

NVIDIA DGX SUPERPOD COM STORAGE NETAPP ONTAP



Acelerando a inovação em IA em escala



À medida que as organizações adotam iniciativas de inteligência artificial (AI) e aprendizado de máquina (ML), a demanda por soluções de infraestrutura robustas, dimensionáveis e eficientes nunca foi tão grande. O AI Center of Excellence (AI CoE), ou AI Factory, surgiu como um componente crucial na estratégia de IA corporativa, servindo como um hub centralizado para o desenvolvimento e implantação de soluções de IA. No centro dessas iniciativas, está o desafio de gerenciar e treinar modelos de IA cada vez mais complexos, além de garantir a segurança dos dados, a acessibilidade e a otimização de recursos.

A evolução da IA agêntica, que pode tomar decisões, Planejar e se adaptar para atingir metas específicas sem muita contribuição humana, criou demandas sem precedentes de infraestrutura computacional e de storage. Agora, as organizações precisam lidar com grandes conjuntos de dados, dar suporte a vários workloads de treinamento simultâneos e manter ambientes de computação de alta performance, protegendo os dados e cumprindo com as regulamentações.

As soluções de infraestrutura tradicionais podem ter dificuldades para atender a essas demandas, levando a ineficiências operacionais e atraso no retorno do investimento dos projetos de IA.

A solução

O NVIDIA DGX SuperPOD™ com o storage NetApp® ONTAP® combina a performance de computação dos sistemas NVIDIA DGX™ com o recurso de IA empresarial de classe mundial dos sistemas de storage da NetApp AFF A90 conectados à nuvem. Assim, é possível realizar workflows orientados pelos dados para ML, IA e computação técnica de alta performance (HPC). Essa solução integrada oferece às organizações uma plataforma dimensionável, segura e eficiente para desenvolvimento e implantação de IA. A arquitetura pode eliminar silos de infraestrutura tradicionais, permitindo uma colaboração otimizada entre cientistas de dados, engenheiros e outras partes interessadas.

Acesso e gerenciamento de dados

O NVIDIA DGX SuperPOD com storage NetApp ONTAP implementa uma tecnologia avançada de redução de dados que otimiza a utilização do storage e mantém a performance. Com os recursos do NetApp Snapshot™ e do FlexClone® do ONTAP, as equipes criam instantaneamente cópias com uso eficiente de espaço de conjuntos de dados para desenvolvimento e teste paralelos. Várias equipes de IA podem trabalhar simultaneamente em versões diferentes do mesmo conjunto de dados sem consumir mais espaço de storage nem comprometer a integridade dos dados.

Além disso, com a estrutura unificada de acesso aos dados da solução, as equipes podem integrar facilmente os fluxos de trabalho e as ferramentas existentes. Cientistas e engenheiros de dados podem utilizar protocolos e interfaces familiares enquanto se beneficiam de recursos de gerenciamento e proteção de dados de nível empresarial. A disposição inteligente de dados em categorias da plataforma otimiza automaticamente o posicionamento dos dados entre camadas de storage. Portanto, os conjuntos de dados acessados com frequência permanecem prontamente disponíveis, enquanto os dados usados com menos frequência são migrados para um storage mais econômico. Essa abordagem inteligente para o gerenciamento de dados, combinada às funcionalidades de namespace global do ONTAP, fornece às equipes de IA uma experiência de acesso aos dados eficiente e unificada.

Escalabilidade

O NVIDIA DGX SuperPOD com storage NetApp ONTAP oferece orquestração inteligente de workload que equilibra recursos automaticamente, otimizando a performance em trabalhos simultâneos de treinamento e workloads de inferência. As tecnologias de eficiência de storage da ONTAP, como deduplicação e compactação, maximizam a capacidade de storage à medida que os dados crescem.

Ao quebrar silos de infraestrutura e otimizar o gerenciamento de recursos nos AI Centers of Excellence, as organizações podem eliminar sistemas redundantes e reduzir a sobrecarga operacional. Essa infraestrutura unificada acelera os cronogramas do projeto de IA e melhora o retorno dos investimentos em tecnologia.

PRINCIPAIS BENEFÍCIOS

Gerenciamento de dados empresariais

Uma plataforma centralizada que fornece controle abrangente sobre o ciclo de vida dos dados de IA.

Acesso a dados acelerado

Acesso rápido e seguro aos dados em toda a organização.

Dimensionamento da infraestrutura otimizado

Uma plataforma unificada que cresce de acordo com suas necessidades, da prova de conceito à implantação em toda a empresa.

Proteção completa dos dados

Proteção de dados de nível empresarial que protege informações confidenciais em todo o ciclo de vida da IA

Segurança Zero Trust

Segurança com uma arquitetura Zero Trust que protege sua infraestrutura de IA em todas as camadas

As organizações podem escalar de forma otimizada adicionando nós de computação DGX e sistemas de storage ONTAP sem interrupção. A arquitetura dá suporte a abordagens de escalabilidade vertical e horizontal para poder computacional, memória e armazenamento. A malha unificada mantém alto desempenho e baixa latência em qualquer escala, otimizando a eficiência dos workloads de IA.

Segurança

Com as medidas de segurança robustas da NetApp, como controle de acesso baseado em funções (RBAC), autenticação multifator e logs de auditoria, as organizações podem buscar iniciativas de IA com confiança, ao mesmo tempo em que protegem a propriedade intelectual e mantêm a conformidade. Com essa proteção, somente a equipe autorizada e os agentes de IA podem acessar modelos confidenciais, dados de treinamento e dados empresariais.



A solução protege os dados por meio de criptografia abrangente em repouso e em trânsito, além de fornecer monitoramento de segurança integrado para detecção de ameaças em tempo real e resposta automatizada. Os recursos integrados de Snapshot e backup permitem uma recuperação rápida de incidentes de segurança, minimizando o tempo de inatividade e a perda de dados.

Essa sólida base de segurança permite que as empresas inovem com mais rapidez, colaborem com mais eficiência com os parceiros e escalem as operações de IA, mantendo uma excelente proteção de dados.

Conclusão

O NVIDIA DGX SuperPOD com a solução de storage NetApp ONTAP representa um avanço significativo nas soluções de infraestrutura de IA. Ao enfrentar os principais desafios relacionados à segurança, ao gerenciamento de dados, ao uso de recursos e à escalabilidade, ele permite que as organizações acelerem suas iniciativas de IA, mantendo a eficiência operacional, a proteção de dados e a colaboração. A abordagem integrada da solução elimina gargalos comuns nos pipelines de desenvolvimento de IA, permitindo que cientistas e engenheiros de dados se concentrem na inovação e não no gerenciamento da infraestrutura. À medida que a IA continua a evoluir e impulsionar a transformação digital em todos os setores, a necessidade de soluções de infraestrutura fortes, dimensionáveis e seguras só crescerá. O NVIDIA DGX SuperPOD com o NetApp ONTAP Storage fornece às organizações uma base para iniciativas de IA atuais e futuras, permitindo que elas mantenham vantagem competitiva em um mundo cada vez mais orientado por IA.



Fale conosco



Sobre a NetApp

A NetApp é a empresa de infraestrutura de dados inteligente, que combina storage de dados unificado, serviços de dados integrados e soluções de CloudOps, transformando disrupções tecnológicas em oportunidades para todos os clientes. A NetApp ajuda a criar uma infraestrutura sem silos, aproveitando recursos de observabilidade e IA para oferecer o melhor gerenciamento de dados do setor. Como o único serviço de storage de nível empresarial incorporado de maneira nativa às maiores nuvens do mundo, nosso storage oferece flexibilidade aprimorada. Além disso, nossos serviços de dados criam uma vantagem com resiliência cibernética, e governança e agilidade dos aplicativos. Nossas soluções CloudOps fornecem otimização contínua de desempenho e eficiência por meio de observabilidade e IA. Não importa o tipo de dados, workload ou ambiente, transforme sua infraestrutura de dados para obter ótimos resultados em seus negócios com a NetApp. www.netapp.com

© 2025 NetApp, Inc. Todos os direitos reservados. 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. SB-4321-0225-ptBR