



H.Cloud

通过先进的体系结构
充分发挥GPU算力
让您的智算基础设施获得更大的效能



通过先进的体系结构充分发挥GPU算力 让您的智算基础设施获得更大的效能

随之企业大举迈向人工智能，算力的提升不能再依靠单张GPU卡的性能，更多时候需要多颗GPU卡组合算力资源。在系统内部，多颗GPU之间的通信带宽通常在几百GB/s以上，PCIe总线的数据传输速率很容易成为瓶颈，且传统的PCIe线路接口的串并转换会产生较大的延迟，极大降低GPU之间并行效率和性能。PCIe Switch在接收GPU发出的信号之后进行数据处理，CPU则负责数据的分发和调度，这些都会引发额外的网络延迟，限制了智算系统的性能。因此，为了提升GPU之间的通信性能，当今智算领域的龙头厂商NVIDIA推出了GPU Direct、P2P等先进的通信技术，使GPU相互之间可通过PCIe访问显存，极大的降低了通过CPU与RAM Buffer中转带来的延迟。但受限PCI Express总线协议和拓扑的限制，却无法做到更高的带宽。此后，NVIDIA又进一步提出了NVLink总线协议。

诺亚鸿云致力于智算领域先进的体系结构设计，而非是提供服务器组合GPU卡的整机交付供应商。具备从电子电路系统开发到智算整机产品化。其独特的体系结构设计确保广泛的GPU能够以更高的带宽和效率进行通信，实现大型算力组网的P2P互联特性确保广泛的GPU能够发挥更大性能，同时可基于全国产化的底座运行满足日益严苛的国产化政策需求。

确保快速交付和易用性也是诺亚鸿云产品理念的一部分，我们确保一系列的设备交付之初便适配完成主流的模型和应用程序（类似DPSK、Qwen、Llama、Stable Diffusion、VSCode..），几乎所有厂商的GPU性能压力测试在多个基准下，集成框架、算子、驱动程序的可视化系统平台-AiOS等等。这些研发的目标可确保用户快速的私有化部署人工智能应用程序、模型推理、训练和微调等广泛的场景。



您的GPUs可以做更多

诺亚鸿云创新体系结构设计，可以让您根据需求、预算、喜好选择广泛厂牌的GPUs。通过满足信创政策需求、加速GPU性能、降低成本、实现智算投资价值最大化。摆脱供应商锁定，让您的智算基础设施为满足未来的需求时刻做好准备。

无论您是想利用已经购买的GPUs，还是想快速的构建小型的算力场景，诺亚鸿云都能为您提供帮助。开放性的体系结构能够实现GPU即插即用，而无需开箱即可进行复杂的维护。

深圳市诺亚鸿云信息技术有限公司，成立于2017年7月18日，最初专注于为企业构建高性能的私有云建设方案。为了迎合企业客户向人工智能转型，以及赋能广泛的国产GPU厂商，通过创新型的智算体系结构推出了一系列的AI智算整机等自主知识产权系列产品，同时结合当今AI技术潮流，提供创新型应用方案。旨在通过商业化人工智能模型与应用程序满足智能文档识别，企业级知识库，问答，咨询，快速响应的信息需求，促进生产效率和用户体验，应对人力资源压力等多重因素的需求产生的人工智能方案。

我们不追求降本增效带来的锦上添花，
而是富有激情的去探索科学边界，通过一系列
革新性的科技成果，提升IT组织与企业的生产力水平，
让自主可控更高效！

“诺亚鸿云创新型智算体系结构，有效赋能国产厂商的GPU性能提升，同时提供大规模智算的互联特性。在此基础上还提供了全球独创的“GPU热插拔”、“GPU全局热备”、以及可让GPU实例进行休眠的特性，成为了企业打造全国产化智算基础设施的破冰之船。”

徐女士 - 科翔电子首席架构师,CIO

为什么选择诺亚鸿云智算一体机



每个GPU提供独立的嵌入式电源管理模块，能够在高吞吐Tokens与低频访问之间平衡电力消耗



创新型GPU热备特性，使GPU具备N+1冗余特性。避免关键推理，训练过程不可避免的业务中断时间。



高密度的体系结构设计，最多支持10/20颗GPU扩展成为构建大型云智算首选方案、超大模型运行可规避因为组网带来的延迟和成本开支



革命性的GPU-热拔插特性，使通用GPU可像硬盘一样更换/下线，提供极致的维护便利性。



提供桌面形态小型算力单元，便于移动和静音设计。32B、70B等小语言模型不再依赖于机架式大型算力设备。



具备构建CPU+GPU全国产化的体系结构设计，让GPU算力底座高效运行的同时更加安全可控。

探索产品



Ai WorkStation

移动、桌面型算力单元 AWS系列
特别适合小型算力单元、而无需购入机架式，可适配广泛智能体的应用程序打造全信创交付推理、知识库、Stable Diffusion、Agent

- AWS-P2、AWS-P4 分别支持自由扩展2颗 /4颗通用型标准GPU
- 默认配置X-100通算卡具备完整的计算机体系、打造全国产化体系、GpuDirect Storage 技术可降低数据从内存中转的模型载入延迟
- 便于移动、超静音、免主板的卡式设计方便维护、出厂即预装集成广泛GPU驱动程序、算子与框架于一体的AIOS可视化系统
- AWS系列成为Ai智能体广受青睐的智算载体，我们提前预装主流的语言模型，并且和主流的GPU厂商进行了深度适配和优化
- AIOS不仅支持配置内部磁盘扩展存储，还可以接入公网API算力



Multi-GPU System Platform

独领全球的创新型智算体系
具备全国产化的体系结构，高密度智算单元
专为大规模算力集群互联打造的整机智算
通过一系列的高级特性

- 高达20颗GPU扩展MGP-820ls、以及10颗GPU扩展的MGP-410
- 革命性的GPU-热拔插特性，使通用GPU可像硬盘一样在线热拔插，提供极致的维护便利性
- 每个GPU提供独立的嵌入式电源管理模块，能够在高吞吐Tokens与低频访问之间平衡电力消耗
- 支持GPU异构，即：支持多个品牌国产GPU在同一个节点，服务同一个大模型，并提供丰富的任务调度策略“GPU池化”特性，多个GPU可聚合成算力池专为：气象分析，数学运算，高精度计算，生物分析
- 创新型GPU热备特性，类似硬盘的RAID技术，促使GPU在本地具备N+1冗余特性。一旦GPU故障“Hot-Spare” GPU会立刻接管



Hybrid-Computing Platform

HCP-48 混合算力系统最高支持8颗GPU扩展
致力于构建一套整体的人工智能基建设方案
CPU通算+GPU智算一体机、构建的先进
混合算力体系符合最新理念-
Full-Stack Hyperconverged

- 采用创新型的两段式设计体系，既：通算部分仍然使用高性能国产处理器体系结构，智算部分则采用自研的独立GPU基板
- 通过2颗GPU加速模组，促使GPU在系统内部高速互联，以此来规避PCIe带宽瓶颈，加速模组还提供节点P2P直连特性，满足大型智算组网
- HCP系统耦合了以处理器为核心的通用算力来构建私有云平台，还包含了以GPU为核心的高性能智算体系，符合全国产化设计，让智算基础设施高效运行的同时更加自主可控

H.Cloud.Next

以GPU为核心的创新型架构可快速加速国产GPU推广速度

今天，我们所看到人工智能爆发的规模远远超过了任何时候，对GPU算力需求也越来越大。另一方面国产GPU在符合自主可控的政策下由于生态，兼容性，单卡的性能，以及尚未具备高级互联特性等因素导致在AI浪潮中步履蹒跚，这也给IT团队在进行国产化智算建设时带来了独特而复杂的挑战。现在，需要一种全新的体系结构来打破僵局。

H.Cloud.NEXT 通过创新型的体系结构，不仅让普遍的国产GPU再次加速，还赋予业内首创的高级特性，通过最低的部署成本和更高的性能让国产GPU可以大范围部署推理与训练场景，以GPU为单位的P2P2互联技术成为基于全国产化智算中心建设的中流砥柱。

诺亚鸿云创新不止于此，我们专业的电子电路开发团队仍然在积极探索具备更高带宽的互联技术，如CXL技术，板载光通讯技术等。

我们深知：让科技企业保持竞争力的条件之一、便是具备“持续创新”的能力。这导致我们不断探索与量化、每次“创新”对于客户产生的实际效益！

H.Cloud THE AUTHORITY 技术革新

鸿云智算产品矩阵

- 大规模智算组网
10卡GPU算力单元
MGP-410
- Ai WorkStation
AWS-P2、AWS-P4
- 超大规模智算组网
20卡GPU算力单元
MGP-820Is
- AIOS-vLLM
Workflow RAG
- HCP-48
混合算力平台

迈向国产化信息改造 | 智能体最佳平台 | 服务器性能提升 | 全国产混合算力平台 | Stable Diffusion | 广泛国产GPU适配 | 以GPU为实例体眼

入门级GPU算力单元 | 大型互联算力单元 | GPU“全局热备” | AiOS一体化交付 | Ai Agent | 多种CPU体系可选 | 高密度规组网需求

GPU异构算力池 | GPU卡间互联模组 | GPU整机P2P互联 | AIGC | 多任务智算调度平台 | 全国产化设计 | RAG增强型检索

★★★★★ 业内领先的专家团队

诺亚鸿云汇集了清华微电子，中科院计算所，国科大等顶尖的电子电路设计团队，具备丰富的计算机系统全栈式设计经验。曾经领导过多次国家重点计算机体系建设工程。我们始终强化全球支持和服务价值，提供世界级的客户服务体验支持，确保客户对产品价值的忠诚度，近十年来，我们的客户从中受益匪浅。

0728



Your Computility, Our Priority
您的算力，是我们首要的任务

诺亚鸿云同时拥有强大的软件研发团队（晓软科技），专注于为企业提供前沿的人工智能解决方案。公司以“硬件+软件”双轮驱动为核心战略，在我们的晓软工程院核心研发基地，汇聚了一批充满激情的硬件研发团队。

在存储系统研发，电子电路设计，AI与智算领域有超过15+年研发经验，上线的产品已经经过多年的市场打磨，并承载诸多企业客户的核心业务系统。公司同时拥有强大软件研发团队，专注于为企业提供前沿的智能化解决方案，服务涵盖数智化工厂建设、智能制造系统集成研发、工业数据分析和人工智能应用等方面，致力于人工智能、数字智造深度创新和落地。



诺亚鸿云总部（深圳）

nyhy-cloud.com

400-636-9663