

Blackcomb: Hardware-Software Co-design for Non-Volatile Memory in Exascale Systems

Principal Investigators:
 Jeffrey Vetter, ORNL
 Robert Schreiber, HP Labs
 Trevor Mudge, University of Michigan
 Yuan Xie, Penn State University

Memory, not processing, is the crux of the exascale co-design problem. Exascale machines will push the limits of memory capacity, power, and performance. DRAM, the universal memory technology of today, may not scale to meet the needs of exascale applications. Disk storage, critical for checkpointing and for archiving computational inputs and results, may also fail to provide adequate performance, reliability, and power efficiency by the end of this decade. We confront a memory/storage crisis.

	SRAM	DRAM	NAND Flash	PC-RAM	STT-RAM	R-RAM
Data Retention	N	N	Y	Y	Y	Y
Memory Cell Factor (F ²)	50-120	6-10	2-5	6-12	4-20	<1
Read Time (ns)	1	30	50	20-50	2-20	<50
Write / Erase Time (ns)	1	50	100-10 ⁵	50-120	2-20	<100
Number of Rewrites	10 ¹⁶	10 ¹⁶	10 ²	10 ¹⁵	10 ¹⁵	10 ¹⁵
Power Read/Write	Low	Low	High	Low	Low	Low
Power (Other than R/W)	Leakage Current	Refresh Power	None	None	None	None

The Blackcomb effort seeks to create and understand new memory technologies, develop their roles in exascale systems, adapt applications to them, and assess their relative merits. We focus on emerging nonvolatile memory (NVM) technologies, including spin-torque-transfer RAM (STT-RAM), phase-change RAM (PC-RAM), and memristor (resistive RAM, or R-RAM).

Objectives

- Understand and improve these emerging NVRAM technologies
- Propose new distributed computer architectures that address the resilience, energy, and performance needs of exascale applications. Key ideas:
 - adopt most promising NVM technologies
 - flatten the memory hierarchy
 - place low-power compute cores close to the data
 - replace mechanical disk-based storage with energy-efficient NVM
- Define programmer's APIs as they relate to, and evaluate the impact of the proposed architectures on the performance of critical DOE applications

Approach

The project is structured around five work packages:

NVM Technology identifies and characterizes the most promising NVM technologies. We will assess and improve wearout, error rate, durability, energy, latency.

Memory Architecture explores the architecture space, considering how to assemble NVMs with a space of future processors, and also looks into the uses of NVM for resilience.

System Architecture proposes a novel HPC system architecture. The idea is to explore the use of NVM as a single-level data store co-located with ultra-low voltage processors and balanced network capability. This entails a design-space exploration of the various architecture options, as well as an analysis of the simplification and optimization of the software stack.

System and Runtime Software identifies the most useful programming abstractions of the new NVM architectures. We will look into new programming paradigms that can help to fully take advantage memory nonvolatility. We will re-examine the Message Passing Interface (MPI) and Partitioned Global Address



Space (PGAS) programming models, and respective I/O models, such as MPI-IO and the Hierarchical Data Format (HDF5) in light of the new memory and storage architectures.

Applications identifies, characterizes, and transforms key DoE applications for NVM. The results of the characterization will be made available to the other work packages to provide a quantitative basis for research decisions. New programming and other software techniques will be ported to the selected applications and tested. We will seek to understand the sensitivity of studied applications to faults.

Impact

- Better energy scalability: NV memories have zero standby power
- Increase system reliability: MRAM/PCRAM are resilient to soft errors
- Boost performance: NVM will be much faster than magnetic disk
- Improve programmability and application fault tolerance with enhanced programming models

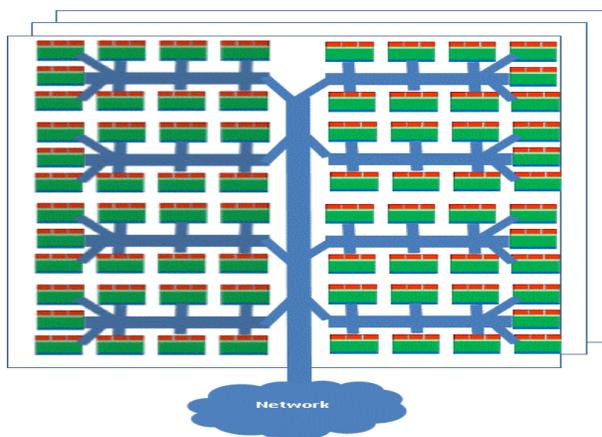
Challenges

- Understand and mitigate limitations of NVMs as a general purpose memory: higher write overheads and lower endurance than SRAM or DRAM
- Need novel analytical/simulation hybrid model to understand tradeoffs between energy efficiency, resilience, and performance
- Evaluate productivity of proposed programming models that exploit NVM to improve fault-tolerance of distributed applications

Research Products & Artifacts

- Identify and characterize a few key applications, making results available to the other areas of the study to provide a quantitative basis for research decisions
- Develop a failure taxonomy, quantify metrics for this taxonomy for important scientific applications, and investigate the use of this information
- Port the techniques developed in other areas of the study to the selected applications and perform testing with realistic workloads
- Broadly explore opportunities for NVM memory, from plug-compatible replacement (like the NV DIMM, below) to radical, new data-centric compute hierarchy (nanostores)

Networked, Stacked Nanostores



NV-RAM DIMM Design

