

Editorial

Grundlagen-
wissen:
NVMe over
Fabrics

Wie sich
Speicher durch
NVMe und
SCM entwickelt

Was DPUs
für Storage
leisten können

Optimieren
mit Optane
und GPUs

NVMe, SCM und DPU: Technologien für schnelleres Storage

Neue Technologien sollen schnelleres Storage gewährleisten, allerdings sollten Anwender wissen, welche Lösungen welche Verbesserungen bringen.

[Home](#)

[Editorial](#)

[Grundlagenwissen:
NVMe over
Fabrics](#)

[Wie sich
Speicher durch
NVMe und
SCM entwickelt](#)

[Was DPUs
für Storage
leisten können](#)

[Optimieren
mit Optane
und GPUs](#)

[Impressum](#)

Technologien für schnelles Storage: Was ist Hype und was bleibt?

von [Ulrike Rieß](#), Senior Online Editor bei TechTarget Deutschland

Storage hatte lange Zeit den Makel, der Flaschenhals in puncto Leistung beziehungsweise Datenübertragungszeiten zu sein: der Hardware-Stack mit zu vielen Komponenten, der Software-Stack zu lang und die Datenplatzierung suboptimal. Wurde an einem Ende eine Komponente verbessert – etwa kapazitätsstarke Festplatten –, dann kam eine andere Komponente nicht hinterher, zum Beispiel die Protokollgeschwindigkeit. Seit Jahren tüfteln die Hersteller daran, hier Abhilfe zu schaffen und Technologien auf den Markt zu bringen, die eine echte technologische Verbesserung von Storage-Systemen gewährleisten. Dabei setzen nicht alle Anbieter auf das gleiche Pferd. So haben sich in den letzten Jahren einige Produkte herauskristallisiert, die mehr als nur ein Marketing-Hype zu sein scheinen. Längst ist klar, kein Unternehmen kann es sich leisten, alle Transfer- oder Latenzprobleme einfach mit DRAM oder anderen Memory-Typen zu „erschlagen“.

So zeigten sich neben neuen Arten an Memory (ReRAM, SRAM oder FeRAM) weitere Lösungsansätze für das althergebrachte Problem. Data Processing Units gesellt sich zu CPU und GPU und kann Storage-Anfragen verarbeiten, was wiederum die Hardware

beschleunigen soll. Storage Class Memory fungiert als Lückenfüller zwischen DRAM und NAND-Flash-Speicher und soll persistentes Storage offerieren, das günstiger als DRAM und leistungsfähiger als SSDs ist. NVMe und NVMe-oF sind Datenübertragungsprotokolle, die weitaus schneller als SAS oder SATA operieren. Die Intel Optane-Technologie zählt zu SCM, basiert auf der 3D-XPoint-Architektur und soll Speichersysteme beschleunigen, entweder als Optane-Festplatte oder als Optane-DIMM. Darüber hinaus bietet der Hersteller seine DAOS-Architektur an, die in Verbindung mit Optane-DIMMs I/Os besser und schneller verarbeiten soll.

All diese Technologien haben es über den reinen Markt-Hype hinaus gebracht, und es scheint, als ob sie nicht nur sinnvolle Optimierungen garantieren, sondern auf lange Sicht auch am Markt koexistieren werden. Da sie aber verhältnismäßig jung sind, lohnt es sich, diese Technologien genauer zu betrachten, ihr Potenzial zu erkennen, aber auch mögliche Probleme zu identifizieren. In diesem E-Handbook beschäftigen wir uns damit, was SCM, DPUs, NVMe und Co leisten können und wie sie sich am besten einsetzen lassen. ■

[Home](#)

[Editorial](#)

[Grundlagen-
wissen:
NVMe over
Fabrics](#)

[Wie sich
Speicher durch
NVMe und
SCM entwickelt](#)

[Was DPUs
für Storage
leisten können](#)

[Optimieren
mit Optane
und GPUs](#)

[Impressum](#)

Die NVMe-oF-Speichertechnologie und ihre Begriffe erklärt

Von *Stacey Peterson*

Erfahren Sie, wie NVMe-oF die Leistung von Speichernetzwerken optimiert und die Latenzzeit senkt. Darüber hinaus erklären wir die wichtigsten Begriffe, die sie kennen müssen.

In den letzten fünf Jahren hat sich die Speichertechnologie mit der Verbreitung von [NVMe](#) und [Persistent Memory](#) rasch weiterentwickelt. Diese Veränderungen haben die Netzwerktechnologie unter Druck gesetzt, nicht zu einem Engpass im Speichersystem zu werden.

[NVMe over Fabrics](#) (NVMe-oF) kann entsprechende Verlangsamungen verhindern, indem die Performance und Low-Latency-Vorteile der NVMe-Technologie im Netzwerk eingesetzt werden, um mehreren [Servern](#) und [CPUs](#) Zugang zu NVMe-Storage zu gewähren.

Das ursprüngliche Ziel von NVMe-oF bestand bei [Storage](#) darin, innerhalb von zehn Mikrosekunden (μ s) der Performance eines lokalen NVMe-Laufwerks zu bleiben. Das erreicht diese Technologie, indem die NVMe-Befehle eingekapselt und über das Netzwerk zu entfernten Speicherlaufwerken gesendet werden.

Branchenexperten gehen davon aus, dass NVMe-oF in den nächsten Jahren richtig durchstarten wird. Allerdings kann die Umsetzung der Technologie kompliziert sein. Um die versprochenen Performance-Verbesserungen zu erreichen, müssen eine Reihe von Details berücksichtigt werden, einschließlich der Netzwerktechnologie und -struktur, den verwendeten [SSDs](#) und schließlich der Konfiguration all dieser Elemente. Im Folgenden finden sich einige der wesentlichen Begriffe und Technologien, die man für einen erfolgreichen Übergang zu NVMe-oF kennen sollte.

DSSD

[DSSD ist eigentlich eine EMC-Technologie](#), die NVMe-oF vorausgeht. Sie teilt einen Flash-Block über verschiedene Hosts in einem Rack und war ein nicht so teurer Ansatz für gemeinsam genutzten (Shared) Low-Latency-Storage im Vergleich zu damaligen konkurrierenden Technologien. Im Jahr 2017, als NVMe-oF auftauchte und mit der 100- μ s-[Latenz](#) von DSSD gleichzog oder sie sogar zu einem Fünftel der Kosten übertraf, stoppte Dell EMC die weitere Entwicklung an DSSD.

[Home](#)[Editorial](#)[Grundlagen-
wissen:
NVMe over
Fabrics](#)[Wie sich
Speicher durch
NVMe und
SCM entwickelt](#)[Was DPUs
für Storage
leisten können](#)[Optimieren
mit Optane
und GPUs](#)[Impressum](#)

HOCH PARALLELE MANY-TO-MANY-ARCHITEKTUR

Mit der geringen Warteschlangen-Tiefe ([Queue Depth](#)) von [SAS](#)- und [SATA](#)-Laufwerken wurden die I/O-Prozesse in eine einzige [Warteschlange](#) gepackt, die schnell zu einem Engpass werden konnte und nicht in der Lage war, Vorteile aus den parallelen Fähigkeiten der [NAND-Flash-Medien](#) zu ziehen, die in SSDs benutzt wurden.

NVME IST EINE LOGISCHE SPEZIFIKATION FÜR EINE GERÄTESCHNITTSTELLE, DIE CONTROLLER UND TREIBER VERWENDET, DIE SPEZIELL FÜR SSDs ENTWICKELT WURDEN.

NVMe-oF-Storage hat das geändert, indem 65.535 Verbindungen zur Verfügung gestellt werden, von denen jede bis zu 65.535 Befehle abarbeiten kann. Dies muss man im Vergleich mit der einzigen Warteschlange von SATA und SAS betrachten, die jeweils 32 und 256 Befehle pro Warteschlange aufweisen.

Die Tiefe der Warteschlange erlaubt es, eine hoch parallele Many-to-Many-Architektur zwischen [Host](#) und Laufwerk zu entwickeln – mit einer eigenen Warteschlange für jeden Host und jedes Laufwerk.

NVME-STORAGE

Non-Volatile Memory Express (NVMe) ist ein offener Standard, der als Alternative zum [SCSI](#)-Protokoll und den Bottlenecks entwickelt wurde, die sich aus der Nutzung von SCSI-basierter Technologie zusammen mit [Flash-Speicher](#) ergeben.

NVMe ist eine logische Spezifikation für eine Geräteschnittstelle, die Controller und Treiber verwendet, die speziell für SSDs entwickelt wurden. Sie verbindet sich über den PCI Express [Bus](#) ([PCIe](#)) mit den Hosts. Weil sich PCIe direkt mit der CPU verbindet, erlaubt diese Technologie einen Memory-ähnlichen Prozessorzugriff, reduziert den I/O-Overhead und ermöglicht es den Geräten, Vorteile aus der Geschwindigkeit, Performance und dem parallelen Zugang von SSDs zu ziehen.

Ein Treiber, der besonders für NVMe-Geräte entwickelt wurde, stellt ihnen auch die schnellere Zugangsgeschwindigkeit der PCIe-Schnittstelle zur Verfügung.

NVM EXPRESS INC.

NVM Express wurde 2014 als ein gemeinnütziges Konsortium mit dem Auftrag für die Entwicklung der NVMe-Spezifikation gebildet. Die Gruppe hat mehr als 100 Mitglieder und konzentriert sich auf die Definition, das Management und die Verbreitung von NVMe als einem industrieweiten Standard, um Speicherzugriff für SSDs mit großer Bandbreite und niedriger Latenz zur Verfügung zu stellen.

Die jüngste Version 1.4 der NVMe-Spezifikation wurde am 23. Juli 2019 veröffentlicht. Die Gruppe ist auch verantwortlich für die Einzelheiten von NVMe-oF, um NVMe-Befehle über eine

Home

Editorial

Grundlagen-
wissen:
NVMe over
Fabrics

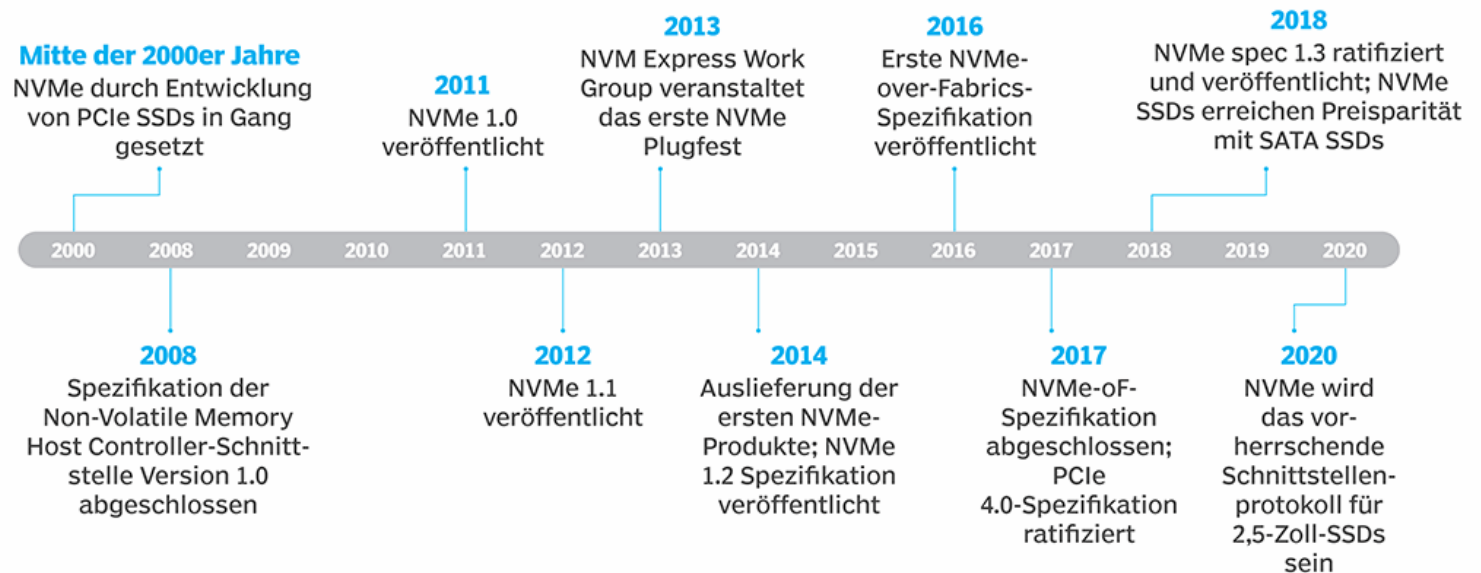
Wie sich
Speicher durch
NVMe und
SCM entwickelt

Was DPUs
für Storage
leisten können

Optimieren
mit Optane
und GPUs

Impressum

NVMe-Meilensteine



SOURCE: "NVM EXPRESS ECOSYSTEM SNAPSHOT" G2M RESEARCH, MAY 2018

©2018 TECHTARGET. ALL RIGHTS RESERVED



Home

Editorial

Grundlagen-
wissen:
NVMe over
Fabrics

Wie sich
Speicher durch
NVMe und
SCM entwickelt

Was DPUs
für Storage
leisten können

Optimieren
mit Optane
und GPUs

Impressum

Netzwerkinfrastruktur und das NVMe Management Interface zu verwenden, das die Verwaltung von NVMe-SSDs in Servern und Speichersystemen erleichtert. Die jüngste Version 1.1 dieser NVMe-oF-Spezifikation wurde am 22. Oktober 2019 veröffentlicht.

DIE NVME-INTEGRATION TESTEN

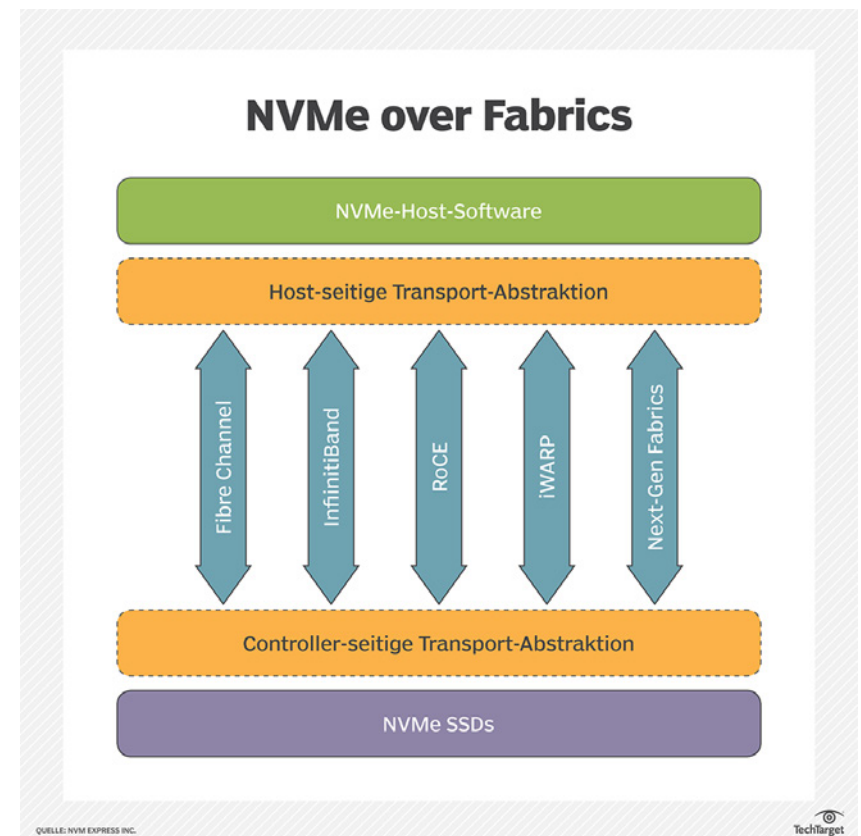
Herauszufinden, ob NVMe-oF-Produkte miteinander kompatibel sind, kann eine Herausforderung darstellen. Aber zum Glück dient das Interoperability Laboratory der University of New Hampshire (UNH-IOL) als eine Clearing-Stelle (sinngemäß: Freigabestelle) für Netzwerkgeräte und die von ihnen unterstützten Netzwerkstandards.

Nach Tests von mehr als zwei Jahren hat die UNH-IOL eine Liste von Produkten, die den Standards entsprechen, zusammengestellt: Sie decken NVMe-Host-Plattformen ab sowie Laufwerke, Switches, Managementschnittstellen und Storage-Hardware – ferner Zielpunkte, Initiatoren und [Switches](#) für [Software](#).

NVME OVER FC

Die ersten Ausführungen der NVMe-oF-Spezifikation unterstützen [Fibre Channel](#) (FC) und [Remote Direct Memory Access](#) (RDMA). Mit diesen beiden Transportmöglichkeiten werden Daten vom Host zum [Array](#) und zurück mittels eines Zero-Copy-Prozesses oder im Fall von NVMe over FC eines annähernden Zero-Copy-Prozesses übertragen, bei dem die Daten nicht mehrere Male zwischen den [Memory](#) Buffers und den I/O-Karten kopiert werden.

Dieser Ansatz reduziert die CPU-Last und die Latenz beim Daten-transport, was die Performance verbessert. Mit dem Credit Buffer Congestion Control System von FC werden die Daten nicht eher transportiert, bis eine Pufferzuweisung (Buffer Allocation) angekommen ist, womit ein sicherer verlustfreier Transport garantiert ist.



[Home](#)[Editorial](#)[Grundlagen-
wissen:
NVMe over
Fabrics](#)[Wie sich
Speicher durch
NVMe und
SCM entwickelt](#)[Was DPUs
für Storage
leisten können](#)[Optimieren
mit Optane
und GPUs](#)[Impressum](#)

FC verringert die Latenz und die Verwendung der Host-CPU, indem [Host Bus Adapter](#) (HBAs) das FC Protocol (FCP) herunterladen. Jeder FC-HBA oder Switch mit 16 Gbps oder 32 Gbps wird NVMe over FC unterstützen, während weitere Services für SCSI over FCP zur Verfügung gestellt werden.

ES ERFORDERT BESONDERE KONFIGURATIONEN WIE ZUM BEISPIEL DIE EINRICHTUNG VON PRIORITY FLOW CONTROL UND EXPLIZITE STAUNACHRICHTEN, UM ABGELEHNE PAKETE ZU ELIMINIEREN.

NVME OVER RDMA

Die Spezifikation von NVMe-oF unterstützt RDMA-Netze für InfiniBand, RDMA over Converged Ethernet (RoCE) und das Internet Wide Area RDMA Protocol (iWARP). RDMA ermöglicht es Computern im Netzwerk, Daten im Hauptspeicher auszutauschen, ohne dass Prozessor, [Cache](#) oder [Betriebssystem](#) jedes Computers involviert sind.

RDMA setzt Ressourcen frei und verbessert Durchsatz und Performance. NVMe over RDMA zu installieren kann komplex sein. Es erfordert besondere Konfigurationen wie zum Beispiel die Einrichtung von Priority Flow Control und explizite Staunachrichten, um abgelehnte Pakete zu eliminieren. Diese Technik kann auch besondere Ausrüstung erfordern wie zum Beispiel RDMA-fähige Netzwerkschnittstellen.

PERFORMANCE-MONITORING VON NVME-OF

In der Vergangenheit konnten ein schlecht konfigurierter Netzwerk-Switch, -Port oder -Adapter in einem Speichersystem leicht übersehen oder unentdeckt bleiben, weil ihre Latenzzeiten im Vergleich zu denen von Speicheranwendungen oder -geräten eher unerheblich waren.

Dies änderte sich mit dem Aufkommen von NVMe-oF. Nun waren es die Netzwerke, die sich schwertaten, mit der Speicherhardware und -software mitzuhalten. Diese Umdrehung des Verhältnisses führte dazu, dass das Monitoring der NVMe-oF-Performance zu einem wesentlichen Bestandteil des Storage-Managements wurde, um so die maximale Performance der Netzwerkkonfiguration abzusichern.

Speichernetzwerke mussten proaktiv überprüft werden, um sicherzustellen, dass das Netzwerkdesign und die eingesetzten Ressourcen mit den I/O-Anforderungen Schritt halten konnten. Das Monitoring der NVMe-oF-Performance hängt von den in Echtzeit aufgezeichneten Telemetriedaten des Netzwerks ab – zusammen mit Tools, die eventuelle Probleme erkennen und sogar lösen, bevor sie tatsächlich eintreten.

Home

Editorial

Grundlagen-
wissen:
NVMe over
Fabrics

Wie sich
Speicher durch
NVMe und
SCM entwickelt

Was DPUs
für Storage
leisten können

Optimieren
mit Optane
und GPUs

Impressum



HURCAI/ADOBE

NVMe/TCP

Die Revision 1.1 von NVMe-oF fügte Unterstützung für die Transportprozesse des [Transmission Control Protocol](#) (TCP) hinzu, das das Internet antreibt und eines der am weitesten verbreiteten Datentransportmittel weltweit ist. NVMe/TCP ermöglicht den Einsatz von NVMe-oF auf einem standardmäßigen [Ethernet](#)-Netzwerk, ohne Änderungen an der Konfiguration oder besondere Ausrüstung zu erfordern.

Es kann auch über das Internet hinweg genutzt werden. NVMe/TCP legt fest, wie TCP-basierte Kommunikation zwischen NVMe-oF-Hosts und Controllern in einem IP-Netzwerk umgesetzt wird. Ein Modell für Messaging und Queuing beschreibt die Kommunikationssequenz.

SCALE-OUT-SPEICHERARCHITEKTUR

Beim [Scaling-Out](#) wird der Speicherplatz durch das Hinzufügen weiterer angeschlossener Arrays, zusätzlicher Server oder [Knoten](#) erweitert, um die Storage-Fähigkeiten zu erhöhen. Bis NVMe-

oF herauskam, war [Scale-Out-Storage](#) komplex zu installieren. NVMe-oF stellt einen direkten Pfad zwischen Host und Speichermedien zur Verfügung, womit es vermieden wird, die Daten durch einen zentralen [Controller](#) zu zwingen – ein Prozess, der Latenz erzeugt und den vollen Nutzen der SSD-Möglichkeiten begrenzt. Der direkte Pfad ermöglicht es den Hosts, mit vielen Laufwerken zu interagieren, was wiederum die Latenz reduziert und die Scale-Out-Möglichkeiten erhöht.

SPEZIFIKATION DER TRANSPORTANBINDUNG

NVMe-oF ist nicht mit irgendeinem Transportprotokoll für Storage verknüpft, weil das Protokoll von dem Transport-Layer, auf dem es läuft, abgetrennt ist. Dies wird als sogenannte Transport Binding Specification bezeichnet. Solche Binding Specs sind für RDMA, einschließlich InfiniBand, RoCE und iWARP sowie für Fibre Channel und seit Neuestem für TCP entwickelt worden. NVMe-oF kann auf jeder Netzwerkverbindung laufen, die über eine Binding Spec verfügt. ■

NVME UND SCM: IHRE AUSWIRKUNG AUF DIE ZUKUNFT DES STORAGE

Von *Marc Staimer*



PANUWAT/ADOBE

HOME

NVMe hat sich in Unternehmen durchgesetzt, angetrieben durch Anwendungen mit hohen Leistungsanforderungen wie [künstliche Intelligenz](#) (KI), [maschinelles Lernen](#) und Analysen und deren Bedarf an Hochgeschwindigkeitszugriff mit niedriger [Latenz](#) auf [Speichermedien](#). Jetzt setzt sich Storage Class Memory (SCM) durch, das Speichermedien mit niedriger Latenz bereitstellt.

Dieser Artikel befasst sich damit, wie NVMe Unternehmen den Weg für den Übergang zu Storage Class Memory (SCM) ebnet. Wenn sich [NVMe](#)- und SCM-Laufwerke durchsetzen, verlagert sich der Fokus von den Speichermedien als Engpass auf Bereiche wie [CPU](#), Interconnect, Speichersoftware-[Stack](#), [Treiber](#) und Protokolle. Lassen Sie uns erkunden, wie NVMe und SCM neue Wege zu schnellerem [Speicher](#) bei geringeren Kosten eröffnen.

EIN KURZER RÜCKBLICK

Bei Speichersystemen ging es schon immer um Kompromisse zwischen Leistung, Kapazität und Kosten. Die höchste Leistung, wie sie zum Beispiel von dynamischem RAM ([DRAM](#)) bereitgestellt wird, hat relativ geringe Kapazitäten und sehr hohe Kosten pro [Terabyte](#).

Das andere Extrem ist die Bandtechnologie. [Band](#) liefert eine sehr geringe Zufallsleistung (random performance), eine gute sequenzielle Leistung und hohe Kapazitäten zu sehr geringen Kosten. Zwischen diesen Leistungs-, Kapazitäts- und Kostenextremen liegen Flash-[SSDs](#) und [Festplatten](#) (HDDs).

SSDs haben eine exzellente Lese- und Schreibleistung, die wesentlich niedriger ist als DRAM, und gute bis hohe Kapazitätä-

Home

Editorial

Grundlagen-
wissen:
NVMe over
Fabrics

Wie sich
Speicher durch
NVMe und
SCM entwickelt

Was DPUs
für Storage
leisten können

Optimieren
mit Optane
und GPUs

Impressum

Vergleich der verschiedenen Speichermedien

MEDIUM	PERFORMANCE	KAPAZITÄT/LAUFWERK	HALTBARKEIT	KOSTEN/TBYTE
Band	sehr gering bei random, gut bei sequentiell	LTO-8: 12 TB nativ, 30 TB komprimiert. LTO-9 soll in Kürze diese Angaben um 50 Prozent erhöhen.	Daten können im ausgeschalteten Zustand bis zu 30 Jahre unverändert bleiben.	sehr niedrig für Medien, abhängig von der Anzahl der Laufwerke
Festplatte	schlecht bei random, gut bei sequentiell	Derzeit 18TB, in Kürze 20 TB	mittelmäßig	gering
Flash-SSD	hoch	bis zu 100 TB	Reicht von schlechten Penta-Level-Cell-SSDs bis zu exzellenten Single-Level-Cell; beeinflusst durch Bits pro Zelle, Error Correction Code und Overprovisioning.	mittleres Niveau
NVMe-Flash-SSD	hoch	bis zu 30,72 TB	wie bei Flash-SSDs	hoch
SCM	sehr hoch	bis zu 1,5 TB	sehr hoch	sehr hoch
DRAM	am höchsten	bis zu 64 GB per DIMM.	unbegrenzt	am höchsten

SOURCE: MARC STAIMER

©2020 TECHTARGET. ALL RIGHTS RESERVED. 

[Home](#)

[Editorial](#)

[Grundlagenwissen:
NVMe over
Fabrics](#)

[Wie sich
Speicher durch
NVMe und
SCM entwickelt](#)

[Was DPUs
für Storage
leisten können](#)

[Optimieren
mit Optane
und GPUs](#)

[Impressum](#)

ten – bis zu 100 TB in einem 3,5-Zoll-Laufwerk –, aber zu einem relativ hohen Preis im Vergleich zu HDDs oder Bändern. Preis, Leistung, Kapazität und Haltbarkeit für Flash variieren in Abhängigkeit von den [NAND-Flash](#)-Bits pro Zelle.

Festplatten liefern relativ schlechte zufällige Lese- und Schreibvorgänge im Vergleich zu DRAM oder SSDs, mit hohen Kapazitäten, guter sequenzieller Leistung, mittelmäßiger Ausdauer und vergleichsweise niedrigen Kosten pro Terabyte.

Aber was ist, wenn es einen Bedarf an höherer Leistung als bei gewöhnlichen Flash-SSDs zu niedrigeren Kosten als bei DRAM gibt? Es sind immer [Arbeitslasten](#) vorhanden – E-Commerce, Online-Transaktionsverarbeitung, [Data Warehouses](#), KI-Analytik, maschinelles Lernen und [Deep Learning](#) –, die eine höhere Leistung von niedrigeren Latenzzeiten und mehr [IOPS](#) oder Durchsatz verlangen. Es ist jedoch schwierig, die Kosten und die begrenzte Kapazität von DRAM zu rechtfertigen. Um diese Produktlücke zu schließen, haben die Anbieter NVMe-Flash-SSDs und NVMe-SCM entwickelt, die von Intel als Optane-SSDs vermarktet werden.

NVME-FLASH-SSDs

NVMe-Flash-SSDs wurden entwickelt, um die Latenzzeiten von Flash-SSDs zu reduzieren. Die Latenz ist der wichtigste Faktor bei Transaktionen und Nachrichtenübermittlung, da sie die Zeit ist, die für den Zugriff oder das Schreiben des ersten [Bytes](#) benötigt wird. Es ist auch die Zeit, die benötigt wird, um auf das letzte Byte zuzugreifen oder für die Tail-Latenz zu synchronisieren. Die Tail-Latenz ist extrem wichtig für Anwendungen, die das im High Per-

formance Computing ([HPC](#)) vorherrschende Message-Passing-Interface-Protokoll verwenden.

NVMe ist die [Open-Source](#)-Standardspezifikation, die nicht flüchtige Medien wie NAND-Flash und 3D XPoint – die Technologie hinter Optane – über den PCIe-Bus anspricht. NVMe beseitigt einen bedeutenden Latenzengpass im [SAS](#)- oder [SATA-Controller](#). Alle [Controller](#) erhöhen die Latenz, da sie zwischen dem [PCIe](#)-Controller und den Laufwerken sitzen. Die Eliminierung dieses Controllers reduziert die End-to-End-Latenz erheblich. NVMe-Flash-Laufwerke reduzieren die SSD-Latenz typischerweise um zwei Drittel oder mehr auf etwa 30 Mikrosekunden oder 0,03 Millisekunden. Das ist eine signifikante Reduzierung der Latenz.

NVMe verbessert auch den Durchsatz und die IOPS. PCIe Gen 4 – mit der doppelten Bandbreite von PCIe Gen 3 – ist erforderlich, um den vollen Leistungsvorteil zu erzielen. Die meisten aktuellen Speichersysteme verfügen über PCIe Gen 3.

NVMe-Laufwerke laufen direkt auf dem [PCIe-Bus](#). Allerdings steht pro Sockel nur eine begrenzte Anzahl von PCIe-Steckplätzen zur Verfügung, abhängig von der CPU. Dies führt zu weniger NVMe-Laufwerken pro Speicher-Controller. Das ist der Hauptgrund, warum so viele NVMe-Flash-Speichersysteme 24 oder weniger NVMe-Laufwerke haben.

Der Start-up-Anbieter Pavilion Data Systems umgeht diese Einschränkung, indem er Architekturen mit mehr Controllern implementiert – bis zu 20 Controllern und 72 NVMe-Flash-Laufwerken. Fungible, ein weiteres Start-up, überwindet die PCIe-Steckplatzbeschränkung mit internen Infrastruktur-Switches über seinen

[Home](#)

[Editorial](#)

[Grundlagenwissen: NVMe over Fabrics](#)

[Wie sich Speicher durch NVMe und SCM entwickelt](#)

[Was DPUs für Storage leisten können](#)

[Optimieren mit Optane und GPUs](#)

[Impressum](#)

anwendungsspezifischen integrierten Schaltkreis der Data Processing Unit (DPU). Die meisten Mainstream-Storage-Anbieter wie Dell EMC, Hewlett Packard Enterprise, Hitachi Vantara und NetApp verwenden [Scale-Out](#)-Clustering-Technologie, um mehr Controller und Laufwerke in einem Speichersystem zu erhalten. Dieser Ansatz verbraucht auch mehr Platz im Rack.

Solange der Kompromiss zwischen den Controllern und dem Platzbedarf im Rack akzeptabel ist und die Verbindung zwischen den Speicher-Controller-[Knoten](#) nicht viel Latenz hinzufügt, reduziert NVMe die Latenz spürbar und erhöht die Leistung. SCM geht noch weiter.

STORAGE CLASS MEMORY

Der Optane SCM von Intel basiert auf 3D XPoint. Er ist schneller als Flash und hat geringere Latenzzeiten, einen höheren Durchsatz, mehr IOPS und eine größere Haltbarkeit. SCM-Laufwerke sind NVMe-basiert. SCM sollte jedoch nicht mit Intels Optane Persistent Memory (PMEM) DIMMs verwechselt werden, die in den DDR4-Speichersteckplätzen eines Servers oder Controllers stecken.

Auf PMEM kann auf zwei Arten zugegriffen werden. Die erste ist die Verwendung des nicht persistenten Speichermodus, der PMEM wie langsames, weniger kostspieliges DRAM aussehen lässt. Die zweite Möglichkeit ist die Verwendung des persistenten Speichermodus App Direct Mode, der eine Modifikation der Anwendung und des Dateisystems erfordert, damit PMEM wie ein Speicher aussieht. Das ist nichts, was die meisten IT-Organisationen nutzen können, aber es gibt Ausnahmen. Anbieter wie

Formulus Black und MemVerge können diese Anforderung vor den Anwendungen und Dateisystemen verbergen. Oracle nutzt den App Direct Mode in seinem Exadata-System für [Oracle Database](#). Auch [SAP S/4HANA](#) nutzt den App Direct Mode für seine Datenbank. Es gibt vier Varianten des App Direct Mode: Raw Device Access, Memory Access, File System und NVM-aware File System.

Für die meisten Speicherimplementierungen und Speichersysteme ist SCM die Optane-Implementierung der Wahl. Sie ist einfach und erfordert keine Änderungen am Speicher-Stack. Es handelt sich einfach um ein schnelleres Laufwerk mit geringerer Latenz. SCM ist in der Regel drei- bis fünfmal schneller als NVMe-Flash-SSDs; allerdings ist es auch drei- bis fünfmal teurer. PMEM-[DIMMs](#) sind sogar noch schneller und teurer pro Terabyte.

AUSWIRKUNGEN VON NVME UND SCM AUF DIE SPEICHERENTWICKLUNG

Die Anbieter von Speicherlösungen stellen sich bereits auf NVMe und SCM ein. Es gibt nur wenige Anbieter von Primärspeichersystemen, die derzeit keine NVMe-Flash-SSD-Speicheroptionen anbieten.

Was passiert, ist die Rückkehr von [Tiered](#) oder [Cached Storage](#). Dies sind nicht die Speichersysteme von vor ein paar Jahren, die Flash-SSDs mit HDDs kombinierten. Diese Systeme verwendeten Flash-SSDs als primären Lese-/Schreib-Tier oder als Lese-/Schreib-Cache und verlagerten die Daten dann mit der Zeit auf kostengünstigere, weniger leistungsfähige Festplatten.

Home

Editorial

Grundlagen-
wissen:
NVMe over
Fabrics

Wie sich
Speicher durch
NVMe und
SCM entwickelt

Was DPUs
für Storage
leisten können

Optimieren
mit Optane
und GPUs

Impressum

DIESE KOMBINIERTEN NVMe- ODER SCM- UND QLC-HYBRIDSYSTEME SIND ZU 100 PROZENT SOLID-STATE MIT EINER VIEL BESSEREN ALLROUND-LEISTUNG IM VERGLEICH ZU FRÜHEREN HYBRID- ODER SOGAR VIELEN AKTUELLEN ALL-FLASH-SYSTEMEN.

Vast Data und StorONE haben Pionierarbeit bei einer neuen Variante von hybriden Solid-State-Systemen geleistet. Sie verwenden NVMe-Flash-SSDs oder neuerdings SCM als primäre Lese-/Schreibschicht mit kostengünstigeren SAS/[SATA-Quad-Level-Cell](#)- (QLC) oder 4-Bit-pro-Zelle-Flash-SSDs mit hoher Kapazität als sekundäre Schicht für kühle und kalte Daten. QLC-Flash-Laufwerke bieten viel größere Kapazitäten zu deutlich niedrigeren Kosten als andere Flash-Laufwerke. Die Technologie bietet eine geringere Leistung und Ausdauer, um mehr Kapazität zu geringeren Kosten zu erhalten. Nimbus Data hat ein 64-TB-QLC-Flash-Laufwerk im 3,5-Zoll-Formfaktor angekündigt, das sich ideal für diese Hybridsysteme eignet.

Während SCM unabhängig von den [Bits](#) pro Zelle eine viel schnellere Leistung und höhere Ausdauer als Flash-SSDs hat, hat QLC eine geringere Leistung und Ausdauer als [Single-Level-Cell](#)-, [Multi-Level-Cell](#)- und [Triple-Level-Cell](#)-Flash-SSDs. Die Ver-

wendung von NVMe-Flash- oder SCM-Laufwerken als primäre Schreibschicht oder [Cache](#) ermöglicht eine viel bessere Systemleistung und erhöht gleichzeitig die Ausdauer der viel preiswerteren QLC-Flash-SSDs.

Diese kombinierten NVMe- oder SCM- und QLC-Hybridsysteme sind zu 100 Prozent Solid-State mit einer viel besseren Allround-Leistung im Vergleich zu früheren Hybrid- oder sogar vielen aktuellen [All-Flash](#)-Systemen. Noch wichtiger ist, dass sie weniger kosten als aktuelle All-Flash-Systeme. Die Wirtschaftlichkeit und die Preis-Leistungs-Metriken zeichnen den Weg dieser Systeme als die nächsten Mainstream-Allzwecksysteme vor.

Aber auch die All-NVMe-Flash- oder SCM-Systeme, die auf extreme Leistung ausgelegt sind, haben ihren Platz auf dem Markt. Sie verwenden nicht nur NVMe-Flash- und SCM-Laufwerke, sondern nutzen auch das [NVMe-oF](#)-Verbindungsprotokoll, das die Latenzzeiten zum und vom Speichersystem minimiert. ■

[Home](#)

[Editorial](#)

[Grundlagenwissen:
NVMe over
Fabrics](#)

[Wie sich
Speicher durch
NVMe und
SCM entwickelt](#)

[Was DPUs
für Storage
leisten können](#)

[Optimieren
mit Optane
und GPUs](#)

[Impressum](#)

Wie Unternehmensspeicher von Data Processing Units profitiert

Von *Brien Posey*

U Obwohl nur eine kleine Anzahl von Hardwareherstellern DPUs anbietet, hat die Technologie erhebliche Auswirkungen auf IT-Speichersysteme und die Administratoren, die sie verwalten.

Während Data Processing Units (DPU, deutsch: Datenverarbeitungseinheiten) ein heißer Trend im IT-Speicherbereich sind, ist das Grundkonzept dahinter nicht völlig neu.

Moderne Computerbetriebssysteme wie [Windows](#) und [Linux](#) führen rechenintensive Aufgaben aus, die nicht direkt mit der primären [Arbeitslast](#) des Systems zusammenhängen. In den letzten Jahrzehnten haben IT-Anbieter kreative Wege gefunden, einige dieser Aufgaben auf spezielle Hardware auszulagern. Diese Hardware setzt [CPU](#)-Ressourcen frei, um die primäre Arbeitslast des Systems auszuführen. DPUs sind eines der jüngsten Beispiele.

In den 1990er-Jahren verwendeten Systeme, die für computer-gestützte Konstruktionsanwendungen entwickelt wurden, häufig einen mathematischen Coprozessor, um mathematisch intensive Operationen auszuführen. Infolgedessen musste die primäre CPU diese Aufgaben nicht bewältigen. In den letzten Jahren haben



VIDEOFLOW/ADOBE

[Home](#)[Editorial](#)[Grundlagen-
wissen:
NVMe over
Fabrics](#)[Wie sich
Speicher durch
NVMe und
SCM entwickelt](#)[Was DPUs
für Storage
leisten können](#)[Optimieren
mit Optane
und GPUs](#)[Impressum](#)

Hardwarehersteller Netzwerkkarten eingeführt, die [TCP/IP](#)-bezogene Aufgaben von der CPU entlasten, sowie dedizierte GPUs, die verhindern, dass die CPU eines Systems Ressourcen für das Rendern von Grafiken verschwendet.

IM GEGENSATZ ZU EINER GPU, DIE EINEM BESTIMMTEN ZWECK DIENT, SIND DATENVERARBEITUNGSEINHEITEN EHER UNIVERSELL EINSETZBAR.

Data Processing Units folgen demselben Grundkonzept, jedoch mit einem entscheidenden Unterschied: Die oben besprochenen Offloading-Methoden konzentrieren sich auf Berechnungen, während DPUs darauf abzielen, die Datenbewegung zu optimieren.

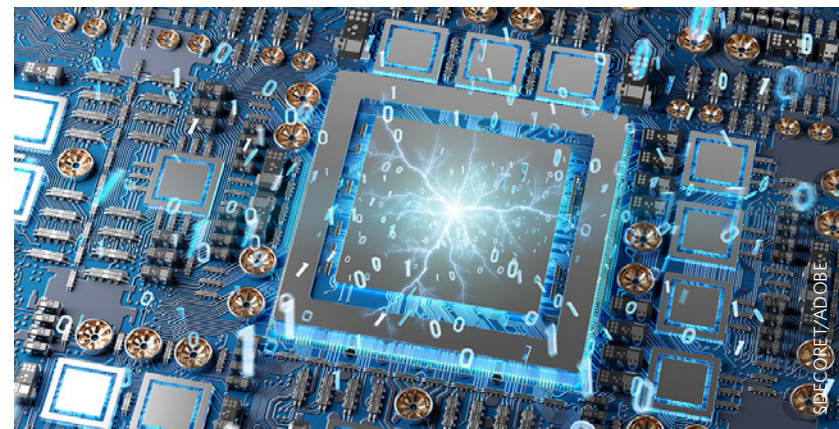
Im Gegensatz zu einer [GPU](#), die einem bestimmten Zweck dient, sind Datenverarbeitungseinheiten eher universell einsetzbar. Eine DPU ist typischerweise mit einer Multi-Core-CPU, Hochgeschwindigkeits-Netzwerkschnittstellen, Speicher-[Controlern](#) und [PCIe](#)-Sockeln ausgestattet. Die DPU verfügt über ein eigenes, vom primären Betriebssystem des Systems getrenntes [Betriebssystem](#). Wenn Administratoren diese Ressourcen mit den nativen Beschleunigern der DPU für Komponenten

wie Kryptowährung und [Storage](#) kombinieren, wird die Vielseitigkeit einer DPU deutlich.

Datenverarbeitungseinheiten können Enterprise-Storage unterstützen. Schließen Sie zum Beispiel ein oder mehrere [NVMe](#)-Speichergeräte direkt an den PCIe-Bus der DPU an, um Hochgeschwindigkeitszugriff auf diese Geräte zu ermöglichen.

Data Processing Units können auch eine Verbindung zu Remote-Speichergeräten herstellen, die [NVMe-oF](#) verwenden. Die DPU kann das entfernte Speichergerät dem System als Standard-NVMe-[Speicher](#) präsentieren. Als solche optimiert die DPU die Konnektivität mit dem Remote-Speicher, aber sie verschleiert auch das Remote-Speichergerät. Die DPU präsentiert es als Standard-NVMe-Speicher, den die meisten modernen Betriebssysteme ohne spezielle [Treiber](#) verwenden können.

Anbieter wie Nvidia und Fungible sind derzeit führend in der DPU-Entwicklung. ■



WIE NVME, OPTANE UND GPUS DEN STORAGE-MARKT VORANTREIBEN

Von *Carol Sliwa*



LASSEDESIGNEN/ADOBE

HOME

Unternehmen, die hochleistungsfähigen Speicher für [künstliche Intelligenz](#) (KI), [Analyse](#) und [Virtualisierung](#) benötigen, können Hilfe in neuen Lösungen wie [NVMe](#) und Optane sowie Prozessortechnologien finden.

Zu den Technologien, die den größten Einfluss auf Unternehmensspeicher haben und in Zukunft noch wichtiger werden könnten, gehören Hochgeschwindigkeits-NVMe, innovative neue Prozessoren sowie Optane und andere neue [Memory](#)-Produkte.

Zu den anspruchsvollen Arbeitslasten, die die Industrie dazu antreiben, schnellere und energieeffizientere Memory- und Speichertechnologien für Unternehmen zu entwickeln, gehören KI, große Datenanalysen, Virtualisierung und [High Performance Computing](#).

Um beispielsweise künstliche Intelligenz zu verwalten, muss die gesamte Rechnerarchitektur überdacht werden. Das gilt sowohl für die Hardware als auch für die Software. Dies lässt sich mit herkömmlichen Speicherlösungen nicht mehr bewerkstelligen.

DAS KI-BIEST FÜTTERN

Zu den Technologien, die beim Füttern des „KI-Biests“ mit Daten eine Rolle spielen, gehören [CPUs](#), [GPUs](#), neue Datenverarbeitungseinheiten (Data Processing Units, DPUs), [Field-Programmable Gate Arrays](#) (FPGAs) und [Application-Specific Integrated Circuits](#) (ASICs). Latenzverringernde Technologien wie NVMe und [NVMe over Fabrics](#) (NVMe-oF) helfen, die Compute-Elemente miteinander zu verbinden.

Ein Problem liegt darin, dass die CPU Höchstleistungen dabei erbringen muss, Daten zwischen den Verarbeitungskomponenten

[Home](#)[Editorial](#)[Grundlagenwissen: NVMe over Fabrics](#)[Wie sich Speicher durch NVMe und SCM entwickelt](#)[Was DPUs für Storage leisten können](#)[Optimieren mit Optane und GPUs](#)[Impressum](#)

zu bewegen. Die drängendste Frage hierbei ist, wie sich diese Arbeit an andere Stellen auslagern lässt, entweder an eine andere Stelle im Netzwerk oder in den Storage-[Controller](#) selbst.

Prozessorgestütztes [Computational Storage](#) kann helfen, die Server-CPU zu entlasten und Aufgaben wie [Verschlüsselung](#), [Komprimierung](#) und Datenbanksuche am [Edge](#) zu erledigen. Auch DPUs sind nutzbringend, indem sie die Intelligenz auf die [Netzwerkkarte](#) verlagern, um einen Teil der Arbeit zu erledigen.

CPU vs. GPU vs. DPU

DPUs sind dafür ausgelegt, die sichere Bewegung von Daten im [Rechenzentrum](#) zwischen CPUs und GPUs zu beschleunigen, wobei der Schwerpunkt auf allgemeinen und beschleunigten Berechnungen liegt.

Viele Jahre lang waren CPUs wirklich das einzige programmierbare Element in Computern, und in jüngster Zeit rücken hier GPUs in den Vordergrund. Ursprünglich wurden diese für die Bereitstellung von Echtzeitgrafiken verwendet, aber dank ihrer Parallelverarbeitungsfähigkeiten sind sie ideal für die Beschleunigung von Rechenaufgaben wie künstliche Intelligenz, [tiefes Lernen](#) und die [Big-Data-Analysen](#).

Für eine Leistungssteigerung könnte der GPUDirect-Speicher von Nvidia die Host-CPU vollständig umgehen und Daten direkt

aus dem Netzwerk in das Memory des Grafikprozessors verschieben, und zwar unter Verwendung der [RDMA](#)-Technologie (Remote Direct Memory Access). Nvidia arbeitet mit wichtigen

Akteuren und vielen Start-ups zusammen, um GPUDirect, das sich Ende 2020 in der Betaphase befand, für die allgemeine Veröffentlichung vorzubereiten.

ANWENDUNGEN MIT OPTANE BESCHLEUNIGEN

Eine der Technologien, auf die sich Intel konzentriert, um besonders anspruchsvolle Anwendungen zu beschleunigen, ist das nicht flüchtige [3D XPoint Memory](#), das Intel ge-

meinsam mit Micron entwickelt hat, um die Leistungslücke zwischen [DRAM](#) und [NAND-Flash](#) zu schließen. Intel verwendet den Markennamen Optane für seine 3D-XPoint-basierten Solid-State-Laufwerke ([SSDs](#)), die seit 2017 ausgeliefert werden, und für [Dual In-Line-Memory Modules](#) (DIMMs), die 2019 auf den Markt kamen.

Insbesondere Oracle-Datenbanken lassen sich durch den Einsatz von Optane-SSDs optimieren und erzielen noch größere Leistungssteigerungen, wenn Optane-DIMMs eng mit Intels neuesten Xeon-Prozessoren gekoppelt würden. Trotz einer eher schleppenden Einführung wird der Markt im DIMM-Geschäft auf mehr als 10 Milliarden Dollar innerhalb der kommenden Jahre geschätzt.

VIELE JAHRE LANG WAREN CPUs WIRKLICH DAS EINZIGE PROGRAMMIERBARE ELEMENT IN COMPUTERN, UND IN JÜNGSTER ZEIT RÜCKEN HIER GPUS IN DEN VORDERGRUND.

[Home](#)[Editorial](#)[Grundlagenwissen: NVMe over Fabrics](#)[Wie sich Speicher durch NVMe und SCM entwickelt](#)[Was DPUs für Storage leisten können](#)[Optimieren mit Optane und GPUs](#)[Impressum](#)

VIELVERSPRECHENDSTE AUFKOMMENDE MEMORY-TECHNOLOGIEN

Darüber hinaus spielen andere aufkommende, integrierbare Memory-Technologien künftig eine Rolle. Dazu gehören [Magneto-resistive Random Access Memory](#) (MRAM), [Resistive RAM](#) (ReRAM oder RRAM) und Ferroelectric RAM (FeRAM oder FRAM).

Experten sehen bis 2030 ein Marktpotenzial von 36 Milliarden Dollar bei diesen neuen Memory-Technologien, angeführt von 3D XPoint und MRAM. Das größte Unternehmen, das eigenständige MRAM-Produkte herstellt, Everspin Technologies, gibt an, bislang mehr als 120 Millionen Chips ausgeliefert zu haben.

Durch all diese Innovationen wird ein breites Angebot an Memory-Typen entstehen, das den gesamten Markt verändern wird. Da [NOR-Flash](#) und [SRAM](#) in puncto [Skalierung](#) an ihre Grenzen stoßen, könnte sich hier eine Verschiebung von flüchtigen zu nicht flüchtigen Memory-Architekturen abzeichnen.

Das wiederum kann großen Einfluss auf das Gerätedesign haben, sowohl für integrierte als auch große Systeme.

Mit Optane Persistent-Memory-DIMMs lässt sich die Memory-Kapazität weitaus günstiger vergrößern als dies mit DRAM mög-

lich ist, um [In-Memory-Datenbanken](#) wie [SAP HANA](#) zu beschleunigen und mehr [virtuelle Maschinen](#) pro Server in Virtualisierungsumgebungen zu ermöglichen. Ein weiterer entscheidender Vorteil der Verwendung von persistentem Memory gegenüber DRAM ist eine Reduzierung des Stromverbrauchs.

Auf der Softwareseite ist die Entwicklung, von Open Source Distributed Asynchronous Object Storage (DAOS) bedeutsam. Diese wurde von Intel konzipiert, um mit den Optane PMem-DIMMs zu arbeiten und kleine und falsch ausgerichtete I/Os zu

handhaben. DAOS dient als ein Beispiel dafür, wie eine Mischung aus der Verwendung von persistentem Memory zusammen mit Speicher wirklich vorteilhaft für HPC- und KI-Arbeitslasten sein kann.

Viele Analysten erwarten, dass Optane-DIMMs langfristig einen größeren Einfluss haben werden als Optane-SSDs, und die Preise für beide werden sinken, da die Hersteller mehr Chips produzieren. Allerdings könnten sie auch zunächst nur eine Nische bedienen. [Storage Class Memory](#) muss sich und sein Poten-

zial am Markt erst noch langfristig beweisen. Derzeit ist das größte Wachstumspotenzial wohl eher in Bereichen wie Speichersysteme mit NVMe-SSDs und [All-Flash-Arrays](#) mit NVMe-oF zu sehen. ■

**VIELE ANALYSTEN ERWARTEN,
DASS OPTANE-DIMMs
LANGFRISTIG EINEN GRÖßEREN
EINFLUSS HABEN WERDEN
ALS OPTANE-SSDs.**

[Home](#)

[Editorial](#)

[Grundlagen-
wissen:
NVMe over
Fabrics](#)

[Wie sich
Speicher durch
NVMe und
SCM entwickelt](#)

[Was DPUs
für Storage
leisten können](#)

[Optimieren
mit Optane
und GPUs](#)

[Impressum](#)



NVMe, SCM und DPU: Technologien für schnelleres Storage
ist eine Publikation von [ComputerWeekly.de](https://www.computerweekly.de)

Michael Eckert | Editorial Director
Malte Jeschke | Senior Online Editor

Ulrike Riess | Senior Online Editor
Tobias Servaty-Wendehost | Senior Online Editor

Franziska Holzfurtner | Junior Editor
Becky Wrigley | Production Editor

Brent Boswell | Herausgeber
BBoswell@techtarget.com

Hauptsitz:
TechTarget, 275 Grove Street, Newton, MA 02466
www.techtarget.com

Deutsche Redaktion:
TechTarget Deutschland, Landwehrstraße 61, 80336 München
www.techtarget.de

© 2021 TechTarget Inc. Kein Teil dieser Veröffentlichung darf ohne
vorherige schriftliche Genehmigung des Verlages in irgendeiner Form
oder auf irgendeine Weise weitergegeben oder reproduziert werden.
Nachdrucke von TechTarget-Publikationen sind verfügbar über The YGS Group.

Über TechTarget:

TechTarget publiziert Informationen für Profis im Bereich Informationstechnologie. Mehr als 100 Themen-Websites ermöglichen schnellen Zugriff auf ein reichhaltiges Angebot an Nachrichten, Ratgebern und Analysen über die Technologien, Produkte und Prozesse, die entscheidend sind für beruflichen Erfolg. Unsere Live- und virtuellen Veranstaltungen vermitteln direkten Zugang zu den Einschätzungen und Ratschlägen unabhängiger Experten. IT Knowledge Exchange, unsere soziale Community, bietet die Möglichkeit, um Rat zu fragen und sich mit Kollegen und Experten über Lösungen auszutauschen.

VIDADY/ADOBE

