



Bewältigung der Herausforderungen von KI-Workloads mit dem Dell KI-Portfolio

Ein Vergleich des Dell KI-Portfolios mit ähnlichen Angeboten von HPE

Es besteht kein Zweifel daran, dass künstliche Intelligenz (KI) die Geschäftslandschaft transformiert und Branchen und Unternehmen jeder Größenordnung ermöglicht hat, tiefere Erkenntnisse aus ihren Daten zu gewinnen, Geschäftsprozesse zu automatisieren, maßgeschneiderte Kunden- und Nutzererlebnisse bereitzustellen und in ihrer Branche besser wettbewerbsfähig zu sein. Um die Leistungsfähigkeit von KI effektiv nutzen zu können, benötigen Unternehmen einen Infrastrukturanbieter, der ein umfassendes und integriertes Lösungsportfolio für den gesamten KI-Lebenszyklus anbieten kann.

Infrastrukturanbieter wie Dell Technologies und Hewlett Packard Enterprise (HPE) helfen Kunden dabei, die wachsenden Anforderungen von KI zu erfüllen und die damit verbundenen Komplexitäten zu bewältigen. Diese Anbieter präsentieren KI-fähige Portfolios und bieten verschiedene KI-Lösungen, die leistungsstarke On-Premise- und Cloud-Infrastrukturlösungen mit strategischen Partnerschaften und einer Reihe von Support- und Beratungsservices vereinen.

Dieser Bericht untersucht öffentlich verfügbare Informationen über die KI-Portfolios von Dell und HPE* mit dem Ziel, spezifische Architektur-, Leistungs- und Supportvorteile hervorzuheben, von denen Kunden profitieren könnten, wenn sie Dell Technologies für ihre KI-Anforderungen wählen. Wir vergleichen Details der Server, die Dell zur Unterstützung von KI-Bereitstellungen entwickelt hat, und verweisen auf die Ergebnisse von Branchenbenchmarktests von ML Commons®. Wir untersuchen außerdem zusätzliche Software- und Serviceangebote, die Kunden in jeder Phase ihrer KI-Reise unterstützen.

* Hinweis: PT hat alle Recherchen am oder vor dem 5. Dezember 2023 abgeschlossen, sodass dieser Artikel keine Angebote oder Änderungen berücksichtigt, die Dell oder HPE nach diesem Datum veröffentlichen.



Die Herausforderungen der Einführung von KI

Die Einführung einer KI-Strategie bringt viele neue Herausforderungen für Rechenzentren und die IT-MitarbeiterInnen mit sich, darunter:

- Behebung der Qualifikationsdefizite bei den derzeitigen MitarbeiterInnen entweder durch interne KI-Schulungen oder externe Einstellungen
- Verständnis der Datenaufbereitungsanforderungen von KI, einschließlich der Qualität und Quantität, des Standorts und des aktuellen Zustands der Unternehmensdaten
- Bewertung der spezifischen KI-Ziele des Unternehmens, um besser bestimmen zu können, welche KI-Modelle und -Implementierungen Vorteile bieten.
- Bewertung der Computing-, Netzwerk- und Storage-Anforderungen der geplanten KI-Systeme und Festlegung eines Beschaffungsplans

Dies sind nur einige Beispiele für die vielen, oft erheblichen Hürden, mit denen Unternehmen konfrontiert sind, wenn sie die Vorteile der Implementierung von KI in ihren Rechenzentren nutzen möchten.

Das Dell KI-Portfolio soll Kunden dabei unterstützen, diese Herausforderungen durch professionelle und beratende Services zu bewältigen, die KundInnen dabei unterstützen, Implementierungs-Roadmaps zu erstellen und ihre Daten für KI-Modelle vorzubereiten.¹ Das Portfolio umfasst außerdem Schulungskurse, die Konzepte wie maschinelles Lernen (ML) und andere Bildungsthemen abdecken, und bietet validierte Designs für KI, um den Erfolg der Implementierung sicherzustellen.² Darüber hinaus arbeitet Dell mit Drittanbietern zusammen, um KundInnen zusätzliche KI-Tools bereitzustellen, z. B. ein kundenspezifisches Dell Portal innerhalb der Hugging Face-Community mit dedizierten Containern und Skripten für die Bereitstellung von Open-Source-KI-Modellen³ und die einfache Bereitstellung des großen Meta Llama 2-Sprachmodells (LLM).⁴ Neben einer großen Auswahl an Compute- und PC-Angeboten, von mobilen Workstations bis hin zu Servern, die bis zu 8 High-End-NVIDIA-GPUs unterstützen, bietet Dell auch KI für den Storage unstrukturierter Daten mit einem Portfolio an leistungsstarken Datei- und Objektspeicher-Arrays. Diese Storage-Angebote, einschließlich Dell PowerScale, ObjectScale, ECS und integriertem Storage, können die unstrukturierten Daten verarbeiten, die KI-Workloads häufig nutzen.⁵ Dell ist außerdem eine Partnerschaft mit Snowflake eingegangen, um eine hybride Cloud-Storage-Lösung für Dell KundInnen anzubieten.⁶ Laut einer Analyse von Dell bietet das Unternehmen ab August 2023 das „breiteste generative KI-Portfolio“ an, das über Server und Storage hinausgeht und Ressourcen für die gesamte KI-Implementierung bereitstellt.⁷

KI-Performance und beschleunigte Compute-Optionen: Dell im Vergleich zu HPE

KI-Workloads können CPUs, GPUs oder beides je nach Größe oder Art der Workload als Computeresourcen verwenden. Einige CPUs bieten KI-spezifische Accelerators wie Intel Advanced Matrix Extensions (Intel AMX) in den neuesten Intel Xeon Scalable Prozessoren.⁸ GPUs sind oft besser für größere und/oder anspruchsvollere Workloads, aber der GPU-Formfaktor kann sich auf die Performancelevel auswirken. Einige NVIDIA A100- und H100-GPUs sind beispielsweise in universellen PCIe- oder proprietären SXM-Formfaktoren erhältlich, wobei letztere die leistungsfähigere NVIDIA SXM-Architektur verwenden.⁹ Große Arbeitsspeicherkapazitäten und Serverdesignfunktionen wie Kühlungsarchitektur und Energieeffizienz wirken sich ebenfalls auf die Performance aus. Die meisten Rechenzentren nutzen weiterhin Luftkühlung, was bedeutet, dass für High-Performance-Compute-Workloads (HPC) Server erforderlich sind, die so effektiv wie möglich mit Luft gekühlt werden. Im Folgenden werden wir die PowerEdge-Serverangebote in Bezug auf Komponenten, Kühlungsoptionen und mehr sowie die veröffentlichten MLCommons® MLPerf®-Bewertungen vorstellen.

Benchmarkperformance des KI-Modells: MLPerf-Ergebnisvergleich

MLPerf® ist eine Benchmarksuite, die die KI-Performance in Bezug auf Training und Inferenz testet. Damit ein Unternehmen offizielle MLPerf®-Ergebnisse veröffentlichen kann, müssen die Ergebnisse den spezifischen Bedingungen entsprechen, die vom Benchmarkentwickler MLCommons® festgelegt wurden.¹⁰ Diese Compiancerichtlinien enthalten Standards, die den Leistungsvergleich vereinfachen. Für Inferenztests verwendet MLPerf® Datacenter-, Edge-, Mobile- und Tiny-Datenvolumen und meldet KI-Bewertungen und Watt des Stromverbrauchs während der Tests. Die Inferenz-Benchmark-Suite umfasst Tests für viele gängige KI-, ML- und DL-Modelle; siehe Tabelle 1.

Tabelle 1: KI-, ML- und DL-Modelle, die in MLPerf®-Tests enthalten sind, und typische Anwendungsbeispiele für jedes Modell. Quelle: Principled Technologies

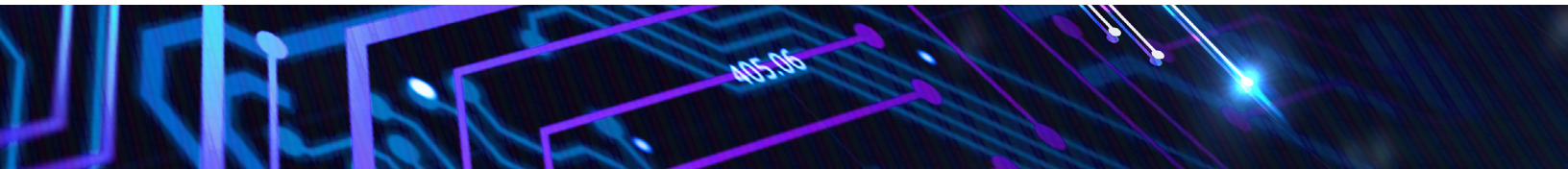
Gängige KI-Modelle	Typisches Einsatzbeispiel
ResNet	Ein Bildklassifizierungsmodell, mit dem Computer verschiedene Bilder für Anwendungsfälle wie medizinische Bildgebung, Moderation von Social-Media-Inhalten und Gesichtserkennung lernen, sich merken und identifizieren können
RetinaNet	Eine Art der Objekterkennung, die im Vergleich zu RESNET zusätzliche Komplexität bewältigen kann. Sie hilft Computern, Objekte in Bildern oder Videoframes zu identifizieren und zu finden und sie nach ihrer Bedeutung zu klassifizieren. Einsatz in Bereichen wie autonomes Fahren, automatische Fahrzeugassistenzsysteme, Überwachung, Gesichtserkennung
3D-UNet	Spezifisch für die Segmentierung medizinischer Bilder
RNN-T	Spracherkennung für Anwendungsfälle wie automatisierte Sprachübersetzung
BERT	Natürliche Sprachverarbeitung für Anwendungsfälle wie Textzusammenfassung, Sprachübersetzung und automatische Fertigstellung von Aufgaben
DLRM-v2-99.9	Empfehlungsmodell für Anwendungsfälle wie gezielte Werbung und personalisierte Produktempfehlungen
GPTJ-99 und 99,9	LLM für die Verarbeitung natürlicher Sprache, die sich bei der Texterstellung für Anwendungsfälle wie Chatbots und chatbasierte KI-Tools auszeichnet

MLPerf

MLPerf®-Ergebnisse umfassen neben den KI-Modellen auch mehrere Parameter, was dazu führen kann, dass in einem einzigen Diagramm oder einer Tabelle eine große Datenmenge analysiert werden muss. Hier finden Sie eine Kurzübersicht zu diesen Parametern:

- 99,0 und 99,9: Diese Zahlen beziehen sich auf die Genauigkeit, mit der das Modell trainiert wurde. Je genauer die Ausgabe sein muss, desto komplexer ist das Modell und desto länger kann die Verarbeitung der Daten dauern.
- Offline-Samples/Sek.: Modus, in dem der Benchmark alle Abfragen zu Beginn des Tests sendet, um Daten zu simulieren, die bereits auf dem System vorhanden sind.
- Serverabfragen/Sek.: Modus, in dem der Benchmark Abfragen während der Testdauer sendet, um die Analyse eines Livestreams von Daten zu simulieren.

Weitere Informationen zu MLCommons®- und MLPerf®-Ergebnissen finden Sie unter <https://mlcommons.org/benchmarks/inference-datacenter/>.



Die Ergebnisse in diesem Bericht stammen aus MLPerf® v3.1 Inference Datacenter-Ergebnissen, die auf der MLCommons®-Website vom November 2023 veröffentlicht wurden.¹¹ Diese Ergebnisse umfassen Beiträge von Technologieherstellern und Cloud-Serviceanbietern und decken eine Reihe von Konfigurationen ab. Im Vergleich zu öffentlich verfügbaren Daten von HPE erzielten Dell Server bei bestimmten KI-Modellen bessere Ergebnisse. (Hinweis: Unterschiedliche GPU-Konfigurationen auf den Servern können direkte Vergleiche erschweren.) Weitere Details finden Sie in Tabelle 2.

Tabelle 2: Dell und HPE-Server, die in den MLCommons® MLPerf® 3.1-Ergebnissen enthalten sind (Stand: 29.11. 2023).
Quelle: Principled Technologies

Anfragesteller	Servermodell	Nr. und Modell der GPUs	Beschreibung
Dell ¹²	PowerEdge XE9680	8 x NVIDIA H100 SXM	Für KI-Training und Inferenz bei großen Workloads wie großen Sprachmodellen
	PowerEdge XE9640	4 x NVIDIA H100 SXM	Für das Training großer KI-Modelle in Rechenzentren mit hoher Dichte und Flüssigkeitskühlung
	PowerEdge XE8640	4 x NVIDIA H100 SXM	Für den Betrieb herkömmlicher KI-Trainings-, HPC- und Datenanalyseanwendungen in einem 4-HE-Formfaktor für luftgekühlte Rechenzentren
	PowerEdge R760xa	4 x NVIDIA H100 PCIe	Für eine Vielzahl von rechenintensiven Workloads, einschließlich KI-ML/DL-Training und Inferenz, für die keine Hochleistungs-GPUs erforderlich sind
HPE ^{13,14}	ProLiant XL675d Gen10 Plus	8 x NVIDIA A100 SXM	Für High-Performance Computing und KI
	ProLiant DL380a Gen11	4 x NVIDIA H100 PCIe	2-HE-Server für moderate KI-Workloads

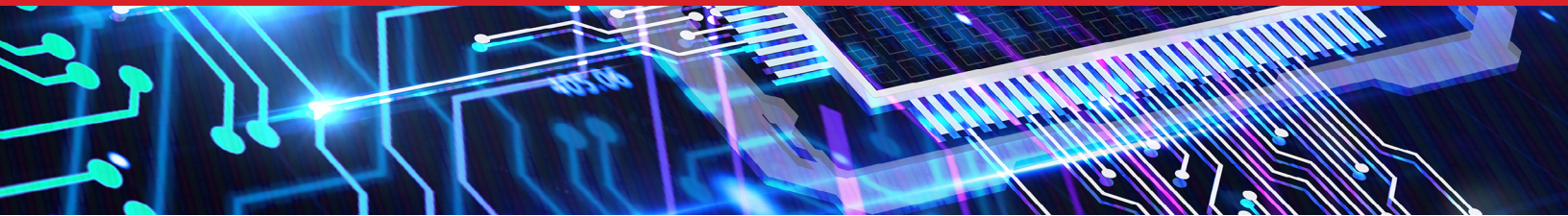
Direkter Vergleich zwischen Dell und HPE Servern

Zu einer umfassenden KI-Strategie gehört zwar mehr als nur die Hardware, aber die Gewährleistung der bestmöglichen Hardwareleistung ist einer der wichtigsten Faktoren für den Erfolg von KI-Workloads. In dem Maße, wie neuere Grafikprozessoren und andere Technologien verfügbar werden, entwickeln sich auch die KI-Workload-Funktionen weiter. Zum Zeitpunkt der ersten Veröffentlichung der MLPerf® v3.1-Ergebnisse war die beste verfügbare NVIDIA-GPU der H100 Tensor Core, mit dem Dell MLPerf®-Ergebnisse in mehreren seiner Server sowohl in PCIe- als auch in SXM5-Formfaktoren veröffentlicht hat.¹⁵ Die öffentlich verfügbaren Daten von HPE beinhalteten nur eine H100-Einreichung und nur den PCIe-Formfaktor. Unsere Forschung hat gezeigt, dass nur wenige der verfügbaren GPU-fähigen HPE-Server den SXM5 H100-Formfaktor für die beste NVIDIA-GPU-Performance unterstützen, und keiner der HPE ProLiant-Server tut dies.¹⁶ Wie wir unten zeigen, verbessert sich die Leistung von KI-Workloads in der Regel durch bessere GPUs.

MLPerf-Ergebnisse mit acht GPUs

Der Dell PowerEdge XE9680 bietet Support für bis zu acht NVIDIA H100 SXM5-GPUs für KI-Beschleunigung und bis zu zwei skalierbare Intel® Xeon® Prozessoren der 4. Generation. Die PowerEdge XE-Produktreihe verfügt über eine modulare Architektur, die SXM4- oder SXM5-NVIDIA-GPUs oder OAM-GPU-Baugruppen (Open Compute Project Accelerator Module) von AMD unterstützt, wodurch die Leistung im Vergleich zu einer standardmäßigen PCIe-GPU gesteigert werden kann.¹⁷ Der PowerEdge XE9680 nimmt nur 6 HE Platz im Rack ein und ist ein kompakter 8-Wege-NVIDIA H100 SXM5-Server. Die neuesten HPE ProLiant Gen11-Server unterstützen derzeit nicht den H100 SXM-Formfaktor,¹⁸ obwohl einige der HPE Cray Supercomputing-Server dies tun.¹⁹ Da HPE keine MLPerf®-Ergebnisse mit den Cray-Servern eingereicht hat und nur die ProLiant-Server auf der KI-Portfolioseite hervorhebt, konzentrieren wir uns in diesem Whitepaper auf ProLiant-Server. (Siehe Abbildung 1.)





Featured AI products and services

PRODUCT

HPE Ezmeral Unified Analytics Software

Unlock data and insights faster by helping you develop and deploy data and analytic workloads. Provides fully managed, secure, enterprise-grade versions of the most popular open-source frameworks with a consistent SaaS experience.

[Explore more →](#)

PRODUCT

HPE Machine Learning Development Environment

Uncover hidden insights from your data by helping engineers and data scientists collaborate, build more accurate ML models and train them faster.

[Explore more →](#)

PRODUCT

HPE Machine Learning Data Management Software

Uncover hidden insights with a data pipelining and versioning solution that automates data pipelines and accelerates time to ML model production by processing petabyte-scale workloads.

[Explore more →](#)

PRODUCT

HPE ProLiant Servers

Speed time to value with systems that are optimized for computer vision inference, generative visual AI, and end-to-end natural language processing.

[Explore more →](#)

Abbildung 1: Screenshot der vorgestellten KI-Produkte und -Services auf <https://www.hpe.com/us/en/solutions/ai-artificial-intelligence.html> mit HPE ProLiant-Servern vom 05.12.2023.

In den Ergebnissen von MLPerf® v3.1, die erstmals im November 2023 für Server mit acht GPU veröffentlicht wurden, übertraf der Dell PowerEdge XE9680 mit NVIDIA SXM5 H100-GPUs den HPE ProLiant XL675d Gen10 Plus mit NVIDIA SXM4 A100-GPUs um das bis zu 4,25-fache (siehe Abbildung 2).

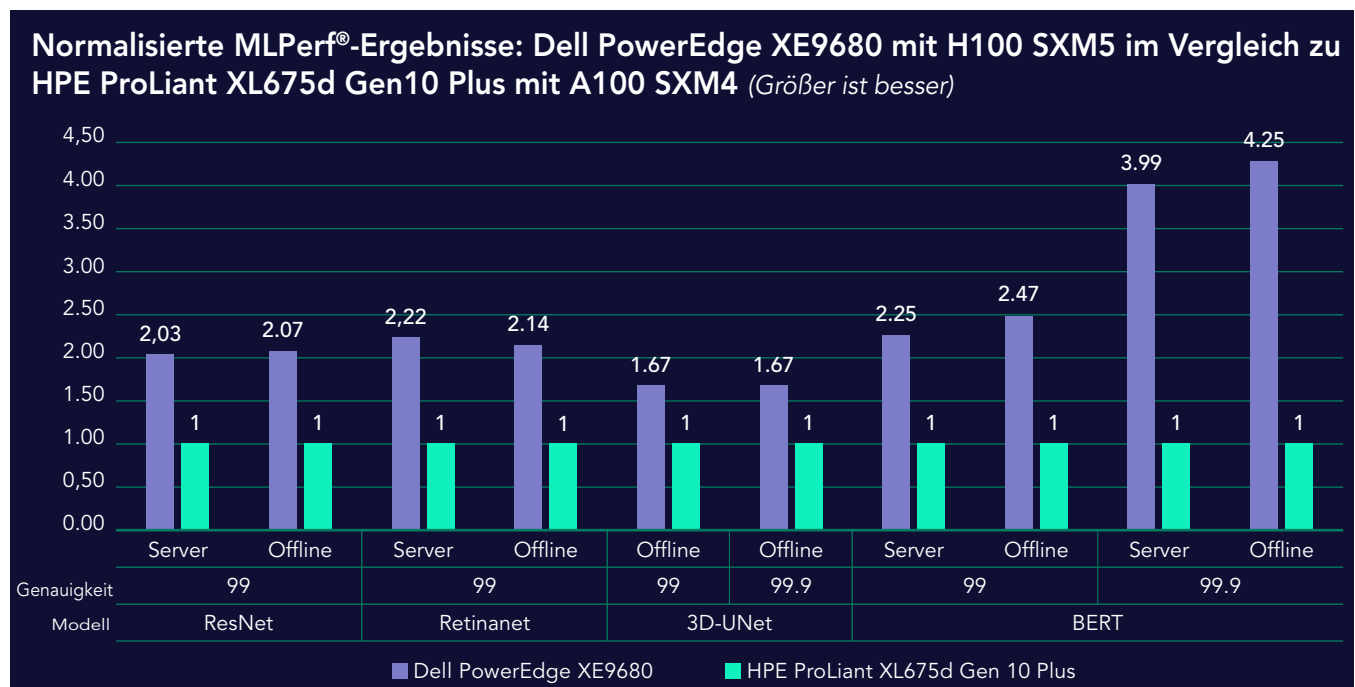


Abbildung 2: Veröffentlichte MLPerf®-Ergebnisse für Dell PowerEdge XE9680 und HPE ProLiant XL675d Gen10 Plus zum 29.11.23. Das Dell System verwendet die NVIDIA H100-GPU, während die GPUs im HPE System eine Generation älter sind. Quelle: Principled Technologies unter Verwendung von Daten von MLCommons®.^{20, 21}

Für einen einfachen Vergleich haben wir die Testergebnisse in den Abbildungen 2 bis 5 normalisiert. Das bedeutet, dass wir jedem HPE ProLiant DL380a Gen 11-Ergebnis den Wert 1 zuweisen und das entsprechende Dell PowerEdge R760xa-Ergebnis in Bezug darauf anzeigen. Wie diese Ergebnisse zeigen, kann selbst ein Generationsunterschied zwischen GPU-Modellen einen signifikanten Unterschied in der Leistung ausmachen, die Sie bei einer Vielzahl von KI-Workloads erwarten können.

MLPerf-Ergebnisse für vier GPUs

Wenn Strom- oder Platzersparnis im Rechenzentrum ein zentrales Anliegen sind, könnte der 2-HE Dell PowerEdge XE9640 die Antwort sein. Mit bis zu vier NVIDIA H100 SXM-GPUs bietet der PowerEdge XE9640 die Hälfte der GPU-Rechenleistung des XE9680 auf zwei Drittel weniger Platz.²² Der dicht gepackte Dell PowerEdge XE9640 ist mit der Dell Smart Cooling-Technologie ausgestattet, die eine Reihe von Wärmetechnologien einschließlich direkter Flüssigkeitskühlung für CPUs und GPUs bietet.²³

Das 2-HE-Gehäuse des PowerEdge XE9640 verfügt über verbesserte Luftstrommechanismen, einschließlich größerer Lüfter und Kühlkörper, um die anderen wichtigen Komponenten wie PCIe-Karten und Arbeitsspeicher zu kühlen.²⁴ Der PowerEdge XE9640 ist derzeit das einzige Angebot von Dell oder HPE, das mit 4-Wege-HGX H100-GPUs in 2 HE ausgestattet ist. Das HPE AI-Portfolio bietet 1-HE und 2-HE Gen11 ProLiant-Server, aber sie sind auf PCIe-Formfaktor-GPUs beschränkt.²⁵

Der Dell PowerEdge XE9640-Server unterstützt auch OAM-GPUs der Intel Max GPU Serie 1550, die eine GPU-Option mit niedrigem Stromverbrauch und hoher Dichte bieten, die eine PCIe-Karte und ein OpenCompute Accelerator Module (OAM) umfasst.²⁶ Wir konnten zwar nicht feststellen, dass HPE ab dem 05.12. 2023 einen Server mit diesen GPUs anbot, aber sie bieten den HPE ProLiant DL380 Gen11 und DL380a Gen11 mit Intel Data Center GPU Max 1100-GPUs an.²⁷ Das bedeutet, dass der Dell PowerEdge XE9640 möglicherweise das einzige aktuelle Angebot mit vier Intel Max 1550 OAM-GPUs in einem 2-HE-Server ist. Für Unternehmen, die Wert auf Platz und Energieeffizienz im Rechenzentrum legen, bietet ein 2-HE-Server mit vier Intel Max 1550-GPUs eine Lösung, die High-Performance-Computing und Energieeffizienz kombiniert, ohne dabei Platz im Rechenzentrum zu opfern.

Der Dell PowerEdge XE9640 mit vier HGX H100-GPUs übertraf den HPE ProLiant DL380a mit vier PCIe H100-GPUs in den veröffentlichten MLPerf® 3.1-Ergebnissen um das bis zu 1,99-fache (siehe Abbildung 3).

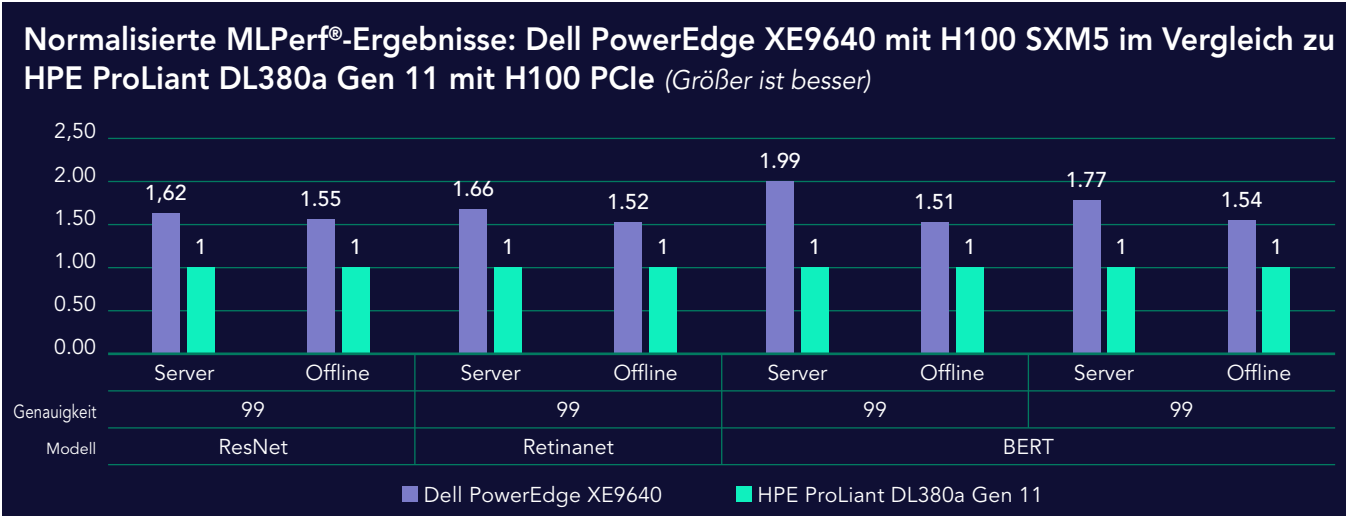


Abbildung 3: Veröffentlichte MLPerf®-Ergebnisse für Dell PowerEdge XE9680 und HPE ProLiant XL675d Gen10 Plus zum 29.11.23. Das Dell System verwendet die NVIDIA H100-GPU, während die GPUs im HPE System eine Generation älter sind. Quelle: Principled Technologies unter Verwendung von Daten von MLCommons®.^{28,29}

Der PowerEdge XE8640 bietet eine 4-Wege-GPU-Konfiguration mit Luftkühlung für Prozessoren und einem flüssigkeitsgestützten Luftkühlungsradiator für die GPUs, der keine Wasserverfügbarkeit im Rack erfordert. Für diejenigen, die kein externes Kühlmittel verwenden oder nicht verwenden können,³⁰ unterstützt der 4-HE-Server Dell PowerEdge XE8640 vier NVIDIA H100 SXM5-GPUs, die dieselbe Rechenleistung wie der PowerEdge XE9640 bieten, ohne dass eine direkte Flüssigkeitskühlung erforderlich ist.³¹

Der Dell PowerEdge XE8640 verfügt über die neuesten skalierbaren Intel Xeon Prozessoren der 4. Generation und bis zu 4 TB Arbeitsspeicher³² für die großen Datenvolumen und komplexen Berechnungen, die in KI und Data Analytics üblich sind. Auch hier bietet HPE die NVIDIA H100 SXM5-GPUs in HPE-Cray-Systemen an, aber HPE ProLiant GPU-fähige Server unterstützen dies nicht.

Im Vergleich zu MLPerf®-Daten, die im November 2023 veröffentlicht wurden, erzielte der PowerEdge XE8640-Server mit vier NVIDIA H100 SXM5-GPUs in neun verschiedenen Kategorien den höchsten KI-Durchsatz unter allen vier eingereichten GPUs. Wie Abbildung 4 zeigt, erzielte er im Vergleich zum HPE ProLiant DL380a-Server eine bis zu 2,07-mal höhere Bewertung.

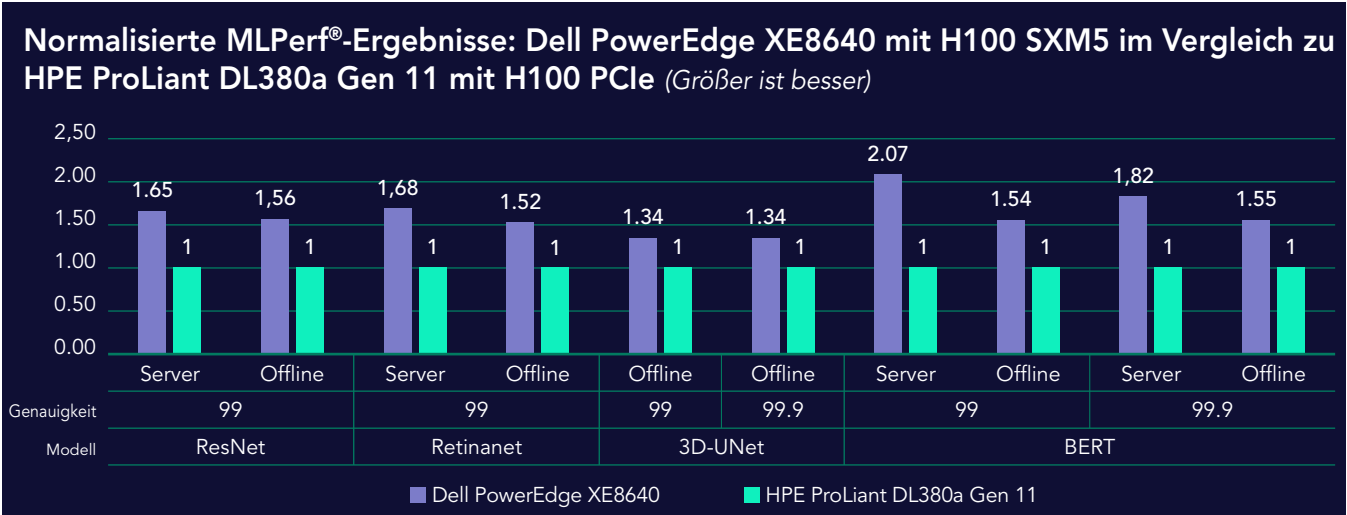


Abbildung 4: Veröffentlichte MLPerf®-Ergebnisse für Dell PowerEdge XE8640 und HPE ProLiant DL380a Gen11 zum 29.11. 2023. Das Dell System verwendet den NVIDIA H100 SXM-Formfaktor, während das HPE System den weniger leistungsstarken PCIe-Formfaktor verwendet. Quelle: Principled Technologies unter Verwendung von Daten von MLCommons®.^{33, 34}

Für Unternehmen, die kleiner anfangen und je nach Bedarf wachsen möchten, bietet der 2-HE-Server Dell PowerEdge R760xa Platz für eine Reihe von Grafikprozessoren von NVIDIA, AMD und Intel mit Unterstützung für bis zu vier PCIe Gen 5-GPUs mit doppelter Breite oder 12 PCIe-GPUs mit einfacher Breite.³⁵ Er verfügt über 32 DIMM-Steckplätze, einen Schacht mit acht Laufwerken für 2,5-Zoll-Festplatten und 12 PCIe-Steckplätze. Damit bietet er skalierbaren Storage, der mit den steigenden KI-Datenanforderungen wachsen kann, sowie Unterstützung für bis zu 12 PCIe-Grafikprozessoren mit einfacher Breite oder vier PCIe-Grafikprozessoren mit doppelter Breite wie die NVIDIA H100 oder L40S.³⁶ Diese Skalierbarkeit bedeutet, dass sich der Server an sich entwickelnde KI-Aufgaben anpassen kann, vom Modelltraining für maschinelles Lernen bis hin zu erweiterter Datenverarbeitung.

Das Luftkühlungssystem des PowerEdge R760xa unterstützt Computing-Umgebungen mit hoher Dichte und kann Accelerators mit höherer thermischer Designleistung (TDP) bis zu 350 W aufnehmen,³⁷ was der IT helfen kann, die Leistung auch bei intensiven Rechenlasten aufrechtzuerhalten. In den Testergebnissen für MLPerf® RESNET, RetinaNet und BERT Server, die im November 2023 im „Server“-Modus veröffentlicht wurden, übertraf der PowerEdge R760xa mit vier NVIDIA H100 PCIe-GPUs den HPE ProLiant DL380a Gen 11, der ebenfalls mit vier H100 PCIe-GPUs ausgestattet ist (siehe Abbildung 5).

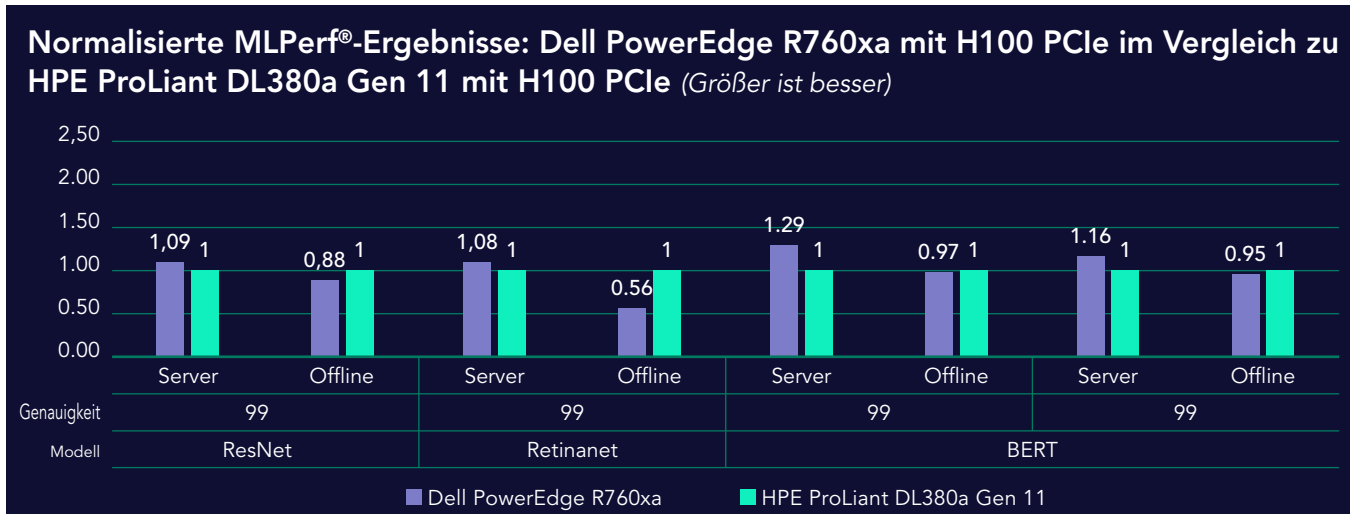


Abbildung 5: Veröffentlichte MLPerf®-Ergebnisse für Dell PowerEdge R760xa und HPE ProLiant DL380a Gen11 zum 29.11. 2023. Beide Systeme verwenden den PCIe-Formfaktor der NVIDIA H100-GPUs. Quelle: Principled Technologies unter Verwendung von Daten von MLCommons®.^{38,39}

Alles in allem zeigen die MLPerf®-Ergebnisse, dass die Leistung zwischen Servern und Komponenten stark variiert, sodass die Auswahl der richtigen Optionen zur Unterstützung Ihrer Workloads und ihrer Performanceanforderungen von entscheidender Bedeutung ist. Dell PowerEdge-Server für KI-Workloads bieten mehrere Optionen für Kühlung und Dichte, um den Rechenzentrumsanforderungen eines Unternehmens gerecht zu werden und gleichzeitig eine starke MLPerf®-Leistung zu bieten.

Detailliertere Abdeckung des Dell KI-Portfolios

Die Rechenleistung ist zwar entscheidend, aber nur ein Aspekt bei der Planung Ihrer KI-Workloads. Sie müssen auch das übrige KI-Portfolio eines Anbieters berücksichtigen, wenn Sie eine KI-Implementierung in Angriff nehmen. Im Folgenden werden weitere Kategorien erörtert, die für diese KI-Portfolios entscheidend sind, darunter Client-Workstations, Cloud-native Produkte, Storage und mehr. Wir heben auch Bereiche hervor, in denen Angebote von Dell im Vergleich zu HPE einen Vorteil bieten können.

Workstations

Für KI-EntwicklerInnen und Data Scientists bieten Dell Precision Data Science-Workstations NVIDIA RTX™-GPUs und Intel Xeon® CPUs sowie eine Reihe von Data-Science-Tools.⁴⁰ Diese Systeme nutzen professionelle Compute-Optionen mit NVIDIA-GPUs, die für mehr als 100 professionelle Anwendungen zertifiziert sind⁴¹, und Intel Xeon Scalable Prozessorbeschleunigern wie Intel DL Boost.⁴² Precision-Workstations sind in Mobile-, Tower- und Rack-Formaten erhältlich und eignen sich für Anforderungen, die von größeren, stationären Datenanalysen bis hin zur wissenschaftlichen Feldmodellierung unterwegs reichen.

Das Angebot an HPE Workstations ist schmäler und umfasst in erster Linie einzelne Workstation-Tower, die mit NVIDIA L4s ausgestattet sind; HPE bietet keine mobile Workstation-Option an.⁴³ Die Workstation-Tower-Angebote eignen sich zwar für viele Aufgaben, bieten jedoch nicht die gleiche Flexibilität und Workload-Abdeckung wie das umfangreichere Angebot von Dell. Die verschiedenen Größen- und Portabilitätsoptionen der Dell Precision-Workstations ermöglichen maßgeschneiderte Lösungen, die den unterschiedlichen Anforderungen von Laboren, Büros und AußendienstmitarbeiterInnen gerecht werden.



Storage

Storage kann für die Ausführung von KI-Workloads genauso wichtig sein wie die Rechenleistung. Mehr Daten verbessern die Genauigkeit des KI-Modells, aber das Speichern und Managen riesiger Datenvolumen kann die Kapazitäten vieler Rechenzentren überfordern. Da Modelle in der Regel mit unstrukturierten Daten trainiert werden, müssen KI-fähige Storage-Systeme viele verschiedene Datentypen problemlos verarbeiten können.⁴⁴ Zur Bereitstellung von Kapazität und Skalierung für KI-, ML- und DL-Datensätze bietet Dell die PowerScale™-Serie für Datei-Storage und Elastic Cloud Storage (ECS) oder softwarebasiertes ObjectScale für Objektspeicher an.

Das PowerScale-All-Flash-NAS-Portfolio bietet Kapazitätsoptionen von 3,84 TB bis zu 720 TB Rohkapazität pro Node mit geclusterten All-Flash-Kapazitäten von 186 PB Rohkapazität. Die Flexibilität und Skalierung von PowerScale kann eine Vielzahl von Kunden- und KI-Anwendungsfällen unterstützen.⁴⁵ Im Cluster kann der PowerScale F900 bis zu 186 PB an Rohspeicher erreichen.⁴⁶ Alle drei All-Flash PowerScale-Modelle – F200, F600 und F900 – beinhalten Inline-Datenkomprimierung und -Deduplizierung, um die Storage-Effizienz zu verbessern.⁴⁷ Jedes PowerScale-Storage-Modell verwendet das Dell OneFS™-Dateisystem, das Polycys für das Tiering von Storage verwendet, um die wichtigsten Daten zur Workload-Optimierung auf den schnellsten Tiers zu priorisieren.⁴⁸ Dell bietet außerdem OneFS-Software auf dem AWS-Markt mit Dell APEX File Storage for AWS an. Kunden können OneFS mit ihren AWS-Compute-Instanzen für ein konsistentes Nutzererlebnis mit denselben Funktionen nutzen, die in On-Premise-OneFS-Arrays verfügbar sind.⁴⁹ HPE bietet zwar eine Public-Cloud-Integration für hybride Speicherlösungen an, aber wir haben in dem Angebot keine Cloud-native Option wie Dell APEX File Storage for AWS gefunden.

Zu den Objektspeicheroptionen von Dell zählen Dell Enterprise Object Storage (ECS), der „speziell für die Speicherung unstrukturierter Daten in Public-Cloud-Skalierung entwickelt wurde“.⁵⁰ Neben der integrierten Kompatibilität mit Amazon S3-Objektspeicher für Hybrid-Cloud-Funktionen bieten ECS-Storage-Nodes Kapazitäten von bis zu 14 PB pro Rack.⁵¹ HPE bietet ebenfalls unstrukturierten Storage mit Datei- und Objektspeicheroptionen, obwohl das Objektspeicherangebot über eine Partnerschaft mit Scality erfolgt. Kunden können HPE Solutions for Scality von HPE erwerben.⁵²

Dienstleistungen

Dell bietet ein breites Spektrum an professionellen Services an, darunter Beratung, Datenaufbereitung, Bereitstellung, Support und Managed Services zur Unterstützung von KI-Implementierungen. Für Unternehmen, die nach validierten Architekturen und Lösungen suchen, bietet Dell validierte Designs für KI an, die auf bestimmte Anwendungsfälle abzielen, um das Rätselraten beim Design und der Bereitstellung von KI-Ressourcen zu vermeiden. Diese von Dell validierten KI-Lösungen umfassen Hardware- und Software-Bundles, KI-Konversationsmodelle, maschinelle Lernverfahren und mehr. Durch die Kombination vorkonfigurierter, speziell entwickelter Lösungen mit KI-bezogenen Services bietet Dell eine umfassende KI-Lösung für das gesamte Spektrum der KI-Anforderungen. Diese Angebote könnten im Vergleich zur Entwicklung von Ad-hoc-Lösungen einen schnelleren und einfacheren Weg zum KI-Erfolg bieten.

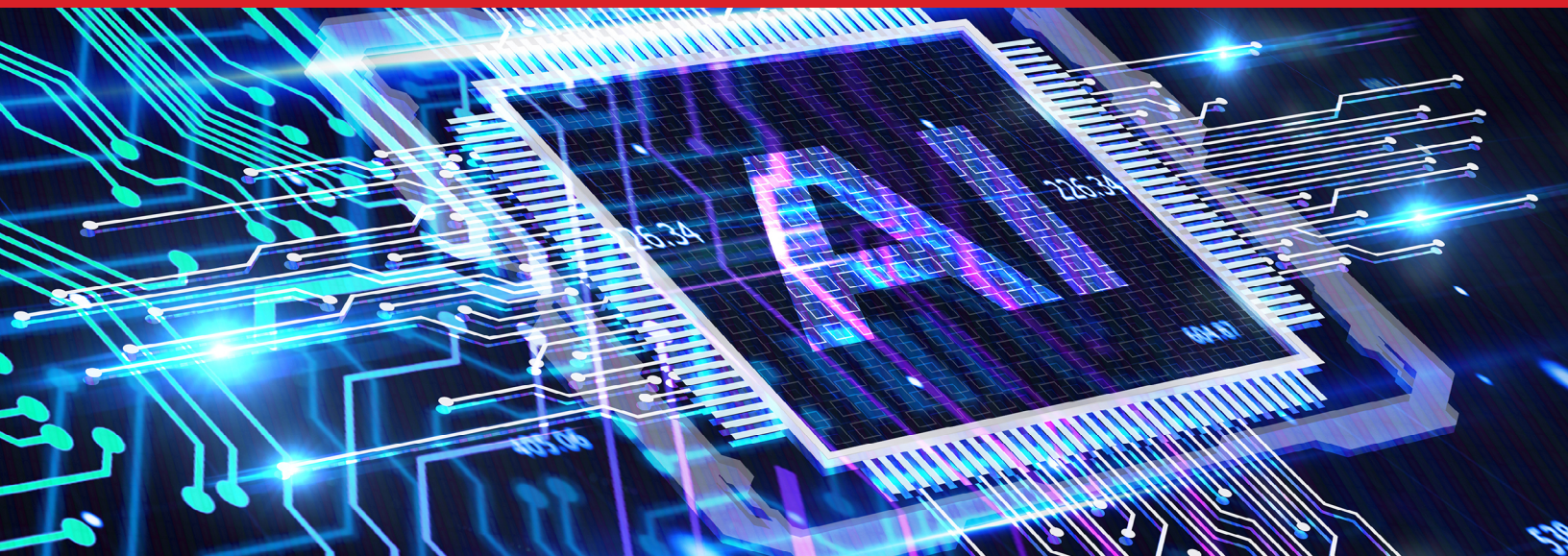
Dell Services können auch Ihre KI-Reise von der Beratung bis zur Implementierung begleiten. Die Dell ProConsult Advisory Services helfen KundInnen dabei, herauszufinden, wo NutzerInnen von der Einführung von GenAI-Prozessen profitieren können, und erstellen eine Roadmap mit den erforderlichen Lösungen und IT-Kompetenzen. Die Services von Dell können Daten für die Integration des Large Language Model vorbereiten und IT-Teams in KI-Wissen schulen. Für eine vollständige Einführung von GenAI überprüfen Dell Teams Ihre spezifischen Anwendungsfälle und ermitteln, implementieren und konfigurieren das beste KI-Modell für Ihre Anforderungen. HPE bietet ebenfalls professionelle Services an, um Unternehmen bei ihren KI-Bemühungen zu unterstützen.^{53,54}

Managementüberlegungen

Server erfordern ein fortlaufendes Management, das Administratorzeit in Anspruch nimmt. Firmware, Software und Treiber benötigen regelmäßige Updates, IT-MitarbeiterInnen müssen Performance und Temperaturen optimieren und aufrechterhalten und vieles mehr. In früheren Tests von Principled Technologies (PT) haben wir die Managementfunktionen von Dell Servern mit Integrated Dell Remote Access Controller 9 (iDRAC9) bewertet.⁵⁵ AdministratorInnen können sich auf automatisierte Online-Updates mit iDRAC9 OpenManage™ Enterprise (OME) mit konfigurierbarer Planung verlassen, um ihre Server auf dem neuesten Stand zu halten, und mithilfe von Profilen schnell und einfach neue Server einbinden, wenn die Arbeitslast wächst. Mit iDRAC und OME können Dell KundInnen auf mehr Remote-Managementfunktionen zugreifen, Server einfacher bereitstellen und Firmwareupdates einfacher durchführen als mit HPE OneView und HPE iLO. Dell PowerEdge-Server umfassen Dell Management und Services, die Unternehmen helfen könnten, „den Zeit- und Arbeitsaufwand für Aufgaben wie das Monitoring der Systemintegrität oder Firmwareupdates zu reduzieren“ und IT-Zyklen für Innovationen und andere Aufgaben freizusetzen.⁵⁶

Tabelle 3: Zusammenfassung des Vergleichs zwischen den Managementtools von Dell und HPE aus einem PT-Bericht vom November 2022.⁵⁷
Quelle: Principled Technologies

	Was ist der Unterschied zu den Managementtools von Dell	Wie viel besser
Mehr Remote-Managementfunktionen iDRAC im Vergleich zu iLO	Mehr HTML5-Konsolen- und BIOS-Konfigurationsfunktionen für mehr Remotefunktionen in iDRAC	Das 2,5fache an HTML5-Konsolenfunktionen und das 13fache an BIOS-Funktionen
Einfachere Serverbereitstellung OME im Vergleich zu OneView	1:n-Profilbereitstellung mit OME	52 % weniger Zeit für die Bereitstellung eines Servers als mit OneView
Einfachere Firmwareupdates OME im Vergleich zu OneView	Automatisierte Online-Updates mit OME	Aktualisieren Sie mehrere Server, indem Sie eine Verbindung zu Dell.com herstellen. So sparen Sie die Zeit, die für die Aktualisierung von Servern durch manuelles Hochladen von Paketen mit OneView erforderlich ist.
Einfachere Warnmeldungen OME im Vergleich zu OneView	Einrichten von Warnrichtlinien in OME und Ausführen automatischer Aktionen auf der Grundlage von Warnmeldungen	Die Automatisierung dieses Prozesses spart Zeit und reduziert das Fehlerpotenzial im Vergleich zur manuellen Ausführung von Aktionen bei jedem Eingang einer Warnung in OneView.
Einfacher zu verwendende Sicherheitsfunktionen (Systemsperrung und dynamischer USB-Anschluss) iDRAC im Vergleich zu iLO	Weniger Schritte, weniger Zeit, keine Neustarts mit iDRAC	¼ Schritte, 91 % weniger Zeit für die Systemsperrung
Mehr robuste Analysen CloudIQ für PowerEdge im Vergleich zu InfoSight	Anpassbare Berichte, mehr Integritätskennzahlen für eine bessere administrative Kontrolle mit CloudIQ für PowerEdge	Mehr als das 15fache an Kennzahlen zur Auswahl im Vergleich zu InfoSight



Entscheidung

Die Nutzung des Potenzials von KI zur Rationalisierung und Verbesserung des Geschäftsbetriebs kann eine schwierige Aufgabe mit erheblichen geschäftlichen Auswirkungen sein. Da die Technologie schneller denn je voranschreitet, ist die Partnerschaft mit dem richtigen Anbieter für KI von entscheidender Bedeutung. Wenn Sie sich für ein Unternehmen wie Dell entscheiden, das nicht nur ein umfassendes KI-Portfolio bietet, sondern auch Planungs-, Vorbereitungs-, Implementierungs- und Managementservices bereitstellt, können KundInnen diese Herausforderungen direkt angehen. MLPerf® Benchmarktests zeigen, dass die Angebote im Dell KI-Portfolio eine konsistente, starke Performance für KI-Workloads bieten. Mit leistungsstarken und flexiblen Serveroptionen sowie einer Vielzahl an Storage-Optionen, validierten Lösungen und speziell auf KI zugeschnittenen Services kann Dell Unternehmen dabei unterstützen, KI und ihre Vorteile zu nutzen.

1. Dell, „Increasing Your Data Value with Dell Generative AI Solutions“, abgerufen am 19. Dezember 2023, <https://www.dell.com/en-us/blog/increasing-your-data-value-with-dell-generative-ai-solutions/>.
2. Dell, „Dell AI Solutions“, abgerufen am 12. Dezember 2023, <https://www.dell.com/en-us/dt/solutions/artificial-intelligence/index.htm#accordion0&tab0=0>.
3. Dell, „Dell Technologies and Hugging Face to Simplify Generative AI with on-Premise IT“, abgerufen am 12. Dezember 2023, <https://www.dell.com/en-us/dt/corporate/newsroom/announcements/detailpage.press-releases~usa~2023~11~20231114-dell-technologies-and-hugging-face-to-simplify-generative-ai-with-on-premises-it.htm#/filter-on/Country:en-us>.
4. Dell, „Dell and Meta Collaborate to Drive Generative AI Innovation“, abgerufen am 12. Dezember 2023, <https://www.dell.com/en-us/blog/dell-and-meta-collaborate-to-drive-generative-ai-innovation/>.
5. Dell, „Dell AI-Ready Data Platform“, abgerufen am 12. Dezember 2023, <https://www.dell.com/en-us/dt/solutions/artificial-intelligence/storage-for-ai.htm?hve=explore+unstructured+storage#tab0=0>.
6. Dell, „Snowflake and Dell Partnership Gases Momentum“, abgerufen am 19. Dezember 2023, <https://www.dell.com/en-us/blog/snowflake-and-dell-partnership-gains-momentum/>.
7. Robert McNeal, „Dell, VMware and NVIDIA Bring AI to Your Data“, abgerufen am 17. Januar 2024, <https://www.dell.com/en-us/blog/dell-vmware-and-nvidia-bring-ai-to-your-data/>. Über den obigen Link: „Basierend auf einer Analyse von Dell, August 2023. Dell Technologies bietet Lösungen, die für Workstations (mobil und stationär) bis Server mühelos KI-Workloads für High-Performance-Computing, Daten-Storage, Cloud-native softwarebasierte Infrastrukturen, Netzwerk-Switches, Data Protection, HCI und Services unterstützen.“
8. Intel, „Accelerate Artificial Intelligence (AI) Workloads with Intel Advanced Matrix Extensions (Intel AMX)“, abgerufen am 12. Dezember 2023, <https://www.intel.com/content/www/us/en/content-details/785250/accelerate-artificial-intelligence-ai-workloads-with-intel-advanced-matrix-extensions-intel-amx.html>.

-
9. Vipera, „NVIDIA's H100 and A100 GPU Cards: Exploring the Intricacies of SXM and PCI-E Connections“, abgerufen am 12. Dezember 2023, <https://www.viperatech.com/unraveling-the-mysteries-sxm-vs-pci-e-connections-in-nvidias-high-end-h100-and-a100-gpus/>.
 10. GitHub, „MLPerf® Results Messaging Guidelines“, abgerufen am 16. Januar 2024, https://github.com/mlcommons/policies/blob/master/MLPerf_Results_Messaging_Guidelines.adoc.
 11. MLCommons®, „MLPerf® Inference: Datacenter Benchmark Suite Results“, abgerufen am 12. Dezember 2023, <https://mlcommons.org/en/inference-datacenter-31/>.
 12. Dell, „PowerEdge XE Servers“, abgerufen am 12. Dezember 2023, <https://www.dell.com/en-us/dt/servers/specialty-servers/poweredge-xe-servers.htm?hve=explore+poweredge+xe#tab0=0>.
 13. HPE, „HPE ProLiant XL675d Gen10 Plus Configure-to-order Server“, abgerufen am 12. Dezember 2023, <https://www.hpe.com/us/en/product-catalog/compute/proliant-servers/pip.1013142988.html>.
 14. HPE, „HPE ProLiant DL380a Gen11“, abgerufen am 12. Dezember 2023, <https://www.hpe.com/us/en/product-catalog/compute/proliant-servers/pip.proliant-dl380-server.1014696168.html>.
 15. MLCommons®, „MLPerf® Inference: Datacenter Benchmark Suite Results v 3.1“, abgerufen am 12. Dezember 2023, <https://mlcommons.org/benchmarks/inference-datacenter/>.
 16. HPE, „NVIDIA Accelerators for HPE ProLiant Servers“, abgerufen am 12. Dezember 2023, https://www.hpe.com/psnow/doc/c04123180.html?jumpid=in_pdp-psnow-qs.
 17. Dell, „PowerEdge XE9680 – Datenblatt“, abgerufen am 19. Januar 2024, <https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/poweredge-xe9680-spec-sheet.pdf>.
 18. HPE, „HPE & NVIDIA Financial Services Solution Sets New Records in Performance“, abgerufen am 12. Dezember 2023, <https://community.hpe.com/t5/alliances/hpe-amp-nvidia-financial-services-solution-sets-new-records-in/ba-p/7197388>.
 19. HPE, „QuickSpecs: HPE Cray Supercomputing XD670“, abgerufen am 12. Dezember 2023, <https://www.hpe.com/psnow/doc/a50004292enw>.
 20. Verifizierter MLPerf®-Score von v3.1 Inference Closed. Abgerufen von <https://mlcommons.org/benchmarks/inference-datacenter/> 5. Dezember 2023, Eintrag 3.1-0069. Der Name und das Logo von MLPerf® sind eingetragene und nicht eingetragene Marken der MLCommons® Association in den USA und anderen Ländern. Alle Rechte vorbehalten. Unerlaubte Nutzung ist strengstens untersagt. Weitere Informationen siehe www.mlcommons.org.
 21. Verifizierter MLPerf®-Score von v3.1 Inference Closed. Abgerufen von <https://mlcommons.org/benchmarks/inference-datacenter/> 5. Dezember 2023, Eintrag 3.1-0085. Der Name und das Logo von MLPerf® sind eingetragene und nicht eingetragene Marken der MLCommons® Association in den USA und anderen Ländern. Alle Rechte vorbehalten. Unerlaubte Nutzung ist strengstens untersagt. Weitere Informationen siehe www.mlcommons.org.
 22. Dell, „PowerEdge XE9640 Rack Server“, abgerufen am 12. Dezember 2023, <https://www.dell.com/en-us/shop/ipovw/poweredge-xe9640>.
 23. Accelsius, „Enabling the AI Revolution with Liquid Cooling“, abgerufen am 12. Dezember 2023, <https://www.accelsius.com/blog/enabling-the-ai-revolution-with-liquid-cooling>.
 24. Dell, „Dell PowerEdge XE9640 – Technisches Handbuch“, abgerufen am 12. Dezember 2023, <https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/poweredge-xe9640-technical-guide.pdf>.
 25. HPE, „HPE ProLiant DL380a Gen11“, abgerufen am 12. Dezember 2023, https://www.hpe.com/psnow/doc/PSN1014696168WWEN.pdf?jumpid=in_pdp-psnow-dds.
 26. Intel, „Intel® Data Center GPU Max Series Technical Overview“, abgerufen am 12. Dezember 2023, <https://www.intel.com/content/www/us/en/developer/articles/technical/intel-data-center-gpu-max-series-overview.html#gs.08874l>.
 27. HPE, „Intel Data Center GPU Max 1100 48GB Accelerator for HPE Data Sheet“, abgerufen am 12. Dezember 2023, <https://www.hpe.com/psnow/doc/PSN1014779728WWEN>.
 28. Verifizierter MLPerf®-Score von v3.1 Inference Closed. Abgerufen von <https://mlcommons.org/benchmarks/inference-datacenter/> 5. Dezember 2023, Eintrag 3.1-0066. Der Name und das Logo von MLPerf® sind eingetragene und nicht eingetragene Marken der MLCommons® Association in den USA und anderen Ländern. Alle Rechte vorbehalten. Unerlaubte Nutzung ist strengstens untersagt. Weitere Informationen siehe www.mlcommons.org.
 29. Verifizierter MLPerf®-Score von v3.1 Inference Closed. Abgerufen von <https://mlcommons.org/benchmarks/inference-datacenter/> 5. Dezember 2023, Eintrag 3.1-0084. Der Name und das Logo von MLPerf® sind eingetragene und nicht eingetragene Marken der MLCommons® Association in den USA und anderen Ländern. Alle Rechte vorbehalten. Unerlaubte Nutzung ist strengstens untersagt. Weitere Informationen siehe www.mlcommons.org.

-
30. Dell, „AI and HPC – with Air or Liquid Cooling“, abgerufen am 12. Dezember 2023, <https://www.delltechnologies.com/asset/en-us/products/servers/briefs-summaries/poweredge-xe9640-and-xe8640-infographic.pdf>.
 31. Dell, „PowerEdge XE8640: Drive AI, HPC Modeling and Simulation Workloads with Superior Performance“, abgerufen am 12. Dezember 2023, <https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/poweredge-xe8640-spec-sheet.pdf>.
 32. Dell, „PowerEdge XE8640 Rack Server“, abgerufen am 12. Dezember 2023, <https://www.dell.com/en-us/shop/ipovw/poweredge-xe8640>.
 33. Verifizierter MLPerf®-Score von v3.1 Inference Closed. Abgerufen von <https://mlcommons.org/benchmarks/inference-datacenter/> 5. Dezember 2023, Eintrag 3.1-0067. Der Name und das Logo von MLPerf® sind eingetragene und nicht eingetragene Marken der MLCommons® Association in den USA und anderen Ländern. Alle Rechte vorbehalten. Unerlaubte Nutzung ist strengstens untersagt. Weitere Informationen siehe www.mlcommons.org.
 34. Verifizierter MLPerf®-Score von v3.1 Inference Closed. Abgerufen von <https://mlcommons.org/benchmarks/inference-datacenter/> 5. Dezember 2023, Eintrag 3.1-0084. Der Name und das Logo von MLPerf® sind eingetragene und nicht eingetragene Marken der MLCommons® Association in den USA und anderen Ländern. Alle Rechte vorbehalten. Unerlaubte Nutzung ist strengstens untersagt. Weitere Informationen siehe www.mlcommons.org.
 35. Dell, „PowerEdge R760xa Rack Server“, abgerufen am 12. Dezember 2023, https://www.dell.com/en-us/shop/dell-poweredge-servers/poweredge-r760xa-rack-server/spd/poweredge-r760xa/pe_r760xa_16902_vi_vp#features_section.
 36. SANStorageWorks, „Dell EMC PowerEdge R760xa: Powerful and scalable for GPU workloads“, abgerufen am 12. Dezember 2023, <https://www.sanstorageworks.com/PowerEdge-R760xa.asp>.
 37. Dell, „Dell PowerEdge Servers and NVIDIA GPUs“, abgerufen am 12. Dezember 2023, <https://infohub.delltechnologies.com/l/design-guide-generative-ai-in-the-enterprise-inferencing/dell-poweredge-servers-and-nvidia-gpus-1/>.
 38. Verifizierter MLPerf®-Score von v3.1 Inference Closed. Abgerufen von <https://mlcommons.org/benchmarks/inference-datacenter/> 5. Dezember 2023, Eintrag 3.1-0064. Der Name und das Logo von MLPerf® sind eingetragene und nicht eingetragene Marken der MLCommons® Association in den USA und anderen Ländern. Alle Rechte vorbehalten. Unerlaubte Nutzung ist strengstens untersagt. Weitere Informationen siehe www.mlcommons.org.
 39. Verifizierter MLPerf®-Score von v3.1 Inference Closed. Abgerufen von <https://mlcommons.org/benchmarks/inference-datacenter/> 5. Dezember 2023, Eintrag 3.1-0084. Der Name und das Logo von MLPerf® sind eingetragene und nicht eingetragene Marken der MLCommons® Association in den USA und anderen Ländern. Alle Rechte vorbehalten. Unerlaubte Nutzung ist strengstens untersagt. Weitere Informationen siehe www.mlcommons.org.
 40. Dell, „Workstations for AI“, abgerufen am 12. Dezember 2023, <https://www.dell.com/en-us/dt/ai-technologies/index.htm?hve=explore+dell+precision+for+ai#pdf-overlay=https://www.delltechnologies.com/asset/en-us/products/workstations/briefs-summaries/ai-industry-brochure.pdf>.
 41. NVIDIA, „NVIDIA RTX in Professional Workstations“, abgerufen am 12. Dezember 2023, <https://www.nvidia.com/en-us/design-visualization/desktop-graphics/>.
 42. Intel, „Intel® Deep Learning Boost (Intel® DL Boost)“, abgerufen am 12. Dezember 2023, <https://www.intel.com/content/www/us/en/artificial-intelligence/deep-learning-boost.html>.
 43. HPE, „HPE ProLiant ML350 Gen11“, abgerufen am 12. Dezember 2023, <https://buy.hpe.com/us/en/compute/tower-servers/proliant-ml300-servers/proliant-ml350-server/hpe-proliant-ml350-gen11/p/1014696172>.
 44. ComputerWeekly.com, „Storage Requirements for AI, ML and Analytics in 2022“, abgerufen am 12. Dezember 2023, <https://www.computerweekly.com/feature/Storage-requirements-for-AI-ML-and-analytics-in-2022>.
 45. Dell, „PowerScale AI-Ready Data Platform“, abgerufen am 12. Dezember 2023, <https://www.dell.com/en-us/shop/powerscale-family/sf/powerscale>.
 46. Dell, „Compare PowerScale“, abgerufen am 12. Dezember 2023, <https://www.dell.com/en-us/shop/powerscale-family/sf/powerscale#compare-module>.
 47. Dell, „Dell PowerScale All-Flash“, abgerufen am 12. Dezember 2023, <https://www.delltechnologies.com/asset/en-us/products/storage/technical-support/h15963-ss-powerscale-all-flash-nodes.pdf>.
 48. Dell, „Dell PowerScale OneFS Software Features“, abgerufen am 12. Dezember 2023, <https://www.delltechnologies.com/asset/en-us/products/storage/technical-support/h18275-onefs-software-features-data-sheet.pdf>.
 49. Dell, „Dell APEX File Storage for AWS“, abgerufen am 12. Dezember 2023, <https://www.delltechnologies.com/asset/en-us/products/storage/briefs-summaries/h19575-so-apex-file-storage-for-aws.pdf>.

-
50. Dell, „Dell ECS Enterprise Object Storage“, abgerufen am 12. Dezember 2023, <https://www.dell.com/en-us/dt/storage/ecs/index.htm?hve=explore+ecs#tab0=0&tab1=0>.
51. Dell, „Dell ECS Enterprise Object Storage“, abgerufen am 12. Dezember 2023, <https://www.dell.com/en-us/dt/storage/ecs/index.htm#tab0=0&tab1=0&accordion0>.
52. HPE, „Storage Solutions for Scalability“, abgerufen am 12. Dezember 2023, <https://www.hpe.com/us/en/storage/file-object/scalability.html>.
53. HPE, „Make AI Work for You“, abgerufen am 16. Januar 2024, <https://www.hpe.com/us/en/solutions/ai-artificial-intelligence.html>.
54. HPE, „HPE AI Services – Generative AI Implementation“, abgerufen am 16. Januar 2024, <https://www.hpe.com/us/en/services/generative-ai-implementation-service.html>.
55. Principled Technologies, „Simplify administrator tasks and improve security and health monitoring with tools from the Dell management portfolio vs. comparable tools from HPE“, abgerufen am 12. Dezember 2023, <https://www.principledtechnologies.com/Dell/Management-tools-vs-HPE-1122.pdf>.
56. Principled Technologies, „Simplify administrator tasks and improve security and health monitoring with tools from the Dell management portfolio vs. comparable tools from HPE“, abgerufen am 12. Dezember 2023, <https://www.principledtechnologies.com/Dell/Management-tools-vs-HPE-1122.pdf>.
57. Principled Technologies, „Simplify administrator tasks and improve security and health monitoring with tools from the Dell management portfolio vs. comparable tools from HPE“

Der Name und das Logo von MLPerf sind eingetragene und nicht eingetragene Marken der MLCommons Association in den USA und anderen Ländern. Alle Rechte vorbehalten. Unerlaubte Nutzung ist strengstens untersagt. Weitere Informationen siehe www.mlcommons.org.

► Lesen Sie die Originalversion dieses Berichts in englischer Sprache unter <https://facts.pt/zPmSx4c>.

Dieses Projekt wurde in Auftrag gegeben von Dell Technologies.



Facts matter.®

Principled Technologies ist eine eingetragene Marke von Principled Technologies, Inc. Alle anderen Produktnamen sind Marken der jeweiligen Inhaber.

GEWÄHRLEISTUNGS-AUSSCHLUSS, HAFTUNGSEINSCHRÄNKUNG:

Principled Technologies, Inc. hat angemessene Anstrengungen unternommen, die Genauigkeit und Richtigkeit der Tests sicherzustellen. Principled Technologies, Inc. schließt jedoch jegliche ausdrückliche und implizite Gewährleistung aus, die sich auf die Testergebnisse und Analysen, deren Genauigkeit, Vollständigkeit oder Qualität bezieht, einschließlich jeglichen impliziten Eignungsversprechens für einen bestimmten Zweck. Alle natürlichen oder juristischen Personen, die sich auf die Ergebnisse der Tests verlassen, tun dies auf eigenes Risiko und stimmen zu, dass Principled Technologies, Inc., seine Mitarbeiter und Auftragsnehmer keinerlei Haftung für Verlust oder Beschädigung jeglicher Art aufgrund vermeintlicher Fehler oder Mängel in einem Testverfahren oder Testergebnis übernehmen.

In keinem Fall haftet Principled Technologies, Inc. für indirekte, spezielle, zufällige oder Folgeschäden in Verbindung mit den Tests, auch wenn das Unternehmen auf die Möglichkeit solcher Schäden hingewiesen wurde. In keinem Fall geht die Haftung von Principled Technologies, Inc. über den in Verbindung mit den Tests von Principled Technologies, Inc. gezahlten Betrag hinaus. Dies gilt auch für direkte Schäden. Die alleinigen und ausschließlichen Rechtsmittel, die dem Kunden zur Verfügung stehen, sind die in diesem Dokument beschriebenen Rechtsmittel.