



采用NVIDIA的NetApp AI Pod NetApp Solutions

NetApp
August 14, 2024

目录

采用NVIDIA技术的NetApp AI	1
采用NVIDIA DGX系统的NetApp AIPod	1
采用 NVIDIA DGX A100 系统的 NetApp ONTAP AI	1
采用 NVIDIA DGX A100 系统和 Mellanox 频谱以太网交换机的 NetApp ONTAP AI	1
采用NVIDIA DGX系统的NetApp AIPod—简介	1
NVA-1151-design: 采用NVIDIA DGX A100的NetApp ONTAP AI系统设计指南	10
NVA-1151-Deploy: 采用NVIDIA DGX A100系统的NetApp ONTAP AI	10
NVA-1153-design: 采用NVIDIA DGX A100系统和Mellanox Spectrum以太网交换机的NetApp ONTAP AI ..	10
NVA-1153-Deploy: 采用NVIDIA DGX A100系统和Mellanox Spectrum以太网交换机的NetApp ONTAP AI ..	11

采用NVIDIA技术的NetApp AI

NetApp 和 NVIDIA 的 ONTAP AI 融合基础架构解决方案概述。

采用NVIDIA DGX系统的NetApp AI Pod

- ["采用NVIDIA DGX系统的NetApp AI Pod"](#)

采用 NVIDIA DGX A100 系统的 NetApp ONTAP AI

- ["设计指南"](#)
- ["部署指南"](#)

采用 NVIDIA DGX A100 系统和 Mellanox 频谱以太网交换机的 NetApp ONTAP AI

- ["设计指南"](#)
- ["部署指南"](#)

采用NVIDIA DGX系统的NetApp AI Pod—简介

本节介绍了采用NVIDIA DGX系统的NetApp AI Pod。

NetApp 解决方案 工程部

NetApp 采用 NVIDIA DGX 的 AI Pod；系统和 NetApp 云互联存储系统、通过消除设计复杂性和猜测、简化了机器学习(ML)和人工智能(AI)工作负载的基础架构部署。采用 NVIDIA DGX 系统的 AI Pod 基于 NVIDIA DGX BasePOD 设计构建、可为下一代工作负载提供卓越的计算性能、并添加了 NetApp AFF 存储系统、支持客户从小规模入手、无故障扩展、同时智能地管理从边缘到核心再到云再到云的数据。NetApp AI Pod 是更广泛的 NetApp AI 解决方案产品组合的一部分、如下图所示-

ALL NetApp AI 解决方案组合 图像: : aipod_nf_bonline.png[]

本文档介绍 AI Pod 参考架构的关键组件、系统连接信息和解决方案规模估算指导。本文档面向有意为 ML/DL 和分析工作负载部署高性能基础架构的 NetApp 和合作伙伴解决方案工程师以及客户战略决策者。

采用NVIDIA DGX系统的NetApp AI Pod—简介

本节介绍了采用NVIDIA DGX系统的NetApp AI Pod。

NetApp 解决方案 工程部

NetApp 采用 NVIDIA DGX 的 AI Pod；系统和 NetApp 云互联存储系统、通过消除设计复杂性和猜测、简化了机器学习(ML)和人工智能(AI)工作负载的基础架构部署。采用 NVIDIA DGX 系统的 AI Pod 基于 NVIDIA DGX BasePOD 设计构建、可为下一代工作负载提供卓越的计算性能、并添加了 NetApp AFF 存储系统、支持客户从小规模入手、无故障扩展、同时智能地管理从边缘到核心再到云再到云的数据。NetApp AI Pod 是更广泛

的NetApp AI解决方案产品组合的一部分、如下图所示-

_ALL NetApp AI解决方案组合_图像：： aipod_nf_bonline.png[]

本文档介绍AIPod参考架构的关键组件、系统连接信息和解决方案规模估算指导。本文档面向有意为ML/DL和分析工作负载部署高性能基础架构的NetApp和合作伙伴解决方案工程师以及客户战略决策者。

采用NVIDIA DGX系统的NetApp AIPod—硬件组件

本节重点介绍采用NVIDIA DGX系统的NetApp AIPod的硬件组件。

NetApp AFF存储系统

NetApp AFF一流的存储系统凭借行业领先的性能、卓越的灵活性、云集成和同类最佳数据管理功能、可帮助IT部门满足企业存储需求。AFF系统专为闪存而设计、有助于加速、管理和保护业务关键型数据。

AFF A900存储系统

由NetApp ONTAP数据管理软件提供支持的NetApp AFF A900提供内置数据保护、可选的反勒索软件功能以及支持最关键业务工作负载所需的高性能和故障恢复能力。它可以消除任务关键型运营中断、最大限度地减少性能调整、并保护数据免受勒索软件攻击。它提供：

- 行业领先的性能
- 不折不扣的数据安全性
- 简化无中断升级

a900 NetApp AFF存储系统 image：： aipod_nf_a900.png[]

行业领先的性能

AFF A900可轻松管理下一代工作负载、例如深度学习、AI和高速分析、以及Oracle、SAP HANA、Microsoft SQL Server和虚拟化应用程序等传统企业数据库。它可以使业务关键型应用程序以最高速度运行、每个HA对的IOPS高达240万次、延迟低至100µs微秒、与以前的NetApp型号相比、性能可提高多达50%。借助基于RDMA的NFS、pNFS和会话中继、客户可以使用现有数据中心网络基础架构实现下一代应用程序所需的高水平网络性能。客户还可以通过对SAN、NAS和对象存储的统一多协议支持进行扩展和扩展、并通过统一的和单个ONTAP数据管理软件为内部或云中的数据提供最大的灵活性。此外、Active IQ数字顾问(也称为数字顾问)和Cloud Insights提供的基于AI的预测性分析可以优化系统运行状况。

不折不扣的数据安全性

AFF A900系统包含一整套NetApp集成的应用程序一致的数据保护软件。它提供内置数据保护和尖端的反勒索软件解决方案、用于抢占资源和攻击后恢复。可以阻止恶意文件写入磁盘、并且可以轻松监控存储异常情况以获得洞察力。

简化无中断升级

AFF A900可作为无中断机箱内升级提供给现有A700客户。NetApp通过其高级可靠性、可用性、可维护性和易管理性(RASM)功能、使更新变得简单、并消除任务关键型操作中断。此外、NetApp ONTAP软件会自动对所有系统组件应用固件更新、因此可以进一步提高运营效率并简化IT团队的日常活动。

对于规模最大的部署、AFF A900系统可提供最高的性能和容量选项、而其他NetApp存储系统(例如AFF A800、AFF C800、AFF A400、AFF C400和AFF A250)则可为较小的部署提供更低成本的选项。

NVIDIA DGX基本POD

NVIDIA DGX BasePOD是一种集成的解决方案、由NVIDIA硬件和软件组件、MLOps解决方案和第三方存储组成。利用采用NVIDIA产品和经验证的合作伙伴解决方案的横向扩展系统设计最佳实践、客户可以实施一个高效易管理的AI开发平台。图1突出显示了NVIDIA DGX BasePOD的各个组件。

_NVIDIA DGX BasePOD solidue_image: : aipod_nf_basepod_layers.png[]

NVIDIA DGX H100系统

NVIDIA DGX H100#8482;系统是AI的动力之源、它通过NVIDIA H100 T能器核心GPU的开创性性能得到了加速。

NVIDIA DGX H100 system image: : aipod_nf_h100_3d.png[]

DGX H100系统的主要规格如下:

- 八个NVIDIA H100 GPU。
- 每个GPU具有80 GB GPU内存、总容量为640 GB。
- 四个NVIDIA NVSwitch™芯片。
- 两个56核英特尔®至强®白金级8480处理器,支持PCIe 5.0。2 TB的系统内存。
- 四个OFP端口、用于支持八个单端口NVIDIA ConnectX-7 (InfiniBand/以太网)适配器和两个双端口NVIDIA ConnectX-7 (InfiniBand/以太网)适配器。
- 两个1.92 TB M.2 NVMe驱动器用于DGX操作系统、八个3.84 TB U.2 NVMe驱动器用于存储/缓存。
- 10.2 kW最大功率。

DGX H100 CPU托盘的后部端口如下所示。其中四个OSFP端口为InfiniBand计算网络结构提供八个ConnectX-7适配器。每对双端口ConnectX-7适配器都提供指向存储和管理网络结构的并行路径。带外端口用于BMC访问。

_NVIDIA DGX H100后面板_图像: : aipod_nf_h100_rean.png[]

NVIDIA网络

NVIDIA昆特姆-2 QM9700交换机

NVIDIA昆士兰-2 QM9700 InfiniBand switch image: : : aipod_nf_QM970.png[]

具有400 Gb/秒InfiniBand连接的NVIDIA昆特姆-2 QM9700交换机为NVIDIA昆特姆-2 InfiniBand BasePOD配置中的计算网络结构提供支持。ConnectX-7单端口适配器用于InfiniBand计算网络结构。每个NVIDIA DGX系统都与每个QM9700交换机建立了双连接、从而在系统之间提供多个高带宽、低延迟路径。

NVIDIA Spectrum 3 SN4600交换机

NVIDIA Spectrum 3 SN4600 switch image: : aipod_nf_sn4600_hires_smit.png[]

NVIDIA Spectrum 3 SN4600交换机总共提供128个端口(每个交换机64个)、可为DGX BasePOD的带内管理提供冗余连接。NVIDIA SN4600交换机可提供1 GbE到200 GbE的速度。对于通过以太网连接的存储设备、还会使用NVIDIA SN4600交换机。NVIDIA DGX双端口ConnectX-7适配器上的端口用于带内管理和存储连接。

NVIDIA Spectrum SN2201交换机

NVIDIA Spectrum SN2201 switch image: : aipod_NV_SN2201.png[]

NVIDIA Spectrum SN2201交换机提供48个端口、可为带外管理提供连接。带外管理可为DGX BasePOD中的所

有组件提供整合的管理连接。

NVIDIA ConnectX-7适配器

NVIDIA ConnectX-7适配器 image: : aipod_nf_cx7.png

NVIDIA ConnectX-7适配器可提供25/50/100/200/400G吞吐量。NVIDIA DGX系统使用单端口和双端口ConnectX-7适配器、可通过400 Gb/秒InfiniBand和100/C200 GB以太网在DGX BasePD部署中提供灵活性。

采用NVIDIA DGX系统的NetApp AIPod—软件组件

本节重点介绍采用NVIDIA DGX系统的NetApp AIPod的软件组件。

NVIDIA软件

NVIDIA基本命令

NVIDIA Base Command[®]；为每个DGX BasePD提供支持、使企业能够充分利用NVIDIA软件创新的优势。企业可以利用成熟可靠的平台充分发挥投资的全部潜能、该平台包括企业级流程编排和集群管理、可加快计算、存储和网络基础架构速度的库以及针对AI工作负载优化的操作系统(OS)。

_NVIDIA BaseCommand solidue_image: : aipod_nf_BaseCommand_new.png

NVIDIA GPU Cloud (NGC)

NVIDIA NGC™提供的软件可满足具有不同AI专业知识水平的数据科学家、开发人员和研究人员的需求。NGC上托管的软件会对一组汇总的常见漏洞和风险(CVE)、加密密钥和私钥进行扫描。它经过测试和设计、可扩展到多个GPU、在许多情况下还可扩展到多节点、从而确保用户在DGX系统上的投资最大化。

_NVIDIA GPU Cloud_图像: : aipod_nf_ngc.png

NVIDIA AI Enterprise

NVIDIA AI Enterprise是一款端到端软件平台、可将生成性AI应用于每个企业、为生成性AI基础模型提供速度最快、最高效的运行时间、这些模型经过优化、可在NVIDIA DGX平台上运行。凭借生产级安全性、稳定性和易管理性、它简化了生成性AI解决方案的开发。DGX BasePOD附带了NVIDIA AI Enterprise、供企业开发人员访问经过预先训练的模型、优化的框架、微服务、加速库和企业支持。

NetApp软件

NetApp ONTAP

ONTAP 9是NetApp推出的最新一代存储管理软件、可帮助企业打造现代化的基础架构并过渡到云就绪数据中心。借助行业领先的数据管理功能，无论数据位于何处，ONTAP 9都可以通过一组工具来管理和保护数据。您还可以将数据自由移动到需要的任何位置：边缘，核心或云。ONTAP 9包含许多功能、可简化数据管理、加快和保护关键数据、并在混合云架构中实现下一代基础架构功能。

加速和保护数据

ONTAP 9可提供卓越的性能和数据保护、并通过以下方式扩展这些功能：

- 性能和更低的延迟。ONTAP 9可以以最低延迟提供最高吞吐量、包括支持使用基于RDMA的NFS的NVIDIA

GPUDirect存储(GDS)、并行NFS (pNFS)和NFS会话中继。

- 数据保护ONTAP提供内置数据保护功能和业内最强大的反勒索软件担保、并在所有平台之间提供通用管理。
- NetApp卷加密(NVE)。ONTAP 提供原生 卷级加密、并支持板载和外部密钥管理。
- 存储多租户和多因素身份验证。ONTAP 支持以最高的安全性级别共享基础架构资源。

简化数据管理

数据管理对于企业IT运营和数据科学家至关重要、这样才能将适当的资源用于AI应用程序和训练AI/ML数据集。以下有关NetApp技术的追加信息 不在此验证范围内、但可能与您的部署相关。

ONTAP 数据管理软件包括以下功能、可简化操作并降低总运营成本：

- 快照和克隆支持协作、并行实验和增强ML/DL工作流的数据监管。
- SnapMirror支持在混合云和多站点环境中无缝移动数据、可根据需要随时随地提供数据。
- 实时数据缩减和扩展的重复数据删除。数据缩减可减少存储块中浪费的空间、重复数据删除可显著提高有效容量。此适用场景数据存储在本机，并分层到云。
- 最低、最高和自适应服务质量(AQoS)。精细的服务质量(QoS)控制有助于在高度共享的环境中保持关键应用程序的性能水平。
- 通过NetApp FlexGroup、可以在存储集群中的所有节点之间分布数据、从而为超大型数据集提供海量容量和更高的性能。
- NetApp FabricPool。可将冷数据自动分层到公有 和私有云存储选项、包括Amazon Web Services (AWS)、Azure和NetApp StorageGRID Storage解决方案。有关 FabricPool 的详细信息，请参见 "[TR-4598 : FabricPool 最佳实践](#)"。
- NetApp FlexCache。提供远程卷缓存功能、可简化文件分发、减少WAN延迟并降低WAN带宽成本。FlexCache支持跨多个站点进行分布式产品开发、并加快从远程位置访问公司数据集的速度。

Future-Proof 基础架构

ONTAP 可通过以下功能满足不断变化的苛刻业务需求：

- 无缝扩展和无中断运行。ONTAP支持向现有控制器和横向扩展集群联机添加容量。客户可以升级到 NVMe 和 32 Gb FC 等最新技术，而无需进行成本高昂的数据迁移或中断。
- 云连接。ONTAP 是云互联程度最高的存储管理软件、可在所有公有 云中选择软件定义的存储(ONTAP Select)和云原生实例(NetApp Cloud Volumes Service)。
- 与新兴应用程序集成。ONTAP 通过使用支持现有企业应用程序的相同基础架构、为下一代平台和应用程序(例如自动驾驶汽车、智能城市和行业4.0)提供企业级数据服务。

NetApp DataOps 工具包

NetApp DataOps工具包是一款基于Python的工具、可简化开发/培训工作和推理服务器的管理、这些工作空间和服务器由高性能横向扩展NetApp存储提供支持。数据操作工具包可作为独立的实用程序运行、在利用NetApp Asta三端存储自动化来实现存储操作的Kubernet环境中、该工具包的效率更高。主要功能包括：

- 快速配置新的高容量JupyterLab工作空间、这些工作空间以高性能横向扩展NetApp存储为后盾。
- 快速配置由企业级NetApp存储提供的新NVIDIA Triton推理服务器实例。
- 近乎即时地克隆高容量JupyterLab工作空间、以便进行实验或快速迭代。

- 近乎即时地为高容量JupyterLab工作空间创建快照、用于备份和/或可追溯性/基线化。
- 近乎即时地配置、克隆和创建高容量、高性能数据卷的快照。

NetApp Astra Trident

Astra Trident是一款完全受支持的开源存储编排程序、适用于容器和Kubernetes分发版、包括Anthos。通过使用NetApp ONTAP等整个NetApp存储产品组合、可以支持NFS、NVMe/TCP和iSCSI连接。Trident 允许最终用户从其 NetApp 存储系统配置和管理存储，而无需存储管理员干预，从而加快了 DevOps 工作流的速度。

采用NVIDIA DGX系统的NetApp AIPod—解决方案架构

本节重点介绍采用NVIDIA DGX系统的NetApp AIPod的架构。

采用DGX H100系统的NetApp AI Pod

此参考架构利用单独的网络结构实现计算集群互连和存储访问、并在计算节点之间建立400 GB/秒InfiniBand (IB)连接。下图显示了采用DGX H100系统的NetApp AIPod的整体解决方案拓扑。

*aipod NetApp*解决方案拓扑图 image: : aipod_nf_a900topo.png[]

网络配置：

在此配置中、计算集群网络结构使用一对QM9700 400Gb/s IB交换机、这些交换机连接在一起以实现高可用性。每个DGX H100系统通过八个连接连接到交换机、其中、偶数端口连接到一个交换机、奇数端口连接到另一个交换机。

对于存储系统访问、带内管理和客户端访问、使用一对SN4600以太网交换机。这些交换机通过交换机间链路进行连接、并配置有多个VLAN以隔离各种流量类型。对于大型部署、可以根据需要为主干交换机添加更多交换机对并添加更多叶片、从而将以太网网络扩展为分支-主干配置。

除了计算互连和高速以太网网络之外、所有物理设备还会连接到一个或多个SN2201以太网交换机、以实现带外管理。有关DGX H100系统连接的更多详细信息、请参见 "[NVIDIA BasePD文档](#)"。

用于存储访问的客户端配置

每个DGX H100系统都配置有两个双端口ConnectX-7适配器、用于管理和存储流量、对于此解决方案、每个卡上的两个端口都连接到同一交换机。然后、将每个卡中的一个端口配置到LACP MAG绑定中、并将一个端口连接到每个交换机、同时在此绑定上托管用于带内管理、客户端访问和用户级存储访问的VLAN。

每个卡上的另一个端口用于连接到AFF A900存储系统、并可根据工作负载要求在多种配置中使用。对于使用基于RDMA的NFS来支持NVIDIA Magrum IO GPU Direct存储的配置、端口配置为主动/被动绑定、因为任何其他类型的绑定都不支持RDMA。对于不需要RDMA的部署、还可以为存储接口配置LACP绑定、以提供高可用性和额外带宽。无论是否使用RDMA、客户端都可以使用NFS v4.1 pNFS和会话中继挂载存储系统、以便能够并行访问集群中的所有存储节点。

存储系统配置：

每个AFF A900存储系统使用每个控制器中的四个100 GbE端口进行连接。每个控制器的两个端口用于从DGX系统访问工作负载数据、每个控制器的两个端口配置为LACP接口组、以支持从管理平台服务器访问集群管理项目和用户主目录。存储系统的所有数据访问均通过NFS提供、其中一个Storage Virtual Machine (SVM)专用于AI工作负载访问、另一个SVM专用于集群管理用途。

工作负载SVM总共配置了八个逻辑接口(LIF)、每个物理端口上配置两个LIF。此配置可提供最大带宽、并允许每个LIF故障转移到同一控制器上的另一个端口、以便在发生网络故障时两个控制器保持活动状态。此配置还支持基于RDMA的NFS、以启用GPU Direct存储访问。存储容量以一个大型FlexGroup卷的形式进行配置、该卷跨越集群中的所有存储控制器、每个控制器上有16个成分卷。可从SVM上的任何LIF访问此FlexGroup、并通过将NFSv4.1与pNFS和会话中继结合使用、客户端可与SVM中的每个LIF建立连接、从而可以并行访问每个存储节点的本地数据、从而显著提高性能。此外、还会为工作负载SVM和每个数据LIF配置RDMA协议访问。有关ONTAP的RDMA配置的详细信息、请参见 "[ONTAP 文档](#)"。

管理SVM仅需要一个LIF、该LIF托管在每个控制器上配置的双端口接口组上。在管理SVM上配置了其 他FlexGroup卷、用于存放集群管理项目、例如集群节点映像、系统监控历史数据和最终用户主目录。下图显示了存储系统的逻辑配置。

a900 NetApp存储集群逻辑配置 image: : aipod_nf_a900logical.png[]

管理平台服务器

此参考架构还包括五个基于CPU的服务器、供管理平台使用。其中两个系统用作NVIDIA Base Command Manager的主节点、用于集群部署和管理。另外三个系统用于提供额外的集群服务、例如、在使用Slurm进行作业计划的部署中、可使用Kubornetes主节点或登录节点。利用Kubnetes的部署可以利用NetApp Asta三端CSI驱动程序为AFF A900存储系统上的管理和AI工作负载提供自动化配置和数据服务以及永久性存储。

每个服务器都会以物理方式连接到IB交换机和以太网交换机、以实现集群部署和管理、并通过管理SVM配置NFS挂载到存储系统、以便如前所述存储集群管理项目。

采用NVIDIA DGX系统的NetApp AIPod—解决方案验证和大小指导

本节重点介绍采用NVIDIA DGX系统的NetApp AIPod的解决方案验证和规模估算指南。

解决方案验证

此解决方案中的存储配置已通过使用开源工具FIO的一系列综合工作负载进行验证。这些测试包括用于模拟由执行深度学习培训作业的DGX系统生成的存储工作负载的读写I/O模式。存储配置已通过一个双插槽CPU服务器集群进行验证、该集群可同时运行FIO工作负载、以模拟一个DGX系统集群。每个客户端都配置了与前文所述相同的网络配置、并添加了以下详细信息。

此验证使用了以下挂载选项：

VERS=4.1	启用pNFS以并行访问多个存储节点
proto = RDMA	将传输协议设置为RDMA、而不是默认TCP
端口=20049	为RDMA NFS服务指定正确的端口
max_connect = 16	启用NFS会话中继以聚合存储端口带宽
write=eager	提高缓冲写入的写入性能
rsize=262144 , wsize=262144	将I/O传输大小设置为256k

此外、还为客户端配置了NFS max_sSession_狭缝值1024。在使用基于RDMA的NFS对解决方案进行测试时、存储网络端口配置了主动/被动绑定。此验证使用了以下绑定参数：

Mode=active-backup	将绑定设置为主动/被动模式
--------------------	---------------

Primary =<interface name>	所有客户端的主接口均分布在交换机之间
mll-monitor-interval=100	指定监控间隔为100毫秒
fail-over-mac-policy=active	指定活动链路的MAC地址为绑定的MAC。要通过绑定接口正确运行RDMA、必须执行此操作。

存储系统按照所述进行配置、配置有两个A900 HA对(4个控制器)、其中两个NS224磁盘架、每个HA对连接有24个1.9 TB NVMe磁盘驱动器。如架构部分所述、所有控制器的存储容量均使用FlexGroup卷进行合并、所有客户端的数据分布在集群中的所有控制器上。

存储系统大小指导

NetApp已成功完成DGX BasePOD认证、经测试的两个A900 HA对可轻松支持由八个DGX H100系统组成的集群。对于存储性能要求较高的大型部署、可以在一个集群中向NetApp ONTAP集群添加更多AFF系统、最多可添加12个HA对(24个节点)。使用本解决方案中所述的FlexGroup技术、一个24节点集群可以在一个命名空间中提供超过40 PB的吞吐量和高达300 Gbps的吞吐量。其他NetApp存储系统(例如AFF A400、A250和C800)以更低的成本为小型部署提供了更低性能和/或更高的容量选项。由于ONTAP 9支持混合模式集群、因此客户可以先减少初始占用空间、然后随着容量和性能要求的增长向集群添加更多或更大的存储系统。下表显示了每个AFF型号所支持的A100和H100 GPU数量的粗略估计。

_ NetApp存储系统规模估算指南_

		Throughput ²	Raw capacity (typical / max)	Connectivity	# NVIDIA A100 GPUs supported ³	# NVIDIA H100 GPUs supported ⁴
NetApp® AFF A900	1 HA pair ¹	28GB/s	182TB / 14.7PB	100 GbE	1 - 64	1-32
	12 HA pairs	336GB/s	2.1PB / 176.4PB		768	384
AFF A800	1 HA pair	25GB/s	368TB / 3.6PB	100 GbE	1 - 64	1-32
	12 HA pairs	300GB/s	4.4PB / 43.2PB		768	384
AFF C800	1 HA pair	21GB/s	368TB / 3.6PB	100 GbE	1-48	1-24
	12 HA pairs	252GB/s	4.4PB / 43.2PB		576	288
AFF A400	1 HA pair	11GB/s	182TB / 14.7PB	40/100 GbE	1 - 32	1-16
	12 HA pairs	132GB/s	2.1PB / 176.4PB		384	192
AFF C400	1 HA pair	8GB/s	182TB / 14.7PB	40/100 GbE	1 - 16	1-8
	12 HA pairs	128GB/s	2.1PB / 176.4PB		192	96
AFF A250	1 HA pair	7.4GB/s	91.2TB / 4.4PB	25 GbE 40/100GbE	1 - 16	1-8
	4 HA pairs	29.6GB/s	364.8TB / 17.6PB		64	32
AFF C250	1 HA pair	5 GB/s	91.2TB / 4.4PB	25 GbE 40/100GbE	1-8	1-4
	4 HA pairs	20 GB/s	364.8TB / 17.6PB		32	8

1 - 1 AFF = 1 HA pair = 2 Nodes. 12 HA pairs = 24 nodes
2 - 100% sequential read

3 - Based on workload testing in NVA-1153
4 - Based on BasePOD validation test results

采用NVIDIA DGX系统的NetApp AIPod—结论和追加信息

本节包括使用NVIDIA DGX系统的NetApp AIPod的追加信息参考。

结论

DGX BasePOD架构是下一代深度学习平台、需要同等高级的存储和数据管理功能。通过将DGX BasePD与NetApp AFF系统相结合、NetApp AIPOD与DGX系统架构几乎可以在一个24节点AFF A900集群上以任何规模扩展到48个DGX H100系统。结合NetApp ONTAP卓越的云集成和软件定义的功能、AFF支持跨边缘、核心和云端的全套数据管道、助力深度学习项目取得成功。

追加信息

要详细了解本文档中所述的信息、请参阅以下文档和/或网站：

- NetApp ONTAP 数据管理软件—ONTAP 信息库
["https://docs.netapp.com/us-en/ontap-family/"](https://docs.netapp.com/us-en/ontap-family/)
- NetApp AFF A900存储系统-
["https://www.netapp.com/data-storage/aff-a-series/aff-a900/"](https://www.netapp.com/data-storage/aff-a-series/aff-a900/)
- NetApp ONTAP RDMA信息-
["https://docs.netapp.com/us-en/ontap/nfs-rdma/index.html"](https://docs.netapp.com/us-en/ontap/nfs-rdma/index.html)
- NetApp DataOps 工具包
["https://github.com/NetApp/netapp-dataops-toolkit"](https://github.com/NetApp/netapp-dataops-toolkit)
- NetApp Astra Trident
["概述"](#)
- NetApp GPUDirect存储博客-
["https://www.netapp.com/blog/ontap-reaches-171-gpudirect-storage/"](https://www.netapp.com/blog/ontap-reaches-171-gpudirect-storage/)
- NVIDIA DGX基本POD
["https://www.nvidia.com/en-us/data-center/dgx-basepod/"](https://www.nvidia.com/en-us/data-center/dgx-basepod/)
- NVIDIA DGX H100系统
["https://www.nvidia.com/en-us/data-center/dgx-h100/"](https://www.nvidia.com/en-us/data-center/dgx-h100/)
- NVIDIA网络
["https://www.nvidia.com/en-us/networking/"](https://www.nvidia.com/en-us/networking/)
- NVIDIA MAGNUM IO GPUDirect存储
["https://docs.nvidia.com/gpudirect-storage"](https://docs.nvidia.com/gpudirect-storage)
- NVIDIA基本命令
["https://www.nvidia.com/en-us/data-center/base-command/"](https://www.nvidia.com/en-us/data-center/base-command/)

- NVIDIA Base Command Manager

["https://www.nvidia.com/en-us/data-center/base-command/manager"](https://www.nvidia.com/en-us/data-center/base-command/manager)

- NVIDIA AI Enterprise

["https://www.nvidia.com/en-us/data-center/products/ai-enterprise/"](https://www.nvidia.com/en-us/data-center/products/ai-enterprise/)

致谢

本文档介绍了NetApp解决方案和ONTAP工程团队的工作：David Arnette、Olga Kornievskaia、Dups f舍尔、Srithan Kaligotla、Motit Kumar和Rajeev Badrinath。作者还要感谢NVIDIA和NVIDIA DGX BasePOD工程团队的持续支持。

NVA-1151-design：采用NVIDIA DGX A100的NetApp ONTAP AI系统设计指南

NetApp公司David Arnette和Sung-Han Lin

NVA-1151-design介绍了一种经过NetApp验证的架构、用于使用NetApp AFF A800存储系统、NVIDIA DGX A100系统和NVIDIA Mellanox网络交换机进行机器学习和人工智能工作负载。它还包括所实施架构的基准测试结果。

["NVA-1151-design：采用NVIDIA DGX A100的NetApp ONTAP AI系统设计指南"](#)

NVA-1151-Deploy：采用NVIDIA DGX A100系统的NetApp ONTAP AI

NetApp公司David Arnette

NVA-1151-deploy包括使用NetApp AFF A800存储系统、NVIDIA DGX A100系统和NVIDIA Mellanox网络交换机的适用于机器学习(ML)和人工智能(AI)工作负载的NetApp验证架构(NVA)的存储系统部署说明。此外、还提供了部署完成后运行验证基准测试的说明。

["NVA-1151-Deploy：采用NVIDIA DGX A100系统的NetApp ONTAP AI"](#)

NVA-1153-design：采用NVIDIA DGX A100系统和Mellanox Spectrum以太网交换机的NetApp ONTAP AI

NetApp公司David Arnette和Sung-Han Lin

NVA-1153-design介绍了经过NetApp验证的机器学习(ML)和人工智能(AI)工作负载架构、该架构使用NetApp AFF A800存储系统、NVIDIA DGX A100系统和NVIDIA Mellanox Spectrum SN3700V 200 Gb以太网交换机。此设计采用基于融合以太网的RDMA (RoCE)作为计算集群互连网络结构、可为客户提供完全基于以太网的架构来处理高性能工作负载。本文档还包括所实施架构的基准测试结果。

"NVA-1153-design: 采用NVIDIA DGX A100系统和Mellanox Spectrum以太网交换机的NetApp ONTAP AI"

NVA-1153-Deploy: 采用NVIDIA DGX A100系统和Mellanox Spectrum以太网交换机的NetApp ONTAP AI

NetApp公司David Arnette

NVA-1153-Deploy提供了适用于使用NetApp AFF A800存储系统、NVIDIA DGX A100系统和NVIDIA Mellanox Spectrum SN3700V 200 GB以太网交换机的经过NetApp验证的机器学习(ML)和人工智能(AI)工作负载架构的存储系统部署说明。此外、还提供了部署完成后执行验证基准测试的说明。

"NVA-1153-Deploy: 采用NVIDIA DGX A100系统和Mellanox Spectrum以太网交换机的NetApp ONTAP AI"

版权信息

版权所有 © 2024 NetApp, Inc.。保留所有权利。中国印刷。未经版权所有者事先书面许可，本档中受版权保护的任何部分不得以任何形式或通过任何手段（图片、电子或机械方式，包括影印、录音、录像或存储在电子检索系统中）进行复制。

从受版权保护的 NetApp 资料派生的软件受以下许可和免责声明的约束：

本软件由 NetApp 按“原样”提供，不含任何明示或暗示担保，包括但不限于适销性以及针对特定用途的适用性的隐含担保，特此声明不承担任何责任。在任何情况下，对于因使用本软件而以任何方式造成的任何直接性、间接性、偶然性、特殊性、惩罚性或后果性损失（包括但不限于购买替代商品或服务；使用、数据或利润方面的损失；或者业务中断），无论原因如何以及基于何种责任理论，无论出于合同、严格责任或侵权行为（包括疏忽或其他行为），NetApp 均不承担责任，即使已被告知存在上述损失的可能性。

NetApp 保留在不另行通知的情况下随时对本文档所述的任何产品进行更改的权利。除非 NetApp 以书面形式明确同意，否则 NetApp 不承担因使用本文档所述产品而产生的任何责任或义务。使用或购买本产品不表示获得 NetApp 的任何专利权、商标权或任何其他知识产权许可。

本手册中描述的产品可能受一项或多项美国专利、外国专利或正在申请的专利的保护。

有限权利说明：政府使用、复制或公开本文档受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中“技术数据权利 — 非商用”条款第 (b)(3) 条规定的限制条件的约束。

本文档中所含数据与商业产品和/或商业服务（定义见 FAR 2.101）相关，属于 NetApp, Inc. 的专有信息。根据本协议提供的所有 NetApp 技术数据和计算机软件具有商业性质，并完全由私人出资开发。美国政府对这些数据的使用权具有非排他性、全球性、受限且不可撤销的许可，该许可既不可转让，也不可再许可，但仅限在与交付数据所依据的美国政府合同有关且受合同支持的情况下使用。除本文档规定的情形外，未经 NetApp, Inc. 事先书面批准，不得使用、披露、复制、修改、操作或显示这些数据。美国政府对国防部的授权仅限于 DFARS 的第 252.227-7015(b)（2014 年 2 月）条款中明确的权利。

商标信息

NetApp、NetApp 标识和 <http://www.netapp.com/TM> 上所列的商标是 NetApp, Inc. 的商标。其他公司和产品名称可能是其各自所有者的商标。