TR-4810: NetApp AFF A400 com Lenovo ThinkSystem SR670 V2 para treinamento de modelos de IA e ML

NetApp artificial intelligence solutions

NetApp
August 18, 2025

This PDF was generated from https://docs.netapp.com/pt-br/netapp-solutions-ai/infra/ai-lenovo-train-intro.html on August 18, 2025. Always check docs.netapp.com for the latest.

Índice

TR-4810: NetApp AFF A400 com Lenovo ThinkSystem SR670 V2 para treinamento de modelos de IA e ML . . . 1
Público-alvo . . . 1
Arquitetura da solução . . . 1

TR-4810: NetApp AFF A400 com Lenovo ThinkSystem SR670 V2 para treinamento de modelos de IA e ML

Sathish Thyagarajan, David Arnette, NetApp Mircea Troaca, Lenovo

Esta solução apresenta uma arquitetura de cluster de médio porte usando armazenamento NetApp e servidores Lenovo otimizados para cargas de trabalho de inteligência artificial (IA). Ele é destinado a pequenas e médias empresas para as quais a maioria dos trabalhos de computação são de nó único (GPU única ou múltipla) ou distribuídos em alguns nós computacionais. Esta solução se alinha à maioria dos trabalhos diários de treinamento de IA de muitas empresas.

Este documento abrange testes e validação de uma configuração de computação e armazenamento composta por servidores Lenovo SR670V2 de oito GPUs, um sistema de armazenamento NetApp AFF A400 de médio porte e um switch de interconexão de 100 GbE. Para medir o desempenho, usamos o ResNet50 com o conjunto de dados ImageNet, um tamanho de lote de 408, meia precisão, CUDA e cuDNN. Essa arquitetura fornece uma solução eficiente e econômica para organizações de pequeno e médio porte que estão começando com iniciativas de IA e que exigem recursos de nível empresarial do armazenamento de dados conectado à nuvem NetApp ONTAP .

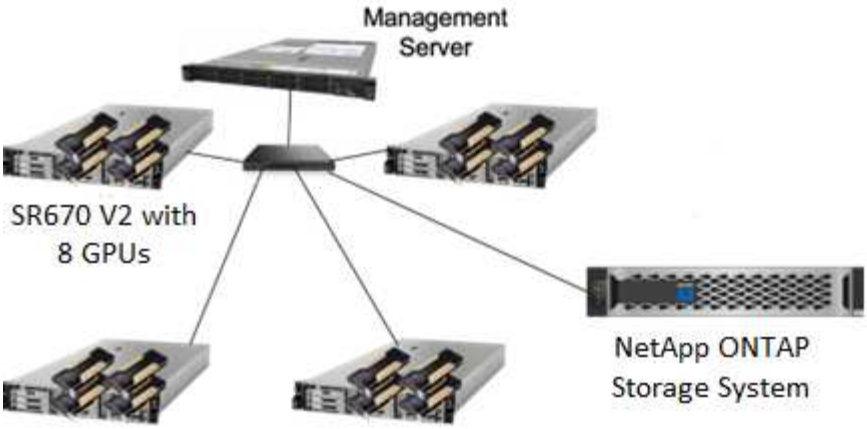
Público-alvo

Este documento é destinado aos seguintes públicos:

- Cientistas de dados, engenheiros de dados, administradores de dados e desenvolvedores de sistemas de IA
- Arquitetos corporativos que projetam soluções para o desenvolvimento de modelos de IA
- Cientistas e engenheiros de dados que buscam maneiras eficientes de atingir metas de desenvolvimento de aprendizado profundo (DL) e aprendizado de máquina (ML)
- Líderes empresariais e tomadores de decisão de TO/TI que desejam atingir o tempo de comercialização mais rápido possível para iniciativas de IA

Arquitetura da solução

Esta solução com servidores Lenovo ThinkSystem e NetApp ONTAP com armazenamento AFF foi projetada para lidar com treinamento de IA em grandes conjuntos de dados usando o poder de processamento de GPUs junto com CPUs tradicionais. Esta validação demonstra alto desempenho e gerenciamento de dados ideal com uma arquitetura de escalonamento que usa um, dois ou quatro servidores Lenovo SR670 V2 junto com um único sistema de armazenamento NetApp AFF A400 . A figura a seguir fornece uma visão geral da arquitetura.



Esta solução da NetApp e da Lenovo oferece os seguintes benefícios principais:

- Desempenho altamente eficiente e econômico ao executar vários trabalhos de treinamento em paralelo
- Desempenho escalável com base em diferentes números de servidores Lenovo e diferentes modelos de controladores de armazenamento NetApp
- Proteção de dados robusta para atender a objetivos de ponto de recuperação (RPOs) e tempo de recuperação (RTOs) baixos sem perda de dados
- Gerenciamento de dados otimizado com snapshots e clones para agilizar os fluxos de trabalho de desenvolvimento

Informações sobre direitos autorais

Copyright © 2025 NetApp, Inc. Todos os direitos reservados. Impresso nos EUA. Nenhuma parte deste documento protegida por direitos autorais pode ser reproduzida de qualquer forma ou por qualquer meio — gráfico, eletrônico ou mecânico, incluindo fotocópia, gravação, gravação em fita ou storage em um sistema de recuperação eletrônica — sem permissão prévia, por escrito, do proprietário dos direitos autorais.

O software derivado do material da NetApp protegido por direitos autorais está sujeito à seguinte licença e isenção de responsabilidade:

ESTE SOFTWARE É FORNECIDO PELA NETAPP "NO PRESENTE ESTADO" E SEM QUAISQUER GARANTIAS EXPRESSAS OU IMPLÍCITAS, INCLUINDO, SEM LIMITAÇÕES, GARANTIAS IMPLÍCITAS DE COMERCIALIZAÇÃO E ADEQUAÇÃO A UM DETERMINADO PROPÓSITO, CONFORME A ISENÇÃO DE RESPONSABILIDADE DESTE DOCUMENTO. EM HIPÓTESE ALGUMA A NETAPP SERÁ RESPONSÁVEL POR QUALQUER DANO DIRETO, INDIRETO, INCIDENTAL, ESPECIAL, EXEMPLAR OU CONSEQUENCIAL (INCLUINDO, SEM LIMITAÇÕES, AQUISIÇÃO DE PRODUTOS OU SERVIÇOS SOBRESSALENTES; PERDA DE USO, DADOS OU LUCROS; OU INTERRUPÇÃO DOS NEGÓCIOS), INDEPENDENTEMENTE DA CAUSA E DO PRINCÍPIO DE RESPONSABILIDADE, SEJA EM CONTRATO, POR RESPONSABILIDADE OBJETIVA OU PREJUÍZO (INCLUINDO NEGLIGÊNCIA OU DE OUTRO MODO), RESULTANTE DO USO DESTE SOFTWARE, MESMO SE ADVERTIDA DA RESPONSABILIDADE DE TAL DANO.

A NetApp reserva-se o direito de alterar quaisquer produtos descritos neste documento, a qualquer momento e sem aviso. A NetApp não assume nenhuma responsabilidade nem obrigação decorrentes do uso dos produtos descritos neste documento, exceto conforme expressamente acordado por escrito pela NetApp. O uso ou a compra deste produto não representam uma licença sob quaisquer direitos de patente, direitos de marca comercial ou quaisquer outros direitos de propriedade intelectual da NetApp.

O produto descrito neste manual pode estar protegido por uma ou mais patentes dos EUA, patentes estrangeiras ou pedidos pendentes.

LEGENDA DE DIREITOS LIMITADOS: o uso, a duplicação ou a divulgação pelo governo estão sujeitos a restrições conforme estabelecido no subparágrafo (b)(3) dos Direitos em Dados Técnicos - Itens Não Comerciais no DFARS 252.227-7013 (fevereiro de 2014) e no FAR 52.227- 19 (dezembro de 2007).

Os dados aqui contidos pertencem a um produto comercial e/ou serviço comercial (conforme definido no FAR 2.101) e são de propriedade da NetApp, Inc. Todos os dados técnicos e software de computador da NetApp fornecidos sob este Contrato são de natureza comercial e desenvolvidos exclusivamente com despesas privadas. O Governo dos EUA tem uma licença mundial limitada, irrevogável, não exclusiva, intransferível e não sublicenciável para usar os Dados que estão relacionados apenas com o suporte e para cumprir os contratos governamentais desse país que determinam o fornecimento de tais Dados. Salvo disposição em contrário no presente documento, não é permitido usar, divulgar, reproduzir, modificar, executar ou exibir os dados sem a aprovação prévia por escrito da NetApp, Inc. Os direitos de licença pertencentes ao governo dos Estados Unidos para o Departamento de Defesa estão limitados aos direitos identificados na cláusula 252.227-7015(b) (fevereiro de 2014) do DFARS.

Informações sobre marcas comerciais

NETAPP, o logotipo NETAPP e as marcas listadas em http://www.netapp.com/TM são marcas comerciais da NetApp, Inc. Outros nomes de produtos e empresas podem ser marcas comerciais de seus respectivos proprietários.