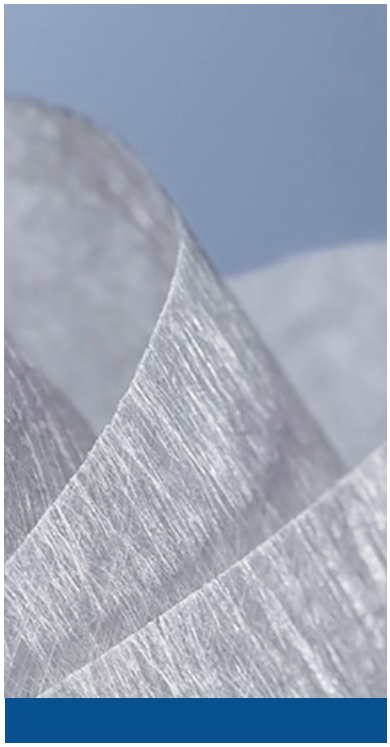

GPU-GESTÜTZTE AI-INFRASTRUKTUR FÜR EFFIZIENTES MLOPS IN EINER GESCHÜTZTEN NETZWERKUMGEBUNG

Freudenberg Performance Materials setzt gemeinsam mit SVA auf leistungsfähige, skalierbare AI-Plattform.

AUF EINEN BLICK

AUFGABE

Integration einer GPU-gestützten AI-Plattform zur effizienten Entwicklung, Skalierung und Bereitstellung von KI-Modellen und Applikationen in einer geschützten Netzwerkumgebung

SYSTEME UND SOFTWARE

> HPE AI Workbench
> HPE ProLiant Server
> VMware ESXi vSphere
> NVIDIA vGPU
> Rancher Kubernetes

VORTEILE

> skalierbare und flexible AI-Infrastruktur zur effizienten Nutzung von GPU-Ressourcen
> automatisierte und optimierte Modelltrainings durch eine zentrale AI-Plattform
> einfache Verwaltung und Integration in bestehende IT-Systeme mit Best Practices für die IT-Administration

FREUDENBERG PERFORMANCE MATERIALS

Freudenberg Performance Materials ist ein weltweit führender Anbieter innovativer technischer Materialstoffe und Textilien für eine große Bandbreite an Märkten und Anwendungen. Das Unternehmen erwirtschaftete 2024 einen Umsatz von mehr als 1,4 Milliarden Euro und hat weltweit 35 Produktionsstandorte in 14 Ländern und beschäftigt rund 5.000 Mitarbeitende. Mit Lösungen in den Bereichen Dichtungen, Schwingungstechnik, Filter, Oberflächenbehandlung, Schmierstoffe, Trennmittel, Spezialtextilien und Medizintechnik treibt Freudenberg Performance Materials die technologische Entwicklung in vielen Branchen voran. Die Nutzung moderner KI-Technologien spielt eine zunehmend wichtige Rolle für effizientere Prozesse, datengetriebene Produktentwicklungen und intelligente Automatisierung.

HERAUSFORDERUNG

Für anspruchsvolle KI-Modelle und deren Entwicklung ist eine leistungsfähige und gleichzeitig flexible Infrastruktur erforderlich. Besonders in einer Proxy-geschützten Netzwerkumgebung sind der Aufbau und die Verwaltung von AI-Clustern eine Herausforderung. Freudenberg Performance Materials benötigte eine skalierbare Lösung, die eine GPU-beschleunigte Entwicklungsumgebung ermöglicht, sichere und effiziente Modelltrainingsprozesse gewährleistet und gleichzeitig in die bestehende IT-Landschaft integriert werden kann. Darüber hinaus war eine Best-Practice-Beratung für den operativen Betrieb sowie eine konzeptionelle Unterstützung zur Anbindung zukünftiger Produktivsysteme gewünscht.

LÖSUNG

Um diesen Anforderungen gerecht zu werden, sollte in enger Zusammenarbeit mit den Experten von SVA eine maßgeschneiderte AI-Infrastruktur aufgebaut werden. Zunächst erfolgte die Integration von NVIDIA AI Enterprise auf HPE Hostsystemen mit VMware ESXi vSphere, um GPU-Ressourcen virtualisiert bereitzustellen und eine flexible Entwicklungs-

FREUDENBERG PERFORMANCE MATERIALS

Hewlett Packard Enterprise Solution Provider PLATINUM PARTNER

Hewlett Packard Enterprise

STABILE UND SKALIERBARE AI-PLATTFORM

umgebung mit hoher Rechenleistung zu schaffen. Anschließend wurde innerhalb der geschützten Proxy-Umgebung ein Kubernetes Cluster mit der Rancher Kubernetes Engine (RKE) ausgerollt, der eine stabile und skalierbare Plattform für AI-Workloads und deren Entwicklung bietet. Die Kernkomponenten der HPE AI-Workbench bieten zahlreiche Vorteile:

> Ressourcenverwaltung: Erzeugung von GPU-Partitionen mehrerer Ressourcengrößen mittels vGPU und deren Verwaltung als Computepools mit prioritätsbasierendem Scheduling, Ad-hoc-Entwicklungsumgebungen und preemtive Tasking in der AI Plattform

> MLOps-Modellverwaltung & -governance: Verwaltung der Experimentergebnisse und Artefakte des MLOps-Zyklus mit Freigabe- & Versionierungsprozessen zur Trennung von experimentellen und produktiven Modellen

> Vereinheitlichung der Runtimes: Standardisierte, containerisierte Laufzeitumgebungen mit Einbindung der Plattformschnittstellen für den gesamtheitlichen MLOps-Zyklus und -Betrieb

> Standardisierung & Erweiterbarkeit der Produkte: Vereinheitlichung und Automatisierbarkeit durch einen zentralisierten API-Katalog mit Schnittstellen für MLOps & Produktbetrieb innerhalb der Plattformlösung und durch exponierte Schnittstellen

Auf dieser Grundlage wurde eine AI-Plattform implementiert und konfiguriert, um eine effiziente Verwaltung und Automatisierung von Modelltrainings zu ermöglichen. Das IT-Team der Freudenberg Performance Materials wurde dazu durch gezielte Beratungen dabei unterstützt, Best Practices im Bereich Modelltraining und -Serving zu etablieren und die langfristige Nutzung der neuen Infrastruktur optimal zu gestalten.

FAZIT

Die Zusammenarbeit mit den Experten von SVA ermöglichte es der Freudenberg Performance Materials, eine leistungsfähige, GPU-optimierte AI-Plattform aufzubauen, die sich sicher in die bestehende IT-Infrastruktur einfügt. Die Kombination aus vGPU-Virtualisierung, Kubernetes-basierendem Cluster-Management und einer spezialisierten AI-Plattform gewährleistet nun die effizientere Entwicklung, ein schnelleres Training von Modellen und eine flexible Skalierung. Mit dieser leistungsstarken Lösung schafft das Unternehmen die Grundlage für eine nachhaltige Nutzung von Künstlicher Intelligenz in Forschung und Entwicklung und setzt auf eine zukunftssichere Infrastruktur.

KONTAKT
SVA System Vertrieb
Alexander GmbH
Borsigstraße 26
65205 Wiesbaden
Tel. +49 6122 536-0
mail@sva.de
www.sva.de

FREUDENBERG
PERFORMANCE MATERIALS

Hewlett Packard Enterprise
Solution Provider
PLATINUM PARTNER

Hewlett Packard
Enterprise