

2024开放计算中国峰会

2024开放计算中国峰会

2024开放计算中国峰会

# AI大模型算力底座实践与思考

中国工商银行数据中心 陈庆

2024开放计算中国峰会

2024开放计算中国峰会

2024开放计算中国峰会

2024开放计算中国峰会

2024开放计算中国峰会

2024开放计算中国峰会

# 金融AI实践背景

国家战略

新质生产力

金融五篇大文章

《“十四五”数字  
经济发展规划》

2023年中央金融工作会议

《算力基础设施高质  
量发展行动计划》

2024年政府工  
作报告

国家金融监督管理总局  
《“五篇大文章”指导意见》

云网协同  
算网融合

科技创新  
产业创新

金融五篇  
大文章

新型生产力

数字经济  
创新发展

守正创新  
风险可控

科技金融质效

# AI大模型基础设施能力要求

算网联动

协同感知

自主可控



# AI大模型集群建设历程

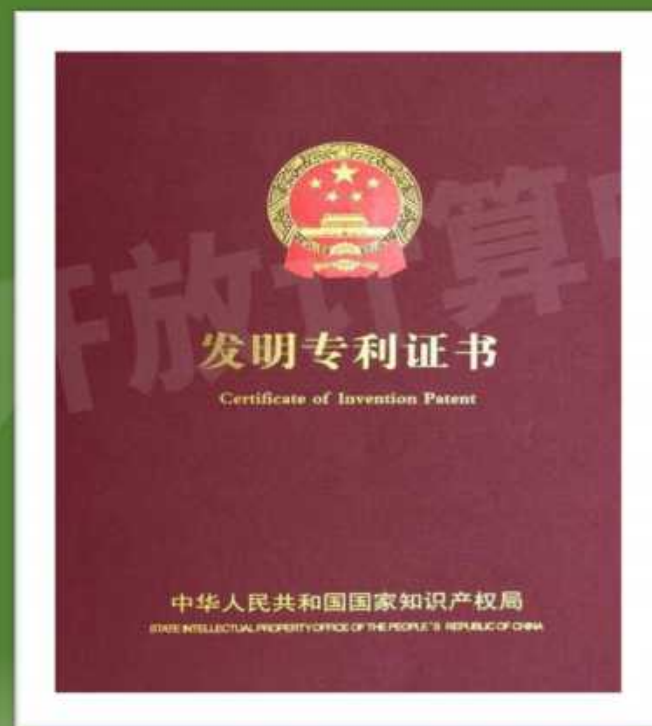


# AI大模型集群技术架构



# AI大模型集群建设成果

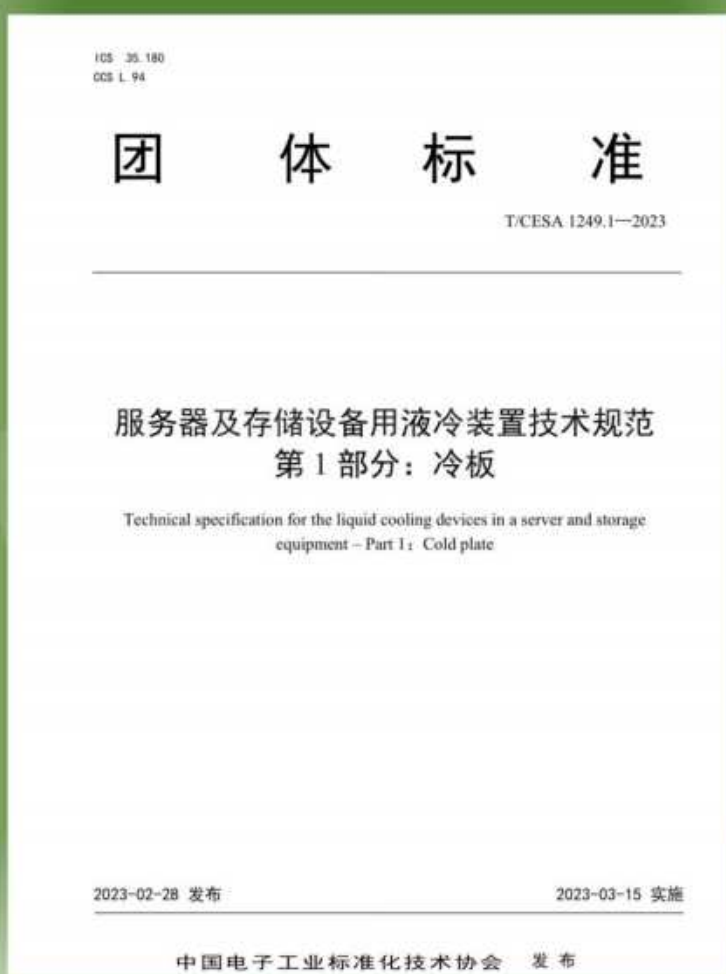
发明专利  
公开25项、授权8项



金融科技产业联盟  
《金融数据中心人工  
智能算力建设指引》



开放计算标准委员会  
液冷标准4项



IDC  
未来数字基础  
架构领军者



国家工信部  
“华彩杯”算力大赛全国二等奖  
中国算力大会  
创新先锋  
算力网络技术创新优秀案例



# AI底座思考—多算力中心互联

算力中心广域互联



可编程SRv6骨干网 | RoCE | 算力卸载 | 分布式算力 | 算力网络感知 | 算力网络调度

端边云协同

# AI底座思考—AI集群运维提升

性能优先



兼顾性能和连续运行





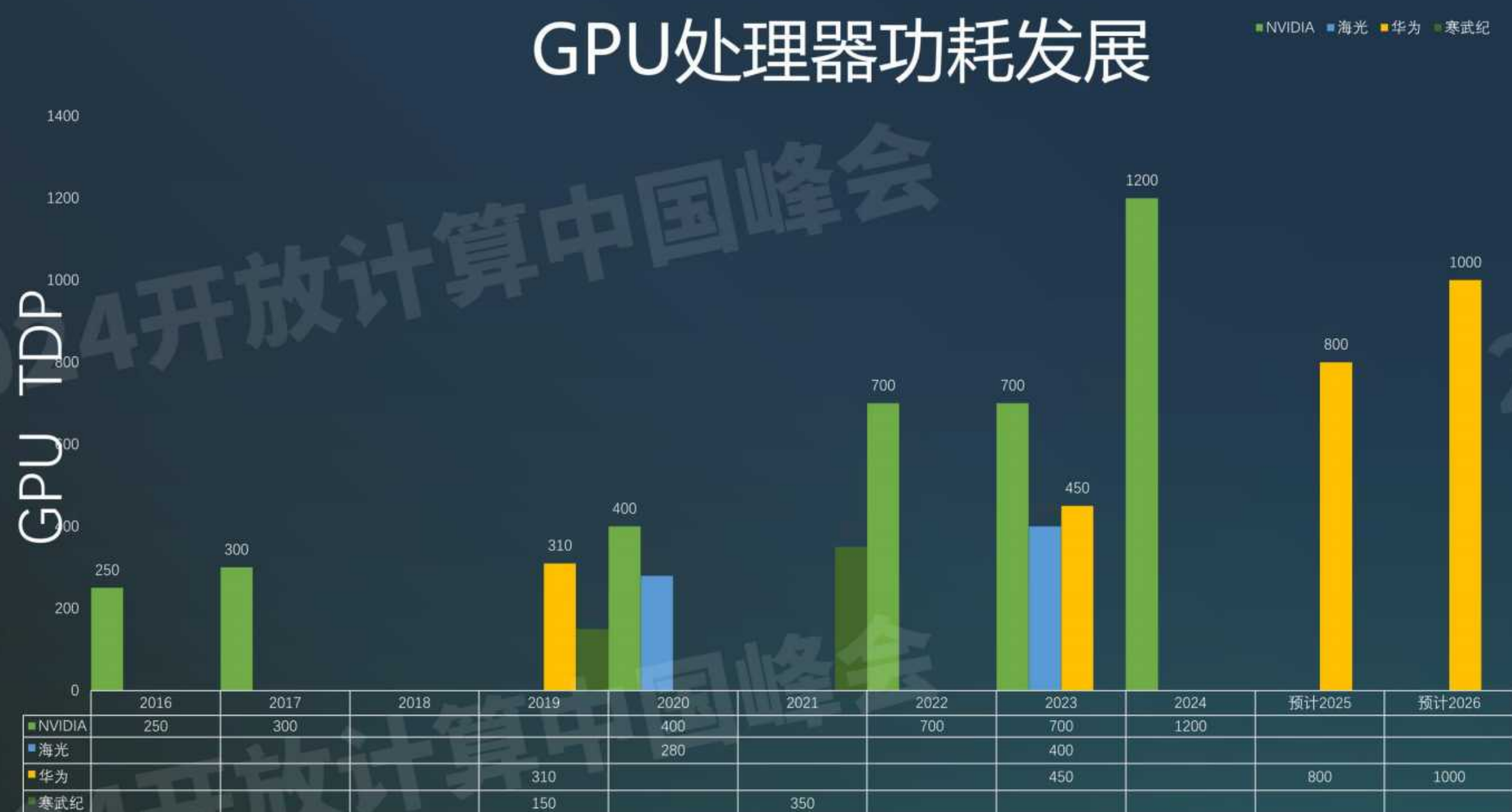
# AI底座思考—散热与供电

高功率、高密度、大规模算力底座趋势

VS

高效散热、高效供电、节能减排

GPU处理器功耗发展



技术路线

安全可靠

液冷技术

兼容

散热性能

安全可靠

按需供给

供电技术

灵活调配

高效节能

2024开放计算中国峰会

2024开放计算中国峰会

2024开放计算中国峰会

2024开放计算中国峰会

2024开放计算中国峰会

2024开放计算中国峰会

**THANKS**

2024开放计算中国峰会

2024开放计算中国峰会

2024开放计算中国峰会

2024开放计算中国峰会

2024开放计算中国峰会

2024开放计算中国峰会