

深度学习计算 Station 技术指标

序号	名称	数量 (台/套)	技术参数
1	深度学习 计算 Station	3	<p>功能描述:</p> <ul style="list-style-type: none"> 1. TFLOPS (混合精度): ≥ 500 ★2. GPU 显存: \geq 整个系统 128GB 3. CPU: \geq 1 颗双路 20 核 E5-2698 v4 2.2GHz 4. CUDA 核心数量: ≥ 20480 ★5. Tensor 核心数量 (基于 V100): ≥ 2560 6. 最大功率: 1500W 7. 系统内存: $\geq 256GB$ DDR4 RDIMM 8. 存储空间: ≥ 1 块 1.92TB SSD 企业级硬盘+3 块 1.92TB SSD 企业级硬盘 RAID 0 9. 网络: \geq 双 10GbE (RJ45) ★10. 散热: 利用创新工程设计和水冷却系统, 实现无噪声运行 11. 运行噪音: $\leq 35dB$ 12. 必须配备人工智能相关的深度学习软件开发工具包需加速广泛应用的深度学习框架, 如: Caffe、CNTK、MXNet、TensorFlow、Theano 和 Torch。 <p>技术参数:</p> <p>★1. GPU: ≥ 4 块 Tesla V100, 支持 Nvlink 通信技术, 配置一个支持拓扑感知的多 GPU 集成通信库。优化 4 个 NVLink 连接的混合立方体网格的 GPU 之间的 集合通信, 单个 NVIDIA Tesla® V100 GPU 即可支持多达六条 NVLink 链路, 总带宽为 200 GB/秒以上, 提供设备制造厂商官方网站证明截图及网址。</p> <p>2. TFLOPS (混合精度): ≥ 500</p>

		<p>★3. GPU 显存: \geq整个系统 128GB</p> <p>4. CPU: \geq1 颗双路 20 核 E5-2698 v4 2.2GHz</p> <p>5. CUDA 核心数量: \geq20480</p> <p>★6. Tensor 核心数量 (基于 V100) : \geq2560</p> <p>7. 最大功率: 1500W</p> <p>8. 系统内存: \geq256GB DDR4 RDIMM</p> <p>9. 存储空间: \geq1 块 1.92TB SSD 企业级硬盘+3 块 1.92TB SSD 企业级硬盘 RAID 0</p> <p>10. 网络: \geq双 10GbE (RJ45)</p> <p>★11. 散热: 利用创新工程设计和水冷却系统, 实现无噪声运行, 提供官方彩页证明</p> <p>★12. 运行噪音: \leq35dB, 提供官方彩页证明</p> <p>13. 显示器支持: 3 个 DisplayPort 接口, 支持 4K 分辨率, 提供官方彩页证明</p> <p>14. 操作系统: Ubuntu Desktop Linux OS 操作系统, 整个系统平台需具有一次同时运行多个深度学习框架的能力, 容器镜像可跨越新的驱动器版本, 支持多个学生在不同框架下同时实验。</p> <p>★15. 软件系统:</p> <p>★15. 1. 配置深度学习框架</p> <p>该平台配备的深度学习软件开发工具包需加速广泛应用的深度学习框架, 如: Caffe、CNTK、MXNet、TensorFlow、Theano 和 Torch。本平台软件堆栈为这些深度学习框架提供经过优化的容器版本。这些框架包括了所依赖的运行环境, 均预先构建、经过测试且随时可以运行。每个框架容器映像还包含框架源代码, 可以让用户自定义</p>
--	--	--

		<p>做修改和功能增强，还包含完整软件开发堆栈。提供针对该平台专门在优化的深度学习框架，包括 Caffe/Caffe2/TensorFlow/Torch/Theano/CNTK/Mxnet 等；并以镜像的方式提供，针对上述镜像文件，提供远程在线仓库存储；提供深度学习训练应用程序、深度学习 SDK；快速多 GPU 集群软件。</p> <p>★15. 2. 平台针对不同的深度学习框架，提供持续优化支持；针对系统平台本身提供在线的更新支持，包括 OS 升级、BIOS 升级、深度学习框架升级、深度学习加速库升级</p> <p>★15. 3. 配置 CUDA 工具包，帮助应用程序开发人员通过 GPU 实现大规模并行处理。实现 GPU 加速应用的平台，为 C 和 C++ 开发人员构建 GPU 加速应用程序提供全面的开发环境。软件包含面向 GPU 的 CUDA C++ 编译器、一套 GPU 加速算法库、调试和分析工具、范例以及综合编程指南和文档。要求软件可以直接安装于系统平台，也可以作为 Docker 容器镜像提供，以用作任何容器化 CUDA 应用程序的基础层。此外，整个软件需要内嵌于各深度学习框架的容器镜像中。</p> <p>★15. 4. 配置 Docker 容器，要求软件将 Linux 应用程序及其所有库、配置文件和环境变量捆绑在一起的机制，无论运行哪种 Linux 系统，应用的执行环境总是相同的，该系统使用 Docker 容器来部署深度学习框架，Docker 容器与平台和硬件无关，通过将用户模式代码（在容器中）与内核模式代码分离来实现，避免这种分离在使用 GPU 这样的专用硬件时出现问题，该软件需是一个开源项目，提供一个命令行工具，用于在启动容器的时候加载驱动程序的用户模式，并将 GPU 加载到 Docker 容器中。</p> <p>★15. 5. 配置深度学习软件开发工具包</p> <p>平台提供构建于 CUDA 并行计算平台之上的全套 GPU 加速库。提供至少两种库可提供用于深度神经网络的 GPU 加速：</p>
--	--	--

		<p>A. CUDA 基础线性代数子例程库 (cuBLAS)： cuBLAS 是完整的标准 BLAS 库的 GPU 加速版，可大幅提高在 GPU 上的运行速度。 cuBLAS 广义矩阵乘法 (GEMM) 例程是深度神经网络应用的关键计算，例如应用于全连接层计算。</p> <p>B. cuDNN：可大幅优化深度神经网络标准例程的实施，例如前向和反向的卷积层、池化层、归一化层和激活层。</p> <p>该平台中配置的深度学习框架容器中，需已自动配置使用针对 Tesla V100 GPU 优化好的并行库。</p> <p>★15. 6. 平台集成通信库 NCCL</p> <p>支持拓扑感知的多 GPU 集成通信库。平台的 Docker 容器中包含一个 NCCL 版本，优化了平台中 4 个 NVLink 连接的混合立方体网格的 GPU 之间的集合通信。当使用平台上的容器进行部署时，深度学习框架（如 Caffe，Torch，Microsoft Cognitive Toolkit 和 TensorFlow）在多个 GPU 上运行时，会自动使用此版本的 NCCL</p> <p>★16. 所有系统要求的试验环境和附带软件需在 1 天内完成部署安装，并调试优化至可以直接使用，软件提供设备原制造厂商至少 36 月的技术支持和售后。</p> <p>★17. 针对性优化：</p> <p>★17. 1 需提供原厂商针对本平台对 Caffe 做出的的优化和变更的详细表格。</p> <p>★17. 2. 需提供原厂商针对本平台对 Caffe2 做出的的优化和变更的详细表格</p> <p>★17. 3. 需提供原厂商针对本平台对 Microsoft Cognitive Toolkit 做出的的优化和变更的详细表格。</p> <p>★17. 4. 需提供原厂商针对本平台对 MXNet 做出的的优化和变更的详细表格。</p>
--	--	---

		<p>★17. 5. 需提供原厂商针对本平台对 TensorFlow2 做出的的优化和变更的详细表格。</p> <p>★17. 6. 需提供原厂商针对本平台对 Theano 做出的的优化和变更的详细表格</p> <p>★17. 7. 需提供原厂商针对本平台对 PyTorch 做出的的优化和变更的详细表格</p> <p>★17. 8. 需提供原厂商针对本平台对 Torch 做出的的优化和变更的详细表格</p> <p>★17. 9. 需提供原厂商针对本平台对 DIGITS 做出的的优化和变更的详细表格</p> <p>★17. 10. 需提供原厂商针对本平台对 TensorRT 做出的的优化和变更的详细表格</p> <p>★18. 培训学习：设备制造厂商需具备深度学习相关的培训学院，可提供自学式在线学习和面对面讲师指导等课程形式，教授学员开启构建、优化 和部署神经网络，从而解决各行业的现实问题，提供官方网站截图及网址。</p> <p>★19. 供货资质：投标人需提供原厂商官方授权的授权书，或原厂商授权的代理商提供的授权书。</p>
--	--	---