

Master-Seminar Hochleistungsrechner: Aktuelle Trends und Entwicklungen

Non-Volatile Memory: Typen und Nutzung für HPC

Markus Loipfinger
Technische Universität München

3. Februar 2016

Zusammenfassung

Eine große Herausforderung von zukünftigen *High-Performance Computing* (HPC) bzw. *Exascale* - Systemen wird das Energiemanagement sein. Da es nötig sein wird, den Verbrauch zu senken und trotzdem genug Rechenleistung aufgebracht werden muss, wird kein Weg an nichtflüchtigen Speichertechnologien (*Non-Volatile Memory*: NVMs) vorbei führen. Diese Speichertechnologien verbrauchen keine Energie, wenn sie sich im Standby-Modus befinden. Deshalb erläutern wir in dieser Arbeit auch die Energieeffizienz der verschiedenen Speichertechnologien. NVMs besitzen allerdings nicht nur Vorteile gegenüber herkömmlichen Speichertypen in derzeitigen HPC-Systemen. Ein Nachteil ist die schlechtere Performanz im Gegensatz zu DRAM (*Dynamic RAM*). Infolgedessen besteht die Lösung der Zukunft womöglich auch aus hybriden Speichertechnologien. Im Laufe der Arbeit werden verschiedene NVMs vorgestellt, miteinander verglichen und deren Nutzung in HPC-Systemen erläutert. Desweiteren werden noch hybride Lösungen erklärt, bevor ein kleiner Ausblick in die Zukunft gegeben wird.

1 Einleitung

Nichtflüchtige Speichertypen sind bezüglich Energieverbrauch sehr effizient, da diese keine Ener-

gie benötigen bzw. verbrauchen, wenn nicht darauf zugegriffen wird. Diesen großen Vorteil gegenüber Speichertechnologien, die in gegenwärtigen HPC-Systemen verwendet werden, wie beispielsweise DRAM, will man sich zu Nutze machen und deshalb soll in zukünftigen Systemen der Fokus mehr auf diese Art von Speichertypen gelegt werden.

Da das weltweit ausgeschriebene Ziel, den Energieverbrauch auf 20 MW zu beschränken, eine drastische Energieeinsparung erfordert, ist eine Umstrukturierung der Speicher notwendig [9]. Jedoch beschränkt jede Möglichkeit den Energieverbrauch zu senken die Kapazität von DRAM. Diese Tatsache zeigt ebenfalls auf, warum in Zukunft auf andere Speichertechnologien gesetzt werden muss.

Phase-change memory (PCM), *spin transfer torque memory* (STTM), *resistive RAM* (ReRAM) und *NAND-Flash*-Speicher sind ein paar der neuen Speichermedien, die in neuen HPC-Systemen verbaut werden könnten. Es gibt auch Ansätze in denen hybride Speicher verwendet werden, d.h. eine dieser Speichertechnologien wird zusammen mit anderen Speichertypen verwendet. So werden die Vorzüge beider Systeme voll ausgenutzt und deren Nachteile gleichzeitig minimiert.

Eine weitere wichtige Frage, die sich Forscher stellen, lautet: Wie integriert man NVMs am besten in die Speicherhierarchie? Auch hier gibt es unterschiedliche Ansätze, von denen ein paar genauer betrachtet werden.

2 DRAM - Hauptbestandteil des Speichers seit fast vier Jahrzehnten

Da in dieser Ausarbeitung Speichertechnologien oft mit DRAM in Beziehung gesetzt werden, folgen hier ein paar Informationen über diesen dynamischen Speichertyp. DRAM findet sich in fast jedem Computer wieder, angefangen von Mobiltelefonen bis hin zu HPC-Systemen. Die Technologie hinter DRAM hat enorme Skalierungsverbesserungen hinsichtlich Kapazität und Performanz erlebt. Zudem war DRAM immer im Überfluss vorhanden, kosteneffizient und zuverlässig. Jedoch sagen Experten voraus, dass in den kommenden fünf bis zehn Jahren keine weiteren Skalierungen mehr möglich sind [13]. Ein großer Nachteil von DRAM liegt darin, dass dessen Speicherzellen periodisch aktualisiert werden müssen. Dieser Vorgang verbraucht Energie, auch wenn keine notwendigen Anwendungsdaten gespeichert werden oder auf DRAM von keinem Prozessorkern aus zugegriffen wird. Für die Zukunft sind drei Lösungen vorstellbar. Zum einen könnte die DRAM-Architektur überarbeitet werden und zum anderen könnten neue Speichertechnologien (z.B. NVMs) DRAM ergänzen oder ganz ersetzen. In Abbildung 1 ist zwar ein statischer RAM dargestellt (keine periodische Aktualisierung notwendig), doch lassen sich diese Eigenschaften auch auf DRAM übertragen. Wie man gut erkennen kann, können die NVMs DRAM bzw. SRAM in Bezug auf Performanz und *Endurance* (Anzahl der Schreibzugriffe, die eine Speichertechnologie verträgt) nicht übertreffen. Dagegen können NVMs ihre Daten auch bei Stromausfall noch speichern und besitzen einen viel geringeren Leckstrom.

3 Typen von NVM

Es werden viele verschiedene neue Speichertechnologien erforscht, die das Potential haben, in naher Zukunft DRAM zu unterstützen oder sogar ganz aus den HPC-Systemen zu verdrängen. Dabei ist NAND-Flash wohl der Speichertyp, der am bekannt-

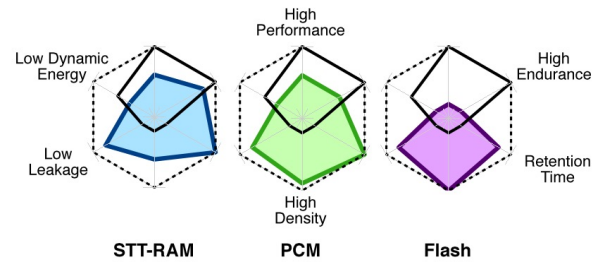


Abbildung 1: Vergleich von SRAM (schwarze durchgezogene Linie) mit STT-RAM, PCM und Flash-Speicher. Die gestrichelte Linie stellt dabei die optimale Speicherausrichtung dar. Die *Retention Time* gibt an, wie lange die Speichertechnologien bei einem Ausfall der Stromversorgung ihre Daten bewahren können. [10].

testen ist. Andere Typen sind PCM, STT-RAM und ReRAM. Diese Speicher benötigen im Ruhezustand keine Energie. Dennoch besitzen diese neuartigen Speichertypen auch Nachteile. Vergleicht man sie mit DRAM, so lässt sich feststellen, dass sie eine höhere Latenzzeit besitzen. Während im Ruhezustand keine Energie benötigt wird, ist diese bei Schreiboperationen hoch. NVMs vertragen außerdem nur eine begrenzte Anzahl an Schreibzugriffen. Diese *Endurance* ist bei NAND-Flash geringer als bei den neueren Technologien. Angesichts dieser Einschränkungen lassen sich NVMs in drei Kategorien aufteilen [8]:

1. NVMs mit langen Zugriffslatenzen für Lese- und Schreiboperationen (z.B. PCM und Flash-Speicher)
2. NVMs mit langen Schreiblatenzen, aber Leselatenzen vergleichbar mit DRAM (z.B. STT-RAM)
3. NVMs, deren Performanz mindestens so gut ist wie die von DRAM (z.B. ReRAM)

Da die Speichertypen der dritten Kategorie alles andere als ausgereift sind, werden wir zwar auch den Vertreter dieser Kategorie kurz ansprechen, aber uns mehr auf die ersten beiden Kategorien

beziehen, zumal diese am ehesten in naher Zukunft in HPC-Systemen eingesetzt werden könnten. Wie bereits erwähnt, besitzen diese Technologien auch Nachteile bzw. Einschränkungen, denen es geschickt auszuweichen gilt. In der ersten Kategorie, in der sich die am meisten verbreitetsten und ausgereiftesten Speichertypen befinden, ist es zwingend erforderlich die Speicherzugriffe zu kontrollieren und im Speziellen die Schreibzugriffe möglichst gering zu halten. Nur so hat deren hohe Latenzzeit keinen so großen Einfluss auf die Systemperformance. Außerdem ist dadurch auch eine annehmbare Lebensdauer, die sich durch viele Schreibzugriffe verringert, sichergestellt. Während in der zweiten Kategorie oft beschriebene Speicherseiten nicht in NVMs untergebracht werden sollten, sollten es oft gelesene Seiten. Es ist allgemein darauf zu achten, dass häufige Schreibzugriffe auf NVMs vermieden werden.

Tabelle 3 zeigt die hier verwendeten Speichertechnologien im Vergleich. Dabei werden die Zellgröße, die Endurance, die Lese- und Schreiblatenzen, sowie der Energieverbrauch im Standby-Modus miteinander verglichen. Die Lebensdauer wird dabei in Abhängigkeit der Schreibzugriffe dargestellt. In der Tabelle lässt sich gut erkennen, dass diese Speichertechnologien um einiges kleiner sind, als beispielsweise der statische RAM (SRAM), der eine *feature size* (Kenngröße) von 125 bis 200 F² besitzen. Dabei bezeichnet F die kleinste Struktur, die zuverlässig hergestellt werden kann.

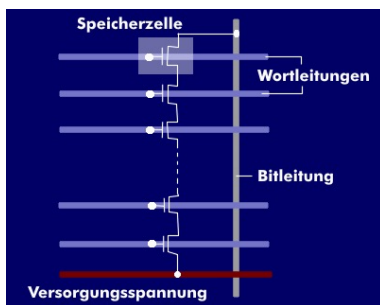


Abbildung 2: Physikalischer Aufbau eines NAND-Flash-Speichers. Das Lesen und Schreiben ist nur in Blöcken möglich [1].

3.1 NAND-Flash

Der wohl bekannteste Vertreter von NVMs ist der NAND-Flash-Speicher, der beispielsweise in mobilen Geräten, Kameras und Smartphones vorhanden ist. Flash-Speicher liegt meist in Form von *solid-state drives* (SSDs) vor, die eine wesentlich bessere Performanz liefern als gewöhnliche *hard disk drives* (HDDs). Außerdem besitzen SSDs im Gegensatz zu rotierenden Geräten wie HDDs keine beweglichen Elemente, die eine mechanische Abnutzung aufweisen können und sind resistent gegenüber Hitze und Erschütterungen. Jedoch besitzen auch SSDs Nachteile gegenüber HDDs, denn die bessere Performanz hängt sehr vom *workload* ab. Es gibt zum Beispiel so gut wie keine Performanzunterschiede, wenn es viele schreibintensive wissenschaftliche Workloads gibt. Hinzu kommt noch, dass SSDs teurer sind als HDDs. Somit lassen sich HDDs nicht komplett durch SSDs ersetzen.

Flash-Speicher basieren auf Halbleiter-Chips, die eine kompakte Größe, einen geringen Energieverbrauch und eine gute Performanz bezüglich zufälligen Datenzugriffen zulassen. Allerdings beeinflussen viele Schreibzugriffe die Speicherzelle negativ, sodass es zu Fehlern und dem Totalausfall kommen kann. Um diese Nachteile aufzuheben, wurden Techniken wie *advanced error correction code* und *wear-leveling* eingeführt. Es wird aber auch auf diesem Gebiet, gerade was das Wear-Leveling betrifft, noch geforscht [9]. Diese Technik beschreibt Verfahren, um die Laufzeit der Speicher zu verlängern.

Da Löschoptionen signifikant langsamer sind als Schreiboperationen, benutzen Flash-SSDs einen sogenannten *flash translation layer* (FTL). Dieser hilft dabei, die Löschoptionen von den oberen Schichten zu 'verbergen' und somit nur Lese- und Schreiboperationen anzubieten. FTL benutzt eine *mapping table*, in dem die virtuellen Adressen von oberen Schichten auf die physischen Adressen des Flash-Speichers abgebildet werden. FTL nutzt diese Abbildung, um Wear-Leveling auszuführen. Desweiteren führt FTL auch *garbage collection* (GC) durch. Diese Funktionalitäten wirken sich kritisch auf die Performanz von Flash-Speichern aus (siehe Abbildung 1) [9].

	Zellgröße	Endurance	Leselatenz	Schreiblatenz	Standby Power
Flash	4 - 6 F ²	10 ⁴ – 10 ⁵	25μs	500μs	0
DRAM	6 - 10 F ²	> 10 ¹⁵	50ns	50ns	Refresh Power
PCM	4 - 12 F ²	10 ⁸ – 10 ⁹	50ns	500ns	0
STTM	6 - 50 F ²	> 10 ¹⁵	10ns	50 ns	0
ReRAM	4 - 10 F ²	10 ¹¹	10ns	50 ns	0

Tabelle 1: Eigenschaften ausgewählter Speichertechnologien [9, 13].

NAND-Flash-Speicherzellen können als *Single Level Cells* (SLCs), als *Multi Level Cells* (MLCs) oder als *Triple Level Cell* (TLCs) vorliegen. SLCs können nur ein Bit speichern, wohingegen MLCs in jeder Speicherzelle zwei Bits speichern können. Letztere erlauben sogar eine Speicherung von drei Bits. Eine SLC besteht aus einem Transistor (MOSFET), der beim Anlegen einer Spannung entweder sperren oder leiten kann. So ist jenachdem, ob Strom fließt oder nicht, das Bit auf 0 oder 1 gesetzt [4]. NAND-Flash-Speicher arbeiten mit sogenannten *Pages*, die aus mindestens 512 Bytes bestehen und zu Blöcken zusammengefasst werden. Jeder Block kann bis zu 256 Pages enthalten. Dabei können Pages nur einmal beschrieben werden und können nur durch vorheriges Löschen eines ganzen Blockes verändert werden. Abbildung 2 zeigt den Aufbau eines solchen NAND-Flash-Speichers. Wegen dieser seriellen Verschaltung ist das Lesen und Schreiben nur in Blöcken möglich. Ein Faktor, der NAND-Flash-Speicher gegenüber DRAM interessant macht, sind die geringeren Kosten pro Bit. Desweiteren werden NAND-Flash-Speicherchips immer kleiner und mit höherer Dichte hergestellt, aber es wird ein Limit der Feature-Size von 15nm vorausgesagt, das Ende des Jahrzehntes erreicht sein wird [8].

3.2 PCM

Der größte Unterschied, der die nachfolgenden NVMs von Flash-Speichern unterscheidet, besteht darin, dass diese Byte-adressierbar sind. Weitere Eigenschaften, die von den folgenden NVMs erhofft werden, sind eine Speicherkapazität und Schreibver-

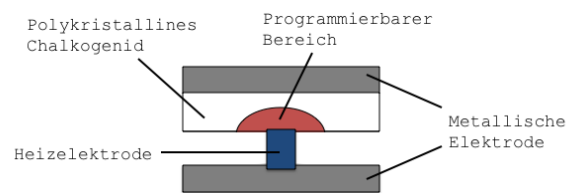


Abbildung 3: Physikalischer Aufbau einer PCM-Zelle. Hierbei werden Daten durch den Kristallisationszustand einer chalkogeniden Glasschicht gespeichert [3].

träglichkeit ähnlich oder besser als Flash-Speicher und Latenzzeiten, die sich mit DRAM vergleichen lassen [5].

PCM speichert Daten im kristallisiertem Zustand auf einer chalkogeniden Glasschicht. Dabei wird das Glas schnell erhitzt, um zwischen unkristallisiertem (hoher Widerstand, binär 0) und kristallisiertem Zustand (niedriger Widerstand, binär 1) zu wechseln. Dieses Aufhitzen erfolgt durch einen kurzen Stromimpuls mittels der Heizelektrode. Abbildung 3 zeigt den physikalischen Aufbau von PCM. Durch die immer kleiner werdende Feature-Size, wird PCM schneller und energieeffizienter. Dies hat auch Auswirkungen auf die Verfügbarkeit. Im Vergleich zu Flash-Speichern, verträgt PCM mehr Schreibzyklen (siehe Abbildung 1). Um den Nachteilen, wie der Verfügbarkeit und der immer noch geringen Anzahl von Schreibzugriffen entgegenzuwirken, benötigt es gute Mechanismen für Wear-Leveling [5]. Da PCM einer der wahrscheinlichsten Vertreter der neuen Speichertechnologien ist, die NAND-Flash-Speicher verdrängen können, finden

sich im Folgenden noch ein paar genauere Informationen. Dieser Speichertyp lässt auch Zustände zwischen kristallisiertem und unkristallisiertem Zustand zu. So ist es möglich, mehr Bits zu speichern. Die Zellen von PCM können mindestens 10^6 mal neu programmiert werden. Preislich reiht sich PCM zwischen DRAM und Flash-Speicher ein [7].

PCM besitzt außerdem eine sehr hohe Kapazität, jedoch zeigt dieser auch asymmetrisches Performanzverhalten in der Hinsicht auf, dass PCM eine höhere Latenzzeit und einen hohen Energieverbrauch hat, wenn Schreiboperationen ausgeführt werden [11].

3.3 STTM

Ein ebenfalls schon erhältlicher Speichertyp ist STTM, dessen Lesegeschwindigkeit der von DRAM gleich kommt. STTM ist folgendermaßen aufgebaut (siehe Abbildung 4). Dieser Speicher verwendet magnetische Tunnelverbindungen (MTJ), um Daten bzw. Bits zu speichern. Jede MTJ besteht aus zwei ferromagnetischen Schichten. Dabei hat eine Schicht eine feste magnetische Richtung (feste Schicht) und die andere Schicht kann ihre Magnetisierungsrichtung frei ändern (freie Schicht). Durch die relative Richtung der beiden Schichten wird eine digitale 0 oder 1 dargestellt. Das Anlegen einer festgelegten kleinen Spannung an die MTJ ergibt eine hohe oder niedrige Stromstärke. Dies ist abhängig von der Ausrichtung der freien Schicht zur festen Schicht. In Abbildung 4 a) sind beide Magnetisierungsrichtungen parallel, was einen niedrigen Widerstand zur Folge hat. In b) sind die Magnetisierungsrichtungen entgegengesetzt gerichtet, was dementsprechend einen hohen Widerstand ergibt.

In dieser Hinsicht ist STTM früheren magnetischen RAM Technologien ähnlich, aber anders als die früheren Technologien, nutzt STTM einen Fluss von polarisierten Elektronen für das Ummagnetisieren [5].

Bei einem Schreibzugriff muss die Schreibstromstärke eine ausreichend lange Zeit aufrecht erhalten werden, um sicher zu gehen, dass die freie Schicht ihren Zustand geändert

hat [10]. Obwohl auch bei diesem Speichertyp ein asymmetrisches Performanzverhalten mit Schreiblatenzen länger als die von DRAM festzustellen ist, besitzt STTM eine gute Performanz sowie niedrigere Energieanforderungen als die meisten NVMs und eine lange Lebensdauer [11].

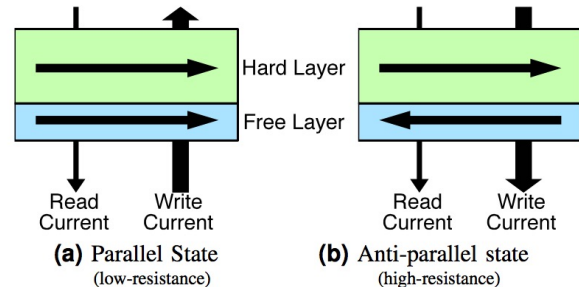


Abbildung 4: Physikalischer Aufbau einer STTM-Zelle. Die Datenspeicherung erfolgt durch die Magnetisierungsrichtung der freien Schicht zur festen Schicht. a) zeigt die parallele und b) die entgegengesetzte Ausrichtung [10].

3.4 ReRAM

ReRAM-Speicherzellen bestehen, wie in Abbildung 5, aus einer Silber- und einer Platinelektrode. Die Ionen können sich an beiden Elektroden lösen und wieder niederschlagen. Somit lässt sich der elektrische Widerstand der Speicherzelle ändern, was für die Datenspeicherung ausgenutzt werden kann. Durch Ablagerungen des Silbers entstehen die sogenannten Silberfilamente. Die Größe dieser Speicherzellen liegt im Nanometerbereich (10nm). Da sich zwischen den beiden Elektroden eine Spannung aufbaut, könnte ReRAM neben dem Datenspeicher auch als sehr kleine Batterie dienen. Dies würde den Stromverbrauch weiter verringern [2].

ReRAM besitzt Zugriffszeiten, die mit denen von DRAM vergleichbar sind und hat im Gegensatz zu NAND-Flash eine Skalierbarkeit, die potentiell besser ist, als das vorhergesagte Limit des Flash-Speichers (siehe 3.1). Doch der Erfolg der ReRAM-Technologie hängt auch ziemlich vom Ko-

stenfaktor ab. Denn es benötigt eine Integrierung in eine konventionelle Basistechnologie [6].

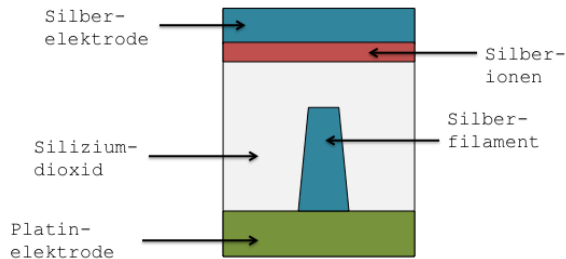


Abbildung 5: Physikalischer Aufbau einer ReRAM-Zelle. Zur Speicherung der Daten werden die Ionen und somit der elektrische Widerstand verwendet [2].

4 Zusammenfassung der Eigenschaften der NVMs

4.1 Allgemeines

Diese Vertreter von nichtflüchtigen Speichertypen haben die Möglichkeit sowohl DRAM als flüchtigen Speicher in Zukunft abzulösen, denn im Gegensatz zu DRAM verbrauchen PCM und STTM weniger Energie, da zum Beispiel deren Speicherzellen nicht aktualisiert werden müssen. Desweiteren benötigen alle NVMs keine Energie, wenn sie sich im Ruhezustand befinden. Eine Einschränkung, der es allerdings entgegenzuwirken gilt, ist die geringe Lebensdauer bei vielen Schreibzugriffen. Dieses Problem kann unter anderem mit einem guten Wear-Leveling gelöst werden.

4.2 Energieeffizienz

Der Energieverbrauch der NVMs im Vergleich zu DRAM für das in [12] verwendete System ist in Abbildung 6 dargestellt. Hier lässt sich erkennen, dass PCM verglichen zu ReRAM und STTM weniger effizient in Hinblick auf den Energieverbrauch ist.

Ab einer Speichergröße von 220 GB verbrauchen alle NVMs weniger Energie als die entsprechende Menge an DRAM. Im Falle von STTM sind dafür nur 65,4 GB nötig [12]. Allerdings ist der Speicherbedarf bei Hochleistungsrechnern hoch genug, dass alle hier aufgeführten neuen Speichertechnologien DRAM in Bezug auf die Energieeffizienz überragen.

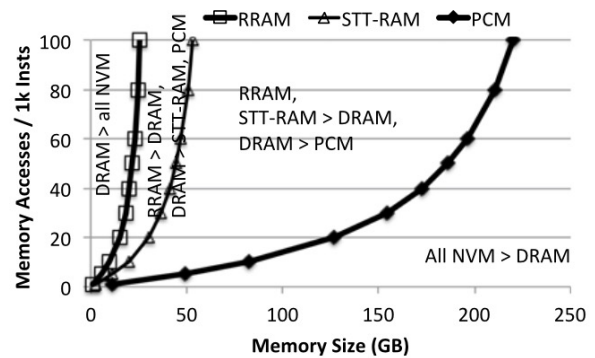


Abbildung 6: Energieverbrauch der einzelnen NVMs in Abhängigkeit des Workloads und der Speicherzugriffe. Dabei bezeichnet $x > y$ einen Bereich, in dem die Speichertechnologie x energieeffizienter als y ist [12].

Um nicht nur den Energieverbrauch zu berücksichtigen, sondern auch die jeweilige Performanz der Speichertechnologien mit einzubinden, wird häufig das *Energy delay product* (EDP) verwendet. Dies multipliziert die verbrauchte Energie (statische + dynamische Energie) mit der für die Ausführung benötigten Zeit. Das bedeutet, dass zwei Technologien hinsichtlich des EDP äquivalent sind, wenn eine schneller ist, aber proportional mehr Energie verbraucht. Somit ergeben sich für die vorher aufgeführten Speichertechnologien auch höhere Werte. Dies ist zu erwarten, da NVMs eine höhere dynamische Energie und, ausgenommen der dritten Kategorie, zumindest höhere Schreiblatenzen besitzen. Dieser Rückstand gegenüber DRAM wird durch deren nicht vorhandenen statischen Energie wieder ausgeglichen. STTM und ReRAM übertreffen DRAM somit bei Speichergrößen von 130,9 bzw. 75,1 GB, wohingegen PCM 1,34 TB benötigt [12].

5 Nutzung von NVMs in HPC-Systemen

Die Integration der NVM in die Hierarchie des Speichers sollte nach den Vor- und Nachteilen des jeweiligen NVMs erfolgen. Wegen der sich verringern Lebensdauer durch viele Schreibzugriffe, sollte eine NVM in keinen Bereich mit vielen Schreiboperationen platziert werden. Die optimale Integrationsstrategie steht allerdings noch nicht fest, sie wird jedoch bereits viel diskutiert.

In Abbildung 7 sind drei verschiedene Möglichkeiten dargestellt, PCM in die Speicherhierarchie einzubinden. Dabei zeigt a) die Integration von PCM als Hauptspeicherersatz für DRAM.

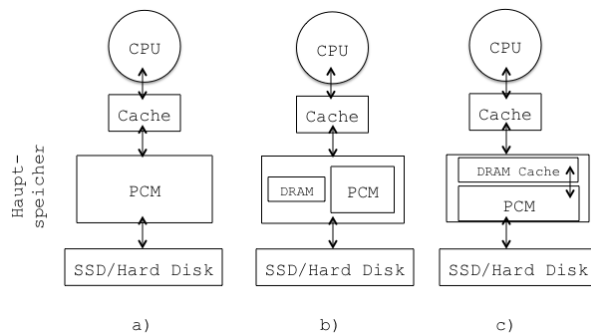


Abbildung 7: Integration von PCM in die Speicherhierarchie auf drei verschiedene Arten. a) zeigt dabei PCM als Ersatz für DRAM im Hauptspeicher, bei b) und c) wird DRAM durch PCM ergänzt.

5.1 NVMs als SSD

Eine Methode, NVMs in HPC-Systemen zu integrieren, ist als Solid-State-Drive (SSD). SSDs sind überall erhältlich und sehr kosteneffizient (2,5 Gbyte/\$). In HPC-Systemen zeigten SSDs bereits eine erhebliche Steigerung der I/O-Performanz, vorallem bei vielen kleinen Ein- und Ausgabeoperationen [13].

5.2 NVMs als globales Speichermedium

Ein anderer Ansatz sieht NVMs als globales Speichermedium vor. In diesem Fall werden NVM-Server gebildet, die von Knoten des Systems oder des angebundenen Netzwerks eingebunden, skaliert oder global adressiert werden können. So könnte außerdem ein *Remote*-Zugriff via verschiedenen Schnittstellen auf diese Server möglich sein. Für den Fall, dass Datensets in unstrukturierter Weise vorliegen und Zugriffsmuster unvorhersehbar sind, ist diese Art der NVM-Nutzung sehr effektiv.

5.3 NVMs als hybrider Cache

Eine weitere Möglichkeit wäre ein hybrider Cache. Dabei wird der NVM mit einem statischem RAM (SRAM) oder einem eingebetteten DRAM kombiniert. Dadurch könnte der große Energie- und Platzbedarf von SRAM gesenkt werden. Jedoch gibt es auch hier nicht zu vernachlässigende Nachteile. Deshalb muss das Speichersystem gute Algorithmen zur Speicherverwaltung besitzen, um nicht zu viele Schreibzugriffe auf den NVM zuzulassen (z.B. Wear-Leveling), da dies die Performanz senken und die Lebensdauer reduzieren würde. Außerdem kommt es dann zu hohen Fehlerraten. Obwohl bereits solche Algorithmen entworfen werden, bleiben noch ein paar Herstellungs- und Entwicklungshürden bestehen [13].

5.4 NVMs als Burst Buffer

Des Weiteren können NVMs auch als sogenannte *burst buffer* integriert werden. Diese helfen dabei, die I/O-Performanz zu verbessern. Dadurch können Anwendungen *Checkpoints* oder andere Ausgabedaten auf die NVMs in die Knoten des HPC-Systems auslagern und so die weitere Berechnung schneller fortsetzen, da sie sich nicht explizit um die weitere Speicherung der Daten kümmern müssen [13].

6 Hybride Speichertechnologien

Eine andere Möglichkeit, NVMs in Zukunft wirksam einzusetzen, ist innerhalb hybrider Speichersysteme mit DRAM oder anderen Speichertechnologien zusammen. Dabei wird der NVM-Typ direkt in die Speicherhierarchie eingebunden. So werden beide Speicher beispielsweise nebeneinander abgelegt. Da es bei einer solchen Zusammenstellung von verschiedenen Speichertypen darauf ankommt, wo welche Daten innerhalb des hybriden Systems platziert werden, bedient man sich beispielsweise an folgenden Metriken [8].

- Verhältnis von Lese- und Schreiboperationen: Eine hohe Zahl resultiert dabei in einer geringeren Schreibbelastung. Dies wird von NVMs favorisiert.
- Speichergröße eines Speicherobjektes: Da die statische Energieeinsparung direkt mit der Speichergröße in Verbindung steht, sollten so viele Anwendungsdaten wie möglich in NVMs gespeichert werden.
- Speicherreferenz-Rate

Mit Hilfe dieser Metriken kann vermieden werden, dass schreibintensive Daten in NVMs abgelegt werden, da dies zu einem Abfall der Performanz und einer geringeren Lebenszeit von den NVMs führt.

6.1 NVM+DRAM

Eine Zusammenstellung von einer oder mehreren NVMs mit DRAM kann zu einer größeren Speicherkapazität bei gleichem Energieverbrauch führen. Jedoch muss auf der anderen Seite wegen der geringen Endurance auch gewährleistet sein, dass dieser Speicher sehr gut verwaltet wird, um die Schreibzugriffe zu kontrollieren.

Da die zwei Hauptkomponenten der Speicherhierarchie heutzutage Caches und Hauptspeicher sind, liegt es nahe, dass NVMs in diese Bereiche entweder horizontal oder vertikal integriert werden. Letzteres würde eine Ebene der Speicherhierarchie entfernen oder eine neue hinzufügen (z.B. ein großer

DRAM Cache im Hintergrund), wohingegen eine horizontale Integration beider Speicherformen eine einheitliche Schnittstelle zur Verfügung stellen könnte, sodass die Details der Speicheradressierung ausgeblendet werden könnten. Dabei müssten Teile des Systems (z.B. Prozessor, *Memory Controller*) davon in Kenntnis gesetzt werden, dass sich zwei verschiedene Speichertypen im System befinden, um die Speicheradressierung effizient verwalten zu können [13].

In Abbildung 7 zeigen b) und c) die Integration von PCM zusammen mit DRAM. In c) wird im Gegensatz zu b) DRAM als zusätzlicher Cache verwendet und in beiden Fällen werden oft beschriebene Daten in DRAM gespeichert.

Ein Beispiel hierfür ist der von IBM im Jahr 2014 vorgestellte hybride Speicher, der bis zu 275 mal schneller ist als SSD. Dieser besteht aus DRAM, herkömmlichen NAND-Flash-Speicher und PCRAM [7].

6.2 Flash+HDD

Eine Möglichkeit, um die Vorzüge von HDDs und SSDs auszunutzen, ist die Verbindung der kosteneffizienten HDDs mit der hohen Leistung von SSDs. Dabei werden die performanzkritischen Blöcke mittels Workload-Zugriffsmustern erkannt und nur diese in die SSD verschoben. Um desweiteren die Performanz bei schreibintensiven Workloads zu verbessern, werden einkommende Schreiboperationen in SSDs mit niedrigen Latenzzeiten zwischengespeichert [13].

6.3 PCM+Flash+HDD

Ein hybrider Speichertechnologieansatz, der aus drei Speichern besteht, wird im Folgenden beschrieben. Bei der Integration von PCM in die Speicherhierarchie mit Flash und HDD wurde festgestellt, dass zwar auf Materialebene die Schreiboperation auf Flash langsamer als auf PCM ist, aber auf Systemebene die Schreiboperationen auf eine Flash-basierende SSD schneller sein können als auf eine PCM-basierende SSD. Dies hängt von verschiedenen Faktoren wie beispielsweise Energieein-

schränkungen ab. Infolgedessen wurden mit *Tiering* und *Caching* zwei Speichernutzungsmöglichkeiten getestet. Erstere bezeichnet eine Methode, bei der Daten entsprechend der Anzahl ihrer Zugriffe auf unterschiedliche Speichermedien abgelegt werden. Bei Tiering stellte sich heraus, dass eine Verbindung von Flash und HDD mit PCM eine Verbesserung im Bereich Performanz pro \$ nach sich zieht. In diesem Fall ist die Verbindung ohne PCM schlechter. Für Caching stellte sich heraus, dass eine Kombination von Flash und PCM-SSDs eine bessere Leselatenzzeit und eine insgesamt bessere Ein-/Ausgabezeit liefert, als eine Konfiguration allein mit Flash [13].

7 Schlussfolgerung und Fazit

Abschließend lässt sich sagen, dass NVMs in zukünftigen HPC-Systemen eine wichtige Rolle spielen werden, da vor allem der Energieverbrauch verringert werden soll. Dafür sind NVMs bestens geeignet, da diese im Ruhezustand im Gegensatz zu DRAM keine Energie verbrauchen. Deshalb können diese auch Daten bewahren, auch wenn die Stromversorgung unterbrochen ist. Bei der Nutzung von nichtflüchtigen Speichertypen muss darauf geachtet werden, dass zu viele Schreiboperationen vermieden werden. Dies würde die allgemeine Performanz des Systems verringern und den Energieverbrauch durch die höhere dynamische Energie bei Schreibzugriffen steigern.

Eine vorerst denkbare Integration in die Speicherhierarchie wäre als hybrider Speicher. So kann man die Vorteile der verwendeten Speicher ausnutzen und deren Nachteile gegenseitig kompensieren. Dabei muss darauf geachtet werden, dass schreibintensive Workloads nicht auf die NVMs ausgelagert werden. Deshalb wurden in dieser Ausarbeitung auch Metriken vorgestellt, um Workloads dementsprechend zu klassifizieren.

Es wird sich zeigen, auf welche NVMs zukünftig in HPC-Systemen gesetzt wird. Alle hier vorgestellten Speichertechnologien haben ihre Vor- und Nachteile. Wie aber in dieser Ausarbeitung gezeigt wurde, wird es in Zukunft nötig sein, vermehrt auf nichtflüchtige Speicher in Hochleistungsrechnern zurück

zu greifen.

Literatur

- [1] *NAND-Flash*. <http://www.itwissen.info/definition/lexikon/NAND-Flash-NAND-flash.html>. – Zuletzt abgerufen am 21.01.2016
- [2] *Nanoelektronik: Bewegliche Teilchen im Festkörper*. <http://www.fz-juelich.de/SharedDocs/Pressemitteilungen/UK/DE/2014/14-06-26-teilchenbewegungen.html>. – Zuletzt abgerufen am 21.01.2016
- [3] *PRAM (phase change RAM)*. <http://www.itwissen.info/definition/lexikon/phase-change-RAM-PRAM.html>. – Zuletzt abgerufen am 21.01.2016
- [4] BENZ, B. : *Die Technik der Flash-Speicherkarten*. <http://www.heise.de/ct/artikel/Erinnerungskarten-290738.html>. – Zuletzt abgerufen am 21.01.2016
- [5] CAULFIELD, A. M. ; COBURN, J. ; MOLL-
OV, T. I. ; DE, A. ; AKEL, A. ; HE, J. ;
JAGATHEESAN, A. ; GUPTA, R. K. ; SNAVE-
LY, A. ; SWANSON, S. : Understanding the
Impact of Emerging Non-Volatile Memories
on High-Performance, IO-Intensive Comput-
ing. In: *Proceedings of the 2010 ACM/IEEE
International Conference for High Performan-
ce Computing, Networking, Storage and Ana-
lysis* (2010), S. 1 – 11
- [6] GOVOREANU, B. ; KAR, G. ; CHEN, Y. ; PA-
RASCHIV, V. ; KUBICEK, S. ; FANTINI, A. ;
RADU, I. ; GOUX, L. ; CLIMA, S. ; DEGRAE-
VE, R. ; JOSSART, N. ; RICHARD, O. ; VAN-
DEWEYER, T. ; SEO, K. ; HENDRICKX, P. ;
POURTOIS, G. ; BENDER, H. ; ALTIMIME, L. ;
WOUTERS, D. ; KITTL, J. ; JURCZAK, M. :
10x10nm² Hf/HfOx crossbar resistive RAM
with excellent performance, reliability and low-
energy operation. In: *Electron Devices Meeting
(IEDM), 2011 IEEE International*, 2011. – IS-
SN 0163–1918, S. 31.6.1–31.6.4

- [7] HRUSKA, J. : *IBM demonstrates next-gen phase-change memory that's up to 275 times faster than your SSD.* <http://www.extremetech.com/extreme/182096-ibm-demonstrates-next-gen-phase-change-memory-thats-up-to-275-times-faster-than-your-ssd>. – Zuletzt abgerufen am 21.01.2016
- [8] LI, D. ; VETTER, J. S. ; MARIN, G. ; MCCURDY, C. ; CIRA, C. ; LIU, Z. ; YU, W. : Identifying Opportunities for Byte-Addressable Non-Volatile Memory in Extreme-Scale Scientific Applications. In: *IEEE 26th International Parallel and Distributed Processing Symposium* (2012)
- [9] MITTAL, S. ; VETTER, J. : A Survey of Software Techniques for Using Non-Volatile Memories for Storage and Main Memory Systems. In: *Parallel and Distributed Systems, IEEE Transactions on PP* (2015), Nr. 99, S. 1–1. – ISSN 1045–9219
- [10] SMULLEN, C. ; MOHAN, V. ; NIGAM, A. ; GURUMURTHI, S. ; STAN, M. : Relaxing non-volatility for fast and energy-efficient STT-RAM caches. In: *High Performance Computer Architecture (HPCA), 2011 IEEE 17th International Symposium on*, 2011. – ISSN 1530–0897, S. 50–61
- [11] SURESH, A. ; CICOTTI, P. ; CARRINGTON, L. : Evaluation of emerging memory technologies for HPC, data intensive applications. In: *Cluster Computing (CLUSTER), 2014 IEEE International Conference on*, 2014, S. 239–247
- [12] VANDIERENDONCK, H. ; HASSAN, A. ; NIKOLOPOULOS, D. : On the Energy-Efficiency of Byte-Addressable Non-Volatile Memory. In: *Computer Architecture Letters* 14 (2015), July, Nr. 2, S. 144–147. – ISSN 1556–6056
- [13] VETTER, J. S. ; MITTAL, S. : Opportunities for Nonvolatile Memory Systems in Extreme-Scale High Performance Computing. In: *Computing in Science and Engineering* special issue 17 (2015), 01/2015, 73-82. <http://dx.doi.org/10.1109/MCSE.2015.4>. – DOI 10.1109/MCSE.2015.4. – ISSN 1521–9615