

Turbo für Ihre KI- Anwendungen

Mit der richtigen Infrastruktur mildern Sie Datenüberlastung und schöpfen das Potenzial Ihrer KI voll aus.

verizon
business

Vorwort

Bringen Sie Ihr KI-Ökosystem auf Hochtouren – mit einem Netzwerk, das Ihnen hilft, das volle Potenzial Ihrer KI-Initiativen auszuschöpfen.



David Bailey
Global Solutions Executive

Seit mehr als zwei Jahrzehnten hilft David Bailey globalen Unternehmen und öffentlichen Einrichtungen, die passende Infrastruktur und Sicherheit für ihre digitalen Anforderungen aufzubauen. Sein Fokus reicht von kabelgebundenen und drahtlosen IoT-Netzen bis hin zu agilen, leistungsfähigen Architekturen für KI-Anwendungen.

Künstliche Intelligenz (KI) wird in rasantem Tempo weiterentwickelt. Lesen Sie in diesem kurzen Whitepaper, warum es so wichtig ist, dass die unteren Technologieebenen mit den KI-Modellen und der Rechenleistung Schritt halten. Eine solche erfolgsentscheidende Technologieebene ist ein leistungsstarkes, latenzarmes Netzwerk.

Wir alle wissen, wie schnell die Entwicklung und Nutzung von künstlicher Intelligenz und maschinellem Lernen Fahrt aufgenommen haben. In vielen, wenn nicht sogar den meisten, Vorständen in allen Branchen wird diskutiert, wie das eigene Unternehmen mit KI wettbewerbsfähiger werden, seinen Umsatz steigern, Betriebsprozesse optimieren und Kundenerlebnisse verbessern könnte. Die Erwartungen an KI sind hoch. Um diese zu erfüllen, sind weiterhin kontinuierliche Investitionen in schnellere, leistungsfähigere Hard- und Software erforderlich.

Hersteller wie Nvidia, AMD, Intel und Qualcomm liefern sich ein Rennen um die besten, für KI-Workloads optimierten Chips. Gleichzeitig haben Cloud-Betreiber wie Google, Amazon und Microsoft begonnen, ihre eigenen Prozessoren zu entwickeln, um stark skalierbare KI-Dienste anbieten zu können und die Anbieterabhängigkeit zu reduzieren. Mit dieser spezialisierten Hardware können Cloud-Betreiber die Workloads für ihre öffentlich zugänglichen KI-Services innerhalb von Millisekunden bewältigen.

Diese Services können eine gute Grundlage für die schnelle Entwicklung und Skalierung Ihrer Lösungen sein. Doch wenn Sie Daten in der Cloud verarbeiten, spielt auch das Netzwerk eine wichtige Rolle. Das Netzwerk, über das Daten von der Quelle in die Cloud oder in Ihr Rechenzentrum übertragen werden – oft das Internet – bestimmt, wie reaktionsschnell Ihre Anwendungen wirklich sind. Ein langsames Netzwerk kann die durch Investitionen in schnellere KI-Chips erzielten Geschwindigkeitsgewinne zunichte machen.

Stellen Sie sich vor, Sie bestellen online einen Burger. Ihre Bestellung könnte innerhalb weniger Minuten fertig sein (wie beim hochleistungsfähigen KI-Chip). Doch wenn der Fahrer (die Internetverbindung) langsam ist und 30 Minuten braucht, um Ihnen Ihren Burger zu bringen, ist es unwichtig, wie schnell er zubereitet wurde: Er kommt trotzdem kalt an.

In diesem Whitepaper erläutere ich, warum die Latenz (die Lieferzeit) eine so wichtige Rolle spielt, wie sie sich minimieren lässt und wie Sie Ihre KI-Initiativen auf Erfolg ausrichten.

Immer mehr Unternehmen entwickeln speziell auf ihre Leistungsanforderungen abgestimmte KI-Chips. Ein Beispiel ist der Meta Training and Inference Accelerator (MTIA). Dieser KI-Beschleuniger der zweiten Generation hat eine höhere Rechenleistung und Speicherbandbreite, was die Leistung der KI-Anwendungen von Meta – wie der Engines, die das Unternehmen für das Ranking und die Empfehlung von Inhalten auf Facebook und Instagram nutzt – insgesamt steigert.

Auch außerhalb von Clouds und Rechenzentren wird maßgeschneiderte KI-Hardware eingesetzt. Beispiele:

- Im Gesundheitswesen müssen CT-, MRT-, PET- und andere Scanner riesige Datenmengen schnell verarbeiten.
- Neue Fahrzeuge sind mit einer Vielzahl an Sensoren ausgestattet, unter anderem für Sicherheitsfeatures wie die Kollisionsvermeidung. Damit Fahrzeuge nahezu in Echtzeit kritische Entscheidungen treffen können, sind diese Sensoren mit extrem leistungsfähigen Chips verbunden.

In den oben aufgeführten Beispielen werden Daten zum Großteil „lokal“ verarbeitet – also im Fahrzeug oder Scanner –, auch wenn einige nach der Verarbeitung in die Cloud übertragen werden.

Doch in vielen Anwendungsszenarien werden Daten extern verarbeitet. Zum Beispiel:

Finanzwesen

- Riesige Mengen an historischen und Echtzeitdaten werden zusammengeführt und analysiert, um Entscheidungen zu treffen, wenn es auf jede Millisekunde ankommt. Das kann zum finanziellen Erfolg beitragen.
- Verhaltens- und biometrische Daten werden analysiert, um Betrugsversuche aufzudecken. Diese KI-gestützten Analysen sind oft effektiver und stören das Nutzererlebnis weniger als traditionelle Methoden wie Sicherheitsfragen.

Fertigung

- Hier wird KI für die automatisierte Analyse von Live-Videoaufnahmen der Produktionslinien genutzt, um jedes Produkt in verschiedenen Phasen auf Qualität zu prüfen. Diese Methode ersetzt traditionelle Stichprobenkontrollen der fertigen Produkte. So können Probleme viel früher erkannt, Verschwendung minimiert und Maschinenschäden vorgebeugt werden.

Einzelhandel

- Mithilfe KI-gestützter Analysen von Kundenverhalten und Kaufhistorien können Einzelhändler Kunden ein besseres, personalisierteres Käuferlebnis bieten. Dies wirkt sich positiv auf das Kundenerlebnis insgesamt und auf den Customer Lifetime Value aus.

Es gibt viele weitere Anwendungsbeispiele aus Branchen wie der Fertigung, der Landwirtschaft und dem Dienstleistungssektor. Maßgeschneiderte KI-Hardware ist mittlerweile allgegenwärtig und viele Unternehmen haben bereits Anwendungen und Rechenleistung in die Cloud verlagert. Das bedeutet jedoch, dass die Orte, an denen die Daten erfasst, verarbeitet und gespeichert sind, durch ein Netzwerk miteinander verbunden sein müssen.

Das erste Problem

In den letzten Jahren hat sich die Kostenstruktur der Unternehmens-IT verändert, vor allem durch die rasant steigende Cloud-Nutzung. Nun versuchen viele interne IT- und Beschaffungsteams, Netzwerkausgaben wo möglich zu straffen. Auf der Suche nach Kosteneinsparungen beziehen viele den Netzwerkzugang (Underlay) getrennt von Netzwerkkomponenten wie Router und Switches (Overlay) und letztere von so vielen verschiedenen Anbietern, dass die Infrastruktur unübersichtlich wird.

Dieser Kostenfokus geht oft zu Lasten der Leistung und Servicequalität, was zum Teil an der weit verbreiteten Annahme liegt, dass aller Internetzugang gleich ist.

In Wahrheit gibt es bezüglich der Zuverlässigkeit, Leistung und insbesondere der Latenz große Unterschiede zwischen den Anbietern.

In dem Begleitdokument unten erfahren Sie mehr zu diesem Thema.



Getrennt oder gemeinsam?

Dieses Whitepaper aus derselben Reihe geht der immer wieder gestellten Frage nach, ob es besser ist, Underlay und Overlay von verschiedenen oder vom selben Anbieter zu beziehen.

Lesedauer: 15 Minuten

Warum Latenz wichtig ist

Die Latenz ist die Verzögerung zwischen einer Netzwerk-anfrage und der Antwort. Sie besteht aus drei Komponenten (siehe Grafik unten) und die vom Nutzer wahrgenommene Latenz ist die Summe dieser drei Komponenten.

Hohe Latenz kann zu verpassten Geschäfts- und Verkaufschancen oder zu spät erkannten Betrugsversuchen führen. In autonomen Fahrzeugen oder medizinischen Anwendungen kann sie sogar Menschenleben gefährden.

Wenn die KI Teil eines geschlossenen Systems (zum Beispiel eines Fahrzeugcomputers oder MRT-Scanners) ist, ist das Netzwerk direkt in das Gerät integriert. In diesen Fällen müssen Daten keine langen Strecken zurücklegen und weil die Leistung nicht von einer externen Infrastruktur abhängig ist, ist die Latenz minimal.

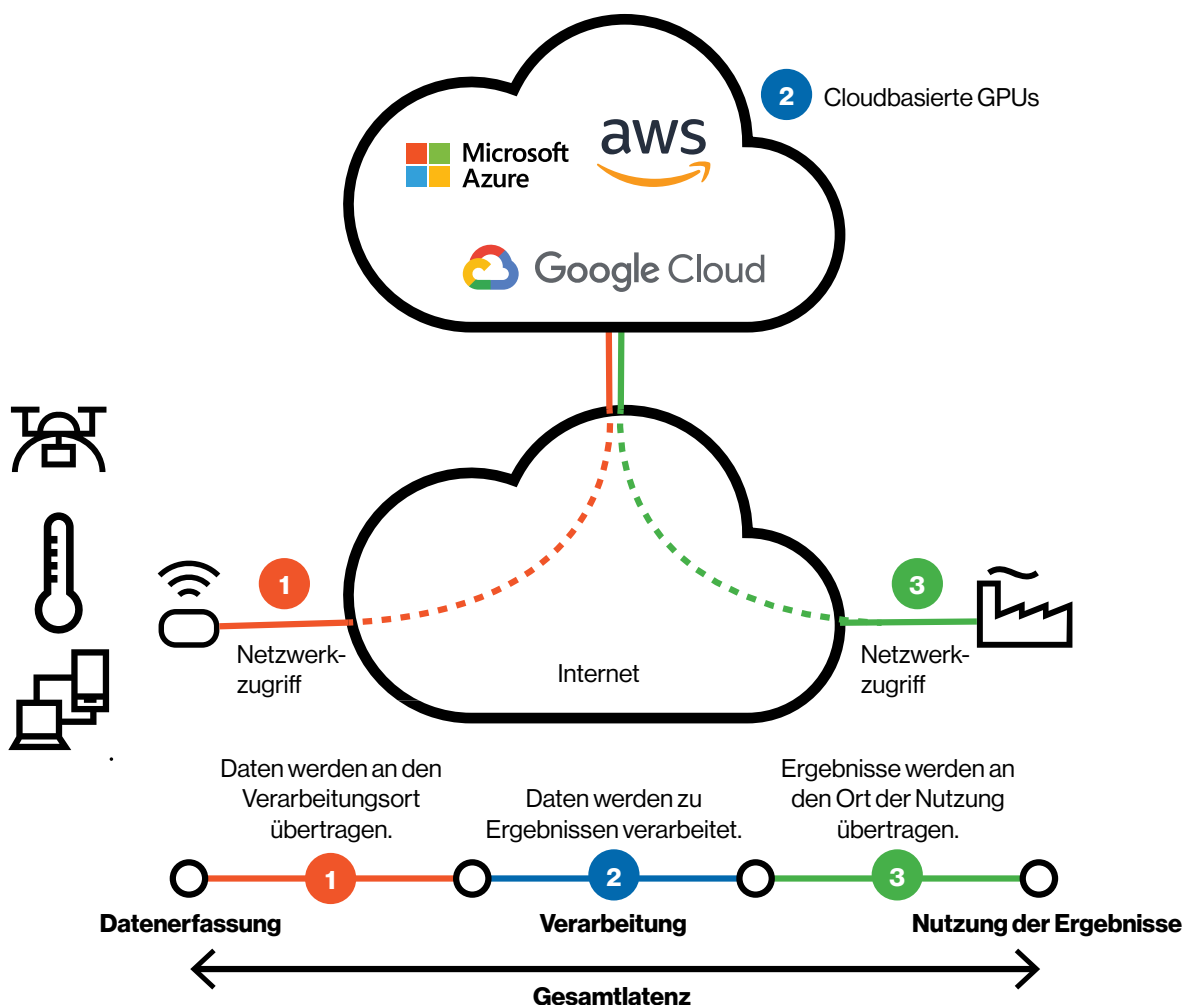
Wenn KI und die nötige Rechenleistung an einem Standort von vielen verschiedenen Geräten genutzt werden – zum Beispiel von Maschinen in einer Fabrikanlage –, muss die LAN-Infrastruktur (ob kabelgebunden oder drahtlos) eine latenzarme Übertragung ermöglichen.

Wenn KI über verschiedene Unternehmensstandorte verteilt genutzt wird, sind die Leistung und Latenz der KI-Anwendungen von den Drittanbietern der WAN-Infrastruktur abhängig.

Unternehmen sollten unter anderem die folgenden Faktoren berücksichtigen, um diese verschiedenen Anwendungsszenarien effektiv zu unterstützen:

- Wie viele Daten von einer Anwendung verarbeitet werden, bestimmt die nötige Rechenleistung.
- Welche Strecken Daten bei der Übertragung zurücklegen müssen, bestimmt die Netzwerklatenz.
- Ob eine Anwendung von Menschen oder Maschinen genutzt wird und wo sie zum Einsatz kommen soll, bestimmt die Auswirkungen von Latenz auf das Unternehmen.

An dieser Stelle sollte ich darauf hinweisen, dass das Internet ein Netzwerk aus vielen verschiedenen Netzwerken ist. Bei der Übertragung über das öffentliche Internet nutzen Daten unterschiedliche Pfade, weshalb es in der Regel keine garantierten Latenzzeiten gibt. Für Unternehmen, die für den Internetzugang in verschiedenen Ländern oder Niederlassungen unterschiedliche Netzwerkanbieter nutzen, ist das Risiko von Leistungs- und Latenzproblemen noch höher. Dasselbe gilt für Unternehmen, die Underlay und Overlay von verschiedenen Anbietern beziehen, die das Leistungs- und Latenzmanagement nicht miteinander abstimmen.



Wenn Sie je für die Verwaltung oder den Betrieb einer IT-Umgebung verantwortlich waren, brauche ich Sie nicht daran zu erinnern, wie wenig es die Endnutzer interessiert, warum eine Anwendung langsam ist. Sie wollen einfach nur, dass das Problem behoben wird. Dasselbe gilt für Unternehmen, die erheblich in kostspielige KI-Chips investiert haben, um die Leistung ihrer KI-Anwendungen zu verbessern. Wenn diese Erwartung nicht erfüllt wird, hilft es wenig, zu wissen warum nicht. Das Problem muss gelöst werden, damit die Investition sich auszahlt und die gesetzten Ziele erreicht werden.

Das nächste Problem

Ein Server- oder Prozessorupgrade oder der Kauf zusätzlicher Rechenleistung erfordert meist nur wenige Sekunden. Ein Netzwerk-Upgrade hingegen kann Monate dauern – bei einem auf viele Standorte verteilten WAN vielleicht sogar noch länger.

Sie haben sicher selbst schon erlebt, dass es manchmal ruckelt, wenn Sie von Zuhause aus an einer Videokonferenz teilnehmen. Das liegt meistens an der vergleichsweise geringeren Leistung der Internetverbindung, die Sie dort nutzen. Wenn Sie kritische KI-Daten über das Internet übertragen, haben Sie bei schlechter Anwendungsleistung keinen Einblick in mögliche Engpässe oder Routingprobleme – dazu finden Sie im unten aufgeführten Whitepaper weitere Informationen. All Ihre Investitionen in KI-Apps, moderne Hosting-Technologien und speziell entwickelte Chips nützen Ihnen also wenig, wenn die Netzwerkleistung nicht mithalten kann und Sie Netzwerkprobleme nicht schnell beheben können.

Je mehr Daten Sie in Ihre KI-Modelle einspeisen, desto höher die Anforderungen an Rechenleistung und Speicherkapazität. Was Sie dabei nicht vergessen dürfen, ist die Netzwerkinfrastruktur. Sie können klein anfangen, doch wichtig ist, für jede überarbeitete Version Ihrer KI-Anwendungen auch die nötige Netzwerkleistung bereitzustellen.

Bei der Auswahl eines Netzanbieters ist Sorgfalt geboten. Stellen Sie jedem potenziellen Anbieter die folgenden Fragen:

- Haben Sie direkten Zugang zu einem Tier-1-Internetbackbone? Oder stellen Sie lediglich den Internetzugang bereit und überlassen Routing für den Rest der Datenübertragung dem Zufall?
- Haben Sie eine direkte Verbindung zu führenden Hyperscalern wie AWS, Microsoft Azure und Google Cloud Platform?
- Bieten Sie garantierte Servicelevel für Latenz und Leistung an?
- Können Sie spezifische, latenzsensible Anwendungen – zum Beispiel KI-Apps – priorisieren?
- Können Sie einen umfassenden Überblick über die Netzwerkleistung bieten, vom Kundenunternehmen bis hin zum Hyperscaler?

Maximale Leistung erzielen Sie erst dann, wenn Ihr Netzwerk ebenso leistungsstark ist wie die KI-Chips in Ihrem Rechenzentrum oder Ihrer Cloud-Umgebung.

Das globale Netzwerk von Verizon ist für die Unterstützung von KI-Anwendungen ausgelegt. Unsere Experten helfen Ihnen beim Aufbau einer Netzwerkinfrastruktur, die Ihre KI-Apps zu Höchstleistungen befähigt und Ihren KI-Initiativen zum Erfolg verhilft.

46 %

Knapp die Hälfte der KI-Projekte, die bei einer Machbarkeitsstudie erfolgreich waren, gelangen nie in die Produktion.¹



Internetzugang: Leitfaden zur Anbieterwahl

In diesem Whitepaper erfahren Sie, worauf Sie bei der Anbieterwahl achten sollten, um die maximale Leistung und besten Ergebnisse für Ihr Unternehmen zu erzielen, und warum das besonders wichtig ist, wenn Sie latenzsensible Anwendungen wie KI-Apps nutzen.

Lesedauer: 15 Minuten



Das KI-gestützte Multi-Service-Internetbackbone von Verizon

Kontinuierliche Verbesserung und Weiterentwicklung gehören zu unseren Top-Prioritäten. Seit 2000 haben wir 176 Milliarden USD in unsere Netzwerkinfrastruktur investiert und planen, im aktuellen Finanzjahr weitere 17,5–18,5 Milliarden USD zu investieren. Unsere Netzwerke nutzen unser globales Multi-Service-Internetbackbone (MSB) und unterstützen somit eine granulare Priorisierung von Anwendungen – auch über das öffentliche Internet.

Dadurch können wir unseren Kunden garantierte latenz- und leistungsspezifische Servicelevelvereinbarungen (SLAs) und eine auf ihre cloudbasierten KI-Leistungsanforderungen abgestimmte Internetanbindung bieten.

Dank unserer Integration mit den Betriebssystemen von Anbietern wie Cisco, Juniper und Aruba können wir Kunden mit Managed Routing und modernen LAN-Technologien mehr Kontrolle über ihr Netzwerk ermöglichen, vom Core bis zum Netzwerkrand. Unsere Netzwerkinfrastruktur nutzt auch ihre eigene KI-Engine, die überwacht, welche Ihrer Anwendungen in welchen Cloud-Umgebungen mehr Rechenleistung erfordern, und Ressourcen Ihren Anforderungen entsprechend dynamisch und intelligent zuweist.

Mit speziellen, in unser MSB integrierten Funktionen können wir latenzsensible Anwendungen priorisieren und nützliche SLAs bieten. Diese leistungsstarke, latenzarme Konnektivität können wir bereitstellen, weil wir:

- unsere eigene Infrastruktur haben und diese selbst betreiben
- Lasten zwischen verschiedenen Cloud-Betreibern ausgleichen können
- eine Infrastruktur für direktes Network Peering mit Cloud-Betreibern aufgebaut haben

Am Anfang dieses Whitepapers habe ich darauf hingewiesen, dass Meta, Google, Amazon und Microsoft alle ihre eigenen für KI optimierten Chips entwickelt haben. Dies stellt eine erhebliche Investition in die Latenzreduzierung und Leistungssteigerung dar. Diese Unternehmen wissen, dass ihr Netzwerk-Backbone entscheidend zum Erfolg ihrer KI-Angebote beiträgt und haben demzufolge in erstklassige Netzwerkarchitekturen investiert.

Bei Verizon arbeiten viele hocherfahrene Netzwerkspezialisten, darunter einige der weltweit führenden Experten. Sie widmen sich der kontinuierlichen Verbesserung unserer globalen Netzwerkinfrastruktur, damit diese nicht nur mit den wachsenden Anforderungen Schritt halten kann, sondern auch auf zukünftige Anwendungen vorbereitet ist. Wir sorgen dafür, dass Anwendungen die für KI-Plattformen erforderliche Leistung erbringen.

Dank unserer hochleistungsfähigen, latenzarmen globalen Architektur profitieren Unternehmen von dem schnellen, sicheren und robusten Netzwerk, das sie benötigen, um Leistungsprobleme hinter sich zu lassen und ihre KI-Initiativen voranzutreiben.

Gespräch vereinbaren

Unser Expertenteam hilft Ihnen dabei, Ihre KI-Anwendungen auf Hochtouren zu bringen und Ihre KI-Ziele zu erreichen.

Von IoT über Echtzeitanalysen bis hin zu KI – Verizon unterstützt Sie dabei, Technologien optimal einzusetzen, um Ihre Prozesse zu überwachen, zu steuern und kontinuierlich zu verbessern. Wenn Sie nach der Lektüre dieses Whitepapers noch Fragen haben, erreichen Sie uns unter

www.verizon.com/business/de-de/contact-us/

Ressourcen rund um smarte Netzwerkbeschaffung

Dieses Whitepaper ist Teil einer Reihe, die sich mit den wachsenden Anforderungen an Unternehmensnetzwerke beschäftigt – und mit den entscheidenden Fragen, die Unternehmen bei der Beschaffung stellen sollten, um sicherzustellen, dass die gewählten Lösungen wirklich unternehmenstauglich sind und sowohl aktuelle als auch zukünftige Anforderungen erfüllen.

Eine neue Daten-Ära

Daten

IoT

KI

In diesem Whitepaper beleuchten wir einige der wichtigsten Faktoren, die das explosionsartige Wachstum der in Unternehmen anfallenden Daten vorantreiben – und welche Auswirkungen das auf die Netzwerkplanung hat.

verizon.com/business/resources/articles/iot-genai-data-explosion.pdf

Internetzugang: Leitfaden zur Anbieterwahl

Leistung

Beim Kauf von Netzwerklösungen müssen viele Entscheidungen getroffen werden. Ein Verständnis der drei Internet-Tiers ist entscheidend, um alle Optionen umfassend beurteilen zu können. Dieses Whitepaper zeigt, wie sich die einzelnen Tiers auf Netzwerkleistung und Sicherheit auswirken.

verizon.com/business/resources/de-de/articles/tier-1-isp-enterprise-connectivity.pdf

Network Peering

Cloud

Leistung

Zuverlässigkeit

Peering ist ein wesentlicher Faktor für die Netzwerkleistung und nimmt dadurch direkten Einfluss auf die Leistung von Unternehmensanwendungen, insbesondere Cloud-Diensten. Dennoch wird es im Beschaffungsprozess kaum je thematisiert. Dieses kurze Whitepaper zeigt, warum Peering so wichtig ist.

verizon.com/business/resources/de-de/articles/network-peering.pdf

Tappen Sie im Dunkeln, was die Leistung betrifft?

Daten

Leistung

Verwaltbarkeit

Lesen Sie in diesem Whitepaper, wie sich die Entscheidung, physische (Underlay) und logische (Overlay) Netzwerke getrennt zu beschaffen, auf Netzwerkleistung, Transparenz und Verwaltbarkeit auswirken kann.

verizon.com/business/resources/de-de/articles/overlay-underlay-network-procurement.pdf

Gemeinsam stärker

Sicherheit

Leistung

Verwaltbarkeit

Cyberangriffe werden immer zahlreicher und komplexer. Dieses kurze Whitepaper gibt sechs Argumente dafür, dass eine engere Verzahnung von Cybersicherheit und Netzwerkbetrieb sinnvoll ist, um den Schutz zu stärken und Aufwand sowie Kosten zu senken.

verizon.com/business/resources/de-de/articles/unified-network-security-services.pdf

Turbo für Ihre KI-Anwendungen

KI

Leistung

KI hat das Potenzial, die bedeutendste technologische Umwälzung seit der Einführung des Internets vor etwa drei Jahrzehnten auszulösen. In diesem Whitepaper erfahren Sie, warum die Leistungsfähigkeit des Netzwerks für KI-Anwendungen eine zentrale Rolle spielt und wie sie hilft, die erwarteten Vorteile tatsächlich zu realisieren.

verizon.com/business/resources/de-de/articles/network-infrastructure-ai-platforms.pdf

1. Verizon, [KI-Nutzung im großen Maßstab](#), 2025

verizon
business