



REFERENZ HPI

KI-Expertise auf Basis neuer IT-Infrastruktur für Hasso-Plattner-Institut

↳ Dank SVA Expertise und NVIDIA-Technologie bietet KI-Servicezentrum Beratung auf hohem Niveau.

AUF EINEN BLICK

Aufgabe

Bereitstellung einer leistungsfähigen KI-Infrastruktur bestehend aus Trainings-Cluster und Inference-Cluster

Systeme & Software

- 1 x Supermicro x NVRAM-Server
- 1 x Supermicro ARM-Server mit NVIDIA L40 GPU
- 1 x Supermicro Medium-GPU-Server mit 8 x NVIDIA L40s
- 4 x Supermicro Virtualisierung-Server
- 5 x Supermicro Inference-Server mit je 8 x NVIDIA A30 GPU
- ...

KI-SERVICEZENTRUM BERLIN-BRANDENBURG AM HPI

Das KI-Servicezentrum Berlin-Brandenburg ist ein vom Bundesministerium für Forschung, Technologie und Raumfahrt gefördertes Projekt am Hasso-Plattner-Institut (HPI) mit dem Ziel, den allgemeinen Zugang zur Schlüsseltechnologie „Künstliche Intelligenz“ zu erleichtern und die Barrieren für den Einsatz von KI in Wirtschaft und Gesellschaft zu senken.

Es unterstützt Start-ups, Unternehmen, öffentliche Einrichtungen, Studierende und die Forschung dabei, KI-Anwendungen zu entwickeln, zu betreiben und einzusetzen. Seit seiner offiziellen Eröffnung im Oktober 2023 konnte das KI-Servicezentrum in über 100 Workshops bereits über 1.000 Teilnehmende zählen und rund 100 Organisationen bei KI-Anwendungsfällen beraten. Die Einsatzbereiche reichen dabei von der Öffentlichen Verwaltung, über Finanzen bis hin zu Pflege und Medizin.

HERAUSFORDERUNG

Das KI-Servicezentrum plante, mit einer KI-Infrastruktur Plattformen für Trainings, Inference und Systemintegration anzubieten und diese KI-Anwenderinnen und -Anwendern in ganz Deutschland kostenlos zur Verfügung zu stellen. Diese Infrastruktur

Systeme & Software

- ...
- 8 x NVIDIA DGX H100
- 1 x VAST-DATA Software Defined Storage Solution mit 1.5 PB Speichervolumen
- 2 x NVIDIA 200 GbE Ethernet Switch
- 2 x NVIDIA NDR Infiniband Switch

Vorteile

- hochmodernes, leistungsfähiges KI-Cluster
- Hardware für alle Einsatzzwecke von Entwicklung, über Training bis Inferenz
- technische Basis für Angebot

sollte sinnhaft, performant und skalierbar sein. Aufgrund befristeter Fördergelder war ein fester Zeitrahmen gegeben, der jedoch durch die allgemeinen Lieferverzögerungen bei der Hardware herausfordernd wurde. Gute Erfahrungen mit NVIDIA, deren State-of-the-Art-Komponenten, die einen Großteil der Use Cases bedienen konnten, sowie deren rasche Lieferbarkeit legten den Fokus auf die Lösungen dieses Herstellers.

LÖSUNG

Der Aufbau des KI-Rechenzentrums begann mit einer detaillierten Analyse der bestehenden Infrastruktur. Dabei setzten sich die SVA-Experten intensiv mit den spezifischen Anforderungen und Zielen des KI-Servicezentrums auseinander, berieten ganzheitlich und entwickelten ein maßgeschneidertes Konzept. Ein wesentlicher Bestandteil der neuen Infrastruktur sind mehrere Server mit leistungsstarker GPU-Hardware, NVIDIA DGX-Systeme und Supermicro GPU-Systeme, die für das Training und Inferencing von KI-Modellen – bspw. Large Language Models (LLMs) – insbesondere mit externen Partnerorganisationen genutzt werden können. Diese Erweiterung der Rechenleistung steigert die Datenverarbeitungs- und Analysefähigkeiten enorm und versetzt das HPI in die Lage, hochentwickelte, KI-basierende Lösungen anzubieten.

Der so geschaffene Zugang zu einem leistungsfähigen und heterogenen KI-Rechenzentrum ermöglicht die Durchführung von Forschungsprojekten und prototypischen Entwicklungsprojekten.

FAZIT

Durch die Implementierung der IT-Infrastrukturlösung konnte das HPI mehrere externe Organisationen im Rahmen des KI-Servicezentrums Berlin-Brandenburg wissenschaftlich beraten und erfolgreich durch Pilotprojekte begleiten, die vorher aufgrund fehlender technischer Möglichkeiten und Expertise nicht durchführbar gewesen wären. Die neue IT-Infrastruktur in Kombination mit der Expertise des HPI ermöglicht es so Organisationen, erste Erfahrungen im Umgang mit KI-Prototypen und Hardware zu sammeln. Damit werden diese Akteure in die Lage versetzt, selbst souveräne Technologieentscheidungen zu treffen.

In einem durchweg positiven Projektverlauf mit reibungsloser Kommunikation übernahmen die SVA Experten eine zentrale Rolle bei der Planung und Umsetzung. Dank ihrer effizienten Projektkoordination konnten der Zeitplan eingehalten und alle Meilensteine termingerecht erreicht werden. Ein weiterer Pluspunkt war die praxisorientierte, klar strukturierte Schulung des HPI-Teams durch SVA Spezialisten, die sicherstellte, dass alle Beteiligten nicht nur die technischen Aspekte der neuen IT-Infrastruktur verstanden, sondern auch in der Lage waren, diese eigenständig zu verwalten und zu optimieren.

„Die Expertinnen und Experten haben durch exzellente Leistungen, ein durchdachtes Konzept und hervorragende Problemlösungsstrategien überzeugt. Wir sind mit den Ergebnissen und der partnerschaftlichen Zusammenarbeit sehr zufrieden“, lobt Lasse Kohlmeyer, AI Engineer & Administrative Project Lead des KI-Servicezentrums am HPI.

Auch die Mitarbeiterinnen und Mitarbeiter des HPI haben sehr positiv auf die zusätzlichen Rechenkapazitäten reagiert. Die implementierte Lösung läuft seit geraumer Zeit stabil und zuverlässig und ihre Flexibilität und Skalierbarkeit ermöglichen es, schnell auf Anforderungen der Stakeholder zu reagieren. Dies hilft dem HPI, exzellente Forschungsprojekte wie das KI-Servicezentrum Berlin-Brandenburg zu betreiben, Nachwuchs- und Fachkräfte zu gewinnen und die digitale Souveränität Europas zu stärken.

“ SVA hat sich als äußerst kompetenter und zuverlässiger Partner erwiesen und die Aufgabenstellung in beeindruckender Weise gelöst. **“**

Lasse Kohlmeyer,
AI Engineer & Administrative Project Lead des KI-Servicezentrums am HPI

Gefördert durch:



Das dieser Darstellung zugrunde liegende Vorhaben wurde mit Mitteln des Bundesministeriums für Forschung, Technologie und Raumfahrt unter dem Förderkennzeichen »KI-Servicezentrum Berlin-Brandenburg« 16IS22092 gefördert.



KONTAKT

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