

MEHR KI, WENIGER WATT

↳ MIT DELL TECHNOLOGIES & SVA ENERGIE UND KOSTEN SPAREN

KI-Server und intelligentes Power-Management senken den Energieverbrauch und die Betriebskosten von KI ohne Performanceverluste.

KI wirtschaftlich betreiben mit Dell Technologies & SVA

EXECUTIVE SUMMARY

Künstliche Intelligenz eröffnet große Chancen – erfordert jedoch teure Hochleistungs-Hardware und verbraucht enorme Ressourcen. Das wird insbesondere dann zur Herausforderung, wenn Sie die volle Kontrolle über Ihre Daten behalten und KI-Modelle On-Premises betreiben möchten. Viele Unternehmen stehen jetzt vor der Frage: Wie können wir die neue Technologie kosteneffizient und nachhaltig einführen? Mit den KI-Servern von Dell Technologies und intelligentem Energiemanagement von SVA behalten Sie den Ressourcenverbrauch im Griff. So können Sie auch als mittelständisches Unternehmen mit überschaubarem Aufwand KI-Projekte umsetzen und schnell von den Potenzialen profitieren. Als erfahrener End-to-End-Partner begleitet SVA Sie von der Beratung und Planung über die Implementierung bis zum Betrieb.

AUSGANGSLAGE

Für Unternehmen lautet die Frage heute nicht mehr ob, sondern wie sie künstliche Intelligenz einführen wollen.

80 Prozent der IT-Entscheider bezeichnen KI als zentralen Bestandteil ihrer Geschäftsstrategie, so eine aktuelle Studie von Dell Technologies. Die meisten befinden sich derzeit am Anfang oder in der Mitte ihrer KI-Journey. Sie identifizieren Use Cases und beginnen, erste Projekte umzusetzen. Das Ziel: Durch die Automatisierung von Routineaufgaben Ressourcen einsparen und die Produktivität steigern.

Besonders hoch im Kurs stehen derzeit generative KI im Allgemeinen und agentische KI im Speziellen. Zu den beliebtesten Anwendungsfällen zählen interne Chatbots, die auf LLMs (Large Language Models) basieren und via

RAG (Retrieval Augmented Generation) auf Unternehmenswissen zugreifen. Mitarbeitende können dann in natürlicher Sprache mit der KI interagieren und schnell Informationen abrufen und verarbeiten, während RAG die Qualität der Antworten sichert und das Risiko von Halluzinationen minimiert. KI-Agenten gehen noch einen Schritt weiter und automatisieren Workflows: Wenn ein Mitarbeiter einem Chatbot eine Aufgabe überträgt, ruft dieser die passenden Tools auf, plant die Aufgabe und führt sie durch. So kann der virtuelle Assistent zum Beispiel alle E-Mails in einem Postfach durchgehen, kategorisieren, an die zuständigen Sachbearbeiter weiterleiten, passendes Unternehmenswissen abrufen und Vorschläge für Antworten machen. Dadurch sparen Mitarbeitende im Arbeitsalltag viel Zeit.

HERAUSFORDERUNG: HOHE KOSTEN

KI-Projekte lassen sich in der Cloud am einfachsten und in den meisten Fällen günstiger realisieren. Viele Unternehmen entscheiden sich jedoch bewusst gegen diesen Weg, weil sie die volle Kontrolle über ihre Daten und KI-Modelle behalten möchten. Auch regulatorische Gründe oder die Wahrung der eigenen digitalen Souveränität sprechen oft gegen einen Cloud-Einsatz. Gefragt sind daher Lösungen, um KI-Anwendungen On-Premises bereitzustellen. Doch das erfordert in der Regel hohe Anfangsinvestitionen, denn die Modelle werden immer ressourcenhungriger: Während vor einigen Jahren noch ein Server mit einer GPU ausreichte, um einen Use Case zu realisieren, setzt heute ein KI-Agent wie der erwähnte E-Mail-Assistent oftmals bereits acht GPUs voraus. In Zukunft werden die Hardware-Anforderungen voraussichtlich weiter steigen.

Das Dilemma zwischen Nachhaltigkeit und Innovation

Zu den Anschaffungsinvestitionen für die KI-Server kommen hohe Betriebskosten und erhebliche Emissionen, die mit der steigenden CO₂-Steuer zum Kostentreiber werden. Nicht nur verbraucht die enorme KI-Rechenleistung sehr viel Strom, auch die Kühlung wird zur Herausforderung, sodass Unternehmen mit einer aufwendigen Rechenzentrumsmodernisierung konfrontiert sind. Laut einer Studie des Öko-Instituts¹ im Auftrag von Greenpeace gefährdet der Energiebedarf von KI-Anlagen die weltweite Energiewende. Demnach wird der Stromverbrauch von Rechenzentren, die KI-Services anbieten, 2030 elfmal so hoch sein wie 2023. Gleichzeitig steigen die Treibhausgasemissionen von 29 Millionen Tonnen CO₂-Äquivalenten auf 166 Millionen Tonnen.

Unternehmen stehen jetzt vor einem Dilemma: Einerseits brauchen Sie KI, um Innovationen voranzutreiben, andererseits müssen sie ihre Nachhaltigkeitsziele erreichen. 84 Prozent der IT-Entscheider sind besorgt über die Auswirkungen von KI-Technologie auf die Umwelt, so die Studie von Dell Technologies, und 88 Prozent halten es für essenziell, die Energieeffizienz in ihren Rechenzentren zu optimieren.

DIE LÖSUNG: KI-SERVER VON DELL TECHNOLOGIES UND POWER-MANAGEMENT MIT SVA

Lösen lässt sich das Dilemma durch die Kombination aus nachhaltiger Infrastruktur von Dell Technologies und intelligentem Power-Management von SVA. So können Unternehmen die Kosten und Energieeffizienz von KI-Projekten optimieren, ihren CO₂-Fußabdruck verringern und Innovation mit Nachhaltigkeit in Einklang bringen. SVA hat dafür eigens im Labor das Energiemanagement situationsabhängig angesteuert, um den Kompromiss von Energieeinsparung bei möglichst geringem Performanceverlust zu automatisieren. Ihre Erfahrung bringen die Experten in die Projektarbeit ein, sodass Kunden direkt von den Ergebnissen profitieren.

¹ <https://presseportal.greenpeace.de/249941-greenpeace-studie-energiehunger-von-kunstlicher-intelligenz-gefahrdet-energiewende/>

So funktioniert das intelligente Power-Management

Die größten Energieverbraucher in einem KI-Server sind die Grafikkarten. Diese werden ab Werk immer für eine Volllastung konfiguriert. Doch Unternehmen brauchen die maximale Leistung anfangs gar nicht. Meist dauert die Entwicklungsphase, bis eine KI-Anwendung produktiv geht, ein bis zwei Jahre. Während dieser Zeit haben die Grafikkarten bei einer geringen Anzahl von Testusern gemittelt nur eine geringe Auslastung, konsumieren aber jede Menge Strom, da der vorhandene Workload den Energiesparmodus blockiert.

Mit intelligentem Power-Management lässt sich der laufende Energieverbrauch erheblich reduzieren. Dafür konfigurieren die SVA-Experten ein passendes Leistungs-Minimum und -Maximum auf jeder Grafikkarte. Basierend auf genauen Berechnungen wählen sie die passende Einstellung, damit die KI-Anwendung genau die richtige Energiemenge erhält, um optimal zu funktionieren. Der Vergleich mit einem Auto verdeutlicht das Einsparpotenzial: Wer gleichmäßig 120 bis 140 fährt, braucht erheblich weniger Benzin als ein Fahrer, der immer Vollgas gibt und an der Ampel eine Vollbremsung hinlegt. SVA konfiguriert die Grafikkarten so für Sie, dass Sie Ihre KI-Anwendung möglichst kosteneffizient betreiben können. Durch die intelligente Leistungsrosselung reduziert sich auch der Kühlbedarf und es lassen sich reale Stromeinsparungen erzielen.



Warum Hardware von Dell Technologies?

Dell Technologies zählt zu den weltweit führenden Anbietern von Hochleistungs-KI-Servern. Die Systeme lassen sich mit GPUs von NVIDIA bestücken und modular an individuelle Anforderungen anpassen. So können Sie Prozessoren, Speicherausstattung, Grafikeinheiten und Netzwerkanbindung genau auf Ihren Use Case abstimmen und problemlos skalieren, wenn sich Ihre Anforderungen ändern. SVA unterstützt Sie dabei, die geeigneten Komponenten auszuwählen und energieeffizient zu konfigurieren. Für eine erste, einfache Anwendung reicht zum Beispiel oft schon eine kleine Einheit aus. Diese können Sie später mit leistungsstärkeren GPUs aufrüsten, sobald Sie anspruchsvollere KI-Modelle trainieren möchten.

Geringer CO₂-Fußabdruck und Kreislaufwirtschaft

Dell Technologies achtet bereits in der Entwicklung auf Nachhaltigkeit. Grundsätzlich werden Produkte von Dell Technologies für eine lange Lebensdauer konzipiert, lassen sich leicht reparieren und am Ende ihres Lebenszyklus gut recyceln. Um unnötigen Elektroschrott zu vermeiden, bietet der Hersteller zum Beispiel verschiedene Reparatur-, Recovery- und Wiederverwendungsservices an. So können Kunden ihren CO₂-Fußabdruck reduzieren und ihre Nachhaltigkeitsziele unterstützen. Bis im Jahr 2030 will Dell Technologies selbst mehr als 50 Prozent seiner Produkte und 100 Prozent aller Verpackungen aus recycelten, nachwachsenden oder kohlenstoffarmen Materialien herstellen.²

Hohe Service-Qualität zahlt sich aus

Was Dell Technologies darüber hinaus auszeichnet, ist die hohe Service-Qualität. Kunden profitieren von deutschsprachigem Support, der in der Regel aus Deutschland erbracht wird. Erfahrene Spezialisten sind über alle Kanäle erreichbar und finden schnell Lösungen. Da die Server mit vielen Selbstheilungsfunktionen ausgestattet sind, lassen sich potenzielle Probleme oft schon im Vorfeld beheben, ohne dass Sie aktiv den Support verständigen müssen. Falls Komponenten ausgetauscht werden müssen, kommt ein qualifizierter Service-Techniker ins Haus. Sie haben die Wahl zwischen verschiedenen Service Levels, sodass eine Fehlerbehebung innerhalb der vertraglich vereinbarten Zeit – je nach gewählter Service-Option im besten Fall vier Stunden – garantiert ist. So lassen sich lange Ausfallzeiten vermeiden.

Gerade bei KI-Servern zahlt es sich aus, auf hohe Service-Qualität zu achten. Denn wenn der Betrieb stillsteht, weil Probleme nicht zeitnah behoben werden können, entpuppen sich vermeintlich günstige Angebote als Kostenfalle.

² <https://www.dell.com/de-de/dt/corporate/social-impact/advancing-sustainability/accelerating-the-circular-economy.htm#anchor>

Warum KI-Projekte mit SVA?

WARUM KI-PROJEKTE MIT SVA?

Auf dem Markt für KI-Infrastruktur herrscht derzeit Goldgräberstimmung. Die verschiedensten Anbieter locken mit vollmundigen Marketing-Versprechen. Ohne Experten-Know-how ist es schwer, die widersprüchlichen Empfehlungen einzuordnen und die geeignete Lösung zu finden.

Tatsächlich benötigen KI-Projekte in der Praxis oft weniger Ressourcen, als so mancher Hersteller proklamiert. Wer unnötige Kosten vermeiden will, sollte sich vor der KI-Einführung unabhängig beraten lassen. SVA zählt zu den Top-3-Infrastruktur-Partnern von Dell Technologies und verfügt zugleich über die höchste Solution-Provider-Zertifizierung von NVIDIA und ist Certified Instructor des NVIDIA Deep Learning Institute (DLI). Unsere Expertinnen und Experten verfügen über langjährige Erfahrung in der Realisierung von KI-Projekten und unterstützen mittelständische Unternehmen von der Beratung über die Auswahl der geeigneten Hardware bis zur Implementierung und zum Betrieb. Wir finden die Lösung für Sie, die am besten für Ihren Use Case geeignet ist, und konfigurieren die Server so, dass Sie Ihr Ziel möglichst kosteneffizient erreichen – sowohl in der Anschaffung als auch im laufenden Betrieb.

USE CASES

Einführung eines KI-Agenten bei Windmüller & Hölscher

Der Maschinenbauer Windmüller & Hölscher hat gemeinsam mit SVA eine innovative Lösung entwickelt, um den Zugriff auf seine komplexe Prozesslandschaft deutlich zu vereinfachen. Das Unternehmen fertigt hochspezialisierte Maschinen für flexible Verpackungen – mit tausenden kundenspezifischen Varianten und entsprechend vielschichtigen Abläufen, die in SAP Signavio modelliert werden. Mit dem neuen KI-Agenten können Mitarbeitende Prozessinformationen jetzt in natürlicher Sprache abrufen. Der Chatbot versteht den Kontext, greift auf die Prozessdatenbank zu und liefert innerhalb von Sekunden die passenden Abläufe. Mithilfe von RAG werden relevante Informationen aus den Prozessmodellen gefiltert, angereichert und für das LLM bereitgestellt. Die Lösung läuft aktuell auf Microsoft Azure, kann aber mit identischem Leistungsumfang auch On-Premises umgesetzt werden. SVA verantwortete das gesamte KI-Engineering – von der Architektur und Modellbewertung über das Feintuning bis zur Implementierung der Cloud-Infrastruktur und Datenpipelines.

Aufbau einer effizienten KI-Infrastruktur bei BarmeniaGothaer

Die BarmeniaGothaer Versicherungsgruppe nutzt künstliche Intelligenz, um die Effizienz in der Sachbearbeitung zu steigern und innovative Produkte zu entwickeln. Dafür benötigte sie eine leistungsfähige, flexible Infrastruktur, die den kontinuierlichen Betrieb verschiedener KI-Modelle ermöglicht. Gemeinsam mit SVA hat das Unternehmen eine GPU-beschleunigte Private-Cloud-Umgebung aufgebaut. Im Mittelpunkt stand die Transformation von monolithischen Anwendungen zu Kubernetes-nativen Microservices, um Entwicklungszyklen zu verkürzen und Ressourcen optimal zu nutzen.

Die SVA-Experten begleiteten den Aufbau einer modularen MLOps-Architektur, die den Prozess der Modellbereitstellung automatisiert. Dank GPU-Virtualisierung kann die BarmeniaGothaer heute mehrere Modelle oder Workloads auf einer GPU parallel betreiben – ohne Performance-Einbußen. Durch die präzise GPU-Frakturierung werden Ressourcen bedarfsgerecht zugeteilt, wodurch sich die Auslastung der Beschleuniger signifikant erhöht und unnötiger Leerlauf entfällt. Das Ergebnis: maximal effiziente Ressourcennutzung, geringere Betriebskosten und eine nachhaltigere MLOps-Umgebung, unterstützt durch GitOps-gesteuerte Deployments und ein erweitertes GPU-Monitoring.

Integration einer KI-Plattform bei Freudenberg Performance Materials

Der Werkstoffspezialist Freudenberg Performance Materials setzt KI gezielt in Forschung und Entwicklung ein. Gemeinsam mit SVA hat das Unternehmen eine leistungsfähige, GPU-gestützte KI-Infrastruktur aufgebaut, um die Entwicklung und den Betrieb von KI-Modellen effizienter zu gestalten. Ziel war es, in einer geschützten Netzwerkumgebung eine skalierbare Plattform zu schaffen, die sich nahtlos in die bestehende IT-Landschaft integriert und den gesamten MLOps-Zyklus – von der Entwicklung über das Training bis zum Deployment – unterstützt. Durch die Kombination aus GPU-Virtualisierung, Kubernetes-Cluster-Management und einer spezialisierten KI-Plattform kann Freudenberg KI-Modelle heute schneller entwickeln, trainieren und flexibel erweitern.

JETZT KI-PROJEKT STARTEN!

Sie möchten auch ein KI-Projekt realisieren? SVA unterstützt Sie ganzheitlich von der Konzeption bis zum Betrieb. Gemeinsam mit Ihnen identifizieren wir Use Cases und erarbeiten Vorschläge für die Umsetzung. Mit einer bedarfsgerechten, nachhaltigen Infrastruktur und intelligentem Power-Management können Sie künstliche Intelligenz kosteneffizient einführen und schnell erste Erfolge erzielen. Nehmen Sie Kontakt mit uns auf – wir beraten Sie gern!

ÜBER UNS

Die SVA System Vertrieb Alexander GmbH ist einer der führenden deutschen System-Integratoren. Das unternehmerische Ziel des Unternehmens ist es, hochwertige IT-Produkte der jeweiligen Hersteller mit dem Projekt-Know-how und der Flexibilität von SVA zu verknüpfen, um so optimale Lösungen für die Kunden zu erzielen. Kernthemen des Unternehmens sind Datacenter-Infrastruktur – wie Speicher-, Server und Netzwerk-Lösungen sowie Virtualisierungs-Technologien – und auch Business Continuity, Digital Process Solutions und SAP. Darüber hinaus bietet SVA nicht nur Unterstützung im Betrieb der Infrastruktur durch Operational Services, sondern auch Managed Services. Mit dem Angebot zu Strategischem IT-Consulting stehen auch Experten für eine ganzheitliche und umsetzungsorientierte Beratung zur Verfügung.



Ich freue mich
über Ihre Kontaktaufnahme!

Christian Nagel

Partnermanager Dell Technologies

Mobil: +49 151 44845333

Mail: christian.nagel@sva.de

SVA System Vertrieb Alexander GmbH

Borsigstraße 26

65205 Wiesbaden

www.sva.de

