NANTERO

NANTERO, INC.

(Industry Sector: Microelectronics – Memory Chips)

25-B Olympia Avenue, Woburn, MA 01801 USA https://nantero.com
CAGE Code: 36GRO DUNS Number: 101731128 SBC

POC: Robert Snowberger, CEO
(202) 263-9143, rsnowberger@nantero.com

EXECUTIVE SUMMARY

Nantero, Inc.'s NRAM® memory technology is a solution to the non-volatile, byte-addressable fabricattached memory tier to address the need for a better Optane™ replacement. NRAM is a disruptive replacement for DRAM and NAND Flash that can beat the cost, have much better power, latency, and performance characteristics, deliver EMP protections and RadHard capabilities, and provide futureproofing against the system architecture and platform enhancements of 2025-2030 and beyond. NRAM uses less energy than legacy memory technologies, and thus lowers carbon emissions, while also supporting the computing capability changes of the future such as CXL, In-Memory Processing, Disaggregation, Edge Computing, etc. Most power in a memory system today goes to refresh; which NRAM can eliminate with a DDR5-compatible part, providing an immediate huge win to DoE, the USG as a whole and all of industry. Nantero needs government support dollars for technology innovation and access to the new EUV fabs built with government money for researchers and small companies to encourage innovations like Nantero's. With this vital support, the fab access and readiness gap can be bridged between what innovators like Nantero need today and limited reality of what the large established companies regularly provide to innovators without government participation and oversight. Once this fab access and support is provided, Nantero's NRAM memory technology will deliver its cost and performance benefits to disrupt DRAM and NAND Flash, providing the extensive capabilities DoE, the USG and industry needs now and in the coming years.

The costs of computing capabilities continue to rise, both in terms of the energy required to run the systems themselves and to run the facilities in which their housed, as well as costs to the environment via the carbon emissions the ever-increasing computing demands are dictating. Today's datacenters consume over 3% of the entire global power supply and account for over 2% of total carbon emissions. By 2030, the number of datacenters is expected to grow by 10x, demanding 10% of all future anticipated global power supply and producing over 11% of the world's total carbon emissions. Datacenter energy usage and carbon emissions are doubling every four years as data creation expands exponentially from the spread of internet usage to the 50% of the global population not currently using IT systems. At the same time, the current internet users are increasingly embracing 5G, new IoT devices and thriving cryptocurrency ecosystems. Consumption and performance demands are clearly not going to diminish, and so the world must find a way to provide more efficient datacenters for the future.

Energy costs make up nearly 80% of the operational expenses of a datacenter. The largest contributors to the energy expense in a datacenter are the cooling systems, followed by logic chips and then memory chips. While cooling system efficiency work is underway, logic chip power efficiencies are largely ignored in deference for the continued market pressures to provide ever more powerful logic chips; which has the net result of increasing the overall power demands of the datacenter and offsetting much of the cooling system improvements being formulated. Legacy DRAM memory chips have similar performance

Copyright © 2022 Nantero, Inc. All Rights Reserved.

improvement demands as their logic chip counterparts, but with the critical distinction being that a memory chip's performance improvement is fundamentally different from a logic chip's in that memory performance increases do not have to only come from adding more memory chips into a system. Memory performance can be enhanced in other ways that do not result in greater energy usage. One such memory solution is Nantero's NRAM® carbon nanotube-based memory, which delivers performance exceeding DRAM while also delivering 32% energy savings. A single 300 MWH datacenter using NRAM can conservatively expect to save 59 MWH of power per year and reduce its annual carbon emissions by 4.1 metric tons. It is critical to embrace a memory technology like NRAM that delivers energy savings while not sacrificing performance, form factor ubiquity or the ability to future-proof against the ever increasing needs of the global microelectronics ecosystem.

While addressing the costs of energy consumption and the subsequent carbon emissions, computing capabilities of the future must also continue their march toward ever-great performance capabilities. As Artificial Intelligence ("AI") and Machine Learning ("ML") systems expand across the enterprise, mobile, edge, critical infrastructure, IoT, SCADA/IIoT, big data, graph analytics, modeling & simulation, exascale supercomputing, and upcoming quantum architectures, so too must the computing environments and the foundational components themselves mature to deliver the capabilities AI/ML and other systems need to perform. To fully realize the capabilities of new logic processors, Memory Chips must become more energy efficient, increase their capacity and increase their speed in magnitudes of order, not merely incrementally. Additionally, Memory Chips must decrease their latency, provide nonvolatility, and be adaptable to critical use cases requiring Rad-Hardness, EMP protection, environmental tolerances, and resiliency. Memory Chips must also provide future-proofing against continually changing computing capabilities as system architectures and computing platforms embrace next-gen capabilities like CXL, InMemory Processing, Disaggregation, Edge Computing, etc. Memory Chips must also be built to higher security and operational standards with designed-in capabilities that contribute to a system's overall security posture and performance capabilities. Memory vulnerabilities such as rowhammer, and quickrecovery capabilities like check-point and nonvolatile functionality, must become inherent qualities of the US's future Memory Chips.

Current legacy memory technology consists primarily of DRAM and NAND Flash. Both technologies have been around for decades and are now quickly nearing their end of lives where they cannot be innovated fast or far enough to deliver the critical performance requirements needed by future computing systems. Legacy memory technologies are nearing their absolute limit, and they can no longer keep pace with either the needs of today, nor the projected needs beyond 2025.

Nantero, Inc. is a US domestic Memory Chip company with over 20 years of development and hundreds of memory patents to its name that offers the NRAM® CNT-based Nonvolatile Memory which meets and exceeds all of the new critical performance requirements stated above. Despite its US roots, including being founded and funded with support by the US Government, having delivered critical tech in support of the US Government, and having attempted numerous times to get larger support from the US Government to address its own Valley of Death dilemma, Nantero was ultimately forced to find a foreign partner to support its early stage innovation and low-level fabrication needs. This partner was Fujitsu, a Japanese company that helped to facilitate both the development of Nantero's NRAM Memory Chips, and also supported Nantero's technology becoming one of the pillars of a ¥88.5 Billion (~$730M USD) fund to establish "Next Generation Green Datacenter Technology Development" for Japanese technology companies. The Japanese national program kicked off in 2022 and has a primary focus of integrating NRAM across a sundry of products and markets via a consortium of Japanese companies. Nantero wants to facilitate U.S. adoption in support of agencies with critical missions such as the DoE, the same as Nantero is doing in Japan working with its Japanese industry and government partners.

Copyright © 2022 Nantero, Inc. All Rights Reserved.

The Nantero Team includes members, including foundry and design partners, to address all work efforts required to supply Nantero's latest NRAM carbon nanotube memory chip design to replace DRAM in exascale and green datacenters. Anchored by DoE exascale computing system providers HPE and AMD, and backed by foundry partners such as GlobalFoundries, the Nantero Team has the unique capabilities to meet and exceed the performance needs of the DoE. Letters of Commitment for all team members are available upon request.

NRAM® MEMORY TECHNOLOGY

Nantero is the world leader in Carbon Nanotube (CNT)-based semiconductor technology, with current locations in Woburn Massachusetts, Sunnyvale California, and Wilmington Delaware. We are the inventors, owners, designers, and suppliers of the only American-based emerging memory technology, NRAM®, which is our patented memory technology based on semiconducting CNT technology. We are also experts in CNT chemical integration into foundry and fabrication processes and are currently adapting our methodology and best-practices in this field to build CNT-based transistors and logic devices.

The current market-leading semiconductor memory technology, DRAM (Dynamic Random Access Memory), is widely understood to be unable to keep up with Moore's law and is nearing the end of its useful life, and Nantero's NRAM® (Non-Volatile Random Access Memory) is a highly recognized emerging technology that is expected to disrupt and capture the memory market as the demand for better and faster memory technologies continues to grow. Nantero's NRAM® enables the following advantages over DRAM:

- <u>Green Memory</u> – 32% more energy efficient operation for data farms using NRAM®-based servers (Green Technology).
- <u>Performance</u> – Same performance as DRAM, with ability to achieve capacities far beyond wellunderstood DRAM limitations (<5nm).
- <u>Cost</u> – Current 32Gb NRAM project of record has lower cost than DRAM, with roadmap to be substantially cheaper than DRAM at even higher densities beyond the DRAM roadmap. With a dedicated fab, NRAM® cost can be reduced further.
- <u>Measurably Secure Supply Chain</u> – Nantero's NRAM materials and fab facilities will enable the U.S. to create an alternative electronics ecosystem, starting from raw material and including fab, test, assembly/packaging and system build including software all in one physical US facility, to counter-balance current silicon-based electronics supply chain dominated by foreign players.
- <u>Highly Secure</u> – Enhanced cybersecurity features including "Rowhammer" invincibility and the enablement of secured main memory.
- <u>Rad-Hard</u> – Space-tested and proven Radiation Hardening (RAD hard) inherent properties of CNTs enable memory that is secure from EMP attacks.
- <u>Extreme Environmental Tolerances</u> – Extreme heat/cold range tolerance and does not corrode in salt water or acids.
- <u>Improves Artificial Intelligence and Machine Learning</u> – Data persistence avoids loss of learned data on power failure.
- <u>Instant-On</u> – Requires no reboot time.
- <u>Elimination of SSDs and Hard Drives</u> – Operates as both active/high-performance main memory and non-volatile memory cache.

Copyright © 2022 Nantero, Inc. All Rights Reserved.

Nantero's roadmap includes additional products, beyond memory, in the areas of analog/mixed signal/RF and digital logic using CNT-FETs fabricated with Nantero's 300mm fab proven CNT thin film manufacturing technologies. In addition, pathfinding activities for CNT use in quantum computation are in progress with the goal to increase operational temperature and decoherence times compared to other approaches.

Nantero is an established small business concern with over $300m invested through USG/Commercial partner cost-sharing. Nantero has had technology installed in 12 fabs around the world, ranging from research/university-level foundries (~150mm wafers), all the way to high-quality, industrial-scale commercial fabs run by industry giants (300mm wafers). Nantero has 195 granted US patents and an additional 95 internationally granted patents.

Key takeaways for supporting Nantero's NRAM CNT-based Nonvolatile Memory Chip technology include:

- Nantero is a U.S. Domestic company dedicated to support the critical need for a Trusted and Secure Domestic Supply Chain to create an alternative electronics ecosystem that counterbalances the current foreign-dominated microelectronics supply chain.
- As DRAM nears obsolescence, Nantero's NRAM® is the only U.S.-based emerging memory technology that can compete with DRAM performance, while also delivering energy savings and cost parity at lower volumes – 32Gb NRAM cost will be 50% of current DRAM.
- Nantero is duplicating its success in Japan where a $100 million national program was established to specifically design NRAM memory into a wide array of systems, including exascale computers similar to those of the DoE.
- NRAM's energy and cost savings support Green Datacenter goals and the needs of all Critical Infrastructures – NRAM will reduce total datacenter energy costs by 32% (equating to 40 metric tons reduction in carbon emissions).
- NRAM has RadHard capabilities (with proven space heritage) that can help secure the nation's electrical grid and other critical infrastructures while also providing resiliency.
- NRAM has Extreme Environmental Tolerances, such as wide heat/cold range tolerances and no corrosion in salt water or acids, that together greatly extend the footprint of addressable form factors, use cases and markets which NRAM can uniquely support as a true Universal Memory solution.
- CNT chemical solutions technology behind NRAM is becoming established within the United States, which is completely independent of and not reliant upon the supply chain and ecosystem that supports DRAM and Flash.
- Nantero is teamed with DoE exascale incumbents HPE and AMD to support their current offerings and future roadmaps by employing NRAM in both their DoE and other USG/Commercial systems (commitment letters attached).

The following sections summarize the benefits NRAM provides over DRAM and proposes five (5) work packages to be funded by the DoE to deliver a 32% power savings SDRAM replacement for the agency's use in its exascale computing and other missions.

Copyright © 2022 Nantero, Inc. All Rights Reserved.

SDRAM REPLACEMENT WITH 34% POWER SAVINGS

With the advent of Artificial Intelligence, Machine Learning, Simulations and Big Data Analytics todays exascale computers and data centers are striving to overcome the memory wall that direct connect memory bus from DDR5 RDIMMs to CPUs manifest. Along with a stride to lower energy consumption to be environmentally friendly or Green.

Current Synchronous Dynamic Random Access Memory (SDRAM) is stuck at 16Gb memory chips deployed on Registered Dual In-Line Memory Modules (RDIMMs), SDRAMs stacked up to eight high for HBM3 and soon to be installed in Compute Express Link (CXL) memory modules. 24 Gb SDRAM memory chips have been announced from Samsung and SK Hynix for 4Q of 2022.

Nantero proposes to replace these three applications of SDRAM with 16, 24 and 32Gb DDR5 NRAM at a 34% power savings with a goal of 40% power savings for all memory devices. These can be accomplished at 14/12nm lithography node. If we move to 7nm node late in the decade we can produce 16Gb, 32Gb, 48Gb, and 64Gb DDR5 NRAM memory chips with 40% reduced power than standard SDRAM.

With the advent of Peripheral Component Interconnect Express (PCIe) 5.0 and CXL 2.0 Composable Disaggregated Infrastructure (CDI) is now viable to enhance and optimize the utility or efficiency of exascale computers and megascale data centers. Disaggregated architectures are shifting the paradigm of innovation by enabling resource composability for more efficient utilization in Cloud, AI, and HPC. The remaining challenge is to disaggregate main memory and allocate to any CPU needing memory.

A.1 Underlying technology trends

Roadmaps

PCI-SIG Standards Release Dates

PCIe 5.0 (May 2019)   PCIe 6.0 (Jan 2022)   PCIe 7.0 (2025)   PCIe 8.0 (2028)

CXL Consortium Standards Release Dates

CXL 2.0   CXL 3.0 (2023)   CXL 4.0 (TBD)

JEDEC Memory Standards Release Dates

JESD238 HBM3 (Jan 2022)   JESDXXX HBM4 (Proposed)

JESD79-5A DDR5 SDRAM (Oct 2021)   JESD79-6 DDR6 SDRAM (Proposed)

JESD79-5-1 DDR5 NVRAM (Proposed)

JESD317 CXL Memory Module (Proposed)

Copyright © 2022 Nantero, Inc. All Rights Reserved.

JESD322 Multiplexed Rank DIMM (Proposed)

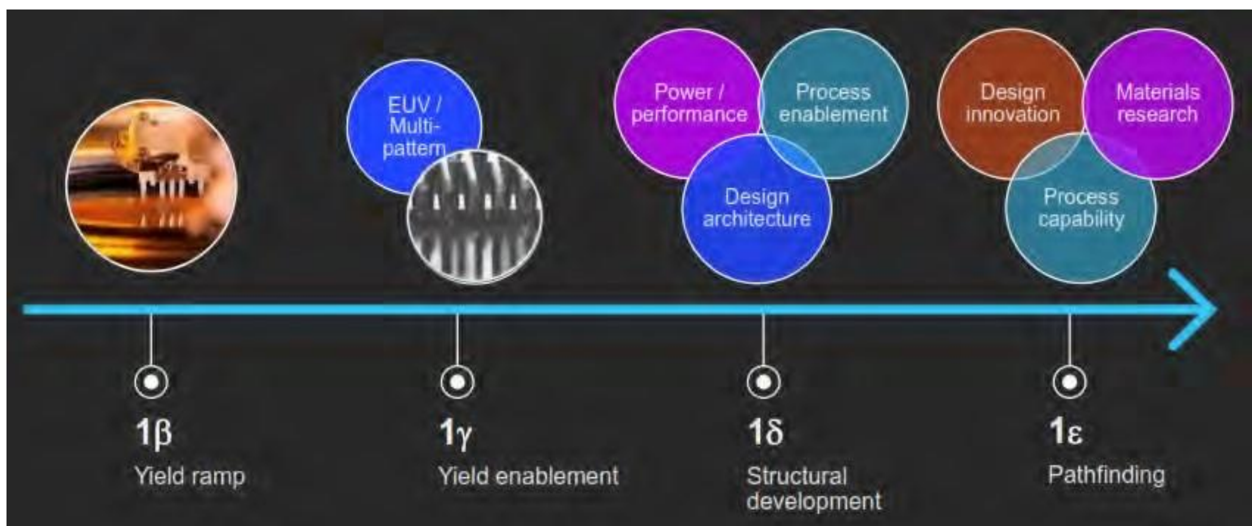
SNIA Standards Release Dates

EDSFF (Enterprise and Data Center Standard Form Factor) E1.S, E3.S, E1.L, E3.L. E3 also has double wide widths for Short (S) and Long (L). These replace M.2 and 2.5" HDDs and SSDs.

A.4 Memory

- What are your expectations for memory technologies, capacity, latencies, and bandwidths during 2025-2030 timeframe?

Answer: Today (1st Quarter 2022) SDRAM is manufactured mainly by three companies, Samsung (43.5%), SK Hynix (27.3%) and Micron (23.8%). Together they dominate about 95% of the ~$100 billion SDRAM market. Both Samsung and SK Hynix have announced 24Gb SDRAMs by 4th quarter 2022. These are actually 32Gb designs that are not completely yielding. JEDEC JESD 79-5A allows for up to 64Gb SDRAM chip designs. SDRAM latencies are based on lithography nodes which are stuck at 14nm for these companies. Samsung and SK Hynix are using ASML EUV lithography while Micron is sticking with 193nm DUV lithography for now. In 2025 ASML will be shipping 8nm resolution High Numerical Aperture (NA) EUV systems with 8nm resolution. This can allow further reduction in transistor sizing and increased capacity and smaller latencies. Bandwidths will be in accordance with JEDEC standards.



Manufacturers are targeting higher density, higher capacity, higher speeds, higher energy efficiency and better reliability for DDR applications. Advanced DDR5 technologies are already in development, and the market will continue to grow until next-generation DDR6 memory arrives in late 2024 or early 2025. We believe these dates are very optimistic and more likely 2026 or 2027.

Copyright © 2022 Nantero, Inc. All Rights Reserved.

Regarding the technical specifications of DDR6 memory, the data transfer speed will be doubled compared to DDR5. For example, the JEDEC module can run at approximately 9,600 MB/s for DDR6–9600. In addition, DDR6 will also double the number of memory channels per module, with four 16–bit channels combined in 64 memory banks.

Today 3D NAND is manufacturing 232-layer NAND flash devices with about 1Tb of capacity. These are stacked 116-layer NAND devices. Limit for a monolithic 3D NAND is 128 layers then die stacking is required. 3D NAND is mostly used for non-volatile persistent memory storage. 3D NAND's latencies are way too slow to be used for an extension of main memory even with PCIe I/O. Latencies for 3D NAND technology are measured in thousands of nanoseconds, orders of magnitude slower than DRAM. Bandwidth will be in accordance with JEDEC standards.

Carbon Nanotube based non-volatile random-access memory (NRAM) is the closest emerging technology to universal memory. Capable of SDRAM speed, capacity, and better latencies at 34% lower power per data transferred than SDRAM. Also, about 20% less thermal load than SDRAM. Completely persistent beyond 3D NAND with better speed, lower capacities, lower thermal load and lower power. Currently NRAM is not as fast as SRAM L1/L2 caches which are fabricated at much lower lithography nodes. Capacities are projected to be 8Gb by 2025, 16/24/32Gb by 2026 using GlobalFoundries 12nm process node. By 2028-2030 a 7nm process node can generate 16/32/48/64Gb NRAM memory parts. Read latencies of 2ns are better then SDRAM specifications. Write latencies of initial parts will be one order of magnitude slower than SDRAM with the goal to be equal to SDRAM as the process matures. Bandwidths will be in accordance with JEDEC standards.

Copyright © 2022 Nantero, Inc. All Rights Reserved.

- What fundamentally new memory devices (i.e., **alternatives to CMOS SDRAM and 3D NAND flash**) do you expect to become feasible during 2025-2030 timeframe?

Answer: Nantero carbon nanotube based random-access memory (NRAM) universal memory devices up to 64Gb capacity. See above for details.

- The HBM roadmap or projections, including the following:
  - Will power simply scale linearly with these bandwidths or are there any opportunities for performance or power (i.e., wattage) improvements?

Answer: Due to laws of physics power will scale linearly except NRAM based HBMs will begin with 34% less power and 20% less thermal load. Which should allow us to stack up to at least 12 high. SDRAM may not be able to do this do to their thermal load. Development under the Japan National Green Innovation Fund program has a goal of 40% power reduction by 2030. Performance is improved by eliminating SDRAM refresh, precharge, and rowhammer attacks.

  - How many stacks per SOC (e.g., AMD MI300 APU, AMD EPYC CPU) could be reasonable?

Answer: Currently AMD MI300 APU has 8 stacks of 8 high SDRAM based on 16Gb or 24Gb SDRAMs. AMD EPYC CPUs have four stacks of 8 high SDRAM based on 16Gb or 24Gb SDRAMs. Future projections available directly from AMD. Twelve high stacks are reasonable for 2027-2030 timeframe.

  - Are there new technologies and how might they be used (e.g., processing-inmemory, silicon photonics connecting off-SOC memory modules, new memory devices)?

Answer: The data persistence of NRAM cells solves the primary weakness of SDRAM technology: data loss on power failure. This advantage enables a new generation of memory centric applications, such as machine learning and artificial intelligence, to embed NRAM memory deep into the architectures without concern for the need to checkpoint learned data periodically. Internal control registers may be constructed using NRAM cells to allow for instant-on without requiring time consuming recalibration. NRAM devices may be augmented with enhanced functions such as processing-inmemory (PIM) instruction sets, and also with persistent math units such as multiplyaccumulate circuits where weighting factors are not lost on power failure.

Copyright © 2022 Nantero, Inc. All Rights Reserved.

A.7 Storage

- What are your expected Bandwidths/IOPs for HDD, SSD, PCM, or any others?

Answer: Bandwidths/IOPs for:

HDD Because of how hard drives work—with rotating discs—and the way in which they store and access data, they are usually limited to a transfer rate of about 100 MBps–200 MBps.

For hard drives, an average latency somewhere between 10 to 20 ms is considered acceptable (20 ms is the upper limit). For solid state drives, depending on the workload it should never reach higher than 1-3 ms. In most cases, workloads will experience less than 1ms latency numbers.

NVMe SSDs At present, there are three common SSD data transports used that connect servers to storage media and include SATA, SAS and PCIe® (Peripheral Component Interconnect Express). The NVMe specification uses PCIe as a transport for SSDs.

Generally, the SAS interface provides higher throughput than SATA and is geared toward applications that require data protection and high availability. It has a defined technology roadmap that is expected to support larger capacities and higher performance capabilities in the future with 24Gb/s SAS upcoming.

SAS SSDs enjoy several advantages over SATA SSDs as the latest SAS-3 full-duplex standard can transfer data bi-directionally at speeds up to 12Gb/s (or ~1,200MB/s) – performing reads and writes much faster than eSATA SSDs.

While SATA III theoretically provides 6Gb/s performance, and SAS provides 12Gb/s performance per lane, NVMe SSDs deliver data transfer performance of 1GB/s per lane, or upwards of 4GB/s in a typical x4-lane configuration. SSDs based on NVMe should be more responsive to heavy workloads and more resilient against performance degradation due to too many I/O requests.

The new SAS category is also capable of delivering much faster bandwidth and input/output operations per second (IOPS) performance that include sequential reads at 830MB/s, sequential writes at 650MB/s, random reads at 150,000 IOPS and random writes at 35,000 IOPS.
For 2022 and 2023, Silicon Motion will continue to focus on those PCIe 4.0 SSDs. Budget SSDs like Western Digital's WD Black SN770 SE are only beginning to transition to PCIe 4.0, and according to reviews like this one from Tom's Hardware,

Copyright © 2022 Nantero, Inc. All Rights Reserved.

their controllers and flash memory are not yet fast enough to benefit from the extra bandwidth.

In 2024 PCIe 5.0 NVMe SSDs will be available with read speeds up to 14GB/s and write speeds up to 12GB/s.

Optane PCM Optane DC memory occupies a tier in-between SSDs and DRAM. It has higher latency (346 ns) than DRAM but lower latency than an SSD. Unlike DRAM, its bandwidth is asymmetric with respect to access type: for a single Optane DC PMM, its max read bandwidth is 6.6 GB/s, whereas its max write bandwidth is 2.3 GB/s. Like traditional DRAM DIMMs, the Optane DC PMM sits on the memory bus and connects to the processor’s onboard memory controller. Our test systems use Intel’s new second generation Xeon Scalable processors (codenamed Cascade Lake). A single CPU can host six Optane DC PMMs for a total of 3 TB of Optane DC memory. The memory controller communicates with the Optane DC PMM via a custom protocol that is mechanically and electrically compatible with DDR4 but allows for variable-latency memory transactions.

CXL SDRAM If they're using the full 16 lanes of PCIe 5 that E3.S allows, then they can keep up (bandwidth-wise) with all DDR5 currently available. Certainly, it would keep up with all the standard JEDEC speeds that would be used in the data center. If AMD EPYC Genoa gives us 128 PCIe 5.0 lanes and 12 memory channels for a single socket, then that's only enough for 8 of these things even if you're ignoring all other I/O. The 12 channels of DDR5 will still easily win out on bandwidth. Much better than Optane sure, but without the advantage of being non-volatile.
If instead you take only 8 lanes for each of these, then you can get about 32 GBps and you can have 12 of them (matching the 12 memory channels and leaving 32 lanes for other I/O). But then you're significantly reducing bandwidth in that scenario even compared to the bottom-tier JEDEC DDR5-4800. If you go for 16 of them and ignore all other I/O, and you don't run into any bottlenecks moving that data around at the CPU / among the cores, then it'll keep pace.
But if you're primarily after capacity, you wouldn't be able to double it at comparable speed to DRAM. You'd likely go with a 2-socket setup to get double the memory slots, but by doing so you lose a big chunk of lanes (down to 80 per CPU) because they reserve a bunch of the lanes for inter-socket communication.
Not sure what the current/upcoming Intel offerings give you in terms of lanes.

CXL NVRAM (NRAM)

Note: NRAM® is Nantero’s registered trademark; NVRAM is the equivalent industry term used to generically refer to non-volatile RAM-style memories.

Copyright © 2022 Nantero, Inc. All Rights Reserved.

NRAM will initially be designed with a state of the art 6400 MB/s second interface that follows the DDR5 SDRAM protocol. With current generation processors having 64byte cache lines, and aligning with CXL 64-byte flits, DDR subchannels are 32 bits wide with 8 bits of ECC and metadata. They operate with a burst length of 16 so that each 40-bit subchannel processes one cache line at a time at 25.6 GB/s. Future NRAMs, in conjunction with module-level timing tuning, are expected to achieve DDR5-8000 data rates for a per-subchannel throughput of 32 GB/s peak.

CXL allows for multiple bus widths from 4 data lanes (x4) up to 16 data lanes (x16). These lanes operate at a PCIe 5.0 maximum of 32 Gbps per lane. The DDR5 throughput aligns perfectly with an 8-lane CXL bus throughput of 32 GB/s peak. CXL provides for full duplex operation, i.e., simultaneous reads and writes, for a total throughput of 64 GB/s. A CXL module design with 4 subchannels can transfer 51.2 GB/s on two subchannels, leaving the other two subchannels in idle state. Alternatively, the CXL module can incorporate a dual port architecture with 128 GB/s peak throughput and keep all four DDR subchannels busy; the tradeoffs include higher power (approximately 20 W), additional CXL controller circuitry (an integrated PCIe switch), and more on-module power supply circuitry to support the additional load.

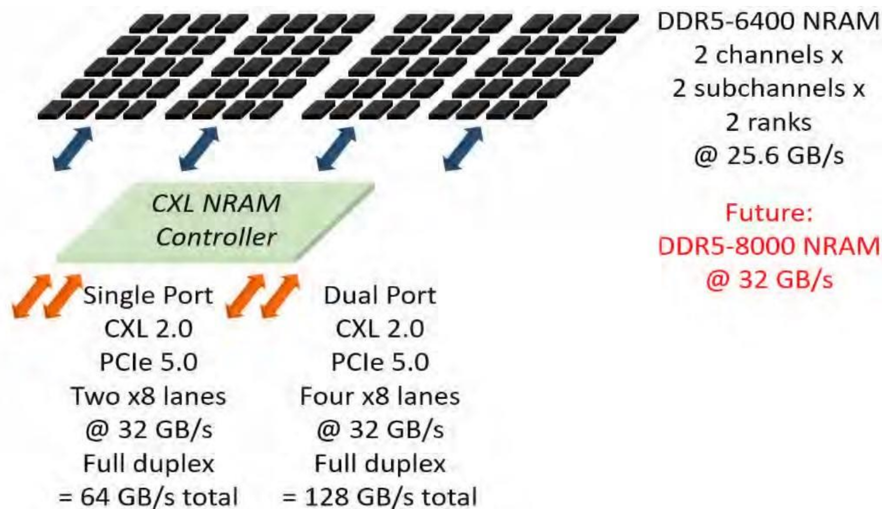


Figure 1: Dual-port CXL 2.0/PCIe 5.0 interconnect

- What capacities will be available for each media type?

Answer: Capacities available for:

HDD Currently 18 to 20TB are available and expected to grow to 50TB by 2026.

Copyright © 2022 Nantero, Inc. All Rights Reserved.

NVMe SSDs Currently Initial offerings will utilize cutting-edge 64-layer BiCS FLASH™ TLC (3-bit-per-cell) 3D flash memory developed by KIOXIA, with capacities ranging from 960GB to 7,680GB.

Optane PCM Currently Series 200 are available in 128, 256, and 512GB. These are mechanically and electrically compatible with DDR4. Series 300 are projected to be mechanically and electrically compatible with DDR5. There is no expectation that capacity will increase to 1TB. Intel is still deciding Optane's future.

CXL SDRAM Currently Samsung is offering a E3.S 512GB unit using 16Gb DDR5 memory chips. I would expect with the availability of 24Gb DDR5 memory chips, a 768GB unit will be next.

CXL NVRAM (NRAM) Non-volatile RAM E3.S unit will be available in 2027-2030 with 256GB, 512GB, 1TB and 2TB. If 7nm 16Gb, 32Gb, 48Gb, and 64Gb NRAM can be made in 2028 then 1TB, 2TB, 4TB, and 8TB could be made.

- Do you see 3D NAND flash drives narrowing the gap in **cost per byte** with HDD?

Answer: No, HDDs currently dominate the cloud exabyte market—offering the lowest cost per terabyte based on a combination of factors including price, cost, capacity, power, performance, reliability, and data retention. SSDs, with their performance and latency metrics, provide an appropriate value proposition for performance-sensitive, highly transactional workloads closer to compute nodes. HDDs represent the predominant storage for cloud data centers because they provide the best TCO for the vast majority of cloud workloads. More than 90% of exabytes in cloud data centers are stored on HDDs, and the remaining 10% are stored on SSDs according to market intelligence firm IDC.

HDD capacity continues to increase, with 18TB drives now widely available. **Increased capacity is driving down the cost per terabyte.** Moreover, new HAMR (heat assisted magnetic recording technology) drives were rolled out in 2020, with an initial capacity of 20TB and are forecasted to increase to 50TB by 2026. Higher capacity HDDs change the dynamics of data center storage. MACH.2 HDDs, for example, use multi-actuator technology to double hard drive IOPS performance, enabling cloud architects to maintain performance as capacity scales.

Higher-capacity drives dramatically reduce the TCO for storage infrastructure. Total cost of ownership improves about 10% for every 2TB of storage added to an HDD. Factors contributing to this TCO reduction are $/TB, Watt/TB, slot cost reduction, densification, and exabyte availability.

The TCO of SSDs today is about six times that of HDDs (ranging from just under 5x to over 7x, depending on application and workload variables), while at the device level SSDs continue to be

Copyright © 2022 Nantero, Inc. All Rights Reserved.

about eight times the cost of HDDs. Over the next decade, increased HDD capacity is expected to offset any decline in SSD costs. As a result, by 2030, the TCO of HDD infrastructure will continue to be about one-sixth the cost of equivalent capacity SSD deployments.

- If so, in what ways?

Answer: N/A

- Are there new technologies and how might they be used (e.g., persistent memory, headless storage, compute-in-storage, new interfaces)?

Answer: CXL memory modules increase the available direct-mapped memory footprint for large scale computing systems. This offsets losses in memory capacity as DDR5 generation memory module solutions (DIMMs) drop from two DIMMs per channel to one DIMM per channel at speeds above 4800 Mbps. Expansion using CXL allows some of total memory capacity as the CXL address space is significantly larger than that supported by the DDR5 DIMM protocol, and the number of sockets can be increased by direct access from CPUs or through CXL switches on the system motherboards.

DAX (direct access) protocols added to major operating systems including Windows and Linux allow smooth transition to support CXL-based main memory for existing and future software applications. DAX allows CXL expansion memory to be mapped as RAM drives for unmodified legacy applications, as memory-mapped files for slightly modified applications, and as direct access large footprint memory for future generations of software including in-memory computing.

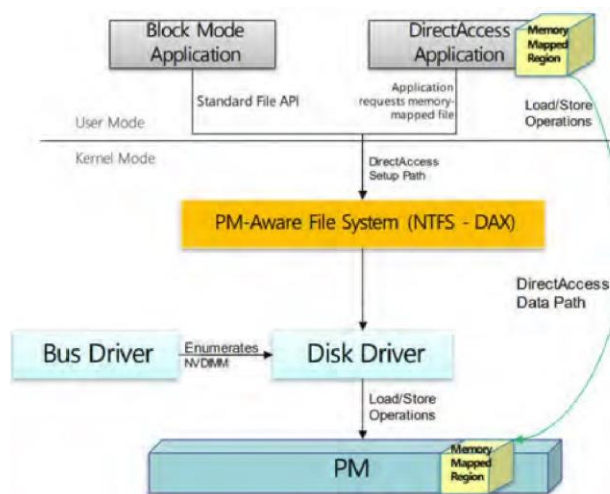


Figure 2: DAX-Mode Software Migration Paths

In phase 1 of DAX migration, systems are likely to mount CXL NRAM modules as hot data storage, with the significant benefit that the non-volatility of NRAM allows this hot data storage

Copyright © 2022 Nantero, Inc. All Rights Reserved.

to be implemented as a persistent data drive like an ultra-fast SSD. In phase 2, key performancesensitive applications will be modified minimally as memory-mapped files, reducing the operating system trap penalties, or optimally modified to use CXL NRAM as direct access memory with data persistence.

- What metadata and data backplane solutions are on your roadmap to better support campaigns running on the future computing ecosystems?

Answer: CXL data packets support user-defined metadata which are both stored and retrieved from CXL memory modules with standard packet access mechanisms.

- If you are an integrator (e.g., HPE), what data tiering and I/O subsystems do you anticipate including, such as a data hierarchy from nodes to center-wide file systems?

Answer: Data provided directly by HPE.

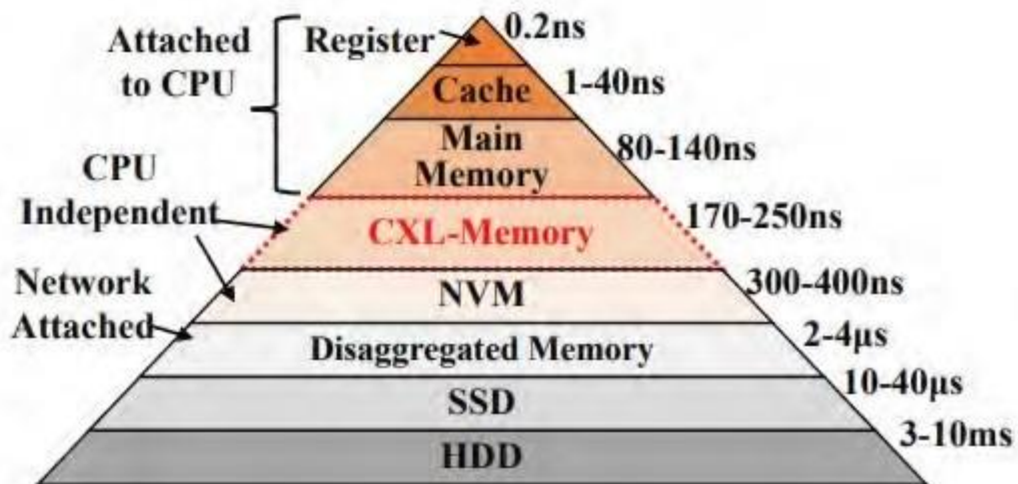


Figure 3: Memory Hierarchy

Summary

Large computer systems typically use an architecture where compute and memory resources are tightly coupled to maximize performance. Components such as CPUs, GPUs, APUs and memory must be placed closely together when connected electrically via copper interconnects. This hardware density results in cooling and energy issues, while persistent bandwidth bottlenecks limit inter-processor and memory performance. These issues are exacerbated in computeintensive applications like HPC, AI, Simulations, and compute-intensive data analytics. Next generation computer systems will use disaggregated system architectures with optical interconnects to decouple a server's elements — processors, memory, accelerators, and storage — enabling flexible and dynamic resource allocation, or composability, to meet the needs of

Copyright © 2022 Nantero, Inc. All Rights Reserved.

each particular workload. Next generation processors, memory and storage must all be lower power and reduced latency to accelerate user applications, especially next generation exascale computers.

C. Address the potential impact of DOE R&D investment and NRE funding on your proposed system(s)

Highlight those innovations that could significantly contribute to accelerating the trajectory of computing capabilities. Responses to this question can include innovations in any or all of your proposed solutions, but particularly the 2026 solutions and solutions that would promote wide ecosystem diversity.

- Hardware innovations (e.g., node, board, interconnect design)
  - Memory system technology, specifically to improve effective memory bandwidth, latency, and/or capacity limitations
  - Improving I/O system performance
  - Resilience and Reliability, Availability, and Sustainability (RAS)
- Innovations that reduce total cost of ownership.
  - Reduced power consumption

PROPOSED WORK PACKAGES

The following technical work packages are proposed to replace State of the Art (SOTA) SDRAM based memory.

WORK PACKAGE 1: DDR5 16Gb NRAM Memory

Currently Nantero is fabricating NRAM on 300mm (12”) silicon wafers at 55nm with Fujitsu in Japan. We have achieved 5 Sigma on demonstration devices. Now we are in the planning stage to fabricate 8Gb to 32Gb NRAMs for exascale super computers and green data centers. Japan’s National Program (The Green Innovation Fund) has selected Nantero’s NRAM as the memory baseline do to its unique properties, especially immunity to Rowhammer attacks and 34% lower power than SDRAM.

Nantero has enter into foundry discussions with GlobalFoundries (GF) and Albany Nanotech Center to demonstrate a single layer 8Gb NRAM memory device on a 300mm wafer. Front End of Line (FEOL) will be performed initially by GF and wafers transferred to Albany for Back End of Line (BEOL) processing. Engineering samples will be packaged and provided for initial testing to interested companies (i.e., AMD, HPE, and Dell Technologies). The 8Gb NRAM will function according to DDR5 protocols for easy system level evaluation by the interested end users.

Copyright © 2022 Nantero, Inc. All Rights Reserved.

Once these demonstration devices have been tested and any design changes made then NRAM with two layers of CNT memory cells (16Gb), three layers (24Gb) and four layers (32Gb) will be fabricated, and engineering samples will be packaged.

Subsequently these NRAM memory devices will be integrated into DDR5 CXL modules, DDR5 RDIMMs, MRDIMMs and DDR5 HBM3 Memory stacks up to eight high, maybe twelve high.

WORK PACKAGE 2: DDR5 32GB RDIMM Module – System Level Solution

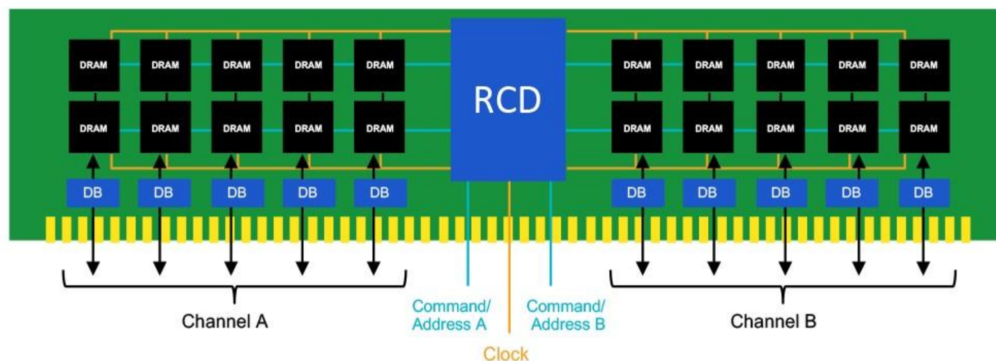


Figure 1: Typical DDR5 DIMM Following JEDEC JESD305 Specifications

Copyright © 2022 Nantero, Inc. All Rights Reserved.

Features	DDR4	DDR5	DDR5 Advantages
Speed	1.6 to 3.2 Gbps data rate 0.8 to 1.6 GHz clock rate	4.8 to 6.4 Gbps data rate 1.6 to 3.2 GHz clock rate	Higher bandwidth DDDR5-4800 initial designs
IO Voltage	1.2 V	1.1 V	Lower power
Power Management	On motherboard	On DIMM PMIC	Better power efficiency Better scalability
Channel Architecture	72-bit data channel (64 data + 8 ECC) 1 channel per DIMM	40-bit data channel (32 data + 8 ECC) 2 channels per DIMM	Higher memory efficiency Lower latency
Burst Length	BC4, BL8	BC8, BL16	Higher memory efficiency
Max. Die Density	16Gb	64Gb	Higher capacity DIMMs

Figure 2: Performance Advantages of DDR5

First demonstration of memory module is proposed as a 32GB DDR5 RDIMM. AMD and HPE have requested an engineering sample as soon as possible to try out. We will team with a major DDR5 DIMM manufacturer to execute this project.

WORK PACKAGE 3: DDR5 32Gb and 64Gb NRAM Memory

Once initial 8Gb NRAM engineering samples are complete and yield. Then 16 Gb (two layers of CNT memory cell arrays), 24Gb (tree layers of CNT cell arrays) and 32Gb (four layers of CNT memory cell arrays) using additional metal layers in the Back End of the Line (BEOL). will be fabricated in the splitfab arrangement with Albany and GlobalFoundries. The packaged NRAM engineering samples will be available to interested end users again for testing and comply to DDR5 protocols.

Toward the later part of this decade, we can move to a 7nm fab node provided a fab partner can be negotiated. Intel would be the best US fab partner. 7nm would provide for 16Gb (single layer), 32Gb (two layers), 48Gb (three layers) and 64Gb (four layers).

WORK PACKAGE 4: DDR5 RDIMM Cards and/or CXL

This task will exploit the availability of 16, 24, and 32Gb NRAM memory chips to populate DDR5 RDIMM cards. There are 40 memory chips per DDR5 RDIMM using two memory channels (A and B). Therefore, if we use 40 NRAM memory chips we get, only 32 for memory and 8 for ECC:

Copyright © 2022 Nantero, Inc. All Rights Reserved.

Monolithic memory chips	Two high memory chips	Four high memory chips
8Gb NRAM x 32 = 32GB	16Gb x 32 = 64GB	32Gb x 32 = 128GB
16Gb NRAM x 32 = 64GB	32Gb x 32 = 128GB	64Gb x 32 = 256GB
24Gb NRAM x 32 = 96GB	48Gb x 32 = 192GB	96Gb x 32 = 384GB
32Gb NRAM x 32 = 128GB	64Gb x 32 = 256GB	128Gb x 32 = 512GB

Eight high memory chips
64Gb x 32 = 256GB
128Gb x 32 = 512GB (Samsung has made a SDRAM version)
192Gb x 32 = 768GB
256Gb x 32 = 1TB

Prototypes need to be made and engineering samples verified by major IT vendors (i.e., AMD, HPE, Dell Technologies). These non-volatile memory chips will satisfy Sandia National Lab's Advanced Memory Technology Program for next generation exascale supercomputer. We are teamed with HPE and AMD to develop this capability.

WORK PACKAGE 5: DDR5 NRAM Multiplexed Rank DIMM (MRDIMM)

Proposed by Intel and AMD within the JEDEC Technical Groups on memory. Will double the bandwidth on traditional DIMMs by ping ponging between ranks of NRAM memory chips.

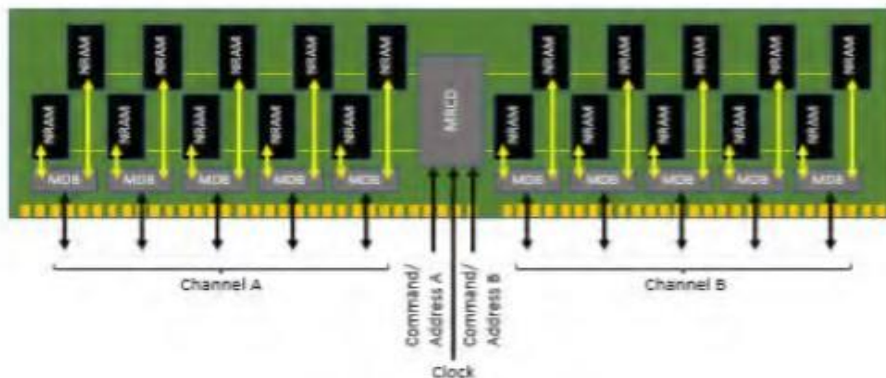


Figure 4: DDR5 MRDIMM for multiplexed rank design per JESD322

Figure 5 below describes the advantages of MRDIMM versus traditional DDR5 DIMM.

Features	DDR5 RDIMM	DDR5 MRDIMM	MRDIMM Advantages
Speed	6.4 Gbps data rate 3.2 GHz clock 25.6 GB/s throughput	8.8 Gbps data rate 4.4 GHz clock 35.2 GB/s throughput	Higher bandwidth, faster future designs to 51.2 GB/s
Channel Architecture	2 each 40-bit subchannels, 2 ranks	4 each 40 bit multiplexed subchannels, Up to 8 ranks	Higher per-module capacity
Burst Length	BL16 → 32 bits + 8 ECC	BL32 → 64 bits + 16 ECC	Higher memory efficiency

Figure 5: Advantages of DDR5 Multiplexed Rank DIMM (MRDIMM)

Copyright © 2022 Nantero, Inc. All Rights Reserved.

WORK PACKAGE 6: DDR5 HBM3 Memory stacks

Leverage 32Gb NRAM chips from Work Package 2 to develop increased capacity HBM3 stacks with a HBM3 interface. This will provide double current HBM3 capacity (so 32GB versus 16GB). In accordance with HBM3 JEDEC 238 specification increase NRAM memory stack from eight high to 12 high once again increasing capacity (48GB versus 32GB). This would provide a 3X increase per memory stack. Currently the El Capitan exascale supercomputer uses eight HBM3 stacks per AMD MI300A/350X APU module. So total memory per MI300A/350X is increased from 128GB to 384GB per AMD APU or 3.072TB of non-volatile random-access memory per EL Capitan server chassis (two nodes).

	HBM3	HBM4 (Proposed)	Remark
Stack configuration	4/8/12/16-high	8/16-high	8N-high stack unit
Die density	16/24Gb	24/32Gb	≤1.5~2X density
Max bandwidth (Max data rate)	820GB/s (6.4Gbps)	1.5TB/s (6Gbps)	≤2X bandwidth
IO width	x1024	x2048	Pin/timing compatibility (BL8, 32DQ/PC, tCK:tWDQS = 2:1)
Channel (IO), Bank	16CH (x64), 16 bank/PC	32CH (x64), 16 bank/PC	
Supply voltage	VDDQL=0.4V VDDC/Q=1.1V	VDDQL=0.6V VDDC/Q=1.05V	C_{IO}/bump pitch/power
Microbump pitch	96μm x 110μm	70μm x 110μm	x2048 IO, 32CH
PHY structure	Offset to Host	Centered to DRAM w/ distributed PHY	Increased flexibility for core arch./power eff.

Figure 6: Proposed JEDEC HBM4 specification

WORK PACKAGE 7: DDR5 CXL Memory Module

This CXL-enabled DDR5 module can scale CPU direct access memory capacity to the terabyte level, while dramatically reducing system latency caused by memory caching.

Dr. Debendra Das Sharma, Intel Fellow and Director of I/O Technology and Standards at Intel said, "Data center architecture is rapidly evolving to support the growing demand and workloads for AI and ML, and CXL memory is expected to expand the use of memory to a new level. DDR5 NRAM-based memory module is designed to meet the high-performance demands of data intensive applications including AI and HPC. This CXL-based module will enable server

Copyright © 2022 Nantero, Inc. All Rights Reserved.

systems to significantly scale memory capacity and bandwidth, accelerating artificial intelligence (AI) and high-performance computing (HPC) workloads in data centers.

This will be the industry's first non-volatile RAM-based memory solution that runs on the CXL interface, which will play a critical role in serving data-intensive applications including AI and machine learning in data centers as well as cloud environments.

Develop Compute Express Link (CXL) DDR5 NRAM for exascale computer main memory extension and green data center servers. If we use a 16Gb NRAM memory chip x 80 memory chips per CXL E3.S module than we can triple RDIMM memory per processor from 6TB to 18TB. (See Below)

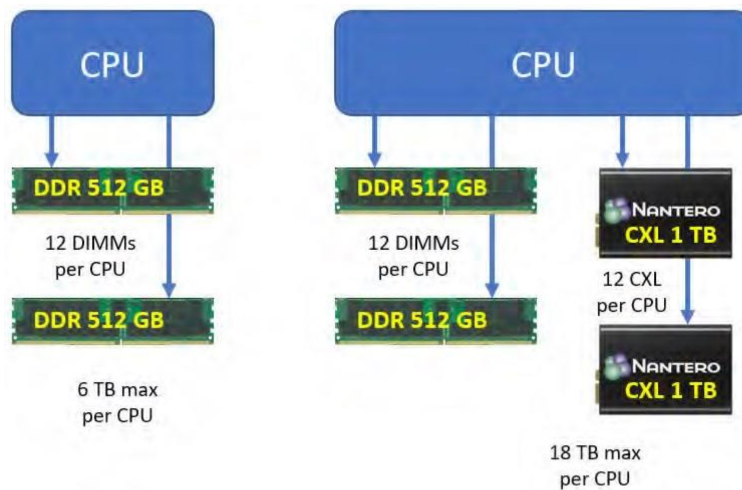


Figure 7: Maximum Direct Access CPU Memory Capacity with CXL

If we use 32Gb NRAM chips, we can quintuple capacity to the maximum of 30TB per processor.

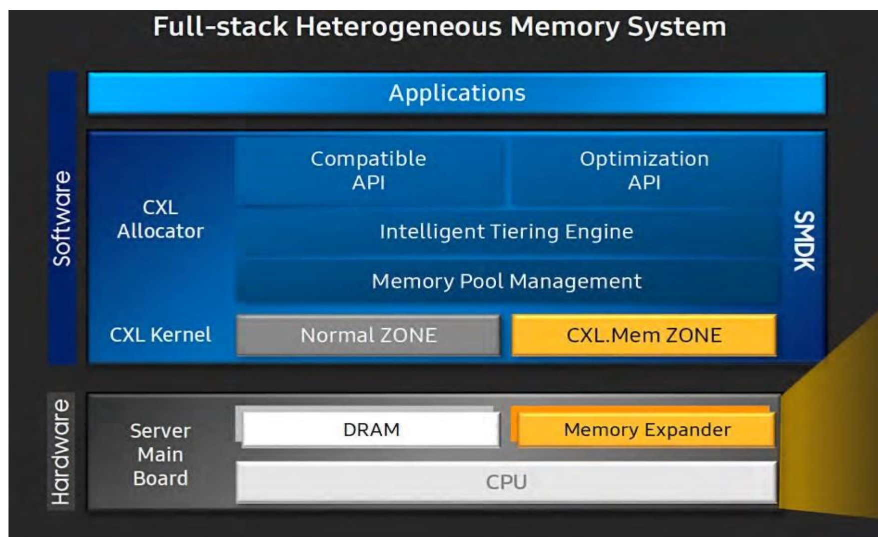


Figure 8: Open CXL Software from Samsung

Copyright © 2022 Nantero, Inc. All Rights Reserved.

The above is a software stack developed by Samsung soon to be release through JEDEC as open software for JEDEC CXL Memory Module standard (JESD317). We plan on leveraging this development along with CXL controller and physical interface chips. We will strategically team with a CXL memory module manufacturer.

C.1 Cost Estimates

Provide cost estimates for NRE (**Proposed Work Packages**) and for the described system(s)

WORK PACKAGE 1: DDR5 16Gb NRAM Memory

ROM Cost of $89M

WORK PACKAGE 2: DDR5 32GB RDIMM Module System Solution

ROM Cost of $52M

WORK PACKAGE 3: DDR5 32Gb and 64Gb NRAM Memory

ROM Cost of $24M per stack

WORK PACKAGE 4: DDR5 RDIMM Cards and/or CXL

ROM Cost of $42M

WORK PACKAGE 5: DDR5 NRAM Multiplexed Rank DIMM (MRDIMM)

ROM Cost of $250K

WORK PACKAGE 6: DDR5 HBM3 Memory stacks

ROM Cost of $50M – (built on work package 1)

WORK PACKAGE 7: DDR5 CXL Memory Module

ROM Cost of $1.5M

(REMAINDER OF PAGE LEFT BLANK)

Copyright © 2022 Nantero, Inc. All Rights Reserved.

Copyright © 2022 Nantero, Inc. All Rights Reserved.

Reference:

DOE Memory Critique Report 2014

Memory Technology

Memory is crucial component in any computing system. To meet a variety of needs and uses, several memory types have found their way into today's systems: NOR Flash for boot storage, SRAM cache for high-speed code and data storage, SDRAM for high density and high-speed code and data storage, 3D NAND Flash for persistent, block-oriented read and write storage, and hard drives for very high density, cost-effective, low speed block-oriented storage. Volatile memory loses data when powered off; nonvolatile memory retains it and can compete with disk for long-term storage. This section addresses the challenges facing both the volatile and nonvolatile memory technology needed for exascale systems.

Today, SDRAM and 3D NAND dominate in the two central roles: volatile, high-density, high-speed code and data memory (SDRAM), and nonvolatile, low-latency, high-density, block-oriented data storage (3D NAND). Some new nonvolatile memory technologies, currently on the horizon, have the potential to make a disruptive impact in the exascale timeframe. Examples include spin-transfer torque (STT-RAM), phase-change (PCRAM), and metal-oxide resistive (ReRAM, or memristor). **Replace with NRAM.**

While these memory technologies are not in volume production today, the issues of exascale are sufficiently challenging those alternative technologies should be examined for their disruptive potential. Notably, one cannot ignore the possibility that a change in memory technology may have a profound effect on architecture and software, and may be the best path to a cost-effective, low energy, high performance system.

A lot has changed since the original **DARPA Exascale report [2008]**, including the consolidation of the memory industry and the emergence of 3D NAND Flash as the predominant memory type. In examining a post-2020 deployment, the pace of Moore's Law will continue to slow as devices reach their physical limits. Before those limits are reached, process complexity and cell interference may impact memory scaling for economic reasons. The problem of cost-effective memory capacity is a central concern that is discussed below.

The Challenges

Memory Capacity: Capacity is critical to applications. It allows numerous problems to be solved in parallel, a powerful form of weak scaling; it allows in-memory checkpoints and message logging/replay for resilience; and it enables algorithms that buy performance by using data structures that may not be minimal in their memory footprint. There are numerous examples of this space/performance tradeoff. The primary challenge will be the continued scaling of memory density on a per-compute-operation basis (bytes per Flop/s).

Figure 1 shows that the machines at the top of the TOP500 do not have sufficient memory to match historical requirements, and the situation is getting worse. This is a big change from the traditional one byte per flop that the NNSA labs prefer in support of their application base. It places the burden

Copyright © 2022 Nantero, Inc. All Rights Reserved.

increasingly on strong scaling of applications for performance, rather than the weak-scaling model which dominated the terascale era. Interestingly, the "tipping point" in the graph is in the 2003-04 timeframe, which coincided with the beginning of the end of Dennard scaling and the start of the multicore era.

One reason is economic, with users often prioritizing arithmetic performance over memory capability. This critical capacity challenge can be addressed by combining volatile and nonvolatile memories into a single programmer-addressable device. However, there are numerous architectural challenges that emerge from integrating the two. These include, but are not limited to, the relatively poor latency, especially write latency, of some of the nonvolatile memories; a lack of abstracted memory protocols which decouple timing and naming from the device interface; differences between the block-oriented access typically favored

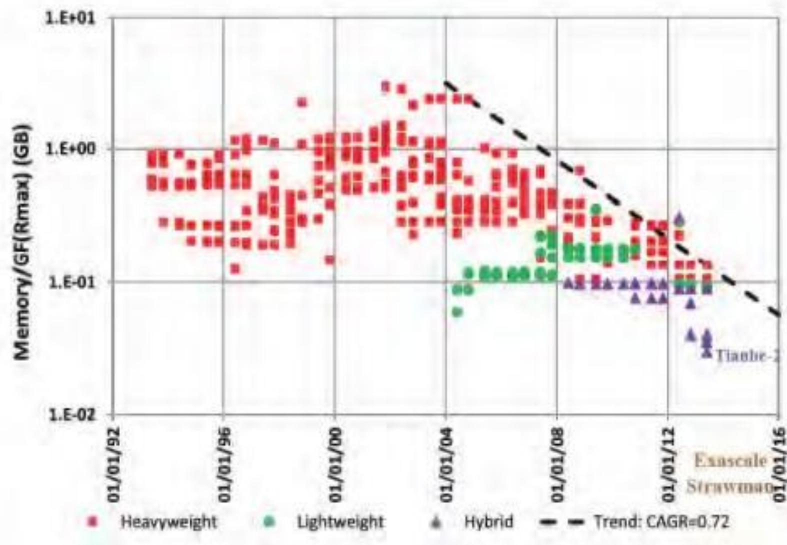


Figure 1: Memory capacity (Bytes) per gigaflop/s, courtesy of Peter Kogge. (For reasons of chip architecture and error management) in a 3D NAND Flash device and the wordoriented access typically favored by CPUs and volatile memory; and differences in read latency, wear out, soft error vulnerability, and other differences between any two classes of memory that are hybridized.

Energy

Energy is the primary technical challenge in supercomputing, but the least likely to be observed directly by the programmer. Memory faces three primary energy challenges:

1. **Processor/Memory Transport**: DDR-style memories will have reached the end of their useful lifespan by the exascale timeframe; more energy-efficient transport will be required

2. **Processor/Memory Protocol**: Since the invention of SDRAM, the interface between the CPU and the memory system has been bound in both naming and timing to the CPU. This produced the least expensive devices possible (from an acquisition cost perspective) at the expense of centralizing a rigid memory-management model and limiting memory evolution. Memory energy consumption, error mitigation, and data resiliency are removed by this model from the memory, and are managed by non-memory vendors, a situation which is poorly considered in the

Copyright © 2022 Nantero, Inc. All Rights Reserved.

complexity of JEDEC JESD DDRx specifications. Instead, distributed control with decoupled naming and timing must be enabled.

3. **Memory Architecture**: Typical commodity SDRAMs are architected for the lowest possible cost per bit, which requires significant "over-fetching" of data for any given memory request. This refer to the granularity of access in the memory devices, in which a "row" of bits is destructively read from a memory array into a row buffer, in each SDRAM chip involved in a cache line read or write. The data are fetched into these row buffers in chunks of 8 kilobits or more. As the number of cores per CPU increases, memory accesses get more random. As a result, data centers employ a "closed page policy" where only 64 bits of the row are used by the CPU, yet all 8 kilobits activated must be written back to the memory array at significant energy cost. Furthermore, these memories often provide a weak banking structure for the same cost reasons. Enabling finer-grained access of a larger number of banks provides significantly improved performance and potentially lower energy profile [See, for example, Jung Ho Ahn, Norman P. Jouppi, Christos Kozyrakis, Jacob Leverich, and Robert S. Schreiber. 2009. Future scaling of processor-memory interfaces. In Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis (SC '09). ACM, New York, NY, USA.]

The High Bandwidth Memory (HBM) roadmap provides examples of how to address each of these problems from a technology perspective, but it is not the complete answer. In particular, HBM demonstrates a path to manufacturable technology with sufficiently low transport energy, an abstracted, packet-based protocol, and a rearchitected SDRAM structure designed for higher performance.

Scaling

Manufacturing at commodity scales and low cost continues to be a problem. Wafer process cost and fabrication capital cost are growing because of the slowing of Moore's Law. This has resulted in a significant slowdown in the building of new SDRAM fabrication capacity.
None of the emerging alternative nonvolatile memory technologies has yet achieved a near-term, costeffective, manufacturable path to replacing commodity memory. Further research in this area could be highly disruptive, especially Nantero carbon nanotube based non-volatile random-access memory (NRAM), which is a desirable outcome for the long-term health of the high-performance computing field.

Resilience

The increasing number of components required for an exascale system contributes to additional complexity in resilience, which will be compounded if memory capacity trends are addressed by a potential exascale program that would require increased memory density. In addition, today's slaved timing and naming require a centralized approach to failure recovery. This implies the following issues:

1. It is becoming difficult for memory controllers designed by CPU vendors to map out bad or problem rows. Recent "**rowhammer**" disturb issues (in which repeated accesses to the same row may alter the state of that row's neighbors) demonstrate that an intimate understanding of internal part topology may be required to address memory error conditions. These kinds of issues will continue to present themselves as technology scales.

Copyright © 2022 Nantero, Inc. All Rights Reserved.

2. Just as 3D NAND is a fundamentally lossy media, which requires significant error correction performed transparently to the user, opportunities exist to enable more resilient memory systems by performing distributed error recovery.

3. Variable Refresh Time (VRT) errors, in which some SDRAM cells require unpredictable refresh times, point towards the benefits of a more distributed approach to resilience and recovery, as well as more intimate knowledge of individual SDRAM parts and processes. Similar issues arise in alternative memory technologies.

On the positive side, the deeply buried capacitor structure used in modern SDRAMs makes each generation of SDRAM **more immune to Single Event Upsets (SEUs) caused by radiation** than prior generations. SRAM, on the other hand, has increasing vulnerability as it scales. This dichotomy has a potentially profound effect on cache architectures, which may be significantly less reliable (and require increased error correction) than main memory. Alternative memory technologies may address some of these problems.

Generous memory capacity can be a boon to system-level resilience through frequent **checkpoints**. **Checkpointing exascale state to disk via a storage network connection** will become impractically slow; in-memory distributed checkpoints are a scalable alternative.

Systems Tradeoffs and Integration:

The government is uniquely positioned to enable a whole-system approach with multi-vendor teams, academia, and the national laboratories to find the best solutions for entire classes of applications. The exascale problem is sufficiently challenging that a business-as- usual approach is not likely to work.

Typical design patterns have forced a component-oriented approach to constructing systems, which is a challenge to the proposed DOE methodology of hardware/software co-design. Examining individual costs (e.g., “cost per bit of capacity”, “cost per bit per second of bandwidth”, etc.) rather than system costs (e.g., “how to design the least expensive system for a given aggregate bandwidth and capacity over the lifetime of the system”) often leads to per-subsystem or per-component tradeoffs resulting in an overall inefficient design.

Device manufacturers are investing in the basic layer of new alternative memory and storage technologies – devices and media controllers. A co-design approach in which the opportunities and implications for the memory channel, cache hierarchy, processors, and algorithms is explored in tandem with the DOE’s applications is required to best assess and then profit from this opportunity, should any of these technologies “make it” as a commercially viable, low-cost alternative.

Research Directions

Given the technology challenges in the early 2020s timeframe and the likelihood of these challenges to intensify over the lifetime of exascale-class systems, it would be beneficial to invest in further development of emerging memory technologies as well as possible process extensions to SDRAM. Several “SDRAM replacement” candidates already exist in the laboratory. While the maturation of these

Copyright © 2022 Nantero, Inc. All Rights Reserved.

to mass production remains uncertain, the possibility of creating valuable alternatives makes them a sound strategic research investment.

The new nonvolatile technologies offer an alternative attack on the capacity wall. The layering of a crossbar memory array on top of physical addressing circuits can produce a highly dense memory structure. Research directed at the optimization of this approach would be helpful.
It is important that several emerging nonvolatile memories will be byte-addressable, making them desirable candidates for main memory usage. For use as main memory, considerable durability is necessary. Research at the device and the circuit/media control levels is needed to achieve simultaneously the density, durability, performance, and energy efficiency needed for use as a main memory technology.
Compared to current memory systems, emerging technologies have some commonality, and some important distinctions in how they relate to exascale computing. First, they share the trend of abstracting the memory system from the CPU, with distributed control including media controllers clustered with NVM physical resources. They also share the benefits of advanced physical packaging such as throughsilicon vias (TSVs) that result in increased volumetric density and intimate local interconnect channels.
As with SDRAM, some emerging memories can be realized across a range of access latencies and densities, depending on device technology and circuit architecture. Some candidates also generally have asymmetric read and write performance or endurance limitations.

Research is also needed to understand overall system optimization accounting for capacity (cost)/latency tradeoffs in order to guide direction in memory architecture. Within the memory module, many low-level aspects of the media controller/ physical media interaction deserve attention. Examples include coding, read/write scheduling, and wear-leveling or mitigation schemes to minimize power consumption and maximize endurance. Research should address the low-level issues, such as manufacturability, and the overall system impacts of changes to the memory technology, which would have a significant ripple effect.

Given the importance of abstracted memory systems (with distributed control and decoupled timing) as outlined above, groups of vendors (processor, memory, integrators) should work on the potential of lowering overall system cost through memory systems that are abstracted, distributed, and more resilient. This would also serve as a vehicle to explore unified memory systems of heterogeneous memory types. Finally, enabling even limited processing capabilities near memory has been demonstrated to significantly improve performance and lower energy by numerous industry, lab, and academic research efforts. Multivendor team exploration of processing-near-memory architectures (e.g., on the logic base of a 3D stack) in the context of overall system performance for DOE applications has a significant potential to yield early fruit.

Impact

Decisions on memory provisioning, interfacing and architecture will ultimately determine the capability, cost, and energy profile of future computing systems. The best of today’s memory interfaces and density need to be further improved if exascale class systems are to deliver the promised performance under real applications. This improvement must be done in an informed manner which maximizes the potential of the machine while simultaneously embracing the capability of new memory technology.

Copyright © 2022 Nantero, Inc. All Rights Reserved.

Reference: ASCAC Subcommittee for the Top Ten Exascale Research Challenges 2014

Copyright © 2022 Nantero, Inc. All Rights Reserved.