

NVMe: Das Protokoll für zukünftige SSDs

Wann brauchen Sie NVMe?

Vielleicht haben Sie gehört, dass Non-Volatile Memory Express oder NVM Express (NVMe) die nächste Must-Have-Storage-Technologie ist. Schauen wir uns an, was NVMe liefert. NVMe ist ein Kommunikationsprotokoll für Hochgeschwindigkeitsspeichersysteme, das auf der physischen Schicht von PCIe läuft und ermöglicht es Hosthardware und Software, die Leistung von High-End-SSDs voll auszuschöpfen. Jeder muss es jetzt schaffen, richtig? Nun, nicht ganz so schnell.

NVMe Hintergrund

Erstens, einige Begriffen. Die physikalische PCIe-Schicht besteht aus dem Silizium-Treiber und dem Empfänger an jedem Ende einer Leitung, der die Spannung und Signalisierungsrate des PCIe-Standards erfüllt. NVMe ist das Protokoll oder die Sprache, die über diese Drähte kommuniziert wird. Diese Leitungen könnten auch das PCIe-Protokoll tragen, aber NVMe ist eine viel besser geeignete Möglichkeit, Informationen zu Speichersysteme zu kommunizieren. Es ist keine Überraschung, dass die Entwicklung von PCIe und NVMe aus der gleichen Gruppe von Sponsoring-Unternehmen kam.



NVMe wurde entwickelt, um die Einschränkungen bestehender Speicherschnittstellen und Protokolle wie PATA, SATA, SCSI und SAS zu beheben. Die Ursprünge dieser Protokolle sind über 30 Jahre alt und haben sich auf der Grundlage der begrenzten Fähigkeiten der traditionellen mechanischen Festplatten entwickelt. Diese Legacy-Schnittstellen haben sich lange gehalten aufgrund der Notwendigkeit für Vorwärts- und Rückwärts-Kompatibilität in Hardware und Treiber. NVMe wurde entworfen, um erweiterbar zu sein, so dass NVMe so lange wie diese traditionellen Schnittstellen verfügbar ist, die es ersetzen wird. Das volle Leistungspotential, das NVMe zu bieten hat, ist genau das, was noch nicht realisiert ist. Intel hat vor kurzem ihre neuen Optane SSD-Laufwerke auf Basis der revolutionären 3D-XP-Speichertechnologie veröffentlicht. 3D-XP ist Größenordnungen schneller als NAND-Flash und NVMe verfügt über umfangreiches Leistungssteigerungspotential nicht nur für die aktuelle Generation von Optane, sondern für viele Generationen von neueren, schnelleren Speichertechnologien.

NVMe ermöglicht es Systemen, die Vorteile der geringeren Datenlatenz und der internen Parallelität des zufälligen Zugriffs, der Speicher-basierten Speicherung, zu nutzen. Die NVMe-Reduzierung von I / O-Overhead- und Performance-Verbesserungen verleiht ihm die Möglichkeit, mehrere, lange Befehlswarteschlangen zu verarbeiten, im Vergleich zu der Möglichkeit, nur eine Warteschlange mit

vorherigen logischen Geräte-Schnittstellen zu verarbeiten. NVMe ermöglichte es, nur eine einzige Nachricht für 4KB Transfers im Vergleich zu zwei verwenden. Dies erhöht die Geschwindigkeit für Server, die viele gleichzeitige Festplatten-E / A-Anfragen verarbeiten.

| High-Level-Vergleich von AHCI und NVMe ^[5] Speicherprotokolle | | |
|--|--|---|
| | AHCI (SATA) | NVMe |
| Maximale Warteschlangen-Tiefe | Eine Befehlswarteschlange; 32 Befehle pro Warteschlange | 65535 Warteschlangen; ^[29] 65536 Befehle pro Warteschlange |
| Uncacheable Registerzugriffe (Jeweils 2000 Zyklen) | Sechs pro nicht quittierten Befehl; Neun pro Warteschlange Befehl | Zwei pro Befehl |
| MSI-X Und Lenkung unterbrechen | Ein einziger Unterbrechung; Keine Lenkung | 2048 MSI-X unterbricht |
| Parallelität Und mehrere Threads | Benötigt Synchronisationssperre Einen Befehl ausgeben | Keine Verriegelung |
| Wirkungsgrad Für 4 KB Befehle | Befehlsparameter erfordern Zwei serialisierte Host-DRAM-Abrufe | Ruft Befehlsparameter ab In einem 64-Byte-Abruf |

Quelle: www.wikipedia.org

Die Systemleistung wurde historisch durch die Leistung des Speichers begrenzt - nicht die Speicherschnittstelle, sondern die Speichermedien. In einer mechanischen Festplatte ist die Medienleistung durch die Mechanik begrenzt. Kein HDD ist in der Lage, eine SATA-III-Schnittstelle auszureizen; Die besten sind bei etwa 25%, was die Leistungsgrenze von SATA-I ist. SATA-II und -III wurden ausgelegt, um einen Bruchteil der potentiellen Leistung von NAND Flash zu verwenden, aber das SATA-Protokoll ist der limitierende Faktor, sobald man bestimmte Durchsatzwerte erreicht hat. Hoher Durchsatz bedeutet, dass eine größere Anzahl von Transaktionen auftreten kann, aber das SATA-Protokoll erlaubt nicht genügend Pointergröße oder Warteschlangengröße, um die Leistung einer SSD zu behandeln. Da SATA eine Brücke zwischen PCIe ist (deren Protokoll auch ein Speicherengpass ist), warum nicht einfach sowohl SATA als auch PCIe eliminieren? Was natürlich genau das ist, was NVMe macht.

Also, warum nicht gleich auf NVMe umsteigen?

Die Verbesserungen in der Speicherschnittstelle, wie etwa die Vervierfachung der SATA-Bandbreite, sind nicht der einzige Grund dafür, dass SSDs so viel schneller sind als Festplatten. SSDs sind schneller, weil sich die Wartezeit (Latenz) bzw. Reaktionszeit der Speichermedien um 1000x von einem mechanischen Antrieb zu einer SSD verbessert hat.

Die beobachtbare Systemleistung, welche der Benutzer sehen und fühlen kann, ist durch einen Engpass begrenzt. Für die längste Zeit ist der Flaschenhals die mechanische Festplatte gewesen. Sobald dieser Flaschenhals entfernt wurde, war der nächste Engpass die SATA-I Schnittstelle. Jetzt mit SATA-III SSDs könnte der Engpass heute ein anderer Teil des Systems sein, vielleicht die GPU für Grafik-Performance oder die CPU oder Speicher für die Verarbeitungsleistung, und in einigen Fällen könnte es auch das Speichersystem sein. Wenn Sie speicherintensive Operationen ausführen, dann könnte NVMe genau das sein, was Sie brauchen. Aber wenn der Performance-Engpass anderswo ist oder wenn Sie sich nicht mit der höheren Leistung belasten wollen, die mit NVMe Performance einhegeht, dann ist NVMe vielleicht nicht das, was Sie gerade suchen.

Die Quintessenz ist, dass die unglaublichen Leistungsverbesserungen, die durch den Wechsel von der mechanischen Festplatte hin zu SSD realisiert werden, nicht erreicht werden, indem sie von SATA zu NVMe wechseln. Wenn die Anwendung nicht die zusätzliche Speicherleistung benötigt, wird kein Vorteil realisiert.

Warum NVMe dominieren wird... eventuell

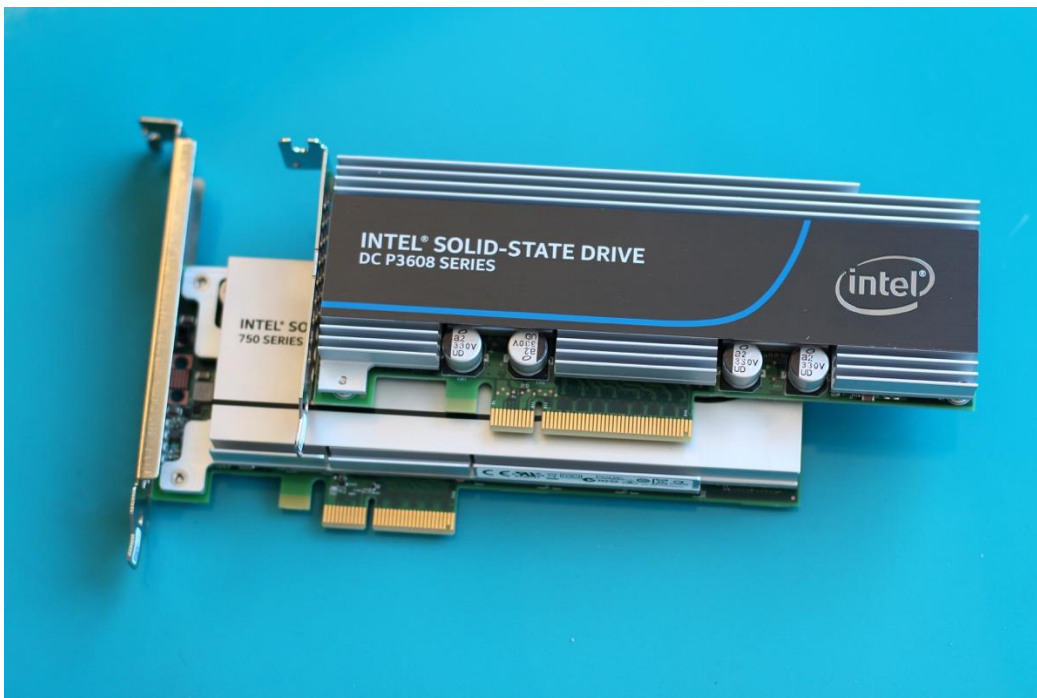
NVMe wird der De-facto-Standard nicht wegen der Leistung, die es bietet, sondern eher wegen der Standardisierung und breiten Nutzung. Diese Prozesse, welche nicht von der Leistung getrieben sind, können sehr lange dauern. Heute erfordert die traditionelle Speicherung einen ACHI (SATA) oder SAS Controller, um mechanische Festplatten mit dem PCIe Bus zu verbinden, und NVMe ist die Lösung, die den Controller eliminiert und stattdessen die direkte Verbindung herstellt. Aber wie so viele andere Standards ist die SATA-Schnittstelle gut etabliert, hat mehrere IP-Anbieter und ist weit verbreitet mit marktbasieren Preisen.

Eine kurze Übersicht für SSDs zeigt, dass NVMe-Preise im Einklang stehen für für Einsteiger-SSD, aber das Hochleistungspotential von NVMe kommt zu einem bestimmten Preis. Betrachtet man eine Reihe von 512GB SSDs zeigt sich, dass die SATA und NVMe M.2s im gleichen Preisniveau sind, aber noch etwas teurer als die 2,5 ". Der höhere Preis ist auf die Leistung und größere Flexibilität des M.2 zurückzuführen.

| Speicherlaufwerke | Schnittstelle | Performance | Preis /GB |
|-------------------------|---------------|-------------|-----------|
| AiC x8 Unternehmen | NVMe | ~40Gb/s | 5.0x |
| U.2 x4 Unternehmen | NVMe | ~18Gb/s | 2.5x |
| M.2 x4 Prosumer | NVMe | ~20Gb/s | 2.0x |
| M.2 x4 Einstiegslevel | NVMe | ~12Gb/s | 1.0x |
| M.2 SATA | SATA-III | ~3Gb/s | 1.0x |
| 2.5" SSD Prosumer | SATA-III | ~4.5Gb/s | 1.5x |
| 2.5" SSD Einstiegslevel | SATA-III | ~3Gb/s | 0.9x |

NVMe SSD Formfaktoren

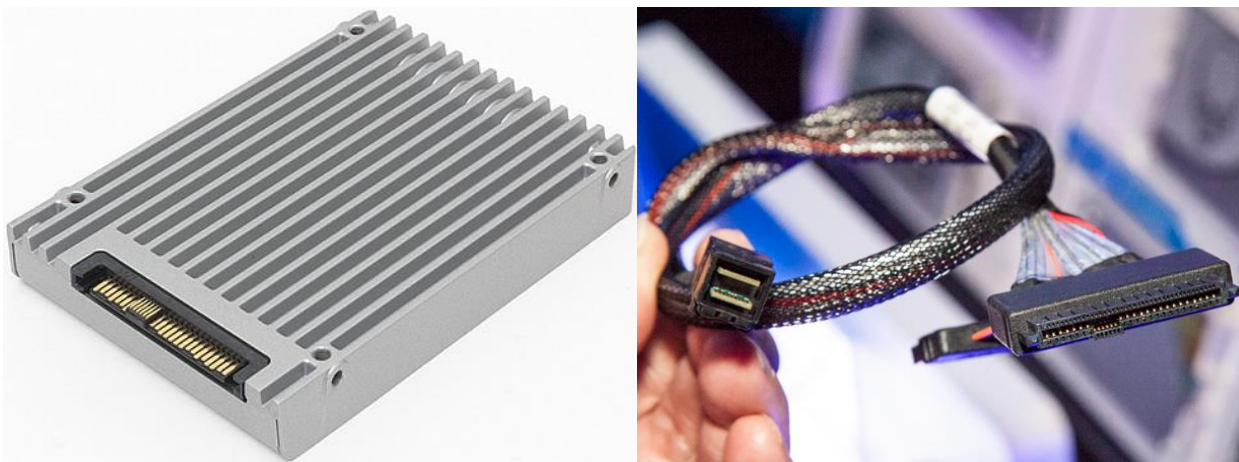
NVMe-Laufwerke kommen in drei grundlegende Formfaktoren, die Add-In-Karte (AiC), die 2,5 "(U.2) und die M.2. Da das NVMe-Protokoll auf der physischen Schicht des PCIe läuft, sind die Verbindungen nahezu identisch mit PCIe. Die AiC-Karten ähneln Standard-Erweiterungskarten, das M.2 ähnelt Mini-PCIe und das 2,5 "soll die Modularitätsanforderungen von Speichersystemen adressieren. Die AiC hat eine hohe Kapazität und kann so konzipiert werden, dass sie eine beliebige Anzahl von PCIe-Skits unterstützt. Diese Karten sind in der Regel für hohe IOPS-Anwendungen, aber Sie brauchen einen eigenen PCIe-Slots für die Karte und sie sind nicht Hot-Swap-fähig. Unten sind eine 4 Lane und eine 8 Lane AiC von Intel.



Eine kleinere Version dieser Art der Implementierung ist der M.2 Formfaktor, der gemeinhin 22mm breit ist und eine Länge zwischen 42mm und 110mm hat. Die M2-Buchse ähnelt der eines SO-DIMMs für den Speicher, so dass die Karte parallel zur Hauptplatine liegt. Natürlich gibt es noch andere Sockel-Orientierungen zur Verfügung, aber die häufigste ist die, die in Laptop-Computern verwendet wird.



Um das Angebot abzurunden ist ein 2,5 "oder 3,5" Formfaktor, der ein U.2 Kabel mit U.2 NVMe Schnittstelle verwendet. Diese Laufwerke sind für Unternehmens- und Rechenzentrumsanwendungen gedacht und sind in der Regel in hohen Kapazitäten und hohen Schreib-Ausdauer-Modellen verfügbar. Ähnlich sind die U.2-Kabel und Host-Anschluss deutlich größer als SATA. Bisher gibt es keine NVMe 2.5 "Laufwerke und mit gutem Grund, da die 2,5" SATA die kostengünstigste Lösung ist und voraussichtlich schon seit längerer Zeit ist.



NVMe Cross-Over

Mit NVMe-Preisen, die ziemlich hoch anfangen und schnell gefallen sind, ist es keine Überraschung, dass die Experten voraussagen, wann es billiger sein wird als SATA, aber meistens kommt es anders. NVMe werden aller Voraussicht nach niemals deutlich günstiger als SATA werden. Die Preise für NVMe sind hoch, weil sie aufgrund ihrer Leistung verkauft werden. Wenn NVMe mit SATA konkurriert, wird es auf der Grundlage von \$ / GB konkurrieren. Da beide den gleichen NAND Flash hinter dem Controller verwenden, gibt es keinen Grund für NVMe, jemals weniger als ein SATA-Laufwerk zu kosten. Dies ist natürlich für die DRAM-Technologie unterschiedlich, da die Speicherchips für einen bestimmten DRAM ein völlig anderes Design sind und unterschiedlichen Marktkräften unterliegen.

Robuste NAND Geräte

Heute haben wir eine Reihe von NAND Flash Geräten mit unterschiedlichen Fähigkeiten und Anforderungen. Die älteste und robusteste NAND-Technologie ist SLC (Single Level Cell), dann MLC (Multi-Level Zelle), jetzt TLC und QLC entstehen zusammen mit 3D-NAND-Architekturen. Diese neueren Technologien sind pro Bit kostengünstiger, aber gleichzeitig haben sie eine niedrigere Zuverlässigkeit und Ausdauer auf Komponentenebene. SLC ist auch heute noch für Systeme konstruiert, die höchstmögliche Verfügbarkeit und Robustheit erfordern.

Die neuesten NVMe-Controller werden entwickelt, um mit den neuesten NAND-Geräten zu arbeiten, dh TLC, QLC und 3D-NAND. Es ist unwahrscheinlich, dass neue NVMe-Laufwerke die zuverlässigeren MLC- und SLC-NAND-Technologien unterstützen. Anwendungen, die eine stabilere und robuste Speicherlösung erfordern, würden am besten mit dem bewährten MLC NAND funktionieren, was die Verfügbarkeit von kostengünstigen NVMe-Laufwerken begrenzt.

Schlussfolgerung

Die etablierten SATA-Speicher sind kostengünstig und für die meisten Anwendungen geeignet. SATA-Laufwerke sind mit zuverlässigen MLC- oder SLC-NAND-Flash-Komponenten erhältlich. Die höhere Performance NVMe werden immer etwas teuer sein, während SATA-Level-Performance sich den Kosten für SATA-Laufwerke annähert. NVMe ist keine revolutionäre (disruptive) Technologie, die die Wirtschaftlichkeit der Speichersysteme erheblich verändern wird - es ist eine leistungsorientierte Lösung für diejenigen, denen es der höhere Preis wert ist.

Quellen

https://en.wikipedia.org/wiki/NVM_Express

https://en.wikipedia.org/wiki/PCI_Express

HAFTUNGSAUSSCHLUSS: Alle Produkt-, Produktspezifikationen und Daten können ohne vorherige Ankündigung geändert werden, um die Zuverlässigkeit, Funktion oder Design zu verbessern oder auf andere Weise. Die hierin enthaltenen Informationen entsprechen dem Besten des Insignis Technology Corporation. Keine Haftung für Fehler, Fakten oder Meinungen wird akzeptiert. Die Kunden müssen sich hinsichtlich der Eignung dieses Produkts für ihre Anwendung befriedigen. Keine Verantwortung für Verluste infolge einer Person, die sich auf das hierin enthaltene Material bezieht, wird akzeptiert.