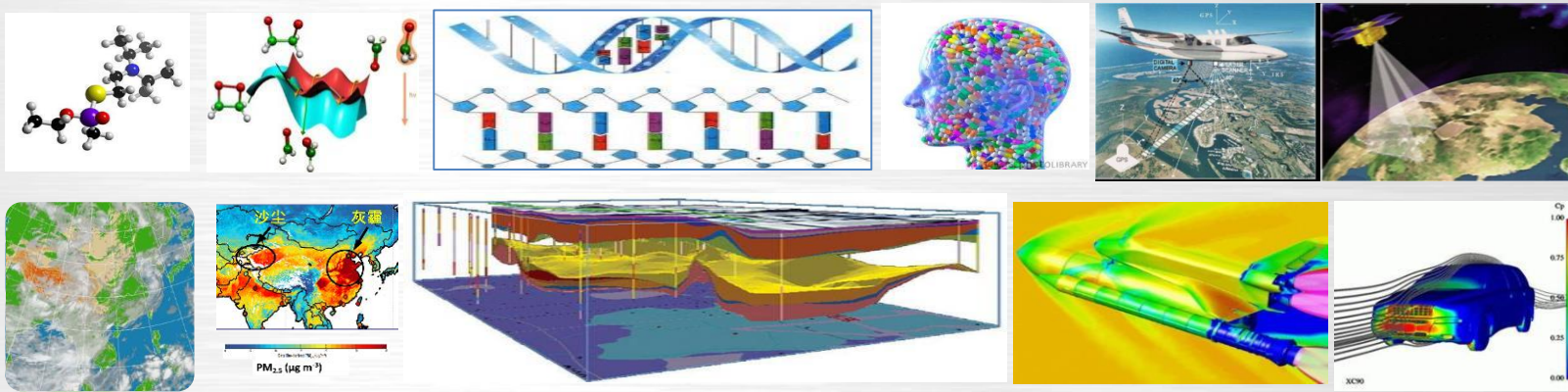


曙光高性能计算并行存储系统技术及应用探究

曙光信息产业股份有限公司

携手成就梦想

广泛的高性能计算应用



应用领域多样，群体广泛

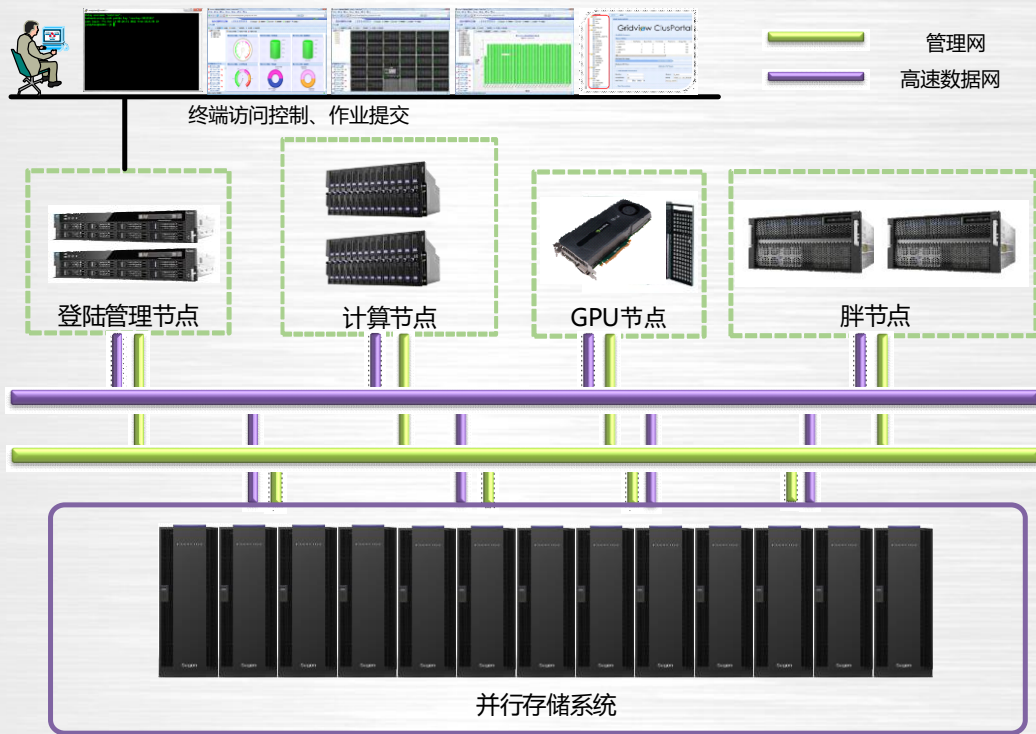
气象环保：数值预报；国家及各省市气象、环保、海洋、地震、环境监测站等

石油地震：地震资料解释；中石油/中海油/中石化及各类油田

生物信息：基因测序&拼接&比对、冷冻电镜等；科研单位、医疗机构等

物理化学、CAE/CAD.....

高性能计算应用对存储系统需求

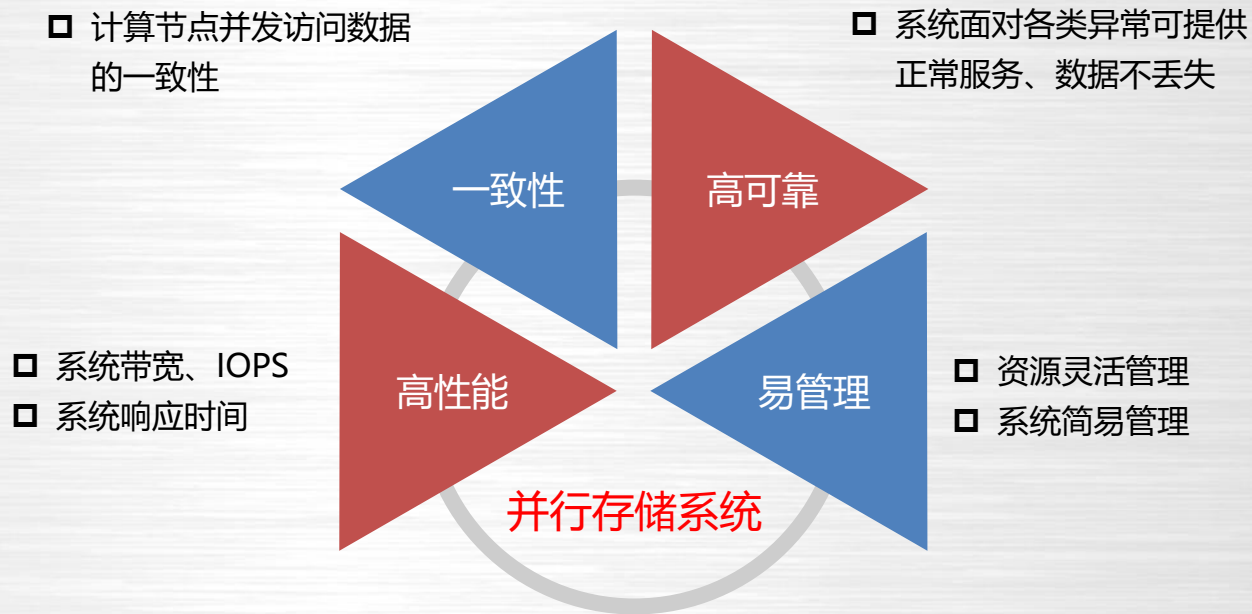


通用应用特点

大量计算节点并发访问

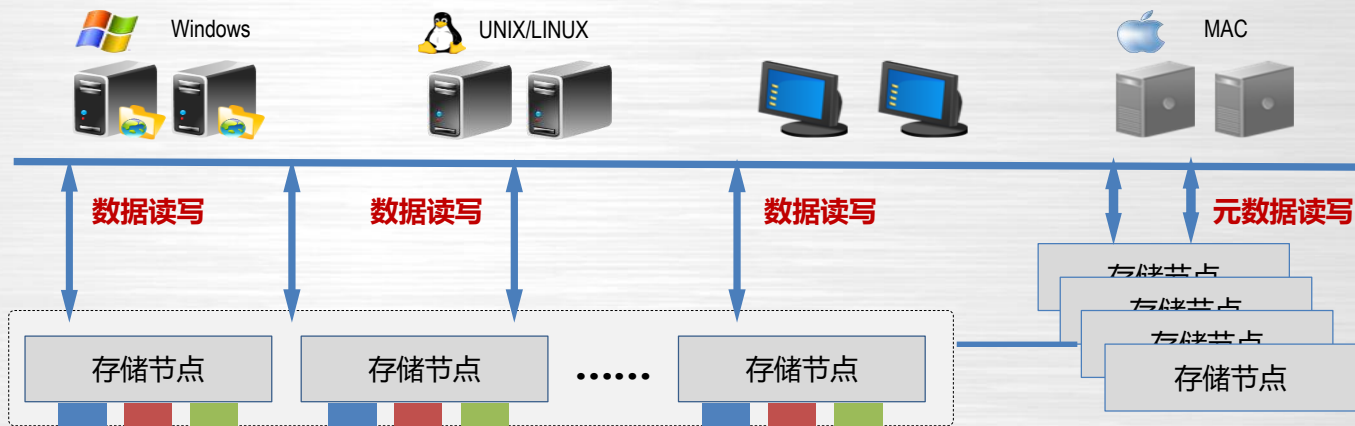
存储关键需求

共享，单一命名空间
极高性能
数据强一致性
稳定可靠



曙光并行存储系统技术实践

先进集群架构，性能优异



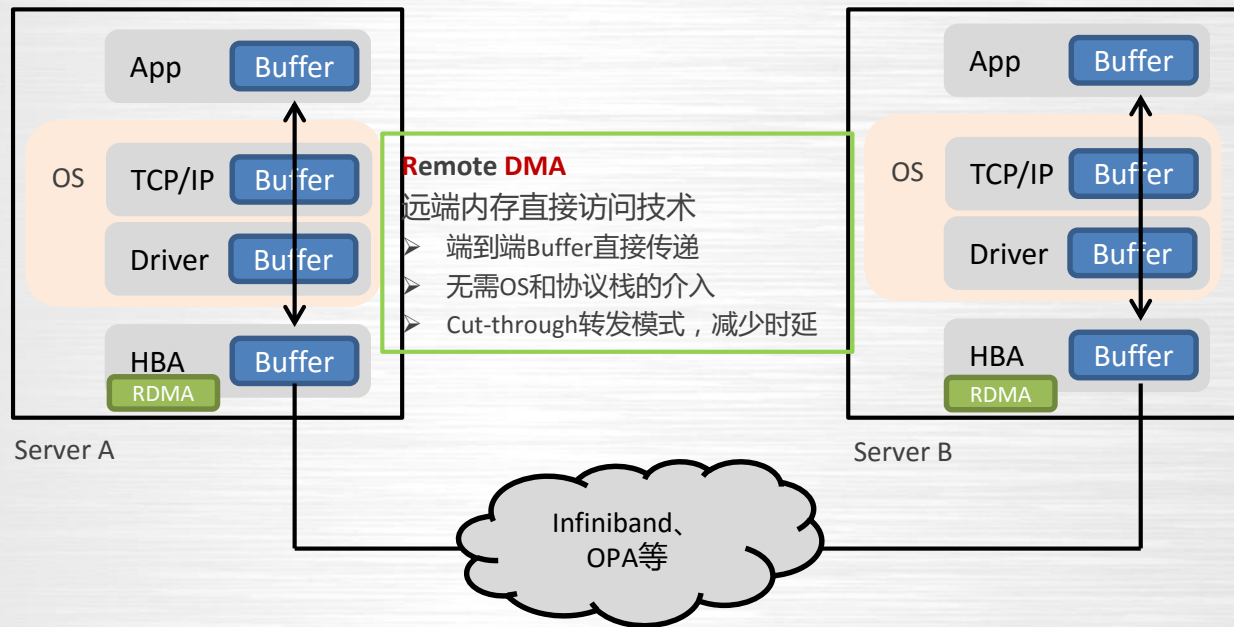
① 全对称分布式架构，满足高IO需求

③ 数据条带化分布，单流带宽性能高

② 多索引集群，元数据并发处理能力强

④ Scale-out横向扩展，聚合带宽线性增加

POSIX+RDMA，性能再提升



POSIX协议

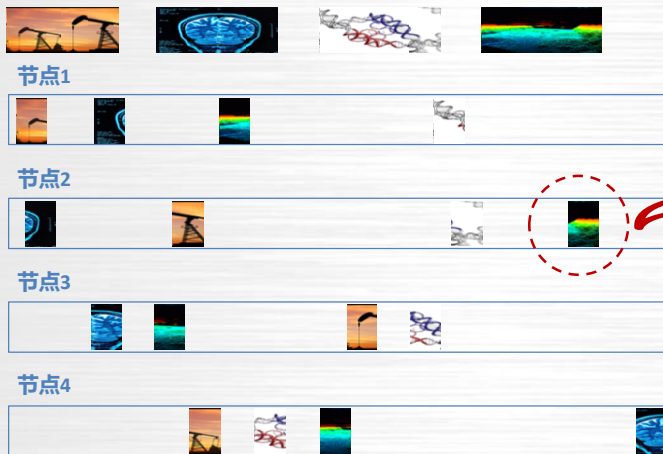
内核态
深度优化
效率更优

RDMA

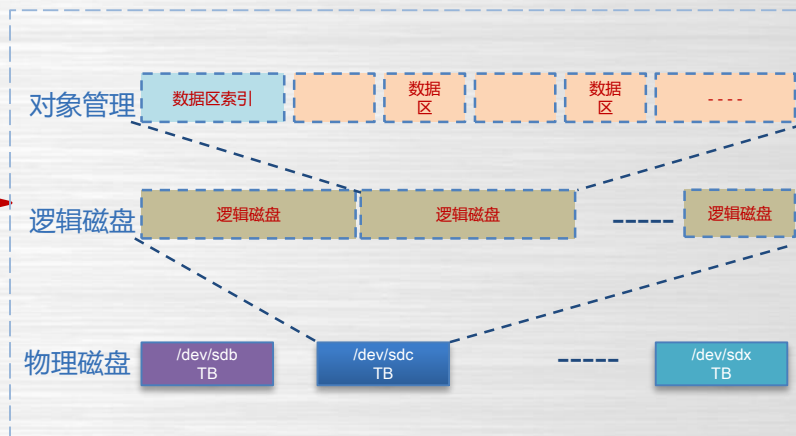
带宽提升
延时降低

自研本地磁盘管理，数据高效存储

文件切成对象，分布到各节点的不同磁盘



对象在磁盘上的存储，由自主研发的OBS模块负责



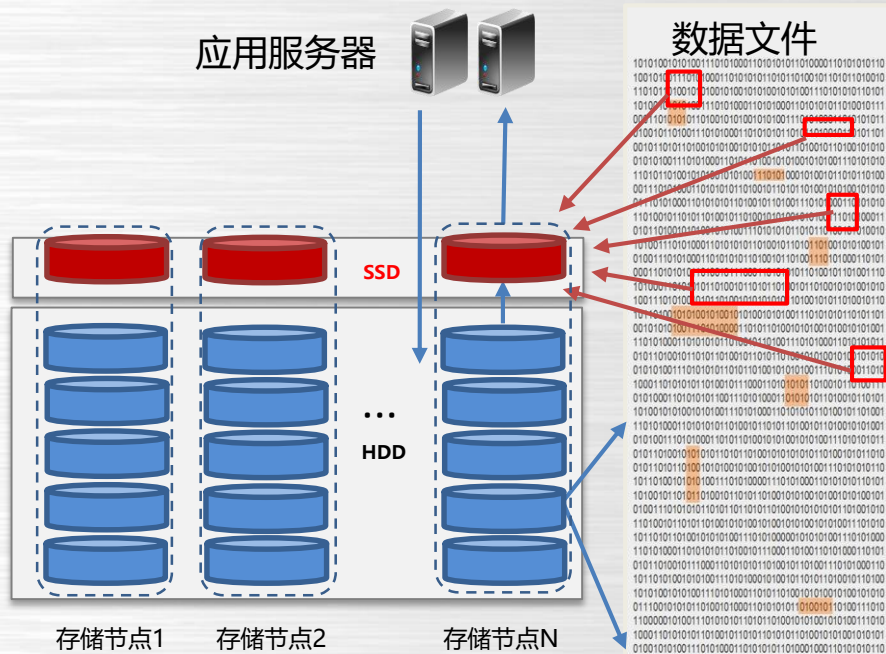
更灵活

更高效

更智能

SSD Cache，提升小文件随机性能

- ❑ 访问频度高的随机数据存入到SSD中，再次读取时直接访问SSD
- ❑ 智能识别连续/随机IO，自动过滤连续数据访问，最大发挥SSD作用
- ❑ 随机、小块数据、读密集型场景



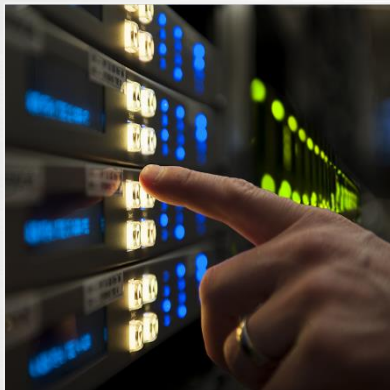
POSIX+强一致性分布式锁，数据访问安全保障

资源共用，即存在争抢的可能，
制定规则，保证秩序

超算系统中，多计算节点并发访问资源，**分布式锁技术**保证
数据的**强一致性**



多维度冗余设计，稳定可靠



部件

电源、风扇、硬盘等
最多允许任意**4**硬盘故障



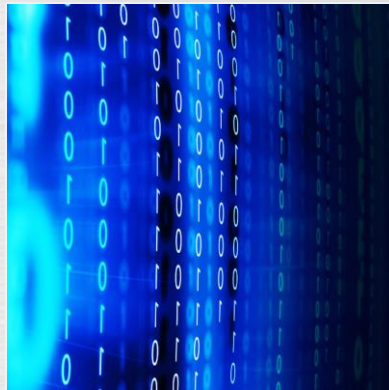
网络

网卡、交换机等
冗余拓扑



节点

多集群冗余设计
最多允许任意**4**节点故障



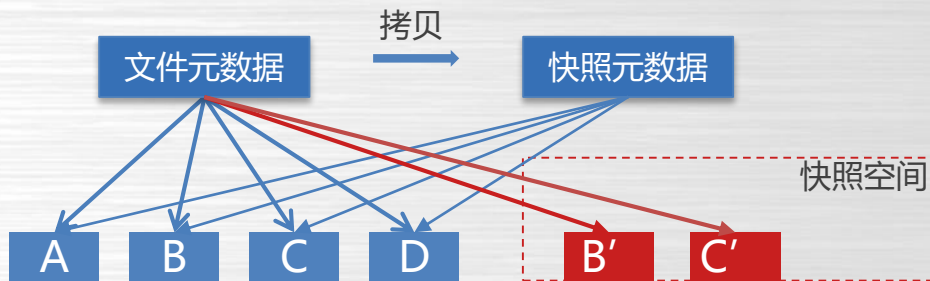
数据

副本/纠删码冗余机制
元数据副本冗余
磁盘分组/节点分区

快照

目录/文件级快照

一次性或策略性快照



掉电内存保护，更高可靠性的保障



- ❑ 正常情况，数据访问DRAM
- ❑ 异常断电，超级电容充当备用电源，数据从DRAM移至NAND
- ❑ 重新上电，数据从NAND移至DRAM

多粒度多层次

- 基于文件系统
- 基于目录
- 基于用户组
- 基于用户

多种配额类型

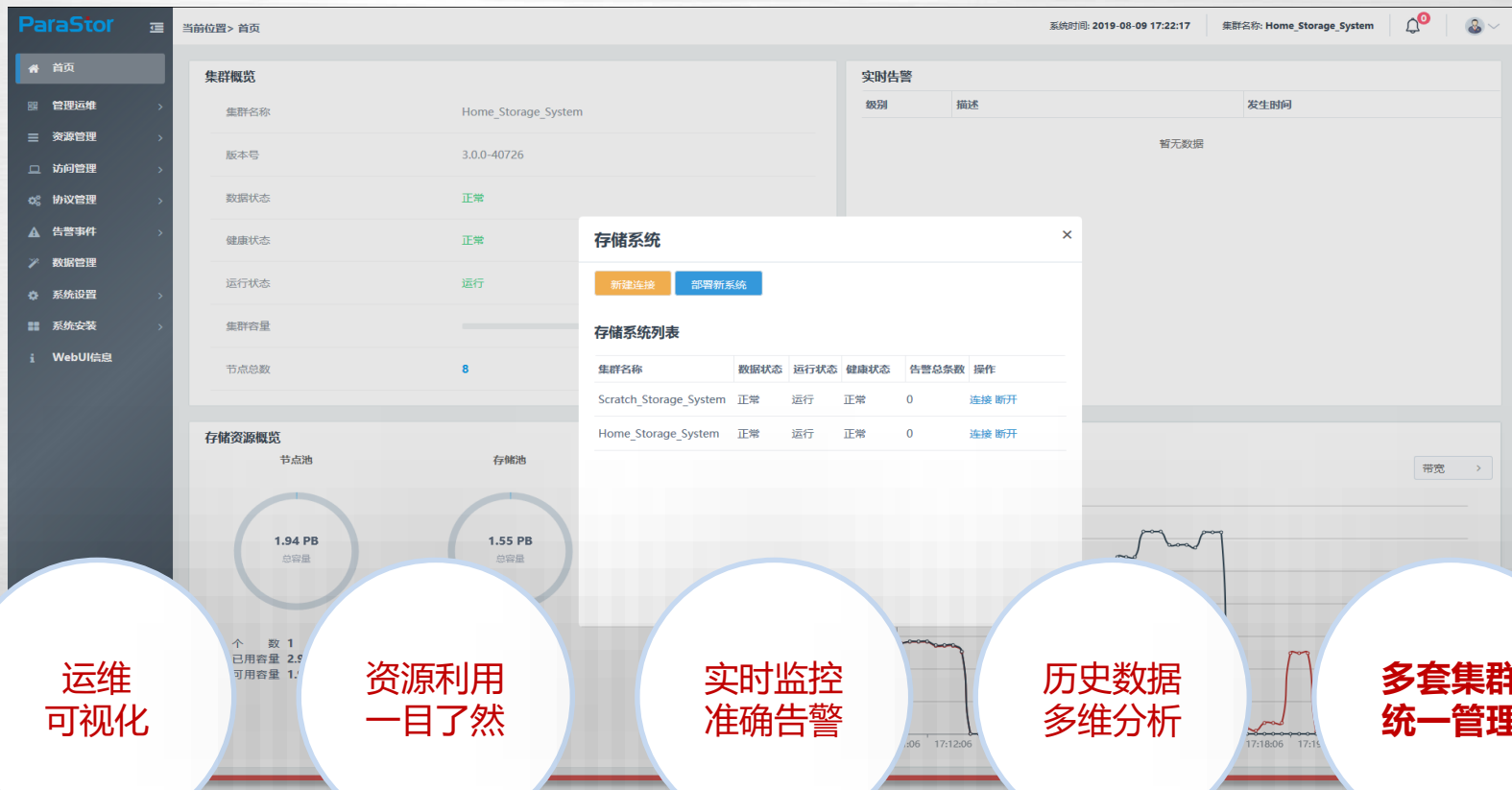
- 统计配额：仅监控，不限制
- 限制配额
 - 文件数/容量
 - 软阈值/硬阈值

写入精准控制

- 软阈值：仅告警
- 硬阈值：无法写入

简易运维，定制专属的管理平台

中科曙光
Sugon



携手成就梦想

谢谢！

IT技术及解决方案的领导者
数据中国百城百行的发起者
中科院产业化联盟的先行者
安全可控信息系统的践行者

Sugon

携手成就梦想