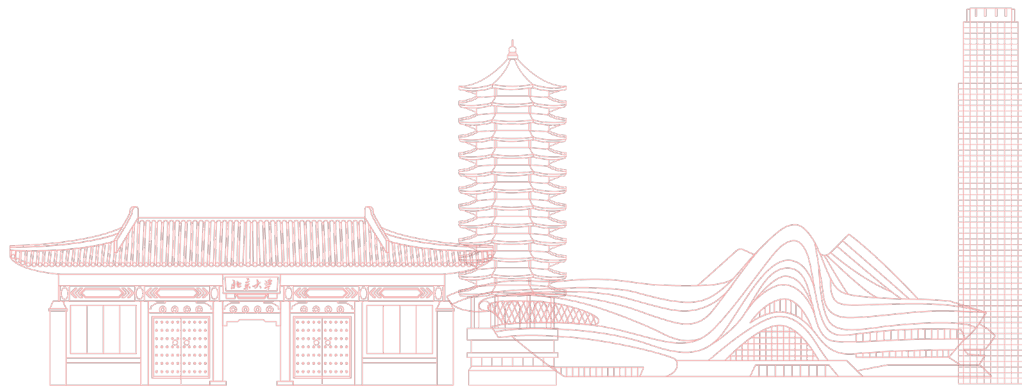
北京大学
PEKING UNIVERSITY

算力网络的探索与实践

李若淼

北京大学计算中心

北京大学长沙计算与数字经济研究院



content

目录

01 | 北京大学高性能计算平台

02 | 算力网络时代的算力平台

03 | 探索与实践

04 | 愿景

北京大学高性能计算校级公共平台

2018年1月3日北京大学高性能计算校级公共平台正式揭牌启用。

“未名一号”、“未名教学一号”和“未名生科一号”等多套集群陆续投入运行，在数学、深度学习、大气海洋环境、新能源新材料、天文地球物理、生物医药健康等领域提供高性能科学与工程计算服务。

集群	节点	核心
5套	755	33584
计算峰值	存储容量	PUE值
4.39PFLOPS	15PB	1.1

高性能计算机系统能效
等级评价1级



国内首个温水水冷集群

超融合以太网解决方案
获解决方案创新奖



首个国产100G高速网

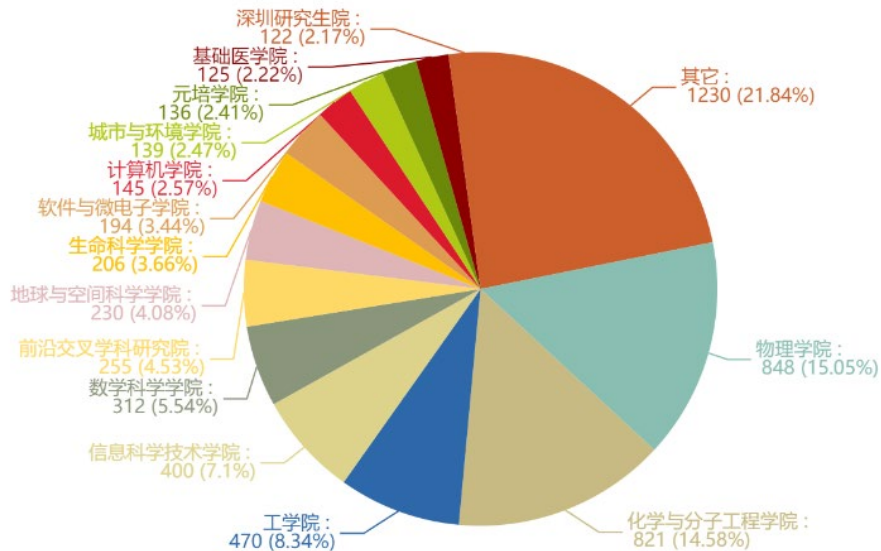


5633人
师生用户

99个
院系单位

19位
两院院士

6位
高校校长



用户来源分布图



平台资源使用状况

支撑高水平论文

1600+ 篇获得用户致谢的论文



17篇



3篇



3篇



35篇



18篇



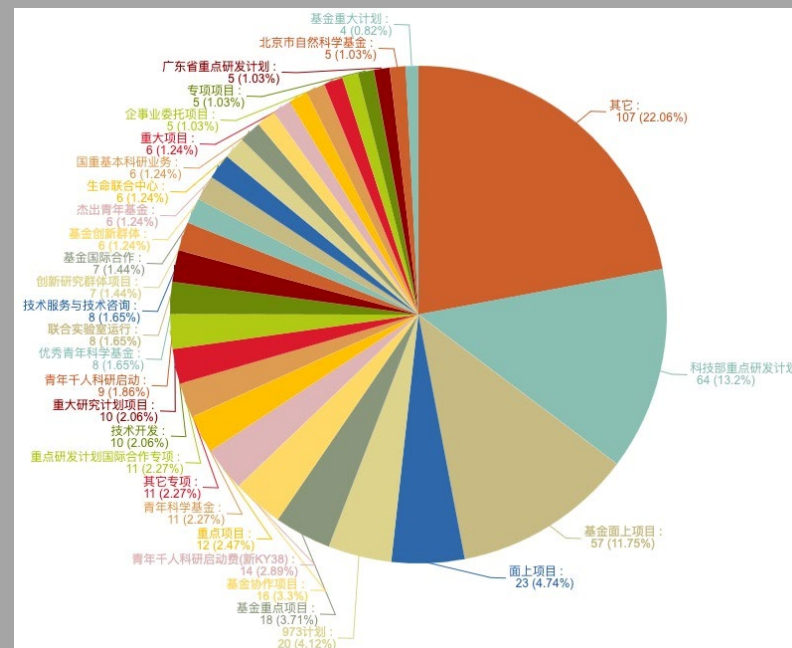
7篇



Nature系列文章100+ 篇

支持科研项目

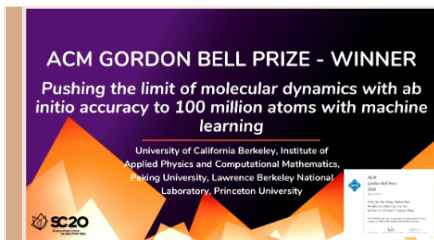
支持项目超545个，总金额达31.36亿元



支持科研项目类型分布



2020年度国家技术发明**一等奖**
超高清视频多态基元编解码关键技术

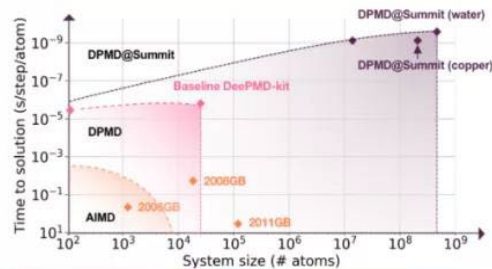
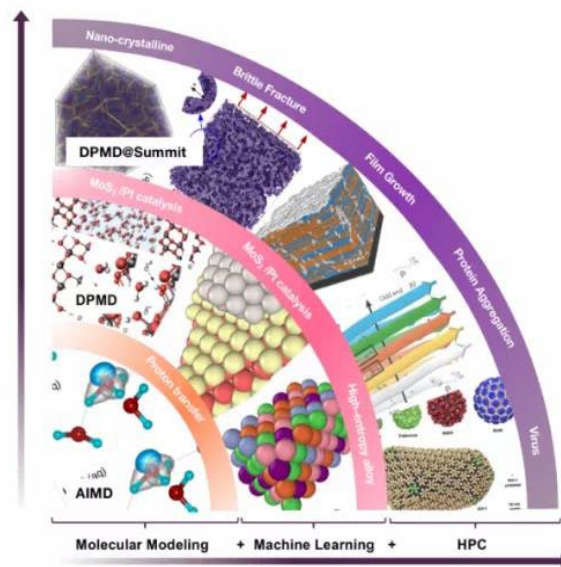


2020年**戈登贝尔奖**
基于深度学习的分子动力学模拟



2020年度国家技术发明**一等奖**
超高清视频编解码技术

分子建模+人工智能+高性能计算



Time and size scales required by important Problems

Problem	Time span [ns]	System size [#atom]
Droplet coalescence	~10	~1e+8
Dynamic fracture	~0.1	~1e+8
Strength of nanocrystalline metal	~0.01	~1e+6
Heterogeneous aqueous interfaces	~100	~1e+6

content

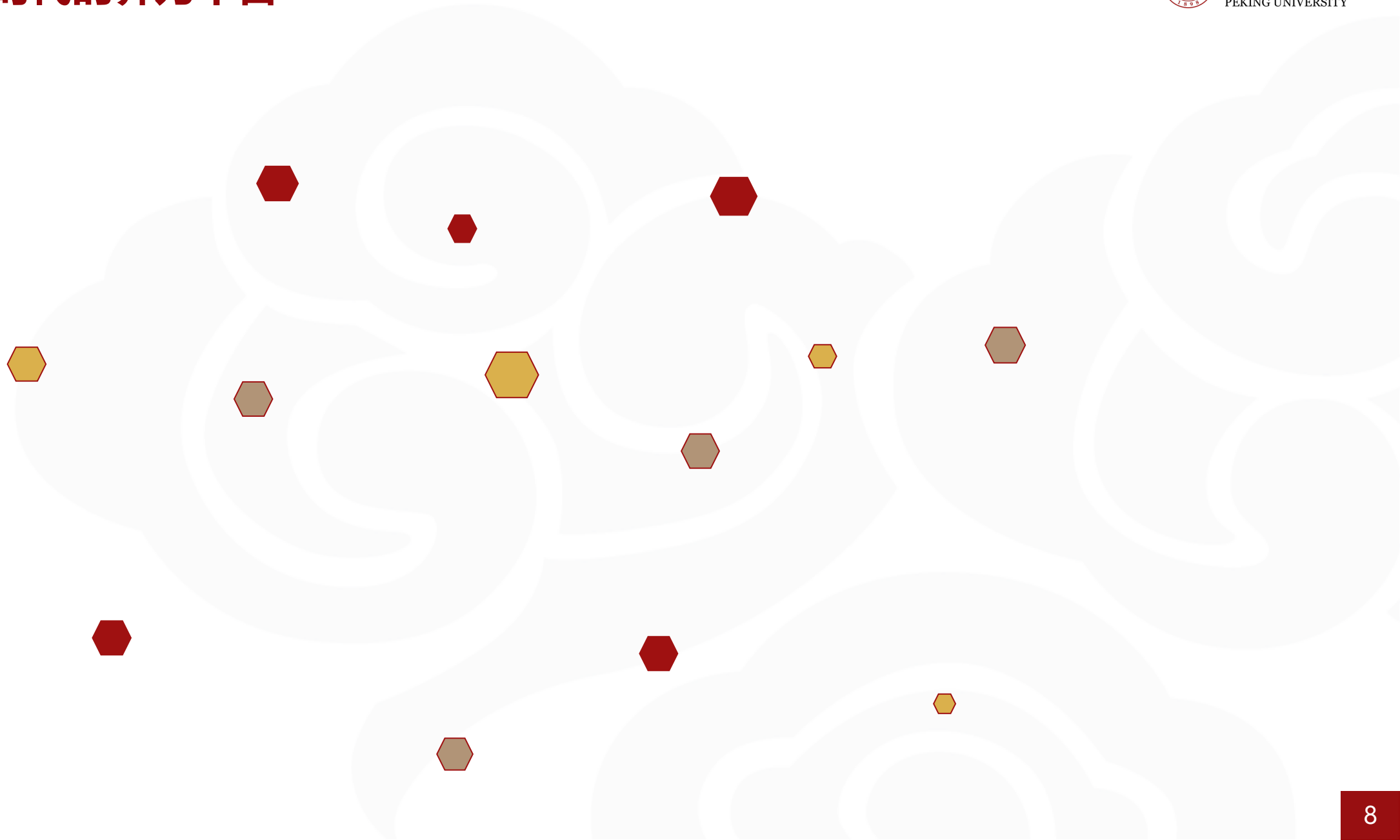
目录

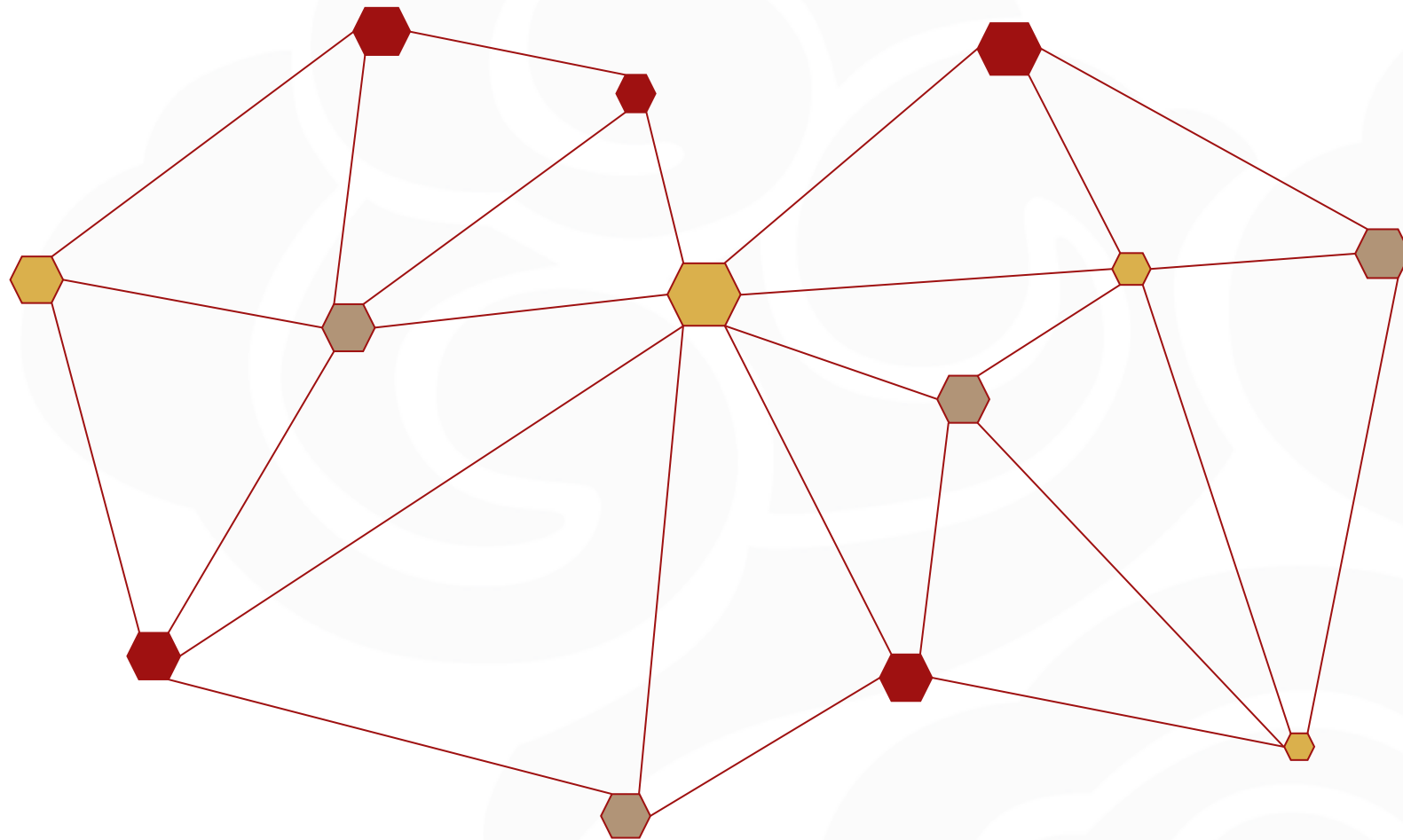
01 | **北京大学高性能计算平台**

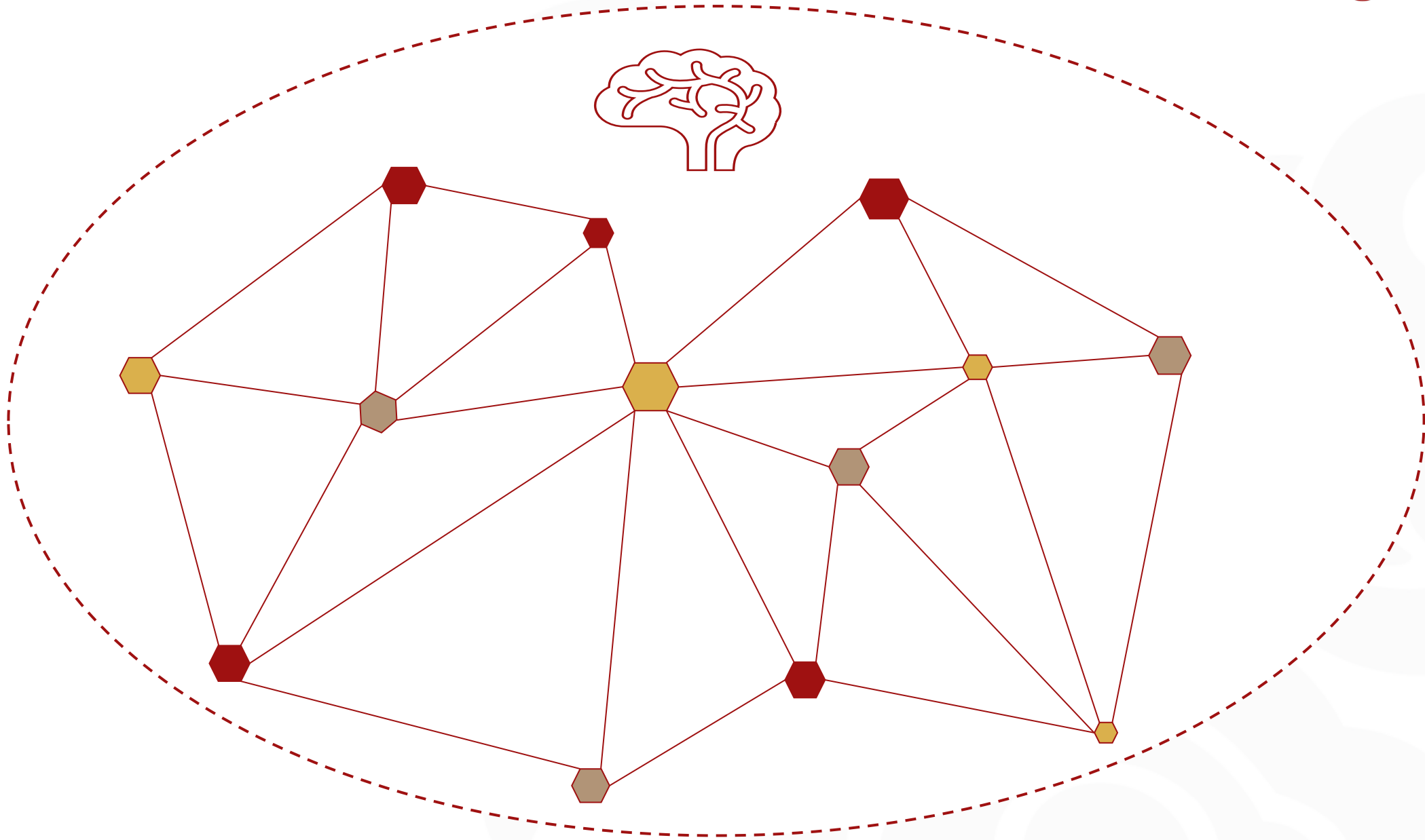
02 | **算力网络时代的算力平台**

03 | **探索与实践**

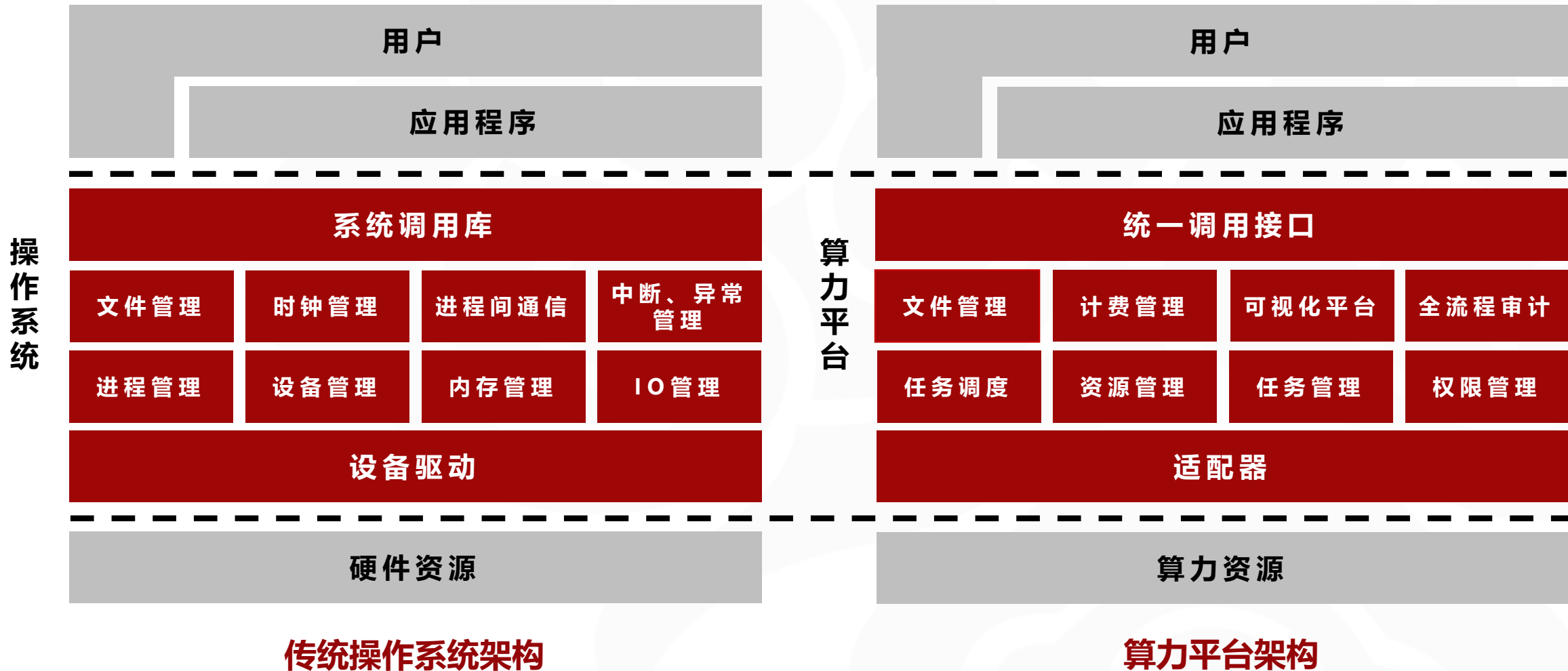
04 | **愿景**







算力平台是帮助用户使用算力资源的**操作系统**，是算力经济时代的**关键基础软件**。



content

目录

01 | **北京大学高性能计算平台**

02 | **算力网络时代的算力平台**

03 | **算力网络的探索与实践**

04 | **愿景**

2022年，我国在用数据中心机架总规模超过590万标准机架，服务器规模约2000万台，算力总规模超过150EFlops，排名全球第二。就算力发展涨幅来说，我国排在世界第一，达到13.5%，美国是5.0%。



运营管理难

- **管理困难**，缺乏统一标准的管理模式，各算力中心需要各自制定管理政策
- **部署困难**，超算集群部署需要各类软硬件配置，缺乏开箱即用的管理平台
- **运营困难**，运维服务人员短缺，需要自动化工具提高运营效率



用户使用难

- 算力终端用户多元化，无法适应基于命令行的传统超算集群使用模式，更习惯图形化、鼠标化的操作
- 可视化交互式应用配置复杂，用户使用门槛高



算力融合难

- 各个算力中心计算**资源独立**，无法互联
- 各算力中心管理模式不同，使用模式不同，**难以融合**
- 各算力中心使用情况不同，算力资源**闲忙有别**，造成算力浪费

SCOW——面向算力网络的算力中心门户和管理平台

建立**面向算力网络的算力中心门户和管理平台SCOW**，通过简化集群软件部署流程、统一平台管理模式、降低用户使用门槛，实现算力中心资源易管理、易使用的目标，提高算力资源使用效率，满足算力中心的管理和维护需求。面向算力网络，构建标准化的平台接口，支撑算力网络平台建设。

SCOW = 平底船 -> 未名湖上的石舫



SCOW = Super COW



SCOW = Super Computing On Web



降低使用门槛

在SCOW门户平台，超算用户无需了解和配置SSH、VNC、命令行等技术，直接在浏览器上就可以使用超算集群。不同于传统超算基于命令行的使用模式，平台门户系统提供基于web页面的各项功能，极大降低了用户使用门槛，让Linux小白用户也能顺利提交作业。

基于Web的多项功能

SSH

Web SSH连接

VNC

Web 远程桌面



Web 文件管理



Web 跨集群数据传输



Web 交互式应用



Web 作业管理



标准化管理模式

SCOW管理平台提供了标准化的模式、模型，能够帮助新建的算力中心快速建立管理和运营制度。管理系统提供了各项管理功能，管理员和运营人员可以很容易在浏览器上实现管理团队的人员和资源，灵活分配人员权限和机时份额，封锁解封团队用户等精细化操作。

通用管理功能

标准化模式、模型



多集群统一管理



用户账户模型
(租户-账户-用户三级模型)



计费收费模型

管理系统多项功能



账户管理



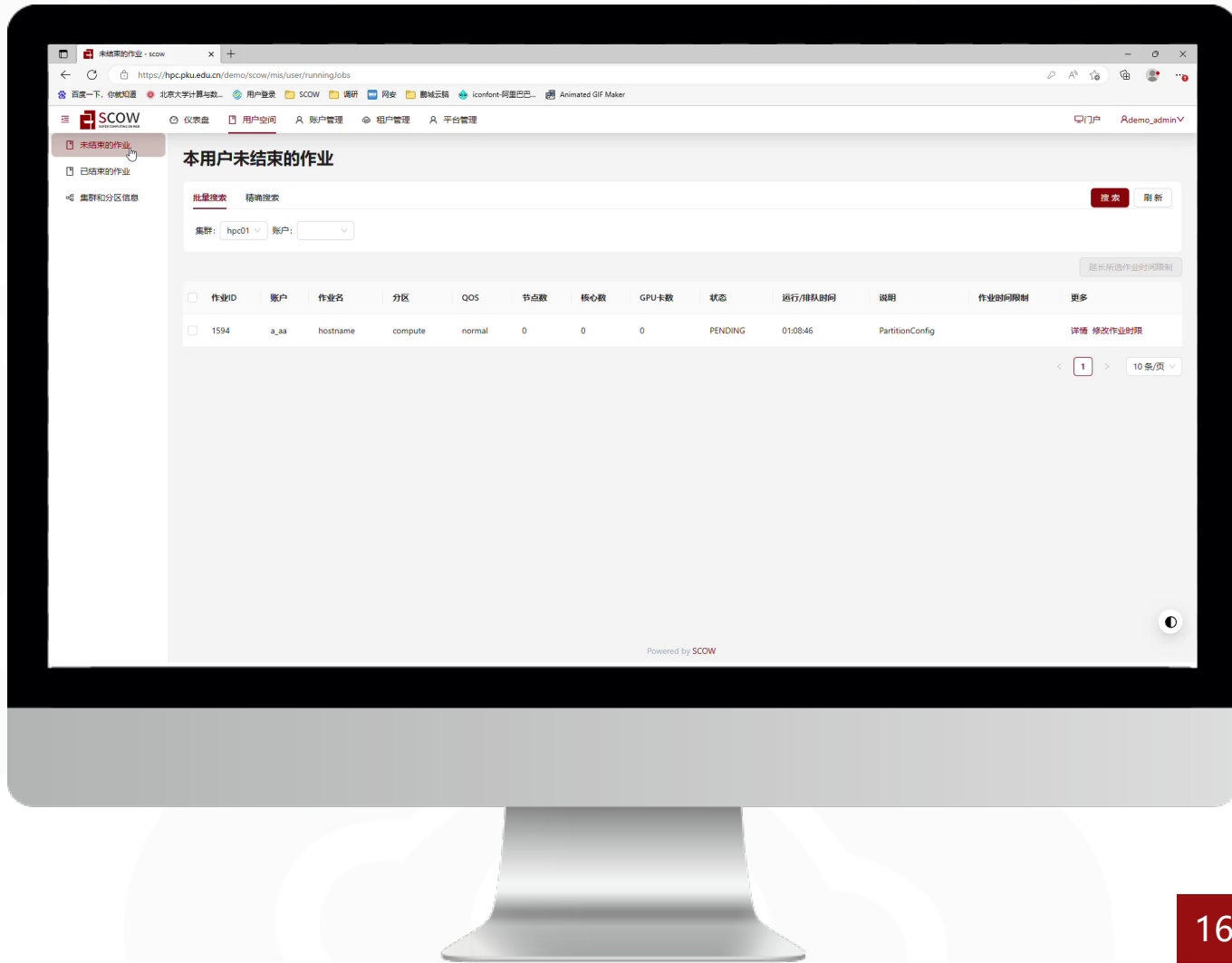
租户管理



平台管理



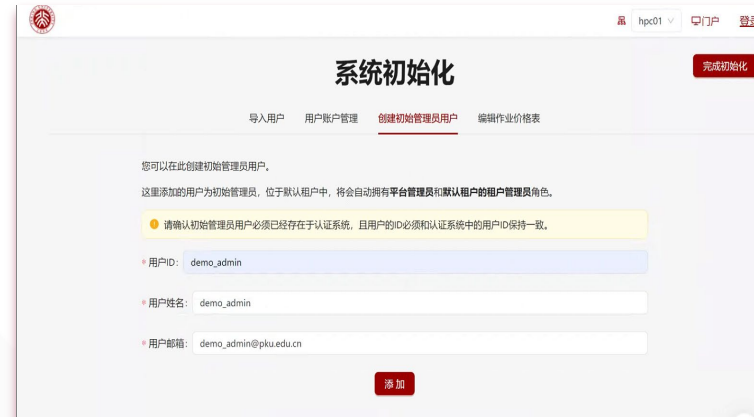
财务管理



SCOW——一体化部署，开箱即用

```
port: 8080
basePath: /demo/scow
image: ghcr.io/pkuhpc/scow/scow
imageTag: master
gateway:
  uploadFileSizeLimit: 20M
portal:
  novncClientImage: mirrors.pku.edu.cn/pkuhpc/novnc-client-docker:master
portMappings: {}
mis:
  dbPassword: must!chang3this
portMappings: {}
log:
  level: debug
pretty: false
fluentd:
  logDir: /var/log/fluentd_scow_logs
auth:
  portMappings: {}
```

```
[root@service scow-deployment]# ./cli compose ps
NAME                                STATUS                                PORTS                                COMMAND                                SERVICE                                CREATED
scow-deployment-auth-1              Up 13 hours                          80/tcp, 3000/tcp, 5000/tcp          "/entrypoint.sh"                    auth                                13 hours
scow-deployment-db-1                Up 13 hours                          mysql:8                             "docker-entrypoint.s..."          db                                13 hours
scow-deployment-gateway-1           Up 13 hours                          3306/tcp, 33060/tcp                "/entrypoint.sh"                    gateway                            13 hours
scow-deployment-log-1               Up 13 hours                          3000/tcp, 5000/tcp, 0.0.0.0:8088->80/tcp, ::8088->80/tcp, "tini -- /bin/entryp..."          log                                13 hours
scow-deployment-mis-server-1        Up 13 hours                          5140/tcp, 0.0.0.0:24224->24224/tcp, 0.0.0.0:24224->24224/udp, ::24224->24224/tcp, ::24224->24224/udp "/entrypoint.sh"                    mis-server                          13 hours
scow-deployment-mis-web-1           Up 13 hours                          80/tcp, 3000/tcp, 5000/tcp          "/entrypoint.sh"                    mis-web                             13 hours
scow-deployment-novnc-1             Up 13 hours                          80/tcp, 3000/tcp, 5000/tcp          "/entrypoint.sh"                    novnc                                13 hours
scow-deployment-novnc-1             Up 13 hours                          mirrors.pku.edu.cn/pkuhpc/novnc-client-docker:master "docker-entrypoint..."          novnc                                13 hours
scow-deployment-portal-server-1     Up 13 hours                          80/tcp                              "/entrypoint.sh"                    portal-server                        13 hours
scow-deployment-portal-web-1        Up 13 hours                          80/tcp, 3000/tcp, 5000/tcp          "/entrypoint.sh"                    portal-web                           13 hours
scow-deployment-redis-1             Up 13 hours                          80/tcp, 3000/tcp, 5000/tcp          "docker-entrypoint.s..."          redis                                13 hours
[root@service scow-deployment]#
```



01 编辑配置文件

配置部署的模块、集群基础信息、认证系统、交互式应用等信息。

02 编排与启动服务

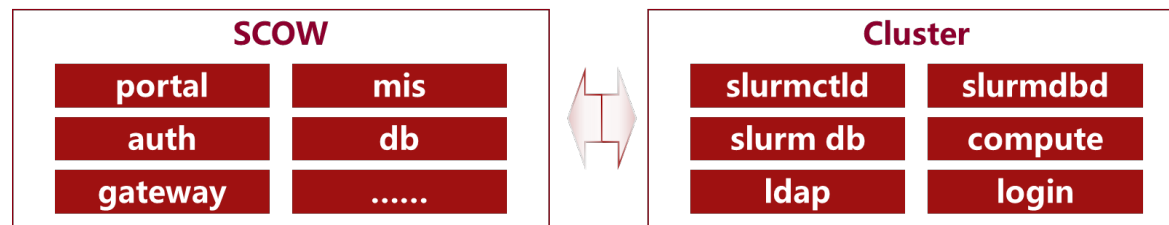
基于Docker Compose，使用自研SCOW-cli工具编排服务，一键拉取镜像，启动服务。

03 初始化管理系统

导入用户、用户账户管理、创建初始管理员、编辑作业价格表，完成初始化。



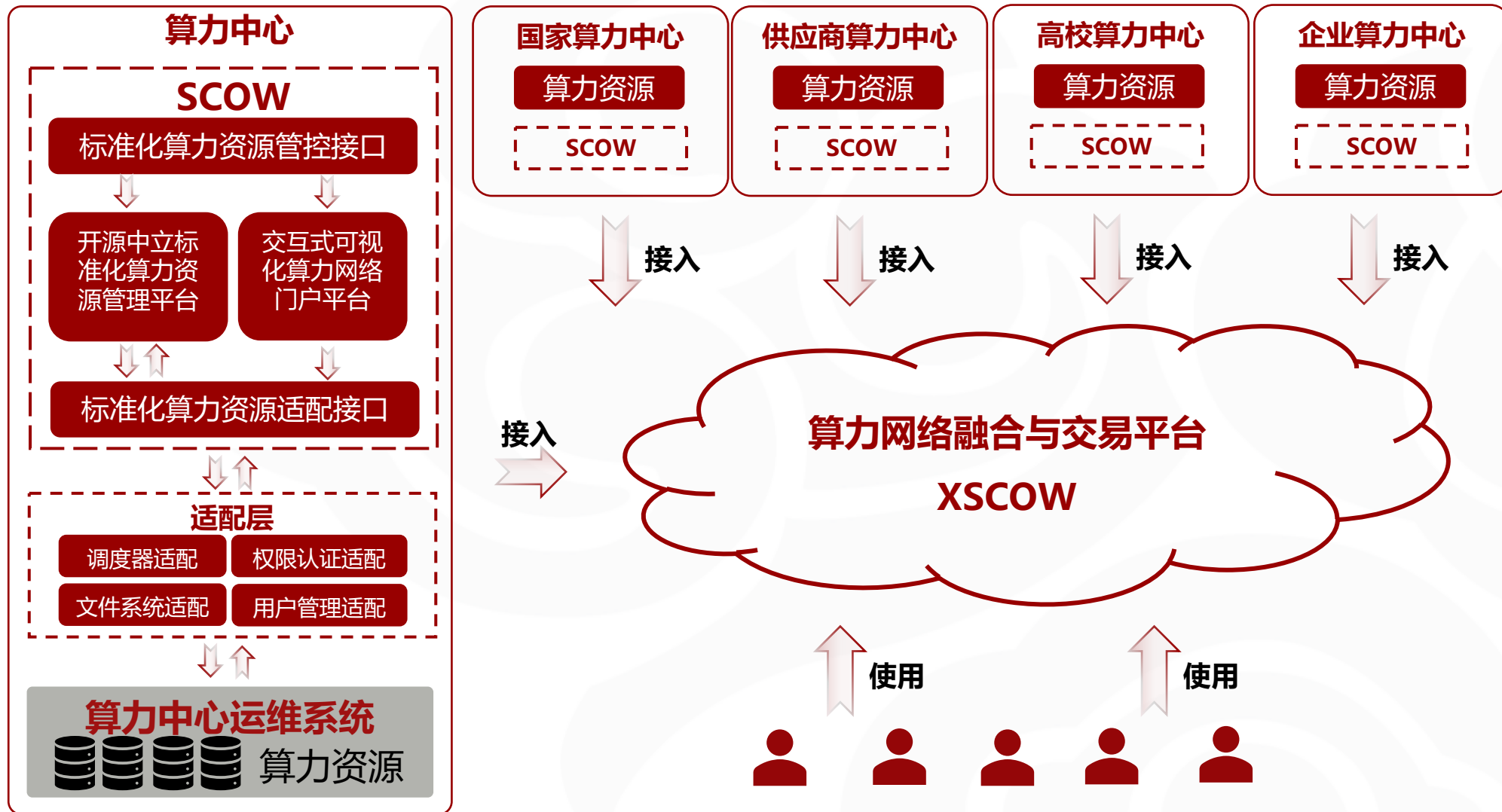
Vagrant 一键部署



Docker 快速部署

SCOW——标准化平台，支持算力融合

SCOW提供了一套标准的平台接口，可以实现的异构平台的统一抽象，在此基础上可以进一步研制跨中心的算力融合平台，打通算力网络中各高性能计算中心的管理、使用、结算通道，连通算力孤岛，实现算力灵活接入、统一调度。



- **项目由北京大学高性能计算平台、北京大学长沙计算与数字经济研究院共同研发**

自主完成系统核心逻辑功能开发，具有自主知识产权；

- **项目采用木兰宽松协议开源**

项目源码地址：<https://github.com/PKUHPC/SCOW>

项目文档地址：<https://pkuhpc.github.io/SCOW/docs/common/deployment>

更多项目信息：<https://icode.pku.edu.cn/info/1134/1137.htm>

- **项目试用地址：**

地址：<https://hpc.pku.edu.cn/demo/scow>

管理员 用户名/密码：demo_admin / demo_admin

普通用户 用户名/密码：demo_user / demo_user

基于透明代理的全流程审计

面向CI/CD的自动化安全
检测机器人



实现基于文件元属性和网络自适应的跨集群智能文件传输

研发低延迟高画质的远程桌面
连接技术

建立标准化算力中心管理模式

相关论文在2022年、2023年中国高性能计算学术年会中连续获得实践类**最佳论文**



2023年9月14日
SCOW 1.0版本正式发布

- **v1.2: 2023年11月19日发布**

- 1.增加在线编辑文件功能
- 2.增加系统基于域名的多风格支持

- **v1.1: 2023年11月6日发布**

- 1.增加跨集群文件传输功能
- 2.增加定时同步账户状态功能
- 3.增加国际化功能，提供英语版本

- **v1.0: 2023年9月14日发布**

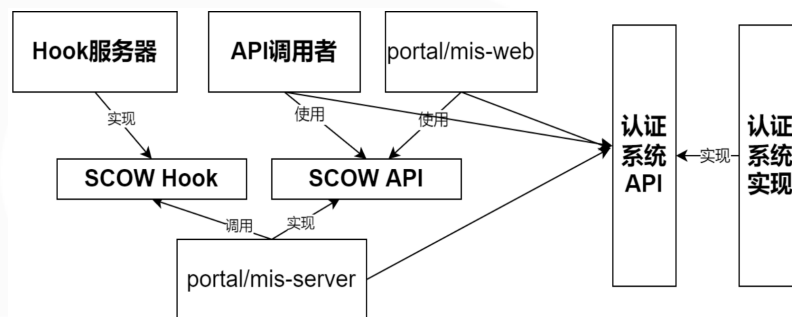
- 1.增加审计功能，不同级别角色可以查看对应操作记录

SCOW——完整功能、完全开放

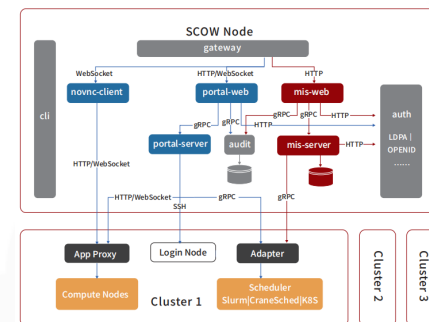
覆盖算力生命周期的完整功能



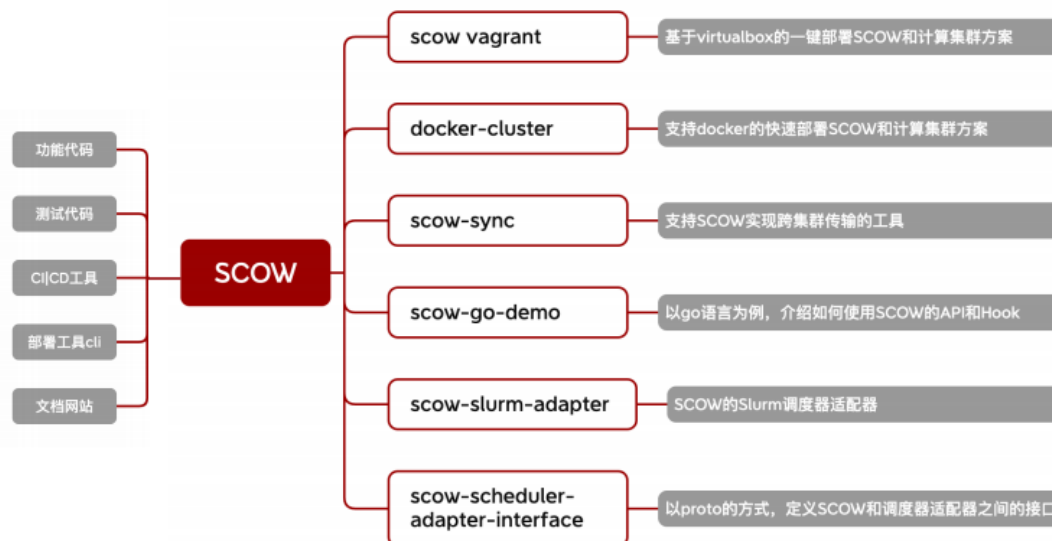
基于Hook和API机制的强大扩展能力



低入侵的部署模式



围绕项目部署开发的丰富代码库



平台部署已点亮14个省市35个算力中心

政府算力中心
4个

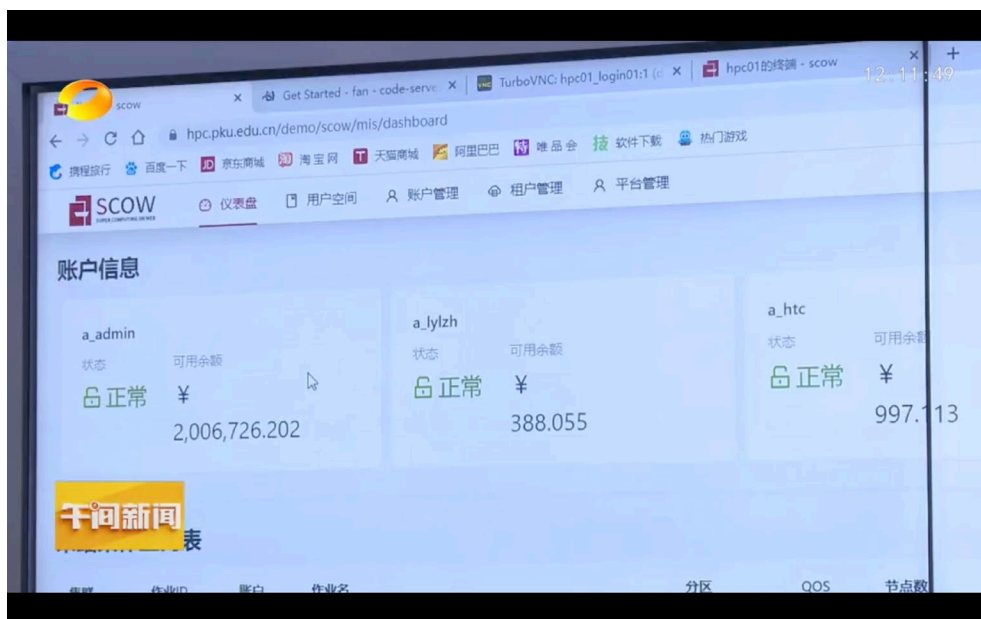
高校算力中心
17个

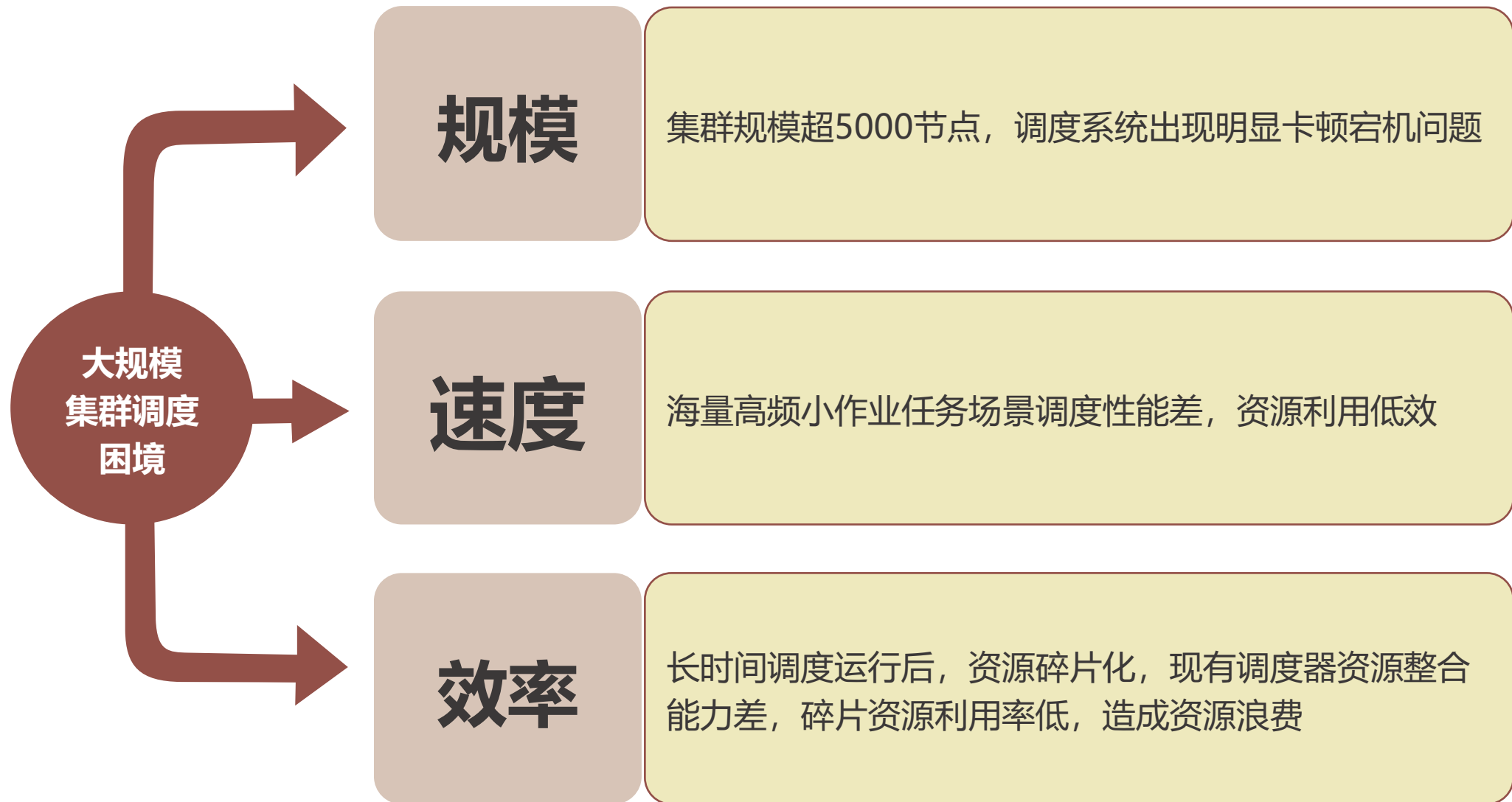
企业算力中心
9个

科研机构算力中心
5个

2023年7月1日以来开源社区下载数量
7000+

影响广泛，SCOW亮相湖南卫视





CraneSched——鹤思算力中心智能调度系统

- 已实现资源管理、任务管理、权限管理、任务智能调度等功能
- 解决了现有调度系统对**高通量计算**支持效率低的问题

应用
领域



支撑

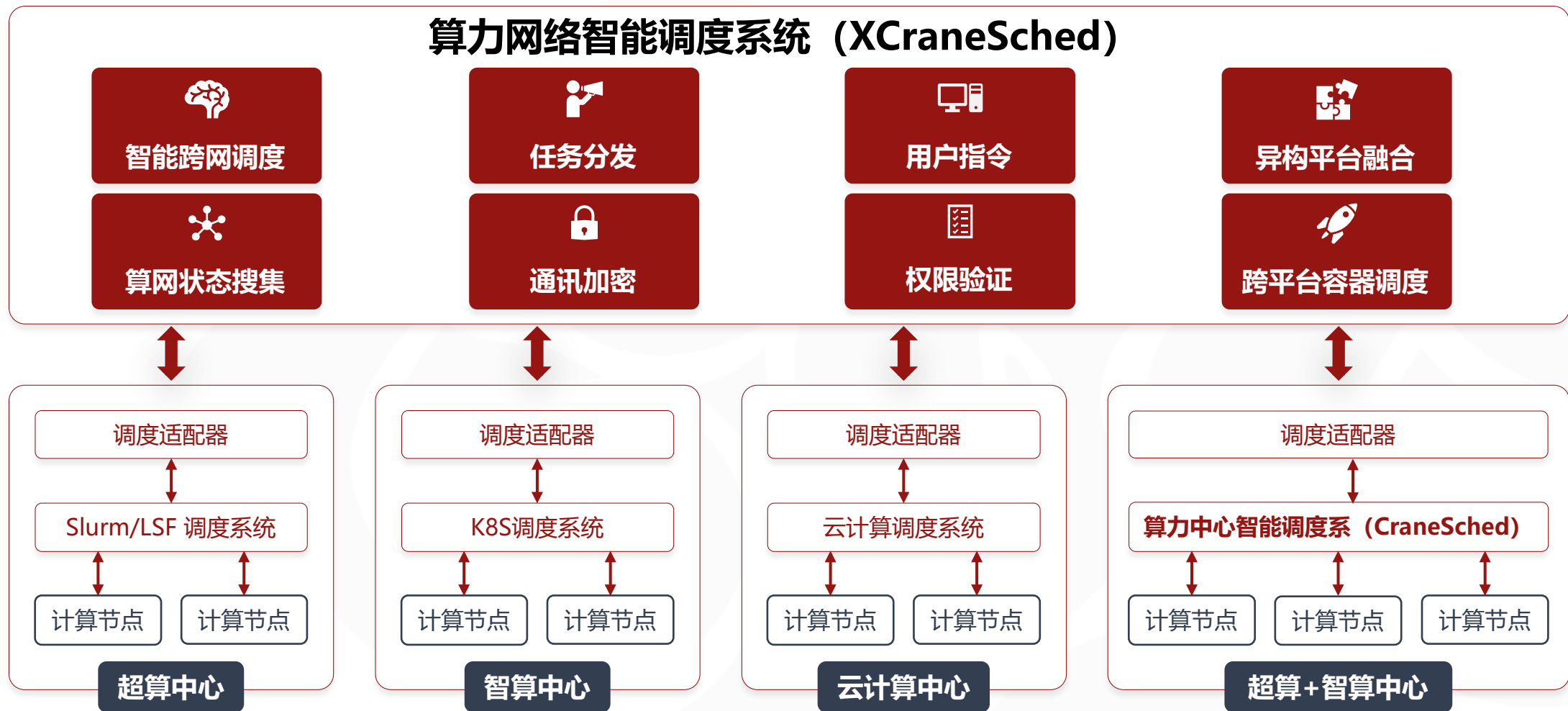
自主
研发
核心
模块



适配

异构
硬件





- 两级调度架构，上层调度为XCraneSched，下层调度为CraneSched
- XCraneSched 连接超算、智算、云计算等算力中心，将用户任务智能调度到最“合适”算力中心
- CraneSched 针对单个算力中心资源调度，调度完成XCraneSched分派的任务

content

目录

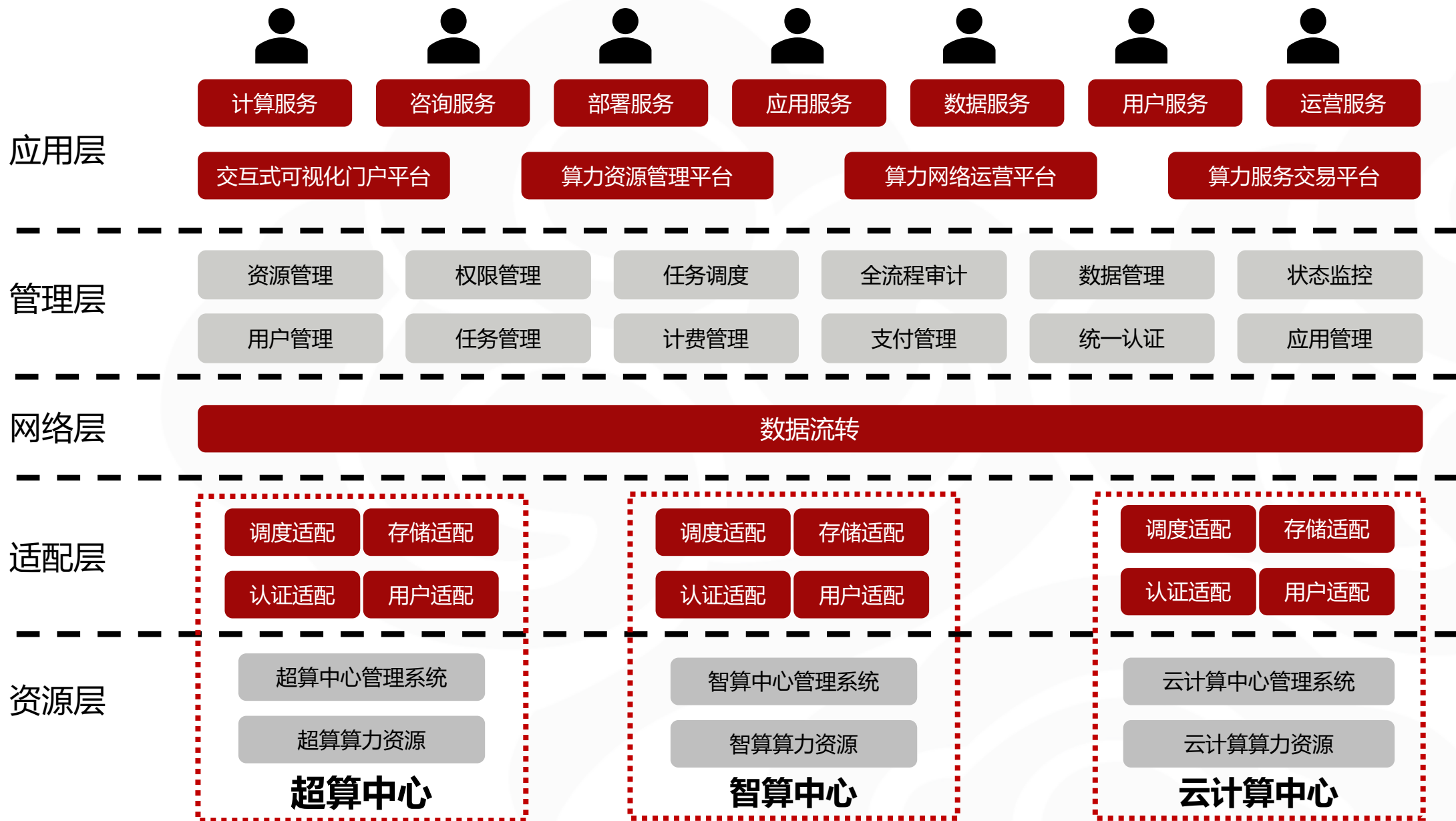
01 | **北京大学高性能计算平台**

02 | **算力网络时代的算力平台**

03 | **算力网络的探索与实践**

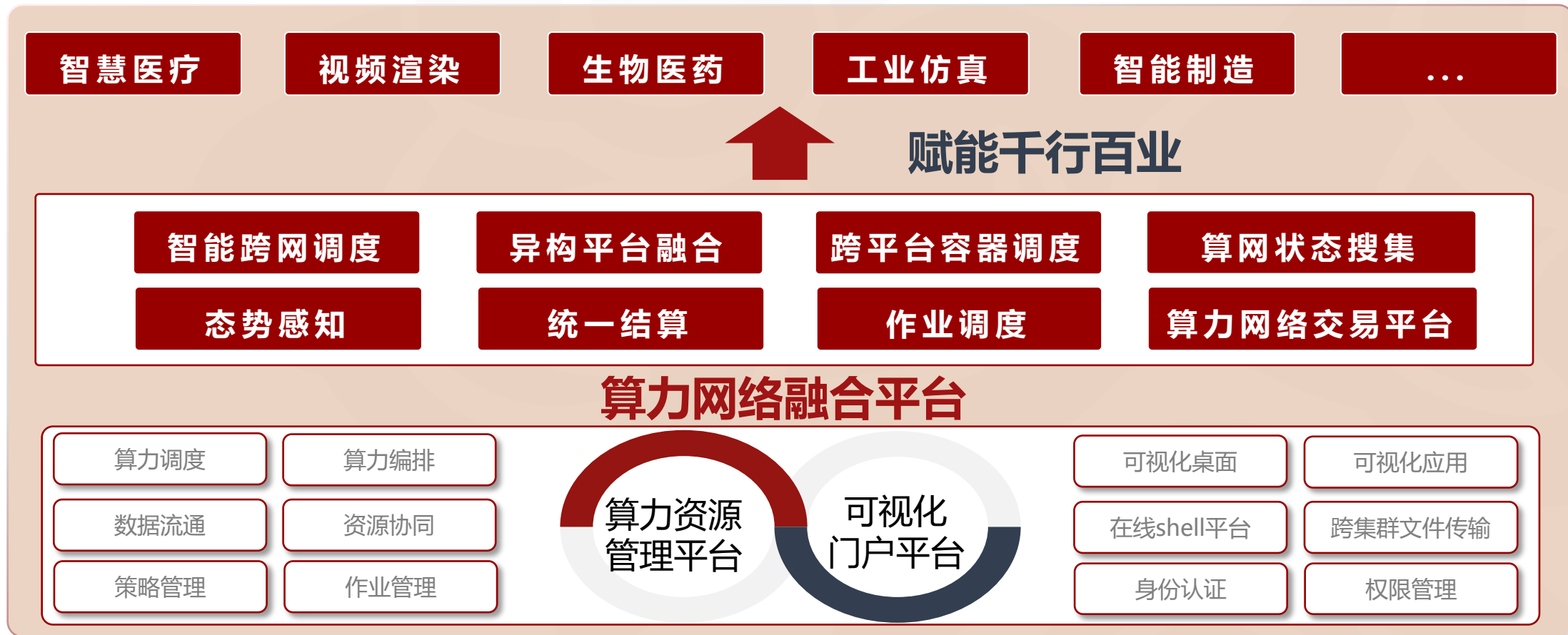
04 | **愿景**

XSCOW——算力网络融合与交易平台



愿景——算力网络融合平台

构建可实现产业化且兼容各类算力资源、对厂商中立的算力网络融合平台。解决算力网络中使用、接入、交易三个关键问题。建成能够覆盖各类算力资源的算力网络统一大市场。





北京大学
PEKING UNIVERSITY

感谢观看

THANK YOU

