

Universität Hamburg
Fachbereich Informatik
Seminar: Speicher und Dateisysteme

3D XPoint
28.09.2016

Name: Patrick Wittke
Email: patrick27089@live.de
Martikelnnummer: 6532784

Inhaltsverzeichnis

1. Einleitung.....	3
2. Architektur.....	4
2.1 Aufbau.....	4
2.2 Phase Change Memory.....	5
3. Schnittstellen.....	6
3.1 NVM Express.....	6
3.2 PCI Express.....	8
4. Aktuelle Dateisysteme.....	9
5. 3D XPoint im Vergleich.....	10
6. Zusammenfassung.....	11

1. Einleitung

3D XPoint, entwickelt von Intel und Micron Technology, wurde im Juli 2015 vorgestellt. Es handelt sich hierbei um einen nichtflüchtigen Speicher, welcher mit Hilfe des Elektrischen Widerstandes die Informationen speichert. Die Speicherchips sind wie ein Gitter angeordnet, bestehen aus 4 verschiedenen Elementen und sind außerdem stapelbar.

Es ist lange her seitdem eine neue Speichertechnologie veröffentlicht wurde. Vorher kamen neu entwickelte Speichermedien in relativ kurzen Abständen auf den Markt. Zuletzt wurde 1989 NAND Flash Speicher veröffentlicht und erst 2015, 26 Jahre später, wurde 3D Xpoint vorgestellt. Intel verspricht das 3D XPoint den Speichermarkt revolutionieren wird.

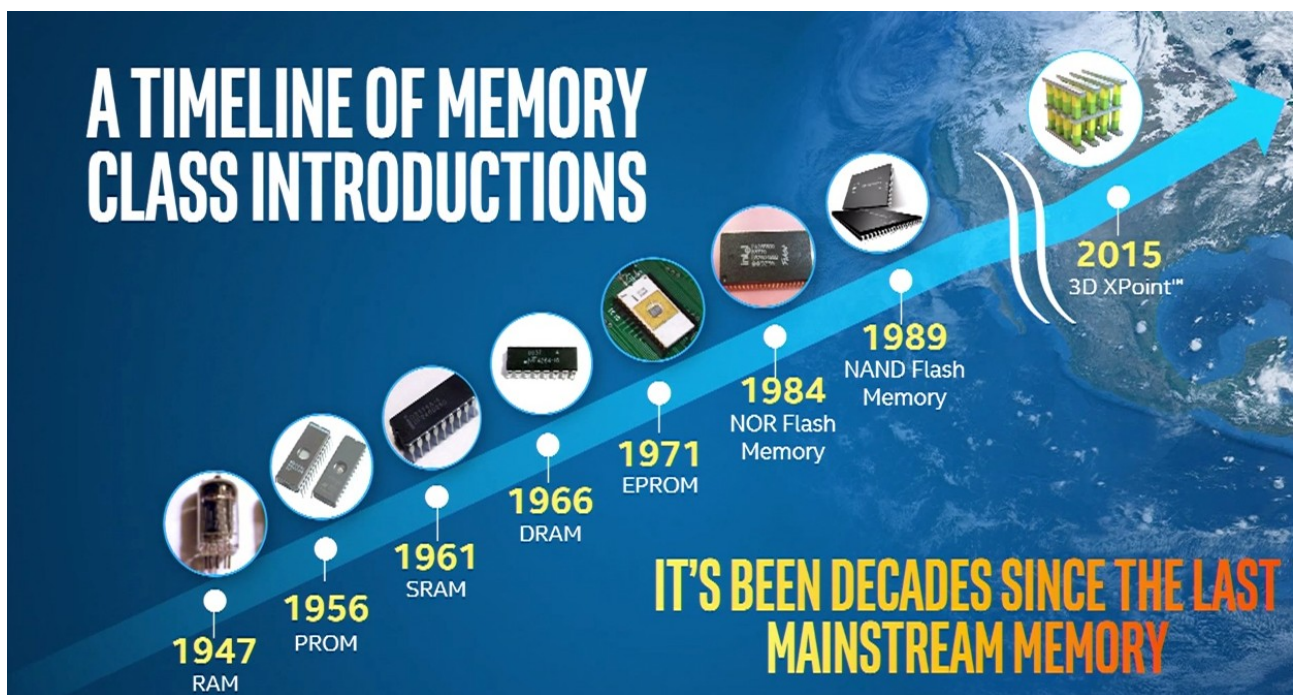


Abbildung: Timeline

Vorteile der neuen Technologie sollen sich vor allem in der geringen Latenz, Haltbarkeit und geringen Herstellungskosten bemerkbar machen.

Es wird keine externe Stromzufuhr und Transistoren benötigt.

Da 3D XPoint noch nicht veröffentlicht wurde, werden sich die Angaben auf die vom Hersteller beziehen. Das bedeutet auch, dass Zahlen möglicherweise beschönigt sind, sodass 3D XPoint möglichst im gutem Licht dargestellt wird.

2. Architektur

2.1 Aufbau

3D-Xpoint setzt sich aus 4 verschiedene Komponenten zusammen. Der Memory Cell, dem Selector, der Word Line und der Bit Line. Die Memory Cell speichert jeweils ein Bit, der Selector ermöglicht das Lesen und Schreiben der Memory Cell und die Kombination aus Word- und Bit Line ermöglicht einzelne Zellen zu erreichen.

Aufgebaut wird das sogenannte Gitter indem unten die Bit Line liegt und diagonal drüber die Word Line. Dazwischen befindet sich die Memory Cell und der Selector. Durch das kreuzen von Word- und Bit Line kann eine bestimmte Speicherzelle nur durch eine bestimmte Word- und Bit Line erreicht werden. Veranschaulicht finden sie dies in Abbildung „ Aufbau 3D-XPoint “.

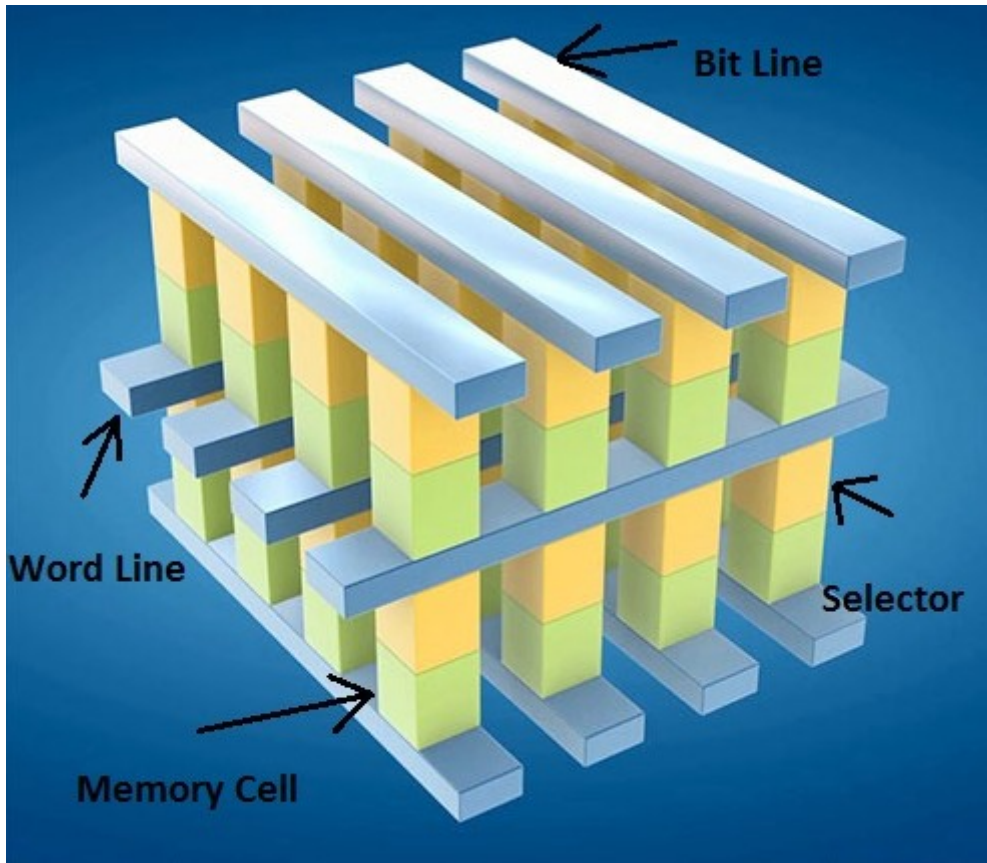


Abbildung: Aufbau 3D XPoint

Wie bereits erwähnt wird die Memory Cell durch das Kreuzen der Word- und Bit Line erreicht. Dies wollen wir uns nun weiter im Detail anschauen.

Zuerst wird die Word Line, auf der unsere Memory Cell sitzt, mit einer gewissen Stromspannung versetzt. Die Stromspannung ist wichtig, da wir durch die Stromspannung entscheiden, ob wir die Speicherzelle auslesen und beschreiben wollen. Nun kommt die Bit Line ins Spiel und wählt die Speicherzelle aus, welche wir uns ausgesucht haben. Die Stromspannung, die wir zurück bekommen, gibt uns dann Auskunft darüber, falls wir lesen, ob eine 1 oder 0 vorliegt.

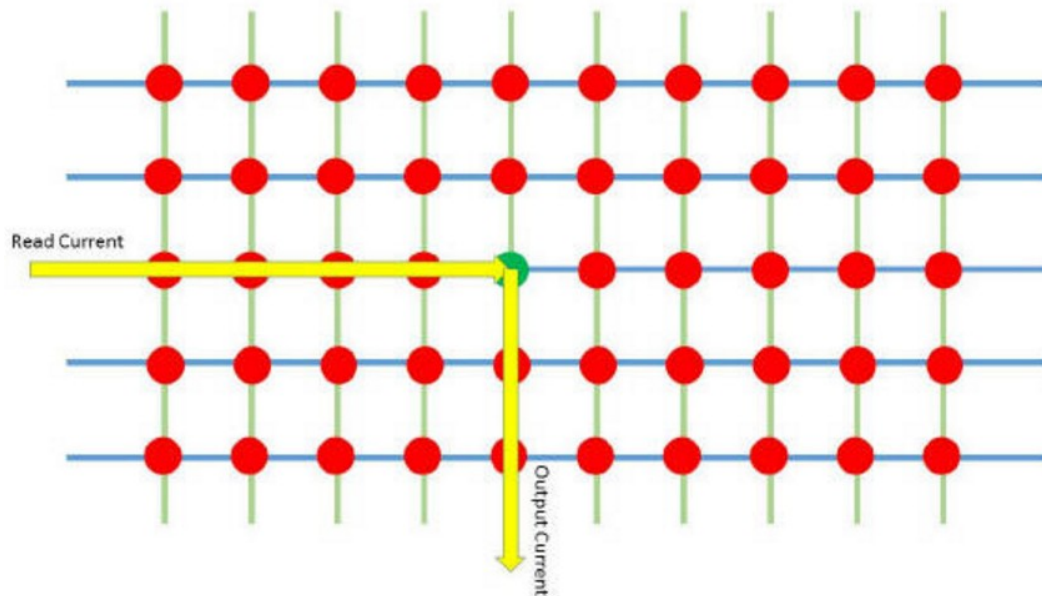


Abbildung: Word-und Bit Line

2.2 Phase Change Memory

Nun gehen wir ins Detail wie die Speicherzelle funktioniert. Die Methode nennt sich „Phase-Change-Memory“ und diese merkt sich Informationen indem sie den elektrischen Widerstands von Materialien verändert. Es wird meist eine Chalkogenid-Legierung verwendet, aber welche Chalkogenid-Legierung bei 3D-XPoint verwendet wird leider noch nicht bekannt. Es gibt insgesamt 2 verschiedene Formen, die amorphe Form, welche einen hohen elektrischen Widerstand aufweist, und den kristalline Form, welche einen geringen elektrischen Widerstand aufweist.

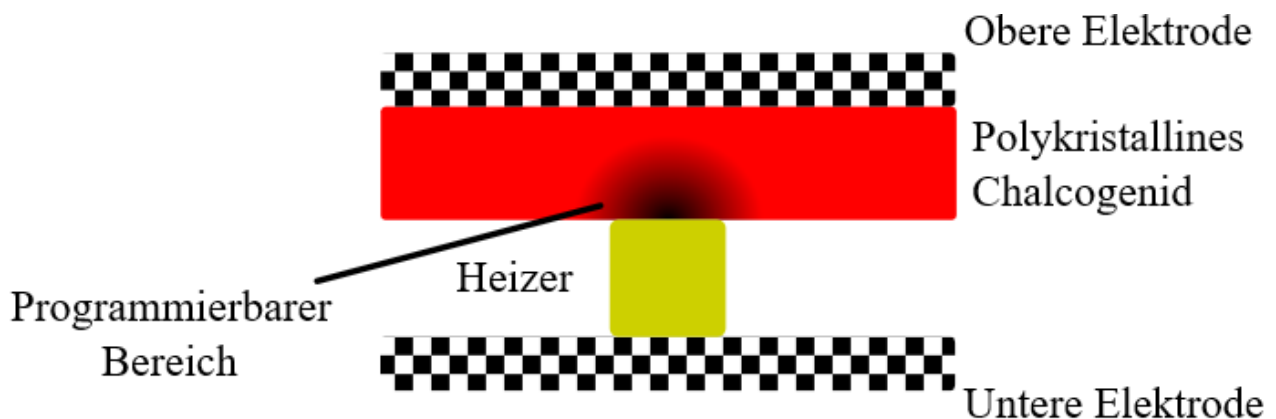


Abbildung: Phase-Change-Memory

Eine Speicherzelle besteht ebenso aus 4 Bestandteilen, der Bottom-und Top Elektrode. Zwischen Bottom-und Top Elektrode liegt das Phasenwechselmaterial und der Heizer, welches das Phasenwechselmaterial erhitzt.

Nun ist aber noch nicht erklärt wie zwischen den amorphen und kristallinen gewechselt wird. Dies erfolgt durch Erhitzung des Materials. Durch einen Stromimpuls höhere Stromstärke (mehrere hundert Mikroampere) und einer geringen Dauer (z.B. 50 Nanosekunden) geht das Material in der amorphen Form über und kühlt rasch ab sodass es auch im amorphen Zustand verbleibt und nicht

kristallisiert.

Um die kristalline Form zu erreichen wird das Material auch erhitzt, nur diesmal mit einem Strompuls längere Dauer (z. B. 100 Nanosekunden) und geringerer Stromstärke (mehrere zehn bis wenige hundert Mikroampere). Die Temperatur muss mindestens die Kristallisationstemperatur erreichen und solange diese Temperatur beibehalten bis die Keimbildung beginnt und Kristallisation eintritt.

Dies alles führt dazu, dass keine Transistoren mehr benötigt werden, welche sich auf einen günstigeren Produktion und eine wesentliche höhere Packungsdichte niederschlägt. Nach aktuellen Angaben schafft ein Die bis her 128 GB zu speichern.

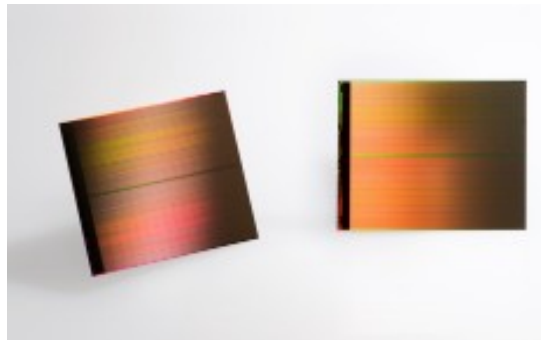


Abbildung: Die

3. Schnittstellen

3.1 NVM Express

Nicht flüchtiger Speicher mit direkter PCI-Express Anbindung bietet ganz neue Möglichkeiten, dass bedeutet auch das herkömmliche Protokolle, wie zum Beispiel AHCI, das Potenzial der neuen Speichertechnologien nicht ausnutzen können. Dieses Problem soll NVM-Express (Non-Volatile Memory Express) beheben und eine höhere Geschwindigkeit garantieren.



Abbildung: NVM-Express Logo

NVM-Express erschien im Jahr 2011 und wurde durch Intel vorangetrieben. Unter anderem waren auch Dell, Cisco, EMC, Marvell, SanDisk, Avago Technologies, Mircon Technologies, HGST, Samsung und Seagate an dem Projekt beteiligt. NVM-Express ist eine Schnittstelle für nicht flüchtigen Speicher welche über PCI-Express verbunden werden.

Ein großer Vorteil besteht unter anderem darin, dass keine Hersteller spezifische Treiber benötigt werden. Seine großen stärken spielt NVM-Express aus bei parallelen Zugriffen, in dem die Latenz und der Overhead verringert werden. Außerdem spart es Strom und schont die CPU-Ressourcen. Da drüber hinaus unterstützt es 64000 Befehle pro Warteschlange und 64000 Warteschlangen. Im Vergleich unterstützt AHCI jegliche nur eine Warteschlange und 32 Befehle.

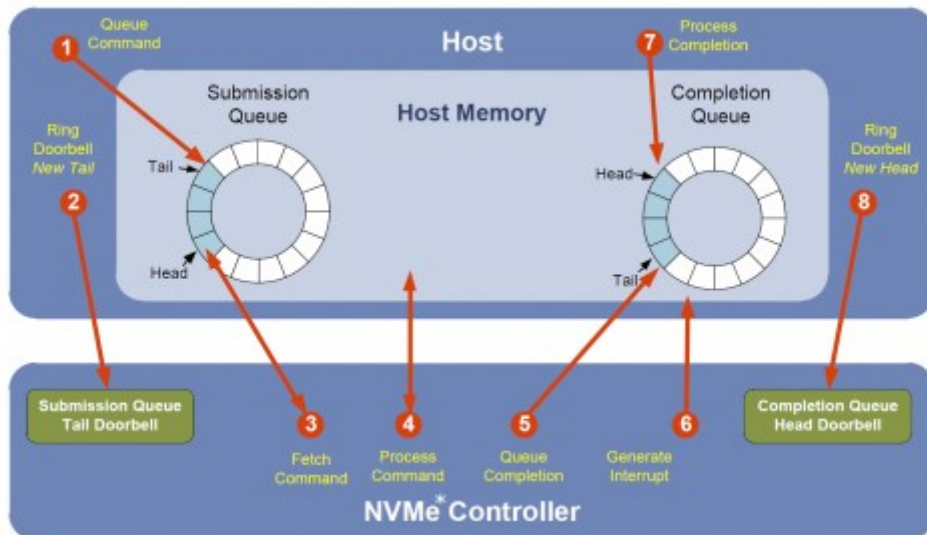


Abbildung: Schematischer Aufbau NVM-Express

Der Host steckt seine Befehle in die Submission Queue, welches sich in der Host Memory befindet und erhöht den Tail um eins. Der NVMe-Express Controller holt sich die Befehle vom Head, verringert den Head ebenfalls um eins und verarbeitet diesen Befehle. Es gibt zusätzlich noch eine Doorbell welche die Position vom Tail übergibt, somit ist der NVMe-Express Controller immer im Wissen wo sich der Tail befindet. Nun kann er sich berechnen wie voll die Submission Queue ist, dies ist wichtig weil es insgesamt 64000 Warteschlangen gibt und somit ist Der NVMe-Express Controller in der Lage Prioritäten zu setzen, welches bedeutet das Submission Queues die voller sind als andere eine höhere Priorität zugeordnet bekommen. Nach dem verarbeiten des Befehls wird das Resultat in die Completion Queue abgelegt und der Tail um eins erhöht. Der Host holt die Ergebnisse sich vom Head aus der Completion Queue und verringert den Head um eins. Genau wie bei der Submission Queue gibt es auch bei der Completion Queue eine Doorbell, welche dem NVMe-Express Controller den Head mitteilt. Somit kann es nicht dazu kommen das die Completion Queue überläuft und der NVMe-Express Controller kann sich erst mal auf andere Warteschlangen abarbeiten.

3.2 PCI Express

PCI Express steht für Peripheral Component Interconnect Express und ist der Nachfolger von PCI und ermöglicht eine höhere Datenübertragung. Der Anschluss dient zur Verbindung vom Chipsatz des Hauptprozessors mit anderem Zubehör.

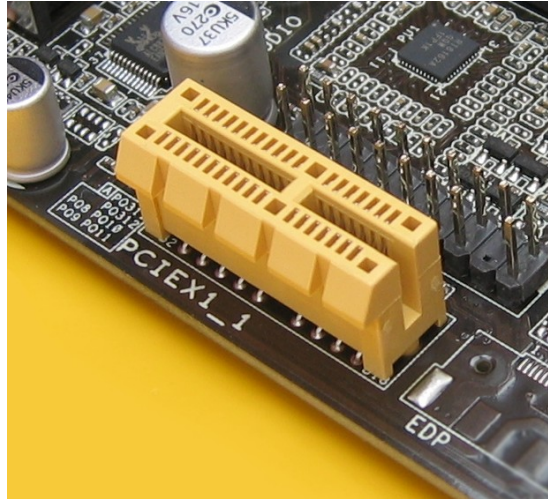


Abbildung: PCI-Express Slot

PCI-Express ist eine separate serielle Punkt zu Punkt Verbindung. Für das Senden/Empfangen von Daten werden Wandler verwendet. Das bedeutet für das Senden werden Parallel zu Seriell Wandler und für das Empfangen Seriell zu Parallel Wandler verwendet. Trotz der großen Unterschiede im Aufbau ist PCI-Express voll kompatibel, auf die Software geschaut, zu PCI. Ein weiterer Vorteil besteht darin, dass PCI-Express voll duplexfähig ist und Datenübertragungen je nach Version von 250, 500 oder 985 MB/s pro Lane und Richtung ermöglicht.

	PCIe 1.0/1.1	PCIe 2.0/2.1	PCIe 3.0	PCIe 4.0
Erschienen	2003	2007	2012	~2017 ^[1]
Transfers/s (je Lane und Richtung)	2,5 <u>GT/s</u>	5,0 GT/s	8,0 GT/s	16,0 GT/s ^[2]
Kodierung	8b10b	8b10b	128b130b	128b130b
Lanes (Breite)				
×1	250 MB/s ^[3]	500 MB/s	985 MB/s	1969 MB/s
×2	500 MB/s	1000 MB/s	1969 MB/s	3938 MB/s
×4	1000 MB/s	2000 MB/s	3938 MB/s	7877 MB/s
×8	2000 MB/s	4000 MB/s	7877 MB/s	15754 MB/s
×16	4000 MB/s	8000 MB/s	15754 MB/s	31508 MB/s
×32 ^[4]	8000 MB/s	16000 MB/s	31508 MB/s	63015 MB/s

Abbildung: PCI-Express Versionen

Speicher	Geschwindigkeit
Festplatte	100 MB/s
SSD (AHCI) mit SATA-Anschluss	500-600 MB/s
SSD (AHCI) mit PCIe-Anschluss	1000 MB/s (1 Lane)
SSD (NVME) mit PCIe-Anschluss	4000 MB/s (4 Lanes)

Abbildung: Tabelle NVM-Express im Vergleich

Die Abbildung „Tabelle NVM-Express im Vergleich“ zeigt, dass NVM-Express in der Lage ist wesentlich mehr Datenübertragung aus der SSD rauszuholen mit 4 Lanes.

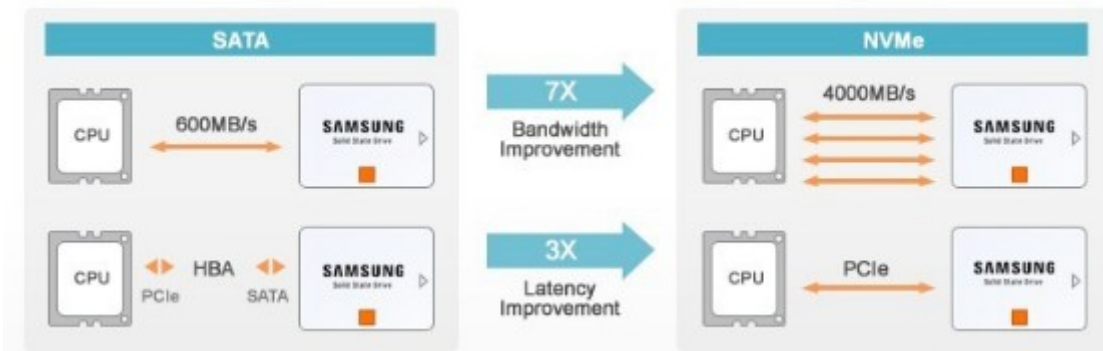


Abbildung: Grafisch NVM-Express im Vergleich

In der Zweiten Abbildung „Grafisch NVM-Express im Vergleich“ sieht man das erst mal die Datenübertragung um ein Vielfaches höher ist und zweitens das die Latenz sich verbessert dadurch das keine Komponente zwischen dem Speichergerät und der CPU zwischengeschaltet sind.

4. Aktuelle Dateisysteme

Die aktuellen Dateisysteme sind leider nicht für 3D XPoint Dimms ausgelegt, welches zur Folge hat, dass nicht das ganze Potenzial der neuen Speichertechnologie genutzt werden kann. Dateisysteme für flüchtigen Speicher für RAM schaffen keine Konsistenz, welche bei nicht flüchtigen Speicher nötig ist.

Man könnte 3D XPoint Dimms als Hauptspeicher verwenden, leider ist 3D XPoint zu langsam um diesen als alleinigen Hauptspeicher zu verwenden. Dies bedeutet, dass man neben dem flüchtigen Speicher DRAM den nicht flüchtigen 3D XPoint Dimm verwendet.

Nun treten folgende Probleme auf, wenn ich den schnelleren flüchtigen DRAM verwende, wenn ich meinen nicht flüchtigen 3D XPoint Dimm verwende und wie schafft man Konsistenz. Dieses Problem soll das Nova Dateisystem entgegenwirken, es wurde speziell für hybride Dateisysteme entwickelt. Es minimiert unnötigen Overhead, kümmert sich um die Aufgabenverteilung und schafft Konsistenz. Zudem werden Operationen in Pakete zusammengefasst anstatt atomar zu verschicken um den CPU

Bus nicht zu blockieren.

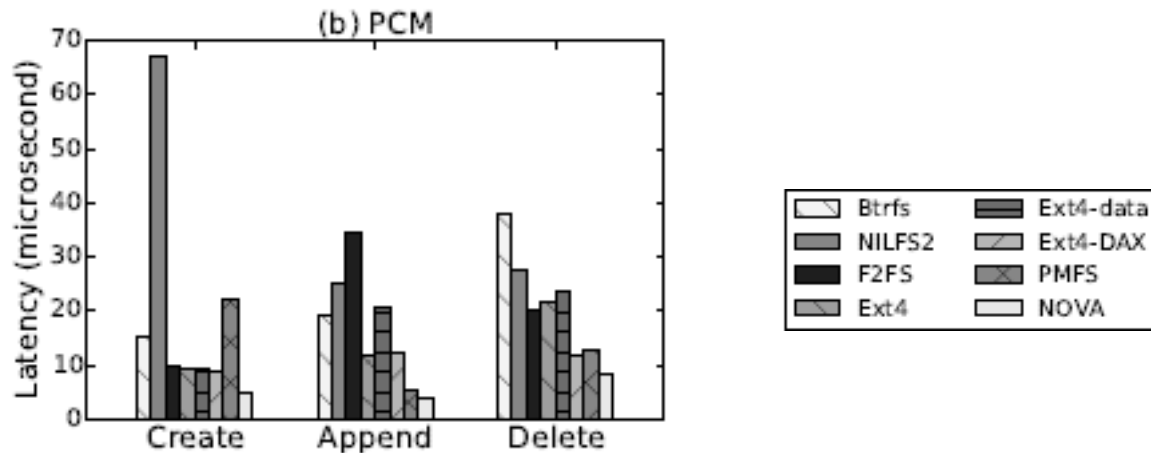


Abbildung: Nova im Vergleich

Die Abbildung „Nova im Vergleich“ zeigt wie stark Nova die Latenz verringern kann. Bei jeder Operation besitzt Nova die geringste Latenz.

5. 3D XPoint im Vergleich

Bei einer Präsentation verglich Intel die Geschwindigkeiten von einer NAND SSD und Optane SSD. Die Optane SSD ist eine SSD mit 3D XPoint Technologie. Die Optane schaffte eine Datenübertragung von 1,6 GByte die Sekunde, wobei die NAND SSD 4 mal so lange brauchte.

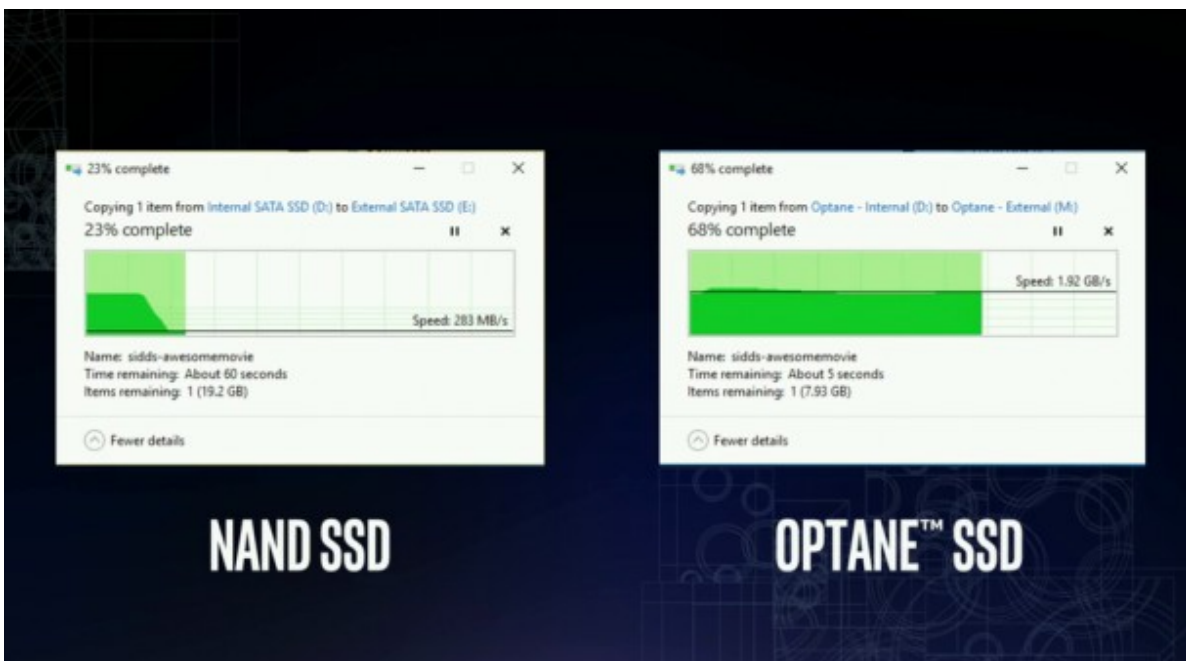


Abbildung: Optane Live Test

Dies ist aber kritisch zu begutachten, denn es wird vermutet das es sich hierbei um eine Mittelklassige TLC SSD handelt, wegen des dramatischen Leistungseinbruch.

Die 3D XPoint Architektur soll wesentlich haltbarer sein als Flash speicher, wobei man nicht weiß ob Intel berücksichtigt hat das die Haltbarkeit bei 3D NAND um einiges verbessert wurde.

Die Adressierung bei 3D XPoint geschieht in Bit und nicht in Blöcken.

Die Latenz soll sich um den Faktor 1000 bei 3D XPoint verbessert im Vergleich zu NAND und ist in der Lage viel mehr Input-/Output Operationen zu bearbeiten

Zudem werden die Chips kleiner, welche bei gleichen Platzverbrauch 10 mal so viel Speicherplatzkapazität anbieten wie DRAM.

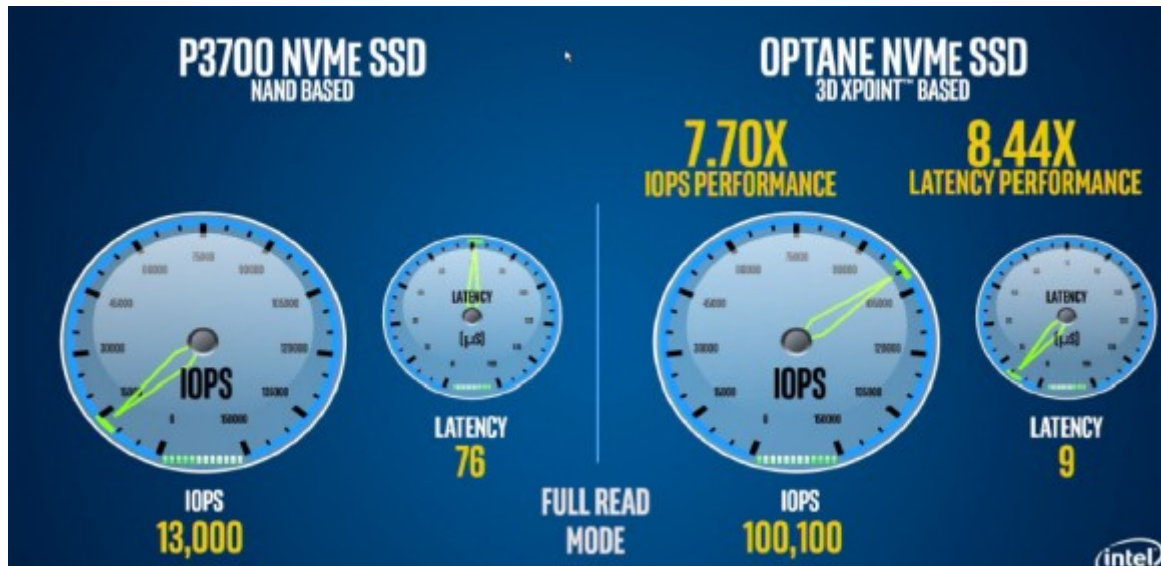


Abbildung: Optane Vergleich

6.Zusammenfassung

Intel's und Micron Technologies's 3D-XPoint spricht für sich, leider kann man nicht nachvollziehen wie viel besser die Optane in Vergleich zu Flash Speicher ist, das es noch keine neutralen Tests gibt. Sicherlich wird die 3D-XPoint Architektur seine Rolle bei großen Servern finden, da dort die Anzahl der Input-/Output Operationen schon sehr wichtig ist.

DRAM und NAND wird sicherlich nicht vom Markt verdrängt werden, da DRAM noch einiges schneller ist und NAND mittlerweile recht preiswert geworden ist und mit 3D NAND auch einen großen Schritt vorwärts gemacht hat.

Literaturverzeichnis

Abbildung Timeline : <https://c.mobilegeeks.de/wp-content/uploads/2015/07/Intel-3DXPOINT-Speicher-Geschichte-Kopie.jpg>

Abbildung Aufbau 3D XPoint : <https://newsroom.intel.com/news-releases/intel-and-micron-produce-breakthrough-memory-technology/>

Abbildung Word-und Bit Line:

http://www.theregister.co.uk/2015/09/23/zeroing_in_on_xpoint_memory/

Abbildung Phase-Change-Memory: https://de.wikipedia.org/wiki/Phase-change_random_access_memory

Abbildung Die: <https://newsroom.intel.com/press-kits/introducing-intel-optane-technology-bringing-3d-xpoint-memory-to-storage-and-memory-products/>

Abbildung NVM-Express Logo: https://de.wikipedia.org/wiki/Phase-change_random_access_memory

Abbildung Schematischer Aufbau NVM-Express: https://www.thomas-krenn.com/de/wiki/NVMe_Grundlagen

Abbildung PCI-Express Slot: https://de.wikipedia.org/wiki/PCI_Express

Abbildung PCI-Express Versionen: https://de.wikipedia.org/wiki/PCI_Express

Abbildung Tabelle NVM-Express im Vergleich: <http://www.giga.de/extra/ssd/specials/nvme-ssd-was-ist-das-und-wie-schnell-ist-sie-einfach-erklaert/>

Abbildung Grafisch NVM-Express im Vergleich: <http://www.giga.de/extra/ssd/specials/nvme-ssd-was-ist-das-und-wie-schnell-ist-sie-einfach-erklaert/>

Abbildung Nova im Vergleich: <http://cseweb.ucsd.edu/~swanson/papers/FAST2016NOVA.pdf>

Abbildung Optane Live Test:

<http://www.hardwareluxx.de/index.php/news/hardware/festplatten/38833-intel-zeigt-benchmarks-der-optane-ssd-mit-3d-xpoint-speichertechnik.html>

Abbildung: Optane Vergleich:

<http://www.hardwareluxx.de/index.php/news/hardware/festplatten/38833-intel-zeigt-benchmarks-der-optane-ssd-mit-3d-xpoint-speichertechnik.html>