



采用联想ThinkSystem SR670V2的NetApp AFF A400人工智能和ML模型培训

NetApp Solutions

NetApp
March 12, 2024

目录

采用联想ThinkSystem SR670V2的NetApp AFF A400人工智能和ML模型培训	1
TR-4810：《采用联想ThinkSystem SR670V2的NetApp AFF A400人工智能和ML模型培训》	1
技术概述	2
测试计划	7
测试结果	7
测试配置	7
测试操作步骤 和详细结果	9
架构调整	11
结论	11

采用联想ThinkSystem SR670V2的NetApp AFF A400人工智能和ML模型培训

TR-4810：《采用联想ThinkSystem SR670V2的NetApp AFF A400人工智能和ML模型培训》

Sathish Thyagarajan、David Arnette、NetApp Mircea Troaca、Lenovo

此解决方案 采用针对人工智能(AI)工作负载优化的NetApp存储和联想服务器、提供了一个中端集群架构。它适用于大多数计算作业为单节点(单GPU或多GPU)或分布在几个计算节点上的中小型企业。此解决方案 可与许多企业的大多数日常AI培训工作保持一致。

本文档介绍了对由八GPU联想SR670V2服务器、中端NetApp AFF A400存储系统和100GbE互连交换机组成的计算和存储配置的测试和验证。为了衡量性能、我们将ResNet50与ImageNet数据集结合使用、批大小为408、半精度、CUDA和cuDNN。这种架构为中小型企业提供了一个高效且经济高效的解决方案、这只是从需要NetApp ONTAP 云连接数据存储的企业级功能的AI计划开始。

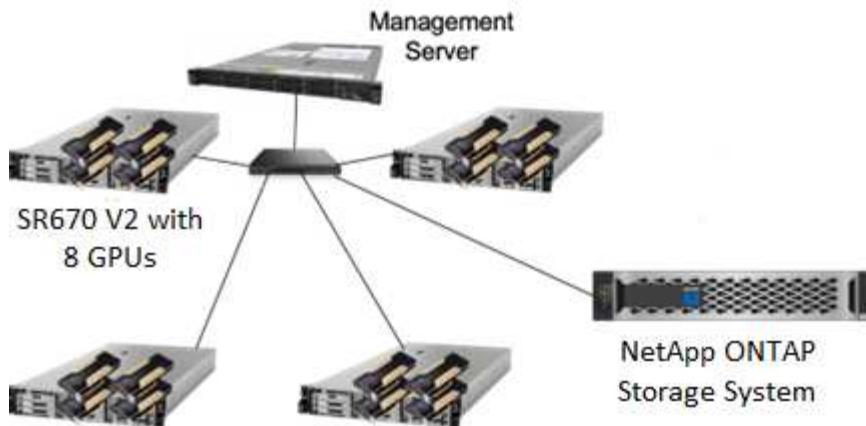
目标受众

本文档面向以下受众：

- 数据科学家、数据工程师、数据管理员和AI系统开发人员
- 设计AI模型开发解决方案的企业架构师
- 正在寻找高效方法实现深度学习(DL)和机器学习(ML)开发目标的数据科学家和数据工程师
- 希望以最快速度将AI计划推向市场的业务主管和OT/IT决策者

解决方案架构

这款采用联想ThinkSystem服务器的解决方案 以及采用AFF 存储的NetApp ONTAP 专为处理大型数据集的AI培训而设计、可利用GPU与传统CPU的处理能力。此验证通过横向扩展架构展示了高性能和最佳数据管理、该架构使用一个、两个或四个联想SR670V2服务器以及一个NetApp AFF A400存储系统。下图提供了架构概述。



NetApp 和联想解决方案具有以下主要优势：

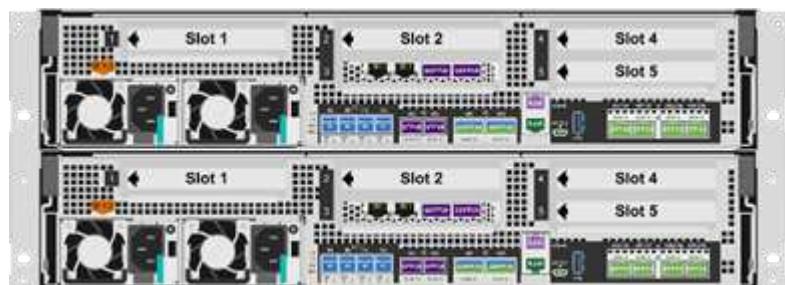
- 并行执行多个培训作业时、性能高效且经济高效
- 可根据不同数量的联想服务器和不同型号的NetApp存储控制器扩展性能
- 强大的数据保护功能、可满足低恢复点目标(RPO)和恢复时间目标(RTO)的要求、而不会丢失任何数据
- 利用快照和克隆优化数据管理、以简化开发工作流

技术概述

本节将更详细地介绍此解决方案的主要组件。

NetApp AFF 系统

借助NetApp AFF 存储系统、企业可以通过行业领先的性能、卓越的灵活性、云集成和一流的数据管理功能满足企业级存储需求。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。



NetApp AFF A400是一款中端NVMe闪存存储系统、具有以下功能：

- 最大有效容量：~20 PB
- 最大横向扩展：2-24个节点(12个HA对)
- 支持25GbE和16GB FC主机
- 通过融合以太网(RoCE)连接到NVMe扩展存储架的100GbE RDMA
- 如果未连接NVMe磁盘架、则可以使用100GbE RoCE端口进行主机网络连接
- 完整的12 Gbps SAS连接扩展存储架
- 提供两种配置：
 - 以太网：4个25 Gb以太网(SFP28)端口
 - 光纤通道：4个16 Gb FC (SFP+)端口
- 100% 8 KB随机读取@.4毫秒400、000次IOPS

适用于入门级AI/ML部署的NetApp AFF A250功能包括：

- 最大有效容量：35 PB
- 最大横向扩展：2-24个节点(12个HA对)
- 440， 000 次 IOPS 随机读取 @1 毫秒
- 基于最新的NetApp ONTAP 版本ONTAP 9.8或更高版本构建
- 两个25 Gb以太网端口、用于HA和集群互连

此外、NetApp还提供了其他存储系统、例如AFF A800和AFF A700、可为大规模AI/ML部署提供更高的性能和可扩展性。

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都可以通过一组工具来管理和保护数据。数据也可以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、以及跨混合云架构打造适应未来需求的基础架构。

简化数据管理

数据管理对于企业 IT 运营至关重要，以便将适当的资源用于应用程序和数据集。ONTAP 具有以下功能，可简化操作并降低总运营成本：

- * 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地，并分层到云。
- * 最低，最高和自适应服务质量（QoS）。 * 细粒度 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。
- * ONTAP FabricPool。 * 此功能可自动将冷数据分层到公共云和私有云存储选项、包括Amazon Web Services (AWS)、Azure和NetApp StorageGRID 对象存储。

加速和保护数据

ONTAP 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：

- * 性能和低延迟。 * ONTAP 可提供尽可能高的吞吐量，并尽可能降低延迟。
- * 数据保护。 * ONTAP 提供内置数据保护功能，并在所有平台之间进行通用管理。
- * NetApp 卷加密。 * ONTAP 提供原生卷级加密，并支持板载和外部密钥管理。

Future-Proof 基础架构

ONTAP 9有助于满足不断变化的苛刻业务需求：

- 无缝扩展和无中断运行。 ONTAP 支持向现有控制器以及横向扩展集群无中断添加容量。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 是云连接最广泛的存储管理软件，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 使用支持现有企业级应用程序的相同基础架构、为OpenStack、Hadoop 和MongoDB等下一代平台和应用程序提供企业级数据服务。

NetApp FlexGroup 卷

培训数据集通常是一组可能包含数十亿个文件的集合。文件可以包括文本，音频，视频以及其他形式的非结构化数据，这些数据必须进行存储和处理才能并行读取。存储系统必须存储许多小文件、并且必须并行读取这些文件、以便执行顺序和随机I/O

FlexGroup 卷(下图)是一个由多个成分卷组成的命名空间、存储管理员可以对其进行管理、使其与NetApp FlexVol 卷类似。 FlexGroup 卷中的文件将分配给各个成员卷，并且不会在卷或节点之间进行条带化。它们支持以下功能：

- 为高元数据工作负载提供高达20 PB的容量和可预测的低延迟
- 同一命名空间中最多可容纳4000亿个文件
- 在CPU、节点、聚合和成分卷之间的NAS工作负载中执行并行操作FlexVol



联想ThinkSystem产品组合

联想 ThinkSystem 服务器采用创新的硬件，软件和服务，可解决客户当前面临的挑战，并提供一种循序渐进的模块化设计方法来应对未来的挑战。这些服务器利用同类最佳的行业标准技术以及联想的差异化创新技术，为 x86 服务器提供最大的灵活性。

部署联想ThinkSystem服务器的主要优势包括：

- 高度可扩展的模块化设计、可随业务发展而扩展
- 行业领先的弹性，可节省数小时的成本高昂的计划外停机时间
- 快速闪存技术可实现更低的延迟，更快的响应速度以及更智能的实时数据管理

在 AI 领域，联想正在采取切实可行的方法帮助企业了解 ML 和 AI 的优势并将其用于工作负载。联想客户可以在联想 AI 创新中心探索和评估联想 AI 产品，以充分了解其特定用例的价值。为了缩短实现价值的时间、这种以客户为中心的方法可以为客户提供解决方案开发平台的概念验证、这些平台已准备就绪、可供 AI 使用并进行优化。

联想SR670V2

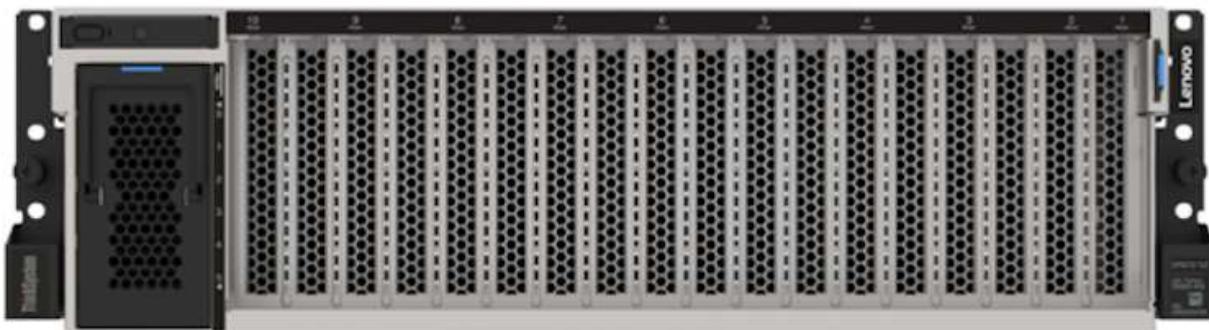
联想ThinkSystem SR670V2机架式服务器可为加速AI和高性能计算(HPC)提供最佳性能。SR670V2最多可支持八个GPU、适合ML、DL和推理的计算密集型工作负载要求。



4x SXM GPUs with 8x 2.5-inch HS drives and 2x PCIe I/O slots



4x double-wide or 8x single-wide GPU slots and 2x PCIe I/O slots
with 8x 2.5-inch or 4x 3.5-inch HS drives



8x double-wide GPU slots with 6x EDSFF HS drives and 2x PCIe I/O slots

借助支持高端GPU (包括NVIDIA A100 80 GB PCIe 8x GPU)的最新可扩展Intel Xeon CPU、ThinkSystem SR670V2可为AI和HPC工作负载提供经过优化的加速性能。

由于越来越多的工作负载使用加速器的性能、因此对GPU密度的需求也在增加。零售、金融服务、能源和医疗保健等行业正在使用GPU通过ML、DL和推理技术获得更深入的见解并推动创新。

ThinkSystem SR670V2是一款经过优化的企业级解决方案、可在生产环境中部署加速的HPC和AI工作负载、在最大程度地提高系统性能的同时、还能为采用下一代平台的超级计算集群保持数据中心密度。

其他功能包括：

- 支持GPU直接RDMA I/O、其中高速网络适配器直接连接到GPU、以最大程度地提高I/O性能。
- 支持GPU直接存储、其中NVMe驱动器直接连接到GPU、以最大程度地提高存储性能。

MLPerf

MLPerf 是用于评估 AI 性能的行业领先基准套件。在此验证中、我们使用了最受欢迎的AI框架之一MXNet的图像分类基准。我们使用了MXNet_Benchmarks培训脚本来推动AI培训。该脚本包含多种常见传统模式的实施、设计速度尽可能快。它可以在一台计算机上运行、也可以在多台主机上以分布式模式运行。

测试计划

在此验证中、我们按照MLPerf v2.0的要求执行了映像识别培训。具体来说、我们使用ImageNet数据集对RESNET v2.0模型进行了培训、直到达到76.1%的准确性。主要指标是达到所需准确性的时间。我们还会报告每秒图像数的训练带宽、以便更好地判断横向扩展效率。

主测试案例评估了多个同时运行的独立培训流程(每个节点一个)。此操作可模拟主要使用情形、即由多位数据科学家使用的共享系统。第二个测试用例评估了横向扩展效率。

测试结果

下表总结了对此解决方案 执行的所有测试的结果。

测试问题描述	结果摘要
映像识别培训：多个并发作业	高效率性能。即使集群已完全使用、所有作业也会以全速运行。NetApp存储系统可提供与本地SSD存储相当的训练性能、同时可在服务器之间轻松共享数据。
图像识别培训：横向扩展	效率高、最多可支持四个节点。当时、横向扩展效率较低、但仍然可行。使用速度更快的计算网络可提高可扩展性。NetApp存储系统可提供与本地SSD存储相当的训练性能、同时可在服务器之间轻松共享数据。

测试配置

本节介绍测试的配置、网络基础架构、SR670V2服务器以及NetApp存储配置详细信息。

解决方案架构

我们使用下表中列出的解决方案 组件进行了此验证。

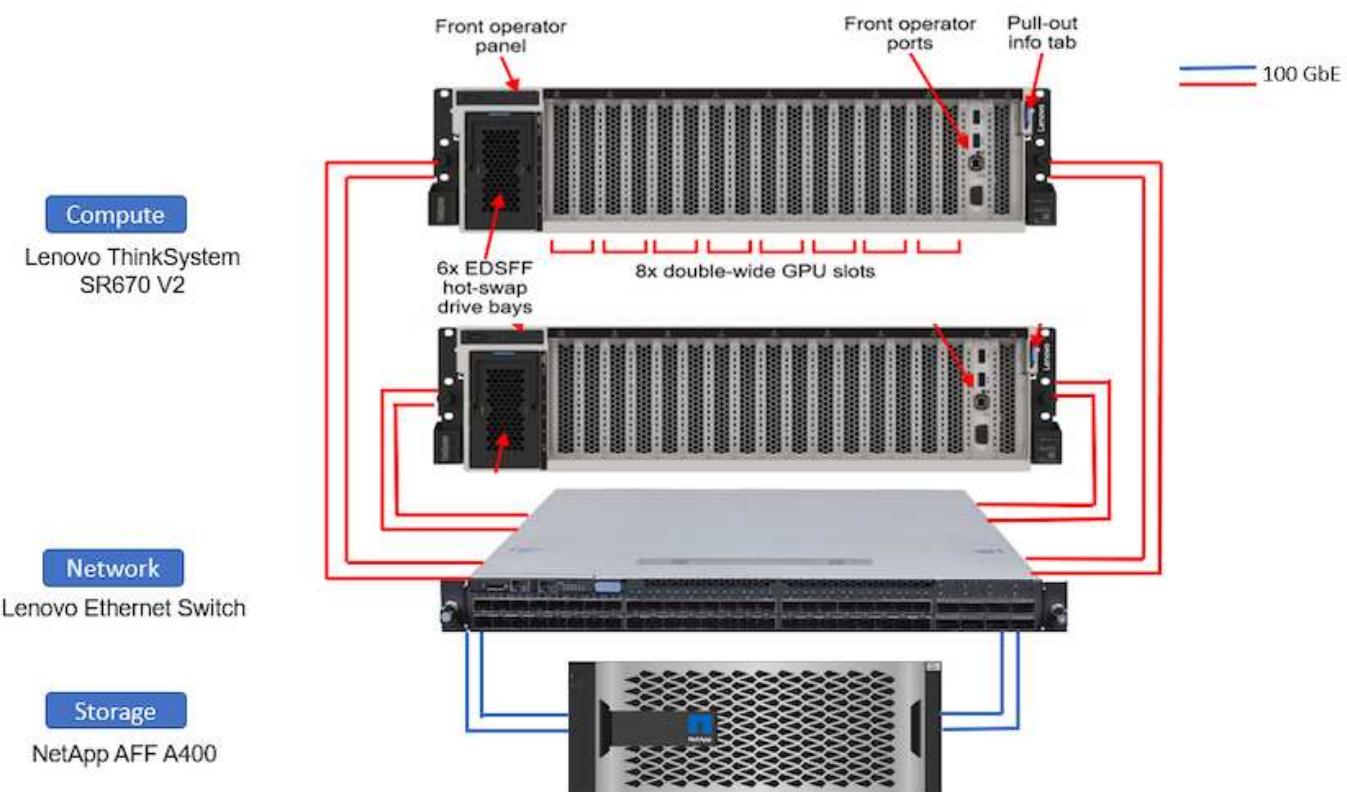
解决方案组件	详细信息
联想 ThinkSystem 服务器	<ul style="list-style-type: none">• 两个SR670V2服务器、每个服务器具有八个NVIDIA A100 80 GB GPU卡• 每台服务器包含2个Intel Xeon Platinum 8360Y CPU (28个物理核心)和1 TB RAM
Linux (Ubuntu—20.04与CUDA 11.8)	

解决方案组件	详细信息
NetApp AFF 存储系统(HA对)	<ul style="list-style-type: none"> • NetApp ONTAP 9.10.1软件 • 24 个 960 GB SSD • NFS 协议 • 每个控制器1个接口组(ifgrp)、挂载点有四个逻辑IP地址

在此验证中、我们使用的是RESNET v2.0、并且ImageNet基础设置为由MLPerf v2.0指定的基础。数据集存储在采用NFS协议的NetApp AFF 存储系统中。SR670s通过100GbE交换机连接到NetApp AFF A400存储系统。

ImageNet是一个常用的映像数据集。它包含近130万个映像、总大小为144 GB。平均映像大小为108 KB。

下图显示了测试配置的网络拓扑。



存储控制器

下表列出了存储配置。

控制器	聚合	FlexGroup 卷	聚合大小	卷大小	操作系统挂载点
控制器 1	aggr1.	/a400-100g	9.9 TB	19 TB	/a400-100g
控制器 2.	Aggr2.	/a400-100g	9.9 TB		/a400-100g



/a400-100g文件夹包含用于RESNET验证的数据集。

测试操作步骤 和详细结果

本节介绍详细的测试操作步骤 结果。

在ONTAP 中使用RESNET进行图像识别培训

我们使用一台和两台SR670V2服务器运行ResNet50基准测试。此测试使用了MXNet 22.04-py3 NGC容器来运行培训。

在此验证中、我们使用了以下测试操作步骤：

1. 在运行脚本之前、我们已清除主机缓存、以确保数据尚未缓存：

```
sync ; sudo /sbin/sysctl vm.drop_caches=3
```

2. 我们在服务器存储(本地SSD存储)以及NetApp AFF 存储系统中使用ImageNet数据集运行了基准测试脚本。
3. 我们使用验证了网络和本地存储性能 dd 命令：
4. 对于单节点运行、我们使用了以下命令：

```
python train_imagenet.py --gpus 0,1,2,3,4,5,6,7 --batch-size 408 --kv-store horovod --lr 10.5 --mom 0.9 --lr-step-epochs pow2 --lars-eta 0.001 --label-smoothing 0.1 --wd 5.0e-05 --warmup-epochs 2 --eval-period 4 --eval-offset 2 --optimizer sgdwfastlars --network resnet-v1b-stats-f1 --num-layers 50 --num-epochs 37 --accuracy-threshold 0.759 --seed 27081 --dtype float16 --disp-batches 20 --image-shape 4,224,224 --fuse-bn-relu 1 --fuse-bn-add-relu 1 --bn-group 1 --min-random-area 0.05 --max-random-area 1.0 --conv-algo 1 --force-tensor-core 1 --input-layout NHWC --conv-layout NHWC --batchnorm-layout NHWC --pooling-layout NHWC --batchnorm-mom 0.9 --batchnorm-eps 1e-5 --data-train /data/train.rec --data-train-idx /data/train.idx --data-val /data/val.rec --data-val-idx /data/val.idx --dali-dont-use-mmap 0 --dali-hw-decoder-load 0 --dali-prefetch-queue 5 --dali-nvjpeg-memory-padding 256 --input-batch-multiplier 1 --dali-threads 6 --dali-cache-size 0 --dali-roi-decode 1 --dali-preallocate-width 5980 --dali-preallocate-height 6430 --dali-tmp-buffer-hint 355568328 --dali-decoder-buffer-hint 1315942 --dali-crop-buffer-hint 165581 --dali-normalize-buffer-hint 441549 --profile 0 --e2e-cuda-graphs 0 --use-dali
```

5. 对于分布式运行、我们使用了参数服务器的并行化模型。我们为每个节点使用了两个参数服务器、并将epochchs的数量设置为与单节点运行相同。之所以这样做、是因为分布式培训往往需要更多的时间、因为各个流程之间的同步不完美。不同数量的时间间隔可能会使单节点案例与分布式案例之间的比较偏差。

数据读取速度：本地存储与网络存储

已使用测试读取速度 dd 对ImageNet数据集的其中一个文件执行命令。具体来说、我们对本地数据和网络数据运行以下命令：

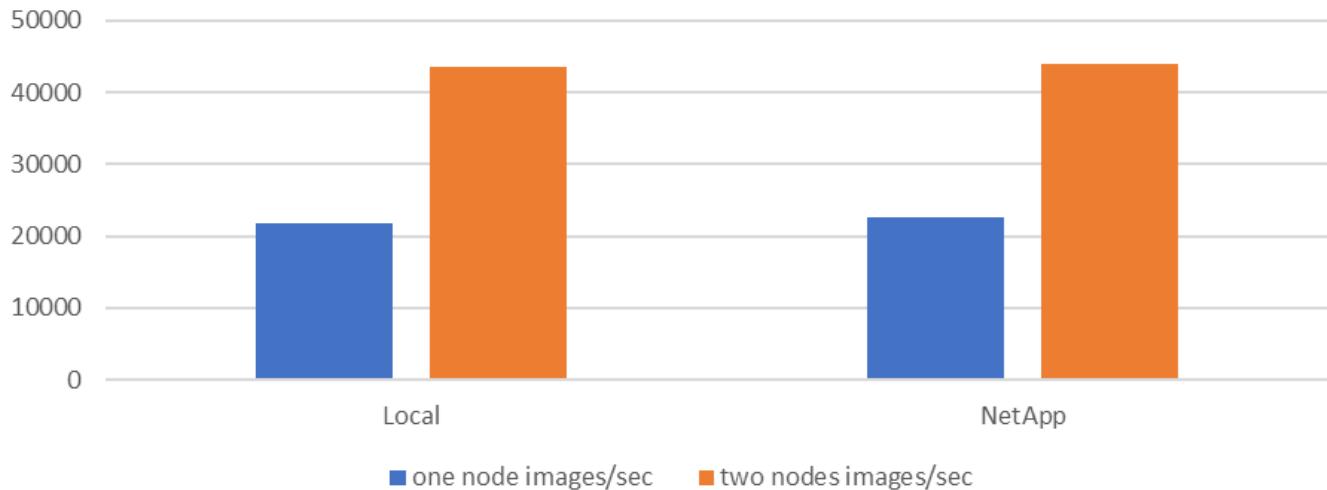
```
sync ; sudo /sbin/sysctl vm.drop_caches=3dd if=/a400-100g/netapp-ra/resnet/data/preprocessed_data/train.rec of=/dev/null bs=512k count=2048Results (average of 5 runs):  
Local storage: 1.7 GB/s Network storage: 1.5 GB/s.
```

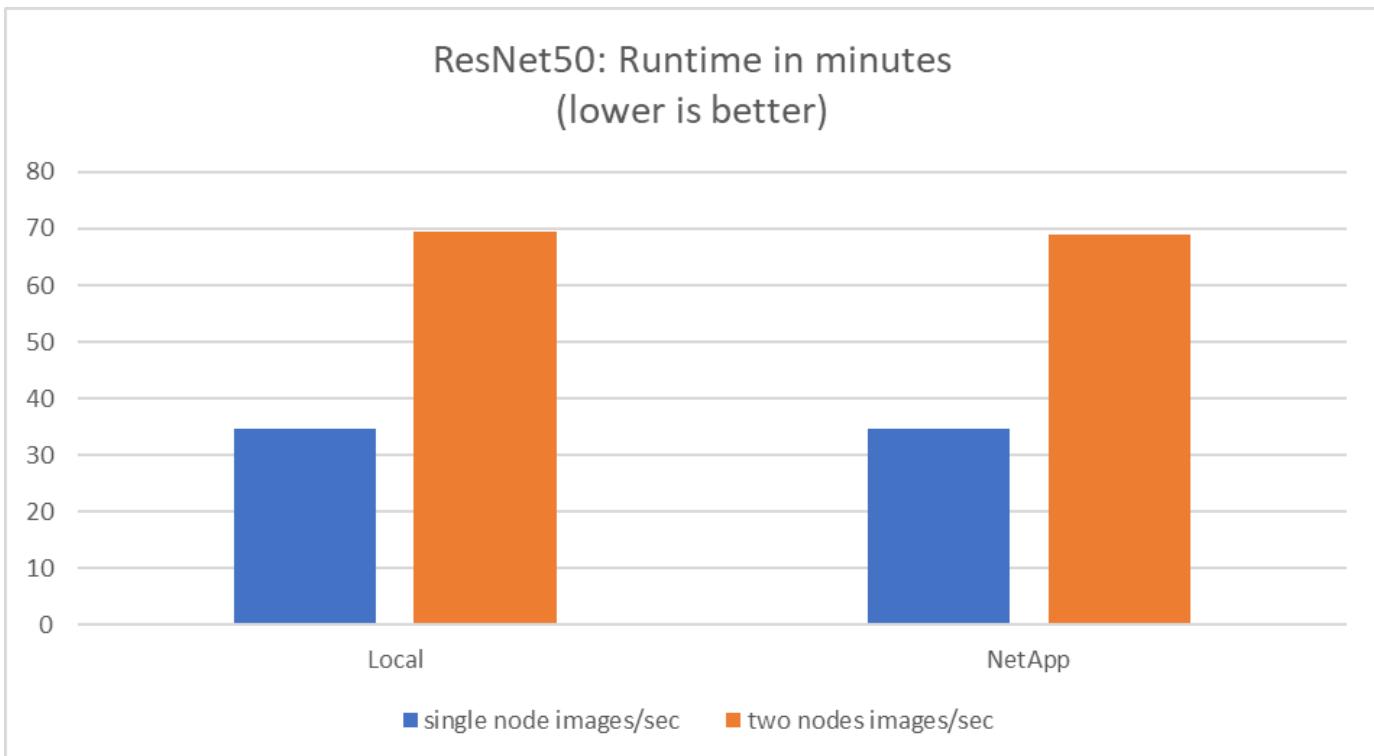
这两个值都相似、表明网络存储可以以类似于本地存储的速率提供数据。

共享使用情形：多个独立的并发作业

此测试模拟了此解决方案 的预期用例：多作业、多用户AI培训。每个节点在使用共享网络存储时都进行了自己的培训。下图显示了这些结果、其中显示了解决方案 案例在所有作业以与单个作业基本相同的速度运行时提供了出色的性能。总吞吐量随节点数线性扩展。

ResNet50: Aggregate Images per second
(higher is better)





这些图以分钟为单位显示了计算节点的运行时间、每秒聚合映像数、这些计算节点在100 GbE客户端网络上使用了每个服务器的八个GPU、并结合了并发训练模型和单个训练模型。此训练模型的平均运行时间为35分9秒。单个运行时间分别为34分32秒、36分21秒、34分37秒、35分25秒和34分31秒。训练模型的平均每秒图像数为22、673个、每秒单个图像数为21、764个、23、438个、22、556个、22、264个和22、548个。

根据我们的验证、一个采用NetApp数据运行时的独立训练模型为34分54秒、每秒显示22、231个图像一个采用本地数据(DAS)运行时间的独立训练模型为34分21秒、每秒显示22、102个图像在这些运行期间、平均GPU利用率为96%、如NVIDIA-SMI上所观察到的那样。请注意、此平均值包括测试阶段、在此阶段、不使用GPU、而使用mpstat测量的CPU利用率为40%。这表明、在每种情况下、数据交付率都足以满足要求。

架构调整

可根据其他使用情形调整用于此验证的设置。

CPU调整

我们按照联想的建议、使用Skylake Intel Xeon Platinum 8360Y处理器进行此验证。我们预计同等的级联湖CPU (Intel Xeon Gold 6330处理器)的性能将相似、因为此工作负载不受CPU限制。

存储容量增加

根据您的存储容量需求、您可以按需增加共享存储(NFS卷)、但前提是您拥有更多的磁盘架和控制器型号。您可以通过CLI或存储控制器的NetApp Web界面以管理员用户身份执行此操作。

结论

此处验证的NetApp和联想解决方案 是一种灵活的横向扩展架构、非常适合入门级企业AI。

NetApp存储可提供与本地SSD存储相同或更好的性能、并为数据科学家、数据工程师和IT决策者带来以下优势：

- 在 AI 系统，分析和其他关键业务系统之间轻松共享数据。这种数据共享可减少基础架构开销，提高性能并简化整个企业的数据管理。
- 可独立扩展的计算和存储、最大限度地降低成本并提高资源利用率。
- 利用集成快照和克隆简化开发和部署工作流、实现瞬时且节省空间的用户工作空间、集成版本控制和自动化部署。
- 企业级数据保护、可实现灾难恢复和业务连续性。

致谢

- NetApp技术营销工程师Karthikeyan Nagalingam
- 联想AI实验室系统部门管理员Jarrett Upton

从何处查找追加信息

要了解有关本文档中所述信息的更多信息，请参见以下文档和 / 或网站：

- NetApp全闪存阵列产品页面
["https://www.netapp.com/us/products/storage-systems/all-flash-array/aff-a-series.aspx"](https://www.netapp.com/us/products/storage-systems/all-flash-array/aff-a-series.aspx)
- NetApp AFF A400页面
["https://docs.netapp.com/us-en/ontap-systems/a400/index.html"](https://docs.netapp.com/us-en/ontap-systems/a400/index.html)
- NetApp ONTAP 数据管理软件产品页面
["http://www.netapp.com/us/products/data-management-software/ontap.aspx"](http://www.netapp.com/us/products/data-management-software/ontap.aspx)
- MLPerf
["https://mlperf.org"](https://mlperf.org)
- TensorFlow 基准测试
["https://github.com/tensorflow/benchmarks"](https://github.com/tensorflow/benchmarks)
- NVIDIA SMI (NVIDIA-SMI)
["https://developer.nvidia.com/nvidia-system-management-interface"](https://developer.nvidia.com/nvidia-system-management-interface)

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，本文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。