上海科学数据分布式存储软件采购项目

任务需求说明

（一） 任务目的

当前科学数据呈现出跨学科覆盖广、数据规模大、类型结构复杂的特点，涵盖生命科学、材料科学、人工智能、集成电路等多个前沿领域，包含科研过程中产生的原始数据、实验过程数据、计算结果数据等多类数据，数据规模已达 PB 级且持续快速增长，传统存储架构在应对此类非结构化数据的高并发读写、低延迟访问及多协议兼容需求时，存在性能不足、扩展性有限等突出问题，难以支撑科学数据的实时接入、融合处理与高效存取。本项目通过采购一套高性能分布式存储软件，构建面向多学科领域科学数据的统一分布式存储底座，解决当前科学数据汇交与使用场景下的存储性能瓶颈问题，支撑多元科学数据的全生命周期管理，全面赋能上海科学数据管理与共享服务平台建设。

（二） 需求描述

1、服务器整合需求

基于平台中心现有服务器，搭建分布式存储软件，现有服务器主要品牌有新华三、浪潮、联想、青云等，各品牌服务器配置也各不相同，需对现有服务器资源进行评估，结合分布式存储特性选择适合的服务器资源进行整合，调整 CPU、硬盘（NVME、SSD、HDD）、内存、网卡等配置，以满足分布式存储软件部署要求，同时针对现有的 14TB、12TB、8TB、6TB 硬盘资源进行组装规划，最终实现 10PB 可用存储空间。

因分布式存储软件部署要求需额外采购的配套配件(如 NVMe 线缆、转换架、硬盘托架、阵列卡、硬盘背板等），由乙方自行负责采购，相关费用已包含在本次项目的总体预算内。

2、分布式存储软件产品需求

本项目核心产品为分布式存储软件。

技术参数	指标	详细要求
基本要求	知识产权	国内知名厂商，非 OEM 产品，非联合产品，不接受使用开源产品，完全独立自主研发，能够提供分布式存储自主知识产权证明，包括但不限于软件著作权，发明专利等。
授权要求	授权容量	本次分布式存储软件需提供冗余配置后可用存储容量≥10PB 的授权许可。
系统架构	系统架构	采用全对称分布式架构，元数据和数据都融合部署在存储节点上，不需要使用独立管理节点（元数据节点或索引节点）；元数据、数据均采用集群方式部署，任何一个节点出现故障，不影响数据的正常访问。
	统一存储	使用三个及以上存储节点组建一个存储集群，同一系统中同时提供文件、块、对象三种存储服务，统一管理，资源灵活分配。
	缓存配比无限制	文件/对象存储无缓存盘与数据盘的配比限制，扩容时无需配置 SSD 做缓存盘，可单独扩容数据盘。
对象存储	对象存储接口	兼容标准对象存储协议接口。
	桶 QoS	配置对象存储 Qos 功能，可以设置不同用户访问某个 bucket 的带宽/请求数上限，从而防止边缘业务的过多资源占用。
	对象数据压缩	为降低数据长期保存成本，对象存储应提供数据压缩能力，支持以桶为单位配置数据压缩策略，可选择节省容量优先和性能优先两种策略，并支持查看压缩的数据量。
文件存储	文件存储接口	同时提供 NFS、CIFS、FTP、HDFS 四种存储接口，无须在应用层安装插件，减少业务系统改造成本，满足不同应用系统对存储接口的要求。
	文件数据保护	**▲配置对任意目录层级打快照的功能，并支持对目录以及该目录下的子目录同时打快照，定时快照间隔最短支持 15 分钟，支持快照数≥20000 个。支持快照重命名功能，支持快照点任意文件数据恢复。（需提供产品功能截图，并加盖公章）**
	访问鉴权	**▲支持在存储控制台配置 Windows NT-ACL 与 Posix-ACL 权限，能有效减少网络时延进而大幅提高文件的权限管理效率。（需提供产品功能截图，并加盖公章）**
	文件压缩	为降低数据长期保存成本，文件存储应提供数据缩减的能力，支持无损压缩，支持以文件目录为单位配置数据压缩策略，可选择节省容量优先和性能优先两种策略，并支持查看压缩的数据量。
	智能数据预读	**▲支持对数据预读功能，将热点数据预读至缓存层，从而提升读性能。（提供带有 CMA、CNAS 标识的检测报告证明，至少包含报告首页，对应功能测试页和报告尾页）**

	文件数量	**▲一套存储支持 100 亿+的文件规模，且在海量数据的条件下保持性能稳定，衰减不超过 10%。（提供带有 CMA、CNAS 标识的检测报告证明，至少包含报告首页，对应功能测试页和报告尾页）**
	目录均衡打散	**▲为使得所投产品在高并发场景下能发挥出更高性能，支持将某一目录下所有的元数据访问请求均衡打散到集群内多个节点的多个元数据服务 MDS 中。（提供带有 CMA、CNAS 标识的检测报告证明，至少包含报告首页，对应功能测试页和报告尾页）**
	NFS 多路径	**▲为解决 NFS 单链路性能瓶颈问题，需支持 NFS 多路径功能，在主机侧和存储侧实现多条链路访问通道，以大幅提升单客户端场景下的访问性能。（需提供产品功能截图，并加盖公章）**
块存储	访问鉴权	支持配置 Chap 认证，支持单向认证、双向认证和不认证多种认证方式。
	存储策略	存储策略，支持以 LUN 为粒度配置副本数、分层 QOS、条带数等存储策略，以实现在性能、成本，可靠性等指标上的平衡兼顾。（需提供产品功能截图，并加盖厂商公章）
可靠性	性能稳定	支持在集群满载（容量使用率达 90%）的条件下仍能保持性能稳定，衰减不超过 5%。
	防勒索	**▲基于业务安全考虑，为避免产生额外的经济损失，存储产品应具备勒索病毒的防护能力，内置免费病毒防护引擎，支持中毒前的勒索病毒防护、中毒时（或检测为疑似勒索病毒时）自动打快照、中毒后进行文件的安全恢复，能有效抵御勒索病毒。（需提供产品功能截图，并加盖公章）**
		▲存储产品内部具备勒索文件扩展名数据库，可实现实时监测勒索文件并提供直接阻止可疑文件创建，可实时监控分析业务系统端的数据行为，智能识别威胁事件，并创建勒索防护快照保护数据，同时及时产生告警通知管理员进行处置，如果有病毒被发现，存储系统可以通过文件审计日志找到受影响的文件路径，借助快照恢复受损数据。（需提供产品功能截图，并加盖公章）
	文件远程复制	**▲为实现存储系统的容灾，保障存储系统的高可用性，此次采购的分布式存储设备需原生支持文件存储的远程复制能力，支持目录粒度的数据保护，当发生灾难时，支特主备集群之间的灾难切换。（需提供产品功能截图，并加盖公章）**
数据采集平台	采集与标签功能	为降低数据管理成本及长期保存成本，所投产品应提供免费的数据采集工具，支持有代理和无代理两种采集方式，同时提供数据上传/下载工具，支持从存储平台上将数据下载至 PC 端。
		支持采集数据自动打标签，可自定义标签规则，产生标签作为条件以便检索

性能	块存储性能	块存储混闪配置（缓存层采用 SSD ，容量层采用 HDD）在 2ms 稳定时延下，单节点可提供至少 15 万 IOPS，三节点集群满足 50 万以上 IOPS 处理能力。
	文件存储吞吐性能	文件存储要求三节点集群读带宽不低于 10GB/s，写带宽不低于 7GB/s。
	文件存储 IO 性能	**▲文件存储要求三节点集群小文件 4K 读 IOPS 不低于 13 万，写 IOPS 不低于 9 万。（提供带有 CMA、CNAS 标识的检测报告证明，至少包含报告首页，对应功能测试页和报告尾页）**
	对象存储性能	对象存储要求三节点集群下载(读)带宽不低于 12GB/s，上传(写)带宽不低于 6GB/s。
运维管理	存储资源状态监控	支持查看文件系统客户端（CIFS/NFS/FTP）的连接个数，以及每个连接的具体信息（OPS 带宽、时延、元数据 OPS、元数据时延），便于管理员查看客户端的接入情况，识别访问热点及快速定位业务访问故障。
	集群健康巡检	支持集群资源环境一键检测，对硬件健康、平台底层服务的运行状态和配置，进行多个维度进行检查，提供快速定位问题功能，确保系统最佳状态。

3、资质要求

投标方应具备较强的综合能力，具有以下认证资质的优先考虑：

✧ CCRC 信息系统安全运维服务资质；

✧ CCRC 信息系统安全集成服务资质；

✧ CCRC 信息安全风险评估资质。

4、服务要求

提供原厂项目管理、现场安装、配置及实施服务，提供三年质保和三年 7*24 小时现场保修服务,包括分布式存储软件保修电话支持、现场支持、软件升级，提供针对本次项目的原厂商授权书和保修服务承诺函。

分布式存储软硬件出现故障乙方 1 小时内响应，4 小时内解决故障，乙方对发生的故障进行总结，并提出今后的维护建议。

在节假日、停电、设备变更等特殊情况时，乙方派工程师到现场提供分布式存储软件护航服务。

5、对项目实施人员要求

分布式存储软件实施人员必须具备以下条件：

- ✧ 具备分布式存储软件独立部署能力；
- ✧ 精通 Linux 系统，能熟练编写 Shell 脚本；熟悉 mysql 数据库；
- ✧ 对网络设备及 TCP/IP 协议，有建设经验和配置能力；
- ✧ 实施团队成员具有分布式存储原厂认证工程师证书、系统集成项目管理工程师认证证书、网络工程师认证证书、信息安全保障人员安全集成(SI)认证证书、信息安全保障人员安全运维(OM)认证证书、注册信息安全管理人员（CISO）认证证书、注册信息安全工程师（CISE）认证证书等。

6、服务时间

自合同签订生效之日起一个月内完成项目实施上线，正式上线验收后免费维保期为三年。

（三） 需求结果

1、 **★实施完成后，分布式存储达到 10PB 可用容量；**

供应商须在标书中对该参数进行响应，不符合该指标要求的视为无效投标。

2、 分布式存储软件具备元数据多活服务与目录分片机制，可实现百亿级小文件的高并发访问。适配高速网络的并行文件系统，达到微秒级数据访问延迟与吞吐性能线性扩展；

3、 整合文件防勒索引擎与故障智能预测能力，赋能分布式存储软件从被动硬件故障保护，升级为主动的数据威胁安全防护，通过文件级、节点级、方案级、系统级的多级可靠机制达到 99.999%的高可用保障。