**放射科高性能并行计算服务器采购项目需求**

# 项目名称

项目名称：放射科高性能并行计算服务器采购项目

# 采购清单

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | 项目名称 | 数量 | 单位 |
| 1 | GPU服务器 | 3 | 台 |
| 2 | 管理服务器 | 2 | 台 |
| 3 | IB线缆 | 5 | 根 |
| 4 | 高性能计算管理平台 | 1 | 套 |
| 5 | 搬迁服务 | 1 | 项 |

# 详细配置参数

3.1、GPU服务器

|  |  |  |
| --- | --- | --- |
| 序号 | 指标项 | 技术规格要求 |
| 1 | 品牌 | 国内外知名品牌。 |
| 2 | 服务器高度 | ≥标准4U机架式服务器，标配原厂导轨。 |
| 3 | 芯片组 | 不低于Intel C740 系列芯片组芯片组，支持Intel至强可扩展处理器家族铜牌、银牌、金牌及铂金处理器产品。 |
| 4 | 处理器型号 | 配置≥两颗Intel至强可扩展处理器，单个Intel至强处理器主频≥2.1GHz，≥32核64线程，270W，L3缓存≥60MB。 |
| 5 | 内存实配规格 | ≥16\*64GB DDR4 4800MHz(4800MT/s) RDIMM。 |
| 6 | 内存可扩展数量 | 支持≥32个DDR4 RDIMM或LRDIMM内存插槽，速率最高支持4800MT/s，支持六通道配置。 |
| 7 | 存储-硬盘 | ≥8\*3.84TB 2.5英寸 SATA 6Gb/s企业级固态硬盘，配置相同数量的硬盘托架。 |
| 8 | 存储-阵列控制器 | ≥1张12Gb PCIe 配2GB闪存 阵列卡，支持0/1/5/10/50 RAID级别，可选RAID 6/60, 最大支持4GB闪存，支持缓存数据保护，配置掉电保护模块。 |
| 9 | GPU | 配置≥8\*48GB GDDR6 w/ECC GPU，单卡FP32值≥91.6 TFLOPS、RT Core值≥212.0 TFLOPS、TF32 Tensor Core值≥366.0 TFLOPS、TGP≤350W。 |
| 10 | 网卡 | 配置≥2个10Gb SFP+光模块，≥4个 1GE接口。 |
| 11 | IB卡 | ≥1个2口100Gb Infiniband HDR/Ethernet适配卡(支持QSFP56光模块)。 |
| 12 | 电源产品认证 | ≥4个2400W 100-240VAC或240HVDC 80+白金牌热插拔冗余电源。 |
| 13 | 冷却系统 | ≥8个热插拔冗余风扇，支持N+1冗余。 |
| 14 | PCI I/O扩展 | ≥14个PCIe插槽，包括8个双宽PCIe5.0 x16插槽，4个单宽PCIe5.0 x16插槽，一个内置阵列卡专用插槽，一个OCP 3.0网卡专用插槽。 |
| 15 | 可管理性 | 配置≥1Gb独立的远程管理控制端口；配置虚拟KVM功能,支持对服务器与操作系统无关且完全控制的远程功能，包括远程的开机、关机、重启、更新Firmware、虚拟光驱、虚拟文件夹等操作，提供服务器健康日记、服务器控制台录屏/回放功能，提供电源监控。 |
| 16 | 带外升级 | 为简化服务器的运维管理，服务器支持不依赖操作系统可带外一次升级多个部件的固件（如网卡部件、存储卡部件等），无需多次升级，要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 17 | 联合管理 | 管理软件联合管理支持对同机房、网络互通的服务器直接管理，简化了小规模服务器的统一管理，要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 18 | 温度监测 | 支持展示各组件温度传感器的分布图，可让用户直观感知服务器整体温感状态，要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 19 | 离线固件更新 | 支持离线固件更新功能，批量更新option的固件；要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 20 | 服务器添加 | 服务器统一管理软件支持通过设备发现/手动添加/批量导入三种方式添加服务器，要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 21 | 服务器配置模板 | 支持通过服务器配置模板批量应用进行BIOS设置、RAID设置、操作系统安装，要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 22 | 密码认证 | 支持对于登录服务器管理软件的用户执行双因素认证(如静态密码和随机动态密码结合的方式)。要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 23 | 诊断日志 | 支持通过服务U盘下载服务器设备故障诊断日志。要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 24 | 应急诊断 | 当服务器因可更换硬件故障导致启动进程挂死在POST阶段时，可以使用应急诊断功能，开启后即可自动诊断，并输出时间戳及诊断结果等信息，要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 25 | 售后服务 | 提供原厂商首次安装服务和3年7\*24小时维保服务。 |

3.2、管理服务器

|  |  |  |
| --- | --- | --- |
| 序号 | 指标项 | 技术规格要求 |
| 1 | 品牌 | 国内外知名品牌。 |
| 2 | 服务器高度 | ≥标准2U机架式服务器，标配原厂导轨。 |
| 3 | 芯片组 | 不低于Intel C621A 系列芯片组芯片组，支持Intel至强可扩展处理器家族铜牌、银牌、金牌及铂金处理器产品。 |
| 4 | 处理器型号 | 配置≥两颗Intel至强可扩展处理器，单个Intel至强处理器主频≥2.3GHz，≥20核40线程，150W，L3缓存≥30MB。 |
| 5 | 内存实配规格 | ≥4\*32GB DDR4 3200MHz(3200MT/s) RDIMM。 |
| 6 | 内存可扩展数量 | 支持≥32个DDR4 RDIMM或LRDIMM内存插槽，速率最高支持3200MT/s，支持六通道配置。 |
| 7 | 存储-硬盘 | ≥2\*960GB 2.5英寸 SATA 6Gb/s 企业级固态硬盘，配置相同数量的硬盘托架。 |
| 8 | 存储-阵列控制器 | ≥1张12Gb PCIe 配2GB闪存 阵列卡，支持0/1/5/10/50 RAID级别，可选RAID 6/60, 最大支持4GB闪存，支持缓存数据保护，配置掉电保护模块。 |
| 9 | PCI I/O插槽 | 最多提供≥11个PCIE4.0 x16插槽。 |
| 10 | 网卡 | 配置≥2个10Gb SFP+光模块，≥4个 1GE接口。 |
| 11 | IB卡 | ≥1个2口100Gb Infiniband HDR/Ethernet适配卡(支持QSFP56光模块)。 |
| 12 | 电源产品认证 | ≥2个800W 100-240VAC或240HVDC 80+白金牌热插拔冗余电源。 |
| 13 | 冷却系统 | ≥4个热插拔冗余风扇，支持N+1冗余。 |
| 14 | 可管理性 | 配置≥1Gb独立的远程管理控制端口；配置虚拟KVM功能,支持对服务器与操作系统无关且完全控制的远程功能，包括远程的开机、关机、重启、更新Firmware、虚拟光驱、虚拟文件夹等操作，提供服务器健康日记、服务器控制台录屏/回放功能，提供电源监控。 |
| 15 | 带外升级 | 为简化服务器的运维管理，服务器支持不依赖操作系统可带外一次升级多个部件的固件（如网卡部件、存储卡部件等），无需多次升级，要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 16 | 联合管理 | 管理软件联合管理支持对同机房、网络互通的服务器直接管理，简化了小规模服务器的统一管理，要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 17 | 温度监测 | 支持展示各组件温度传感器的分布图，可让用户直观感知服务器整体温感状态，要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 18 | 离线固件更新 | 支持离线固件更新功能，批量更新option的固件；要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 19 | 服务器添加 | 服务器统一管理软件支持通过设备发现/手动添加/批量导入三种方式添加服务器，要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 20 | 服务器配置模板 | 支持通过服务器配置模板批量应用进行BIOS设置、RAID设置、操作系统安装，要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 21 | 密码认证 | 支持对于登录服务器管理软件的用户执行双因素认证(如静态密码和随机动态密码结合的方式)。要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 22 | 诊断日志 | 支持通过服务U盘下载服务器设备故障诊断日志。要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 23 | 应急诊断 | 当服务器因可更换硬件故障导致启动进程挂死在POST阶段时，可以使用应急诊断功能，开启后即可自动诊断，并输出时间戳及诊断结果等信息，要求提供CNAS认可的检测机构出具的相关证明材料。 |
| 24 | 售后服务 | 提供原厂商首次安装服务和3年7\*24小时维保服务。 |

3.3、IB线缆

| 序号 | 指标项 | 技术规格要求 |
| --- | --- | --- |
| 1 | IB线缆 | 兼容现有36口EDR100交换机，配置3M 100Gb IB线缆。 |

3.4、高性能计算管理平台

|  |  |  |
| --- | --- | --- |
| 序号 | 指标项 | 具体要求 |
| 1 | 科研计算平台总体要求 | 提供拥有软件著作权的商业版本管理软件。 |
| 兼容现有服务器并授权，提供本次项目所有服务器授权。 |
| 能根据用户的需求或使用习惯进行定制化开发。 |
| 2 | 科研计算平台架构和系统支持 | 统一平台，统一资源池，统一资源调度。支持调度多种任务类型，包括单节点容器、虚拟机、多节点HPC集群、AI训练集群等。 |
| 同时支持两类资源虚拟化技术，包括虚拟机虚拟化和容器虚拟化技术。 |
| 支持银河麒麟高级服务器操作系统V8及以上、Windows server 2012及以上，提供操作系统的部署模板。 |
| 支持Intel、AMD、ARM等架构服务器，支持NVIDIA GPU、FPGA、atlas等异构服务器，支持Infiniband和roce RDMA网络，支持至少3个硬件厂家存储设备。 |
| 3 | 科研计算平台系统管理要求 | 提供不少于三种角色，包括系统管理员、群组管理员和普通用户；支持群组管理员和普通用户创建群组并邀请其他用户加入。 |
| 支持对用户的资源使用量进行限额，便于管理员合理分配集群资源，可以修改每个用户或用户组的最大可用CPU核数、GPU个数、内存使用量、存储空间、实例数等。 |
| 支持管理员对应用运行时所需环境的配置修改，参数包含所属资源池、CPU、内存、GPU、块存储等信息。 |
| 支持设置用户虚拟机实例、容器实例的使用周期，到期自动回收资源并提前自动通知用户。支持到期前手动延长使用周期。 |
| 支持虚拟机实例、容器实例回收站功能。实例释放后自动进入回收站中，可通过回收站对实例进行恢复或彻底删除。 |
| 4 | 科研计算平台数据管理要求 | 平台提供用户目录，项目目录，共享群组目录等数据管理的功能，在项目组中，用户可以通过共享目录进行数据共享，也可以在项目组中有自己私有的数据目录。 |
| 平台提供通过浏览器上传计算配置文件和下载计算结果以及断点续传等功能，支持用户本地挂载平台数据目录，实现用户本地、平台和实例环境的数据同步，支持用户本地是Windows操作系统的数据挂载。 |
| 5 | 科研计算平台场景功能要求 | 平台支持多套环境资源之间灵活调度，不需要重启物理机器，即可实现HPC集群的资源调度到AI集群中进行使用，支持多套HPC集群共存，如已有的x86 HPC集群，支持普通用户在平台上创建自己项目组内的私有HPC集群，HPC集群的性能多节点并行Linpack值与物理机的比相差1%以内。 |
| 支持以透传的方式实现对物理GPU的访问，并设置可使用的GPU数，在环境内，用户有root权限，可自行安装和配置环境。 |
| 支持单机单卡和单机多卡的单容器环境，支持多机多卡容器化分布式GPU训练集群。支持GPU容器虚拟化技术，1块物理GPU卡可同时被多个GPU容器实例共享计算和显存资源，每个容器实例最低的资源共享粒度为1块物理GPU的十分之一。 |
| 支持通过网页界面生成和恢复虚拟机快照；支持Windows系统的远程桌面可视化操作，以及对系统实现硬盘扩容等功能。 |
| 支持通过软件仓库一键创建完整独立的虚拟化环境，包括完整的Linux操作系统环境、Windows桌面环境等。 |
| 6 | 科研计算平台资源统计功能要求 | 平台提供集群运行情况监控功能，包含集群数量、集群应用软件类型、正在运行的用户等，支持统计应用镜像的使用数量。 |
| 支持分别根据CPU、GPU、内存等资源采用不同的费率标准进行分类计费，支持在系统管理中设置默认权值。 |
| 支持不同的节点、不同的资源池设定不同的费用，可以单独配置每一个硬件资源的费用。 |
| 支持用户充值和透支额度设置，支持项目透支额度设置，支持用户转账自己剩余点数到任意一个自己的项目中，在项目中的其他用户可以使用该项目的点数进行扣费，支持对项目中用户使用点数查看和设定限额。 |
| 7 | 软件环境 | GNU C/C++/F77/F90编译器、Java、Intel C/C++/Fortran编译器；Intel MKL、BLAS、LAPACK、FFTW等数学库；并行编程消息库如Intel MPI、OpenMPI等；支持容器化数据分析工具，支持Python, Anaconda, R等，并支持用户自定义软件包安装。 |
| 8 | 在线运维 | 完成对高性能计算集群远程在线监控、异常告警通知等工作。 |
| 在线运维服务监控采购方高性能集群服务器、网络等相关资源状态信息，如发现异常信息、故障信息，则将此类信息第一时间通知用户方，并采取相关措施。 |
| 客户提供的商业软件的安装及调度集成；  客户所需开源软件的安装、编译及优化；  应用软件作业模板的定制及优化；  应用软件的容器模板制作；  应用软件的虚拟机模板制作。 |
| 平台应用商店内软件的升级和更新；  系统GPU驱动/CUDA版本的持续更新和升级；  平台内部组件的持续升级；  平台新版本功能更新的持续安装。 |
| 平台硬件环境的运维监控；  系统报警、平台故障的实施处理；  平台资源使用情况的监控及使用报告；  调度策略的持续优化和调整；  用户配额和QOS的持续优化；  用户微信群进行实时问题解答和使用帮助。 |
| 9 | 售后服务 | 提供3年7\*24小时原厂运维服务，10分钟响应，4小时内处理完。 |
| 原厂提供软件部署实施服务。 |

3.5、搬迁服务

| 序号 | 指标项 | 技术规格要求 |
| --- | --- | --- |
| 1 | 服务要求 | 原机房（广东省广州市越秀区中山一路东山口职工住宅小区内AI实验室）需搬迁11台4U机架式服务器、3台4U塔式服务器、1台1U IB交换机、1台1U千兆交换机、1台1U万兆交换机、若干配套的服务器电源线和网络线缆等至新机房（广东省广州市越秀区中山二路106号广东省人民医院伟伦楼2楼数据中心）。搬迁实施工作中，由承建商提供服务器电源线、网络线缆、PDU、排插等服务器部署配件。 |
| 对数据中心进行基础规划与网络规划，原机房服务器、交换机、本项目3台GPU服务器、2台管理服务器接入至数据中心的UPS，服务器通电通网正常运行，所有服务器的网络互通，部署高性能计算管理平台，实现新数据中心的服务器管理。 |
| 在搬迁全过程，承建商需与院方进行搬迁范围、进度、质量、风险等沟通管理工作。 |

四、服务要求

(一)从合同标的验收合格之日算起，期限为36个月。在免费维护期内，承建商提供7\*24小时服务热线，4小时内安排合格的技术工程师提供技术支持和指导，以及软件的局部改进完善以及故障情况下的现场问题解决。

(二)整机保修；保修期自验收合格之日起计算。

(三)提供原厂安装和3年原厂保修服务，安装及保修服务由承建商负责，费用包含在本项目报价内。

(四)在维护期结束前，须由供货方和院方进行一次全面检查，任何缺陷必须由供货方负责修复，在修复之后，供货方应将缺陷原因、修复内容、完成修理及恢复正常的时间和日期等报告给院方，形成项目总结报告。

(五)超过维护期的，双方另行协商签订维护合同，信息设备（产品）的维护报价不超过合同信息设备（产品）部分金额的5%。

售后服务：提供原厂家7\*24小时维修服务。

响应时间、方式：4小时内响应到院方报修处，供货方需提供备件先行服务。

五、培训要求

项目培训应贯串于整个项目的实施过程中，包括在从项目准备、分析、设计到项目实施运行的全过程中。

(一)设计实施培训即是系统设计和各个阶段的培训，这些阶段包括：项目准备、用户需求分析、系统概要设计、系统详细设计、系统实施、系统运行建立、系统调试、系统维护管理。设计实施培训涉及的相关人员主要是技术人员，分阶段的被培训的人员包括：项目管理人员、需求分析人员、系统分析设计人员和系统管理人员。

(二)运行管理培训是为了使广东省人民医院相关人员掌握有关应用系统的使用、维护和管理方法，达到能独立进行管理、故障处理、日常测试和维护等工作的目的，应进行系统的技术培训，以保证所使用的系统能够正常、安全、平稳地运行。

(三)培训要求派出的培训教员应具有丰富的同类课程的教学经验和应用经验，保证培训的质量，使被培训的人掌握培训的知识，尤其是能独立安装和配置并维护系统；必须为所有被培训人员提供培训用文字资料和讲义等相关材料，如果培训地点在外地，承建方还应为所有被培训人员提供食宿。培训方式：包括课堂讲解、上机操作和实际工作的参与。

六、合同款支付方式

甲方将按如下方式向乙方支付合同费用：

（一）合同签订后，甲方在收到乙方开具相应金额正式发票后，甲方向乙方支付合同总金额的30%。

（二）合同所有硬件设备（产品）运至采购人指定货运详细地址、开箱合格运转正常，所有设备（产品）和软件服务经最终用户签字验收（加电验收）后，且甲方在收到乙方开具相应金额正式发票以及售后服务履约承诺函后，向乙方支付合同总金额的70%。