



赛尔网络
CERNET



华通安和
HUA TONG AN HE

网络服务教育 创新开拓未来

基于教科网的 互联网超算业务探索

集成业务部 赵宇

2023年11月





目录

CONTENTS

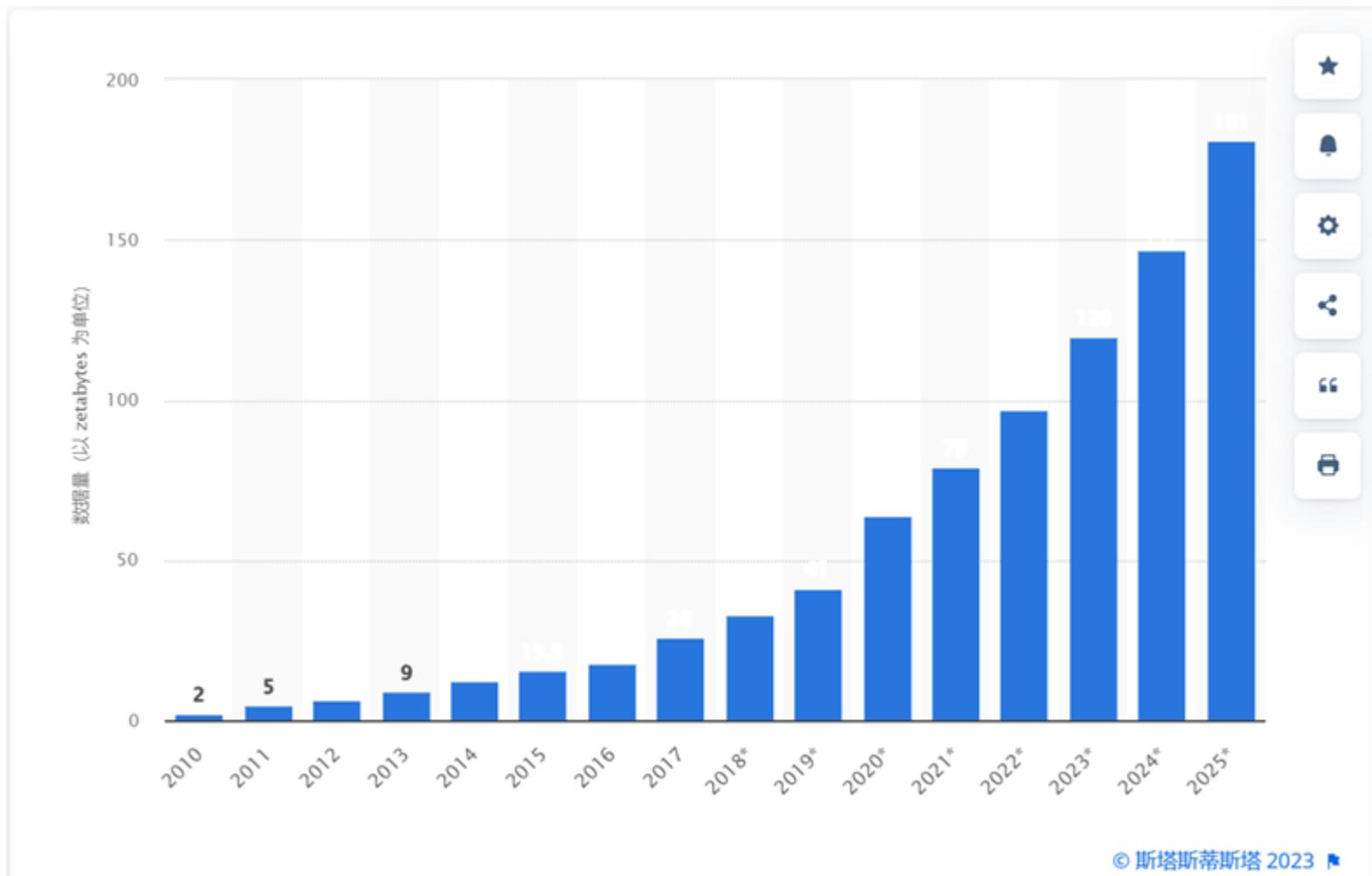
01 算力应用面临的挑战

02 · 互联网超算平台建设的构想

03 · 互联网超算的应用探索

04 · 高校算力建设的心得

三 海量数据及智算时代的时代



- 预计全球创建、捕获、复制和消费的数据总量将迅速增加，预计到2025年，全球数据创建预计将增长到 180 ZB 以上。（单位：ZBTABYTES）
- 每年23%的复合增长率。

以上数据来源：斯塔斯蒂斯塔2023



ChatGPT的前世今生

2018年6月 OpenAI发布GPT-1模型，**1.1亿**参数。

2018年11月 OpenAI发布GPT-2模型，**15亿**参数，但由于担心滥用，不向公众开放模型的全部代码及数据。

2019年2月 OpenAI开放了**GPT-2**模型的部分代码和数据，但仍然限制了访问。

2019年6月10日 OpenAI发布GPT-3模型，**1750亿**参数，并向部分合作伙伴提供了访问权限。（2019年7月，微软向ChatGPT的所有者OpenAI投下**10亿美元**）

2020年5月 OpenAI宣布推出GPT-3模型的beta版本，该模型拥有**1750亿个**参数，是迄今为止最大的自然语言处理模型。

2022年11月30日 OpenAI通过**GPT-3.5**系列大型语音模型微调而成的，全新对话式AI模型ChatGPT正式发布。

2023年1月9日 ChatGPT 第二次更新，改善了回答的真实性，增加了“停止生成”新功能。（2023年1月，微软就再投**100亿美元**）

2023年1月30日 ChatGPT第三次更新，在此提升了答案真实性的同时，还提升了数学能力。

2023年3月15日 OpenAI震撼推出了大型多模态模型GPT-4，不仅能够阅读文字，还能识别图像，并生成文本结果，现已接入ChatGPT 向Plus用户开放。（该模型拥有**1.76万亿**参数）

三 高速增长的算力需求

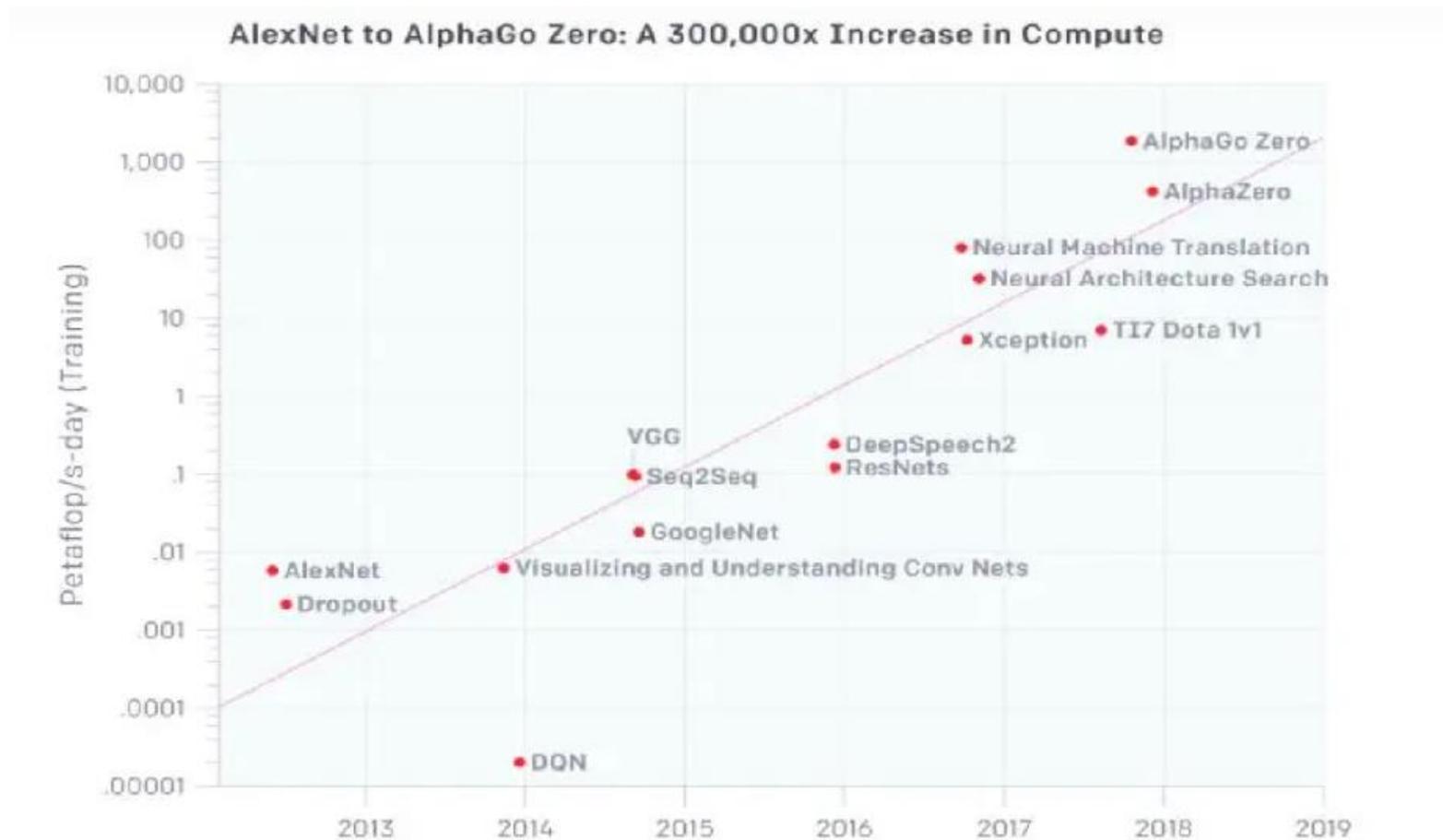


赛尔网络
CERNET



华通安和
HUA TONG AN HE

- OpenAI发布了一份关于AI计算能力增长趋势的分析报告，报告显示：自2012年以来，AI训练中所使用的算力每3.5个月增长一倍。
- 每年AI训练对算力增长需求是10倍以上的增长。



面临的挑战



赛尔网络
CERNET



华通安和
HUA TONG AN HE

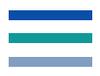
1、数据获取及传输

- 数据中心之间大数据传输，很多靠硬盘邮寄方式
- 大量数据需要从网络上获取，如何依托网络能快速下载数据。

网络带宽 (Gbps)	每小时下载 (TB)	每天下载 (TB)
10	4.5	108



- 64K
- 1M
- 1G
- 100G
- 1.2T



面临的挑战



赛尔网络
CERNET



华通安和
HUA TONG AN HE

2、算力/智算高速迭代

AI迭代速度基本每年需要扩大10倍的算力。

大部分学校建设周期的问题，从立项到实施完毕周期12-24个月，即使从准备招标采购到建设完毕6个月。

产生问题

建设完毕
跟不上需求

校内普及后使用
更跟不上需求

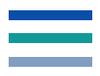
大量使用社会资源
社会资源存在异构问题，社会算力需要采购规模大，溢价问题



面临的挑战

3、国外对GPU限制、超算资源异构及使用中优化问题





面临的挑战

4、计算机类专业人才需求



人才培养

需要培养
超算/AI人才



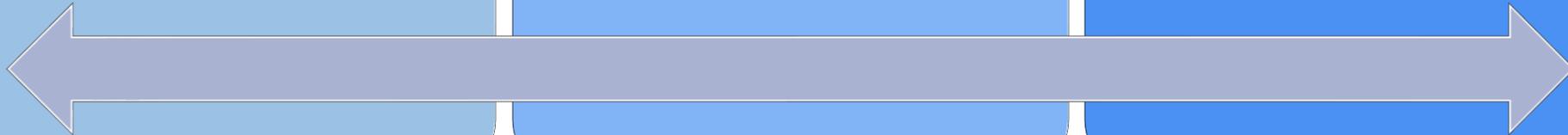
人才使用

需要成熟的软件编
程人员，开展适配、
优化软件工作。



留住人才

需要建立长效机制
留住人才





目录

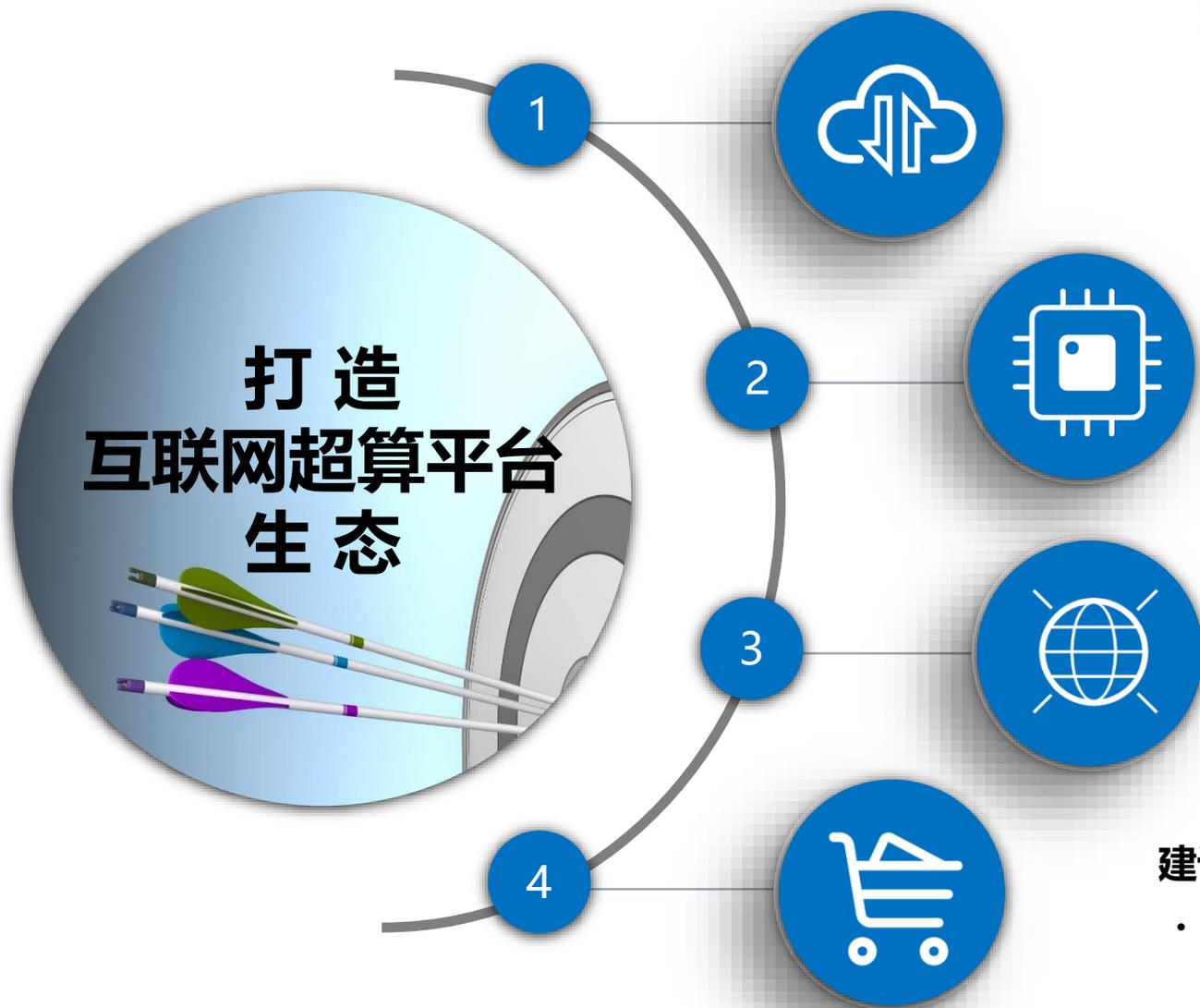
CONTENTS

01 算力应用面临的挑战

02 互联网超算平台建设的构想

03 · 互联网超算的应用探索

04 · 高校算力建设的心得



突破网络传输带宽的限制

- 有遍布全国的骨干网络CERRENT、CERNET2，骨干带宽100G，支持算力资源及用户大带宽接入。
- 1.2T超高速互联网主干通路，“未来互联网试验设施FITI”完成试验并开通，连通了北京、武汉、广州三大核心节点，1.2T超高速互联网主干通路。

引进校外算力满足需求

- 建立结算机制与国家级超算中心一起服务用户。
- 在适当时候接入高校、商业算力资源为用户服务。
- 为高校算力建设提供设计、方案、实施、运维、安全等服务。

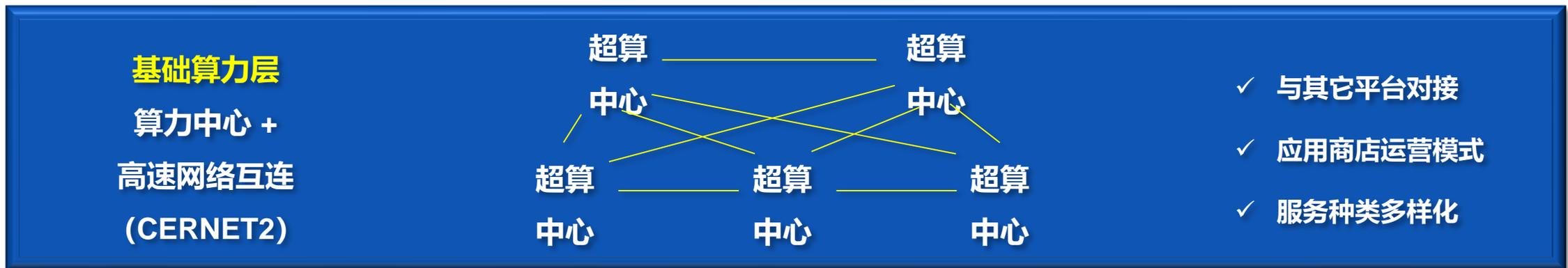
整合软件资源

- 提供网上软件产品店铺平台。
- 提供网上开发者店铺平台。
- 建立适应软件开发采购的结算机制保障店铺利用。

建设长效机制

- 需要建立并有效运行长效机制，培养、留住计算机专业人才，在算力领域发挥作用，让用户在面对各种问题时得到及时和有效的支持。

三 互联网超算算力及服务平台愿景



总体架构



用户层



XX大学



XX实验室\中心



XX企业



.....

拓展增值服务

平台运营

基于公司的集成、软件等经验，提供专业引导，辅助高校一站式购买相关服务。

软件开发服务

应用开发

基于各个开发者提供的应用程序和编译开发服务，为高校提供算力模型等服务。

算力调度服务

算力调度

基于各个超算中心的超算资源，为用户提供算力机时服务、算力调度服务。

网络保障服务

网络管理

基于CERNET2 10/100G的专网和前端机组，提供数据传输、转发和调度服务。

安全保障

运营保障

基于统一身份认证的互联网超算业务管理平台

资源层

XX国家超算中心

XX大学超算中心

XX企业超算中心

...

XX国家智算中心

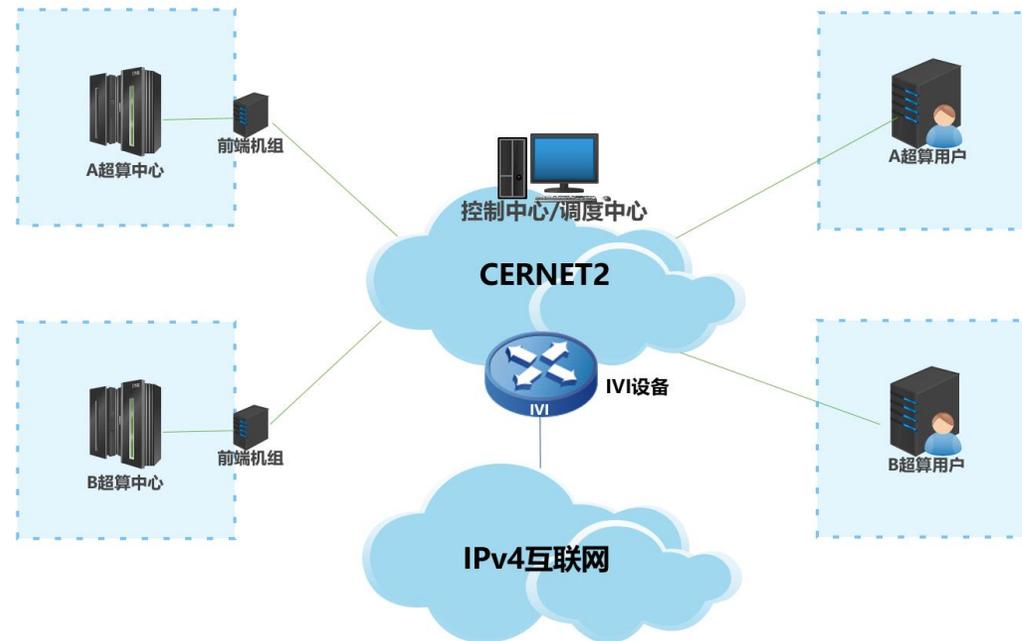
XX大学智算中心

XX企业智算中心

基于CERNET2网络 搭建算力及服务平台

三 总体架构——网络保障服务

- CERNET2网络连接各个超算中心和超算用户，组成超算专网。
- 各个超算中心部署前端机组，建立分布式集群，其作用包括：**数据缓存加速、数据共享效率提高、数据安全性提升**等3个方面。
- 通过统一的控制调度中心，使教育科研网内所有超算用户和超算中心共享一套统一的管理体系。
- 实现数据调度、传输、计算的分离。



图：基于CERNET2对于数据传输的网络保障

三 总体架构——算力调度服务



赛尔网络
CERNET



华通安和
HUA TONG AN HE

任务调度

根据用户提交的计算任务优先级、资源需求和可用性等因素进行任务调度和分配，提高计算效率。

资源分配

集中管理各种计算资源，如服务器、虚拟机、容器等，并按需分配和释放资源，以提供高效的计算环境。

对计算任务的执行过程进行实时监控，并对资源的使用情况进行统计和分析。

运行监控

机时管理

帮助用户灵活选择所需的计算资源，并按照实际需求购买或租用计算资源，满足用户计算需求。



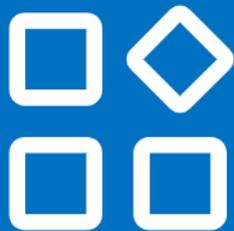
三 总体架构——软件开发服务



赛尔网络
CERNET



华通安和
HUA TONG AN HE



应用软件开发服务

提供各种应用软件开发服务，包括软件的开发、测试、部署和运行等。满足高校在教学、科研和管理等各个方面的需求，提高工作效率。



编译开发服务

提供编译技术的研发、测试和优化等，可以满足高校在高性能计算、人工智能等领域的需求，帮助高校提高计算性能和效率。



算力模型及优化服务

提供算力模型的优化、部署和运行监控等算力模型服务，帮助高校建立和应用更高效的计算模型，从而提高计算性能和效率。



三 总体架构——拓展增值服务



赛尔网络
CERNET



华通安和
HUA TONG AN HE



需求分析与评估

- 帮助高校了解和明确自身的超算服务需求，包括计算能力、数据存储和传输等需求。通过与服务需求进行匹配和分析，为高校提供合适的超算服务方案。



市场调研与比较

- 进行市场调研和比较分析，了解当前超算服务的市场动态和供应商情况，为高校提供全面的信息和咨询服务，为高校提供专业的建议和指导。



技术评估与推荐

- 拥有丰富的经验和专业知识专业团队，深入了解各种超算服务的性能、稳定性和安全性等方面，为高校提供专业的技术评估和推荐意见。



售后支持与维护

- 提供售后支持与维护服务，包括技术咨询、故障排除和系统升级等方面，及时响应高校的需求和问题，并为其提供高效、优质的售后支持与维护服务。

三 总体架构——两大保障



赛尔网络
CERNET



华通安和
HUA TONG AN HE



运营管理保障

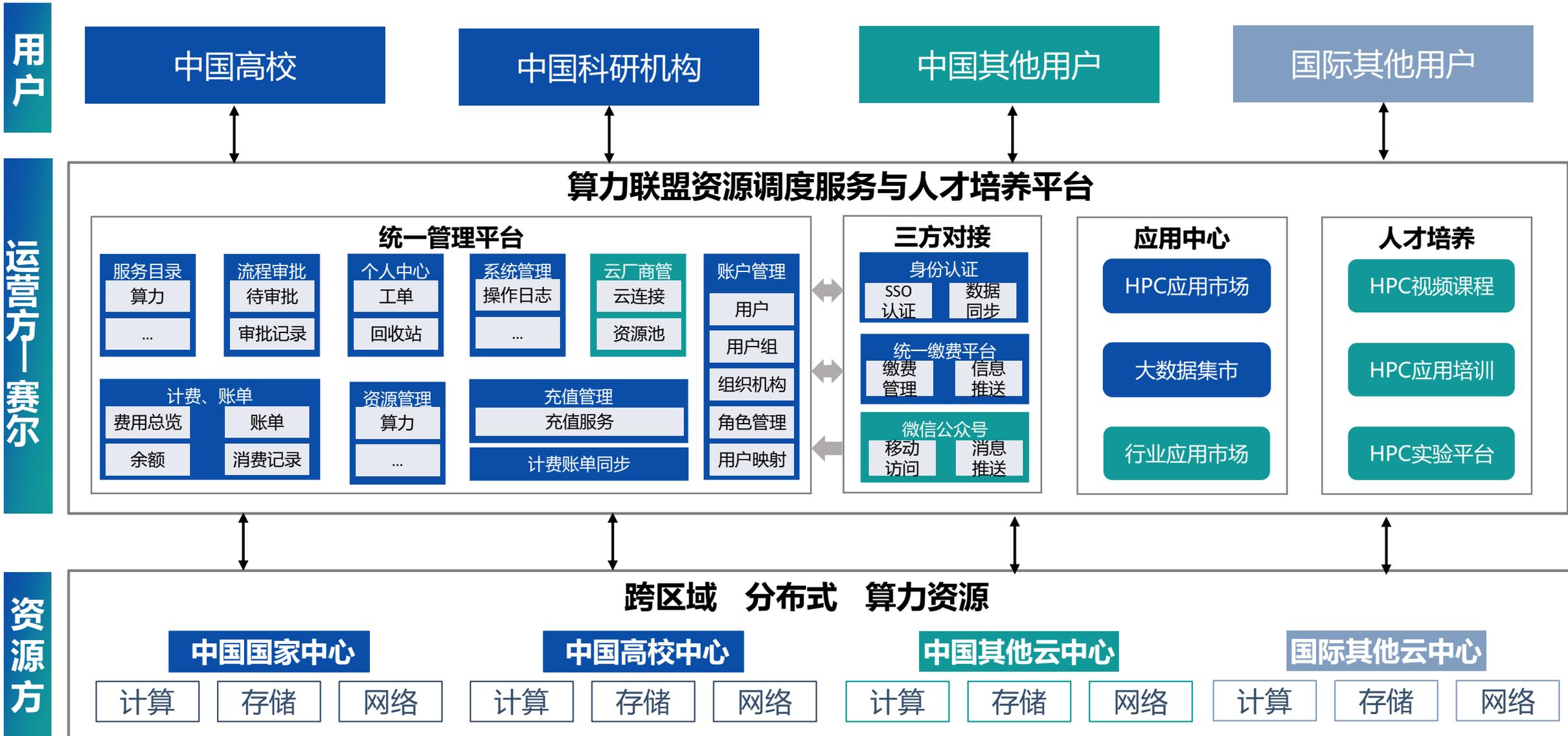
- 1. 资源管理：**对超算资源进行集中管理和调度，确保资源的合理分配和高效利用。
- 2. 任务管理：**对计算任务进行管理和调度，确保任务的按时完成和正确执行。
- 3. 用户管理：**对用户统一认证和管理，确保用户的安全性和权限控制。
- 4. 计费管理：**设定相应的计费标准，如按机时、按核数、按流量等。
- 5. 日志管理：**记录计算任务和资源的使用日志，并提供查询和分析功能。



安全服务保障

- 1. 网络安全：**保障系统网络的安全性，防止网络攻击和数据泄露。
- 2. 数据安全：**对用户的数据进行加密和备份，确保数据的完整性和安全性。
- 3. 认证安全：**使用SSL/TLS等加密协议，确保用户数据在传输过程中的安全性。
- 4. 安全审计：**对系统的访问和使用进行审计和监控，及时发现和处理安全问题。
- 5. 应急响应：**对可能出现的网络安全事件进行应急响应，及时处理和恢复系统。

三 推动教科网高校算力资源共享与人才培养



注：深蓝表示第一阶段工作任务，绿色表示第二阶段工作任务，浅蓝表示第三阶段工作任务。



目录

CONTENTS

- 01 算力应用面临的挑战
- 02 互联网超算平台建设的构想
- 03 互联网超算的应用探索
- 04 · 高校算力建设的心得

科技部“国家高性能计算环境核心资源网”项目



■ 核心网内容

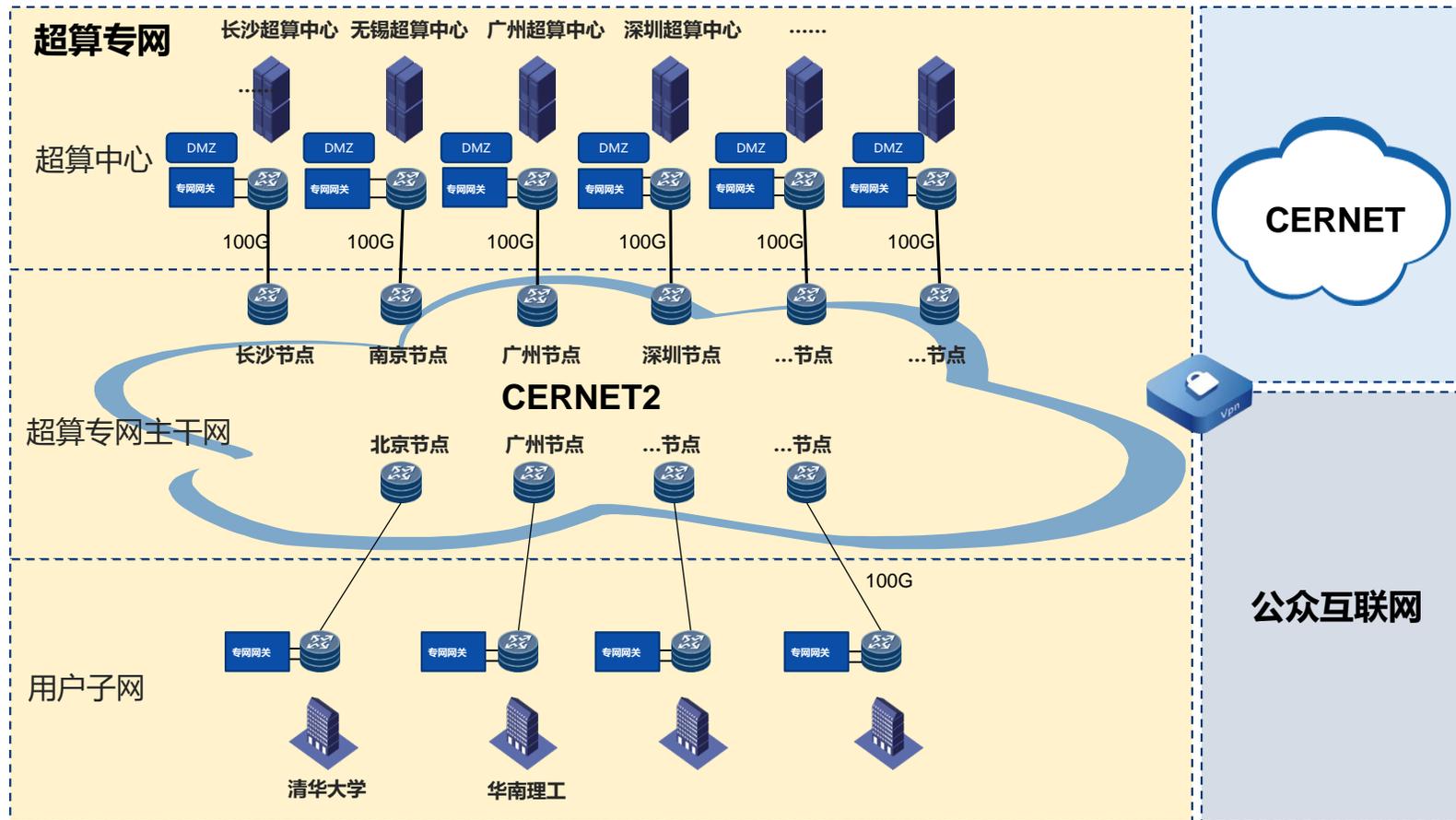
基于CERNET2-IPv6骨干网构建100Gbps的具有足够安全保障的网络互连，实现国家级超算中心间的高速网络连接，形成基于高速网络和超级计算机的国家高性能计算环境核心资源网。

■ 超算中心

无锡、长沙、广州、深圳。。。

■ 重点用户

华南理工大学、清华大学。。。



图：国家高性能计算环境核心资源网

三 本项目—实际完成情况



超算中心	IPv6	CERNET2节点	接入情况
无锡中心	100G	南京节点	已接入100G
广州中心	10G	广州节点	已接入10G
深圳中心	10G	深圳节点	已接入10G
长沙中心	10G	长沙节点	已接入10G
中科院	10G	北京直联点	骨干网北京直联点 (测试中)
郑州中心	10G	郑州节点	测试中
济南中心	10G	济南节点	测试中

➤ 截至目前，4家超算中心已接入10-100G CERNET2 (IPv6) 网络，另外3家超算中心已在测试过程中CERNET2。

注：红色字体代表该项目考核指标中涉及的国家级超算中心。

三 基于CERNET2 (IPv6) 的互联网超算

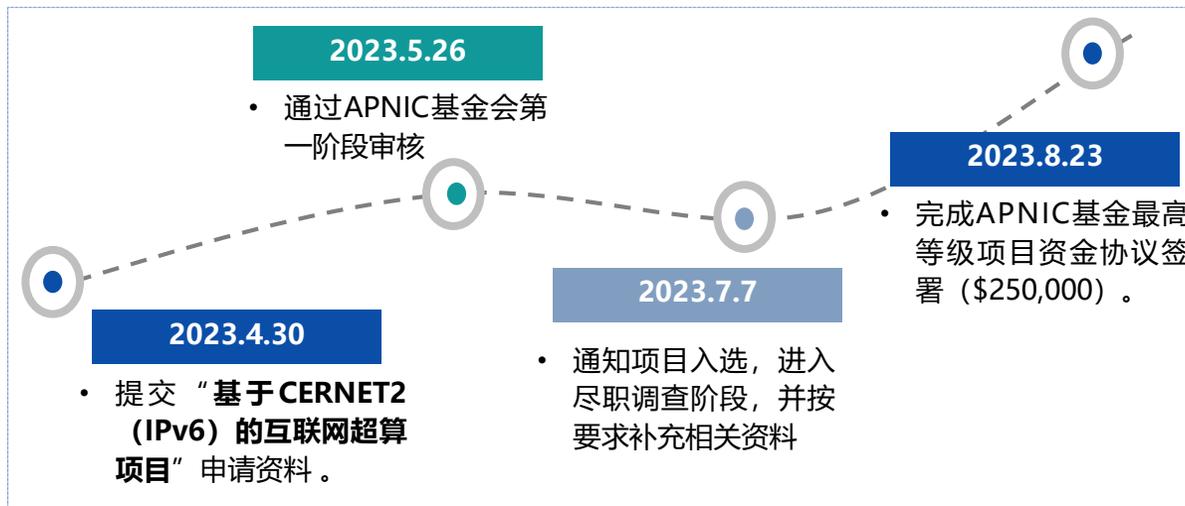


项目简介

在前期参与科技部项目课题的工作基础上，公司向APNIC基金会申请了“基于CERNET2 (IPv6) 的互联网超算”项目资助，资助类型为ISIF Asia IPv6部署项目。

“基于CERNET2 (IPv6) 的互联网超算”项目是基于中国教育和科研计算机网纯IPv6骨干网络CERNET2的超算中心网络高速互联项目，由CERNET网络中心指导，赛尔网络有限公司承担算力资源网络接入、平台运营及用户服务等工作。其目标是广泛连接高性能计算领域的供需端，促进不同超级计算中心之间的异构集成，最终实现大型专业超级计算资源、高校超级计算资源和其他相关社会资源在CERNET2及其纯IPv6基础设施上的部署，通过算力及服务平台合理调度算力资源和使用相关应用。

资助类型
USD 30k 小额资助
USD 85k 大规模资助
USD 150k 影响力资助
USD 30k - 250k IPv6部署
Ian Peter Grants for Internet and the Environment



三 无锡中心与清华大学高速数据共享网络服务



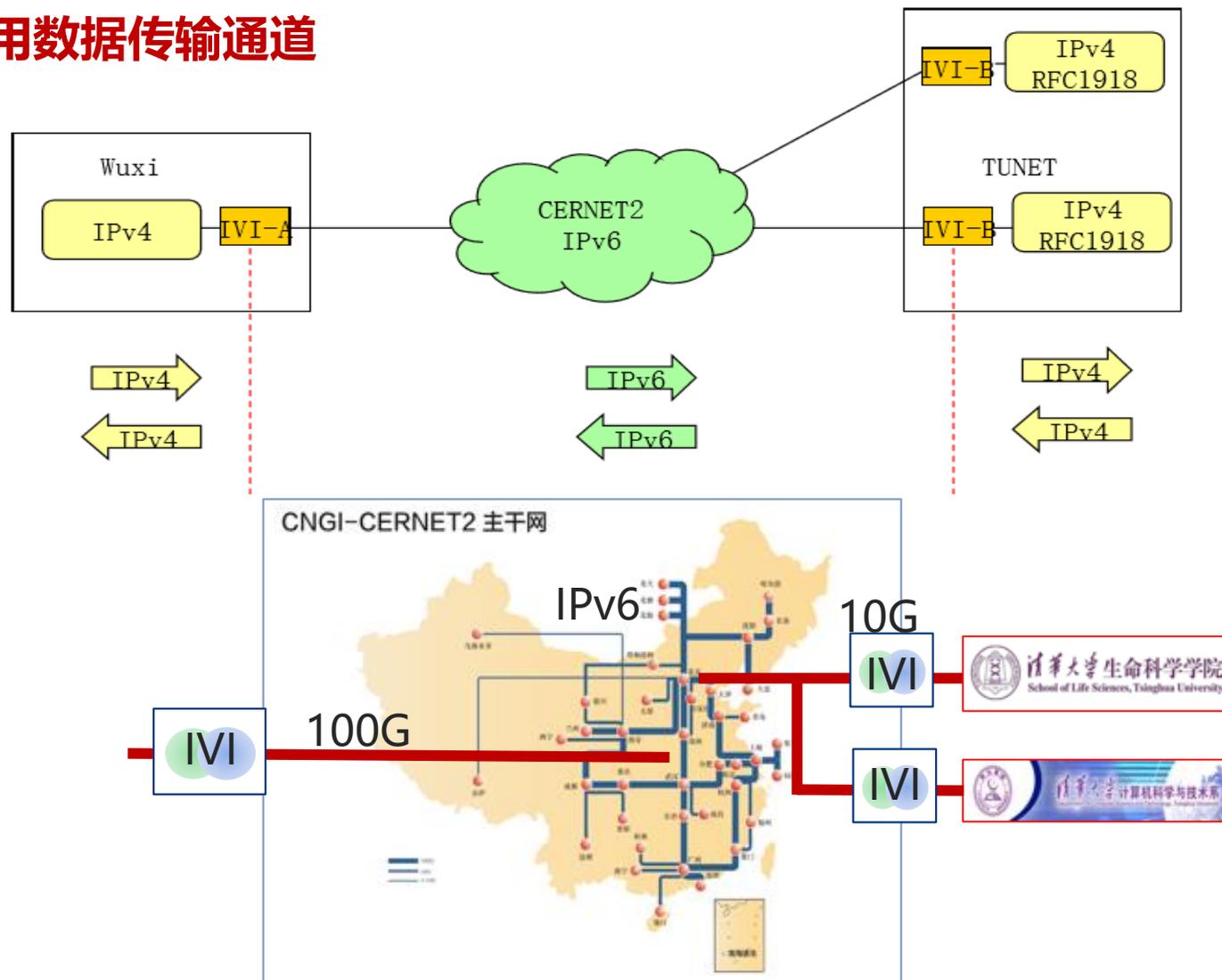
赛尔网络
CERNET



华通安和
HUA TONG AN HE

为超算中心和超算用户搭建高速专用数据传输通道

- 协助有超算需求的教育科研网会员单位诊断并协助处理数据传输瓶颈问题。
- 目前已开通无锡超算 100G 接入 CERNET2。
- 采用两次翻译技术 (IVI) 实现了清华大学与无锡超算 IPv4-IPv6-IPv4 (实测最高 8.3 Gbps)、IPv6-IPv6 的链接。



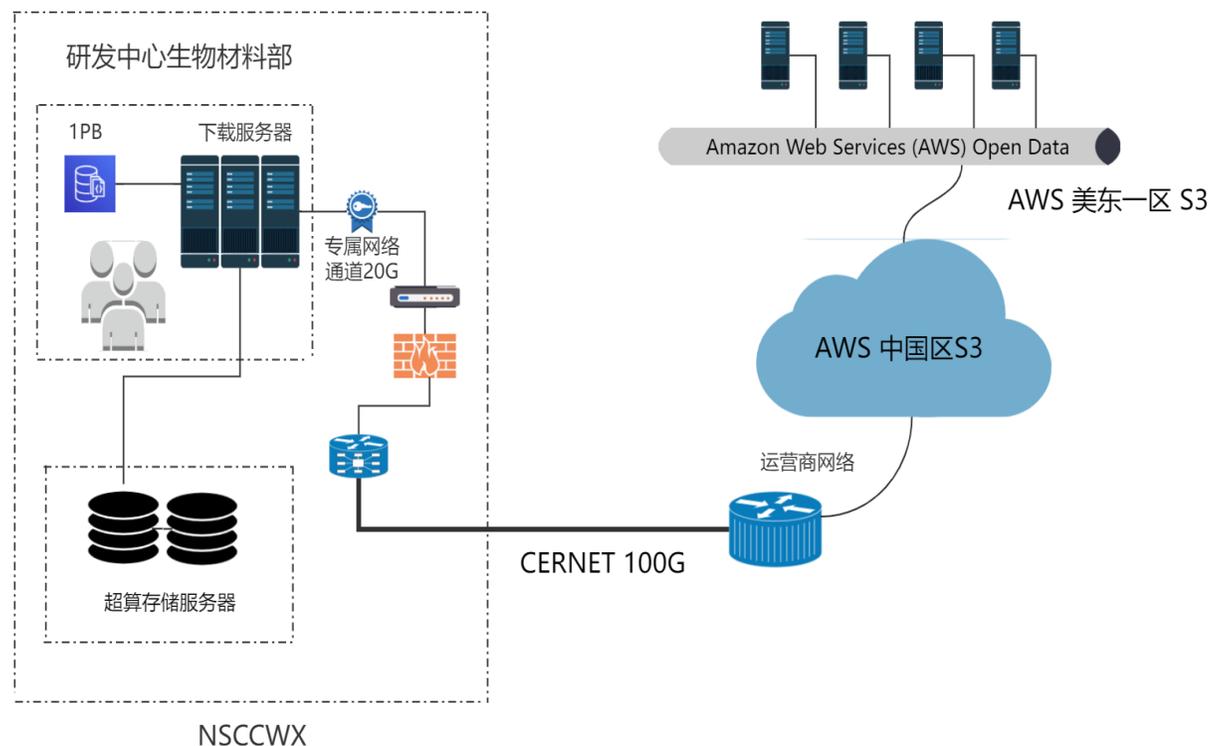
三 无锡超算研发中心生物材料部-亚马逊云大数据传输网络



2023年9月，国家超级计算无锡中心（无锡超算）研发中心生物材料部的研发计划，需要从亚马逊云下载978TB数据。全流程数据部署项目于2023年10月14日开始测试，预计与亚马逊协议签署后1个月内完成。

综合成本、便捷性及后期数据库维护等因素，使用AWS Data Transfer Hub解决方案和AWS CloudFront 服务，实现数据由美东一区分发下载至无锡超算，下载使用IPv4协议。

无锡超算中心独立划分20Gb CERNET带宽进行数据下载，实际下载速度峰值可达**14.06Gbps**（带宽利用率**70.3%**）。



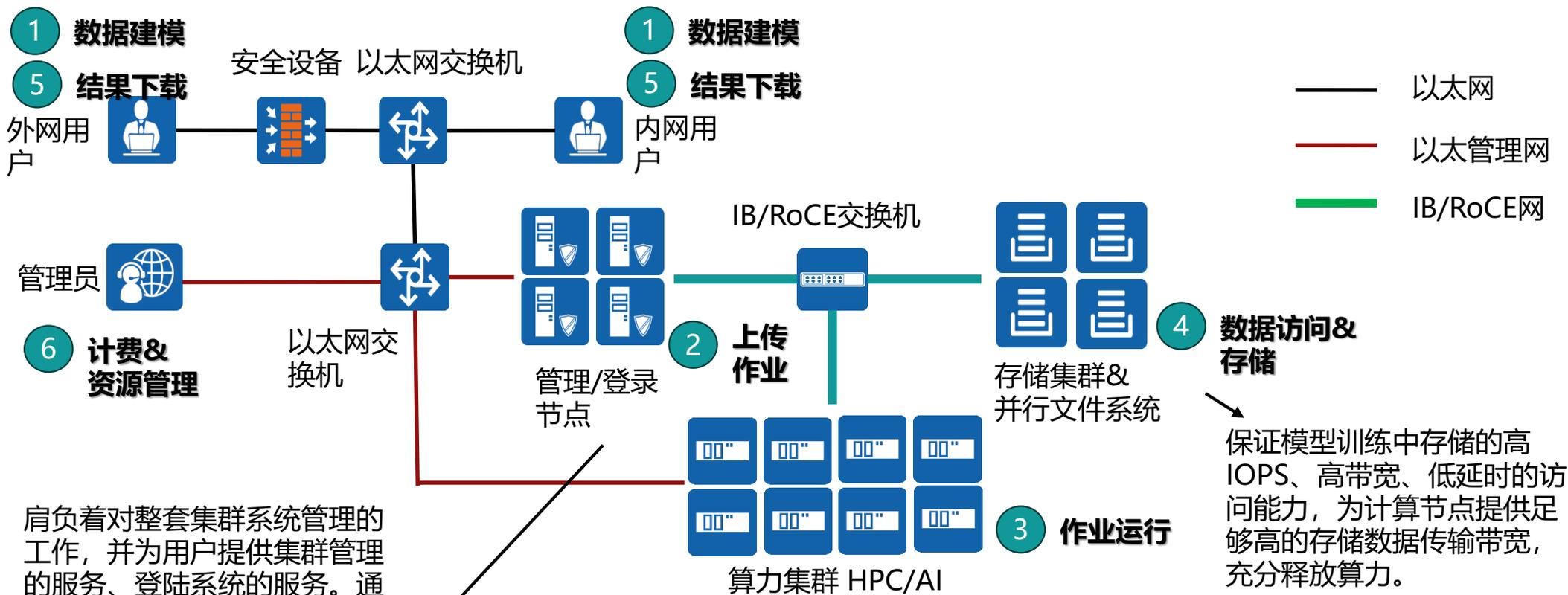


目录

CONTENTS

- 01 算力应用面临的挑战
- 02 互联网超算平台建设的构想
- 03 互联网超算的应用探索
- 04 高校算力建设的心得

三 高性能计算工作流程



肩负着对整套集群系统管理的工作，并为用户提供集群管理的服务、登陆系统的服务。通过节点上管理软件、作业调度软件，能为用户提供统一的用于监控管理的人机交互界面，统一的用于作业调度的人机交互界面，方便用户对集群系统的维护和使用。

主要作用是接收分配的
计算任务，执行实际计算
任务，并将结果发送回
对应节点。

耗电量：	控制运维成本：
● 双碳指标	● 人员
● 电费	● 能耗

三 赛尔网络-技术服务实力与能力



- ISO9001质量体系认证
- ISO20000服务体系认证
- ISO27001信息安全管理体系统认证
- 信息系统建设和服务优秀级-CS4
- 信息技术运行维护标准成熟度二级
- 信息系统安全集成三级
- 信息系统安全运维三级
- 信息安全应急处理三级
- CMMI3
- PMP
- 信息系统项目管理师 (高级项目经理)
- 系统集成项目管理工程师
- IT运维项目经理

- 浪潮战略合作伙伴
- 曙光战略合作伙伴
- 新华三一级代理商+四星服务
- 华为多产品金牌经销商
- 华三认证工程师 (H3CTE, H3CSE, H3CNE)
- 思科认证工程师 (CCIE, CCNP, AM)
- 华为认证工程师 (HCIE, HCNP, HCNA)
- 浪潮认证工程师 (IPPE, IPPT, IPPB)
- CISP信息安全工程师
- 软件研发工程师
- IT服务工程师
- 华为鲲鹏应用认证工程师



三 长期工作-建设互联网超算生态



✓ 持续接入大型超算中心





赛尔网络
CERNET



华通安和
HUA TONG AN HE

网络服务教育 创新开拓未来

感谢各位领导聆听