

# 技术需求

## 一、GPU 服务器（非通用服务器） 1 套

★1. 4U 双路机架式服务器，含 GPU10 块；（供应商须提供证明材料）

▲2.CPU：配置 2 颗 X86 架构处理器，单颗处理器核心 $\geq 48$ ，基本主频 $\geq 2.7\text{GHz}$ ；（供应商需提供证明材料）

▲3.内存：配置 $\geq 16$  条 64GB DDR5 RECC 内存；（供应商需提供证明材料）

★4.GPU: 10\*GPU 卡,每块 GPU 卡要求 CUDA 核心数 $\geq 24064$ ，显存 $\geq 96\text{GB}$ ，FP32 $\geq 120\text{TFLOPS}$ ；支持 FP16、BF16、FP8、FP4 等，FP8 计算能力达到 1P 以上；（供应商须提供证明材料）

★5. 硬盘：配置 $\geq 4$  块 3.84T SSD；（供应商须提供证明材料）

★6.网卡：配置 $\geq 1$  块电口千兆网卡；（供应商须提供证明材料）

7.RAID：1 块 SAS RAID 卡，1G 缓存。

## 二、其他附加功能

### 1.基础资源纳管

★1.1. 以插件的形式支持对不同厂商、异构的 GPU 资源进行标准化，屏蔽异构卡、驱动及接口等的差异；（供应商须提供证明材料）

1.2. 算力节点资源调度指令需要在可信环境进行，调度指令不能直接访问宿主机指令，支持函数服务调用（FaaS）和长任务（task）等模式调用；

▲1.3. 支持高速块存储、对象存储以及网络存储，支持租户网络隔离，支撑租户不同应用场景的需求。（供应商需提供证明材料）

### 2.分布式资源调度与任务编排

2.1. 支持元语化资源调度与编排，将分布式协同任务抽象为微资源、微操作以及微服务等基础三类元调度指令，支持异步分布行为树，简化环境部署和保障流程定义；

▲2.2. 支持多容器等虚拟化任务的编排和保障，支持存储、网络、GPU 等算力资源、模型、镜像、数据等协同保障；（供应商需提供证明材料）

2.3. 支持分布协同环境的自主诊断与故障自动恢复。

### 3.分布协同任务管理

★3.1.分布协同任务由管理容器、若干业务容器、网络以及关联资源构成；（供应商须提供证明材料）

3.2 管理员能够管理和监控内部应用服务、服务代理以及跨租户的 AI 服务与应用的协同；

▲3.3. 管理员能够管控单个容器的起停、恢复等，能够动态管理容器网路，能够主动隔离互联网访问。（供应商需提供证明材料）

### 4.GPU 资源预留与超分配

4.1.支持 GPU 的物理切分软切分

▲4.2.支持 GPU 超分配（供应商需提供证明材料）

▲4.3.支持按租户的 GPU 资源复用（供应商需提供证明材料）

### 5.API 中心

▲通过统一 API Key 动态绑定、调度与调用不同大模型服务，屏蔽模型复杂性和实现权限隔离，使大模型调用简单、安全、高效。（供应商需提供证明材料）

### 6.应用中心

提供一键开通的优质 AI 应用，用户可按需选择共享资源或独占资源，满足不同的使用需求。

### 7.模型推理

7.1.支持 Deepseek 671B 全参数部署，支持 128k 的长度，满足 10 并发以上需求；

7.2.支持多个大模型切换。

说明：证明材料包括但不限于制造商公开发布的印刷资料（产品彩页、产品说明书、产品白皮书、官网截图）、第三方检测机构出具的检测报告、在有效期内的证书扫描件等。

## 二、商务要求

### 1、验收条款（包括验收的指标、验收过程、特殊要求）

- 1) 投标人完成本项目应达到的质量标准应符合国家、地方及相关政府管理部门和行业与本项目有关的各项技术标准、规范要求，并满足采购人实际需求，标准、规范等不一致的，以要求高（严格）的为准。
- 2) 本项目验收将由采购人组织进行或委托第三方进行，设备到货后，由中标方，使用方共同开箱验货，卖方保证货号的型号，规格，数量与合同相符。
- 3) 在系统整体调试完成后，买方认为合格后，签订系统安装验收报告，如连续 3 次验收未获通过，采购人有权解除合同并按照合同约定的违约。
- 4) 卖方应在设备到货前 1 个月，对仪器实验室场地条件，如工作台，水，电，气等配套设施向买方提出建议并出具场地准备书。

### 2、培训方式、要求、计划、大纲等

- 1) 中标方应根据项目实施的计划、进度和使用方的合理要求，及时安排对使用方的相关人员进行培训。培训目标为使受训者能够独立、熟练地完成操作，实现本次采购信息系统的目标和功能。
- 2) 在服务有效期内，提供相关的技术文档和操作手册，帮助用户了解和使用网络设备。系统部署完成后，提供至少 1 次不少于 2 天的培训，培训内容和人数由双方协商确定。

### 3、交货期及售后服务

设备交货期合同签订后 60 天内。质保不少于 3 年。质保期内，乙方提供 7x24 小时故障响应，30 分钟内响应故障报修，4 小时内到达故障现场，8 小时内解决故障或提出解决方案。提供投标产品生产厂家的授权书或投标人为生产厂家的、提供投标产品生产厂家的针对本项目出具的售后服务承诺函的优先考虑。

### 4、实施调试

提供原厂设备安装部署服务，包括设备上架，系统部署，驱动和固件安装升级。供应商需承诺所有本项目物流运输人员、安装施工人员、技术支持人员以及售后支持人员，在项目交付和售后现场期间遵守采购人学校各项相关管理制度要求。