ActualTech Media

FORMULUS BLACK

In-Memory Storage and Formulus Black's FORSA™

Tom Fenton

CONTENTS

IN THIS PAPER

When choosing high-performance storage technology for your data center, not only do you need to consider the key differences between NAND, DRAM, and storage-class memory (SCM), but you want to make sure your storage system provides data integrity, efficiency, and protection.

This paper details the unique characteristics of NAND, DRAM, and SCM, as well as discusses what makes SCM so exciting. Find out how Formulus Black has changed the storage game by incorporating SCM and DRAM technologies to create a high-performance and ultra low latency in-memory storage system, using typical commodity server hardware.

Over the past 20 years, NAND technology has been used—first in SATA and SAS SSDs and then in NVMe devices—to set the standard for high-performance data storage. But with the introduction of Formulus Black's FORSA™, there's a new contender for the high-performance title—one that allows the use of dynamic random-access memory (DRAM) and storage-class memory (SCM) for highly performant data storage.

Choosing high-performance storage technology for your data center involves a consideration of many different factors. This paper will lay out the key differences between NAND, DRAM, and SCM. It will also discuss what makes SCM so exciting, and how Formulus Black has incorporated SCM and DRAM technologies to create a high-performance and ultra-low latency in-memory storage system, using typical commodity server hardware.

DRAM, NAND, and SCM all use semiconductors to store data, but they each have unique characteristics. DRAM connects to the CPU via a memory channel, which makes it very fast—the access time for DRAM is around 51 nanoseconds. Due to its volatility (its data will not survive a reboot) and expense, DRAM has traditionally only been used as primary memory for computers, and not as a data storage solution.

NAND is used in SSD and NVMe devices, but unlike DRAM, it has the capability for data to survive system restarts or power losses. NAND is slower than DRAM, with a read latency of around 47,000 nanoseconds and a write latency of around 15,000 nanoseconds. This places it three orders of magnitude slower than DRAM, but NAND is much less expensive than DRAM, and NAND-based devices have larger capacity than DRAM.

SCM is an emerging technology that's beginning to see use in the data center. However, SCM distinguishes itself from DRAM and NAND in three specific ways: It is persistent data storage, it has larger storage capacities, and it's less expensive than DRAM. Intel claims that although its SCM solution is 10 times slower than DRAM, it's 100 times faster than devices that use NAND.

Many persistent memory (PMEM) devices use SCM, but SCM isn't exclusive to PMEM. Like DRAM, PMEM devices plug into a computer's memory channel rather than a computer I/O channel, as is the case with NAND devices; this is one of the factors that gives PMEM its extremely low latency.

Intel Optane DC Persistent Memory is the most popular PMEM device currently available, which Intel has measured at around 350 nanoseconds. This is roughly 10 times slower than DRAM, but 100 times quicker than NAND. It also has the same storage capacity as NAND devices.

> DRAM, NAND, and SCM all use semiconductors to store data, but they each have unique characteristics.

Let's use an analogy of getting an apple to put the speed difference between DRAM, SCM, and NAND technologies into perspective. If I felt like getting an apple right now, there are three ways I could do so:

- I could leave my desk and grab an apple from my refrigerator in about two minutes (DRAM)
- I could leave my home and walk to the corner store and buy an apple in about 20 minutes (SCM)
- I could leave my neighborhood and walk to a farm to get an apple in about 14 days (NAND)

In other words, it would take me 10 times longer to go to a store, and 10,000 times longer to go to the farm, than it would for me to get an apple from the fridge. These are the relative latency differences in a computer acquiring data from DRAM, SCM, and NAND technologies, respectively.

SCM DEEP DIVE

In order to exploit SCM technology, it's necessary to create new techniques and procedures. The current leader in doing this is Formulus Black, with its software stack FORSA. FORSA allows creation of a virtual storage pool of DRAM or SCM, to create block-level POSIX-compliant storage devices. Each storage device is called logical extended memory (LEM), similar to a LUN, on which you can mount any industry standard filesystem. Alternatively, you can provision KVM-based Windows and Linux virtual machines (VMs), as well as containers that can read and write to FORSA's LEM.

NAND and DRAM have both been around long enough that they're generally understood by IT professionals, but SCM is a new technology; as such, it requires a little bit of explaining to appreciate how transformative it can be to the data center.

SCM isn't the exclusive purview of Intel—in fact, many companies make SCM memory devices. For instance, Intel and Micron joined forces to develop the technology behind 3D XPoint, their SCM offering, and Micron offers an SSD drive that uses 3D XPoint technology. Samsung also has an NVMe device based on its proprietary SCM technology, called Z-NAND.

It's worth noting that not all SCM storage devices are PMEM devices. A prime example of this is the fact that Intel offers both Optane DC Persistent Memory, which is a PMEM device, and Optane DC SSD devices, which are SSD and NVMe drives.

NVMe and SSD devices that use SCM have lower latency than those using NAND technology; however, devices that use NAND have much higher latency then those that use PMEM. One of the limiting factors with both NVMe and SSD devices is that they plug into a computer's I/O channel (for example, PCIe, SATA, SAS) rather than the computer's memory bus via a server's DIMM slots, like PMEM devices do. The memory channel is considerably quicker than the I/O channel.

> NVMe and SSD devices that use SCM have lower latency than those using NAND technology; however, devices that use NAND have much higher latency then those that use PMEM.

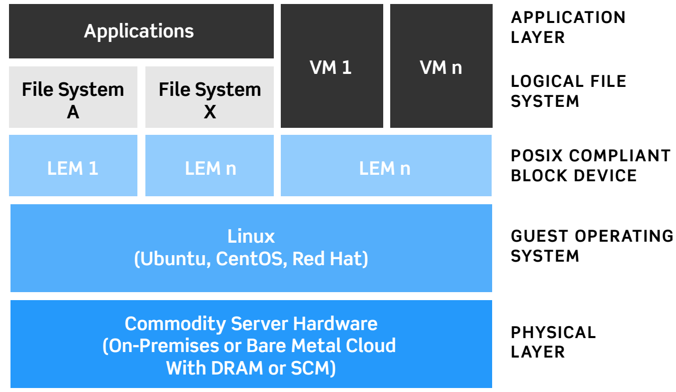

Figure 1: The FORSA LEM architecture.

An interesting characteristic of SCM devices is that they can be addressed at the byte level, whereas NAND can only be addressed at the block level. Moreover, NAND needs to erase and rewrite an entire block/page of data even if only one byte in that block is changed. In contrast, SCM devices, such as 3D XPoint, are byte-addressable, meaning that a single byte can be erased and rewritten. Not only does byte-level access improve performance, it also increases the life span of the device.

ENTERPRISE STORAGE FEATURES OF FORMULUS BLACK USING DRAM AND SCM

Storage has evolved over the last three decades, and enterprise users have come to expect integrity, efficiency, and protection for their data. Any storage system considered for enterprise use and production workloads must offer these abilities, and Formulus Black's FORSA effectively does so.

LEMs have enterprise-level functionality and attributes such as bad block replacement, snapshot, clone, and high availability (HA) functionality. These enterprise-level functions are available to all LEM instances, regardless of whether the LEM is used for the local host or provisioned to FORSA KVM-based Windows or Linux VM instances. See **Figure 1**.

Data integrity is the ability to trust that the data you've stored is accurate and consistent from the time it's written until the time it's erased. To ensure data integrity, FORSA has a Central Fault Tolerance Manager (CFTM) that does memory error checking and Bad Block Replacement (BBR); it will correct these issues if possible, log the event, and, if desired, create an SMTP alert.

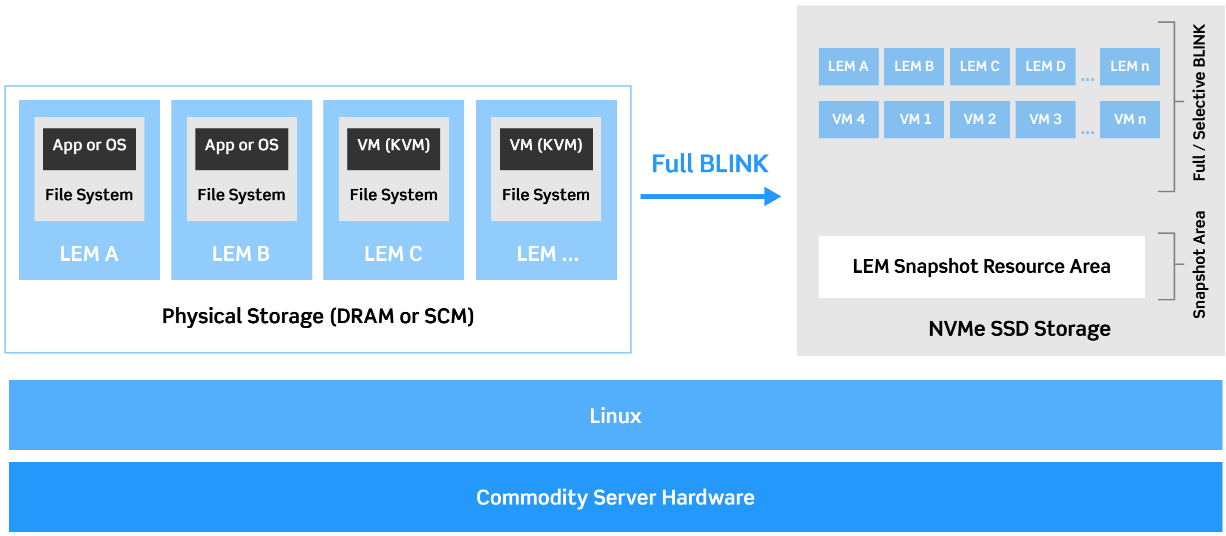

Figure 2: The Formulus Black Full BLINK process.

DRAM and SCM are limited resources and should be used as efficiently as possible. For data efficiency, FORSA uses its proprietary Formulus Bit Marker (FbM) technology to scan and find duplicate blocks of data in real time among all LEM instances. It then uses proprietary encoding to reduce the amount of memory required to store this data.

Bit Marker technology effectively lowers the $/GB cost of using memory media to store data for certain use cases. Early tests by Formulus Black have shown that in typical use, Storage Amplification can increase the effective capacity storage by 30% to 50%, which translates to 100GB of physical storage being able to store 130GB to 150GB worth of data.

> Data integrity is the ability to trust that the data you've stored is accurate and consistent from the time it's written until the time it's erased.

One of the challenges of using DRAM for physical backing of storage is that it's volatile, and the data stored on it will be erased when it loses power. To overcome this obstacle and protect the data on a DRAM device, Formulus Black created a process called BLINK (**Figure 2**). A full BLINK takes an application-consistent memory state capture of all the LEM and VM instances and copies them to a persistent storage device, such as an NVMe or SSD device.

> FORSA overcomes the challenges of using volatile DRAM for physical backing of storage.

BLINK can be used in a more granular manner in Selective mode. Where a Full BLINK operation will back up FORSA and its entire environment, a Selective BLINK will only back up the LEMs and/or VMs that you specify (**Figure 3**).

Both Full and Selective BLINK backups can be scheduled to take place at any time, and you can specify how many backups to keep.

Selective BLINK can be performed non-disruptively if there are applications running on the LEM/VM.

BLINK can be integrated and interfaced with UPS battery backup software, and take a backup once power failure has been detected by the UPS.

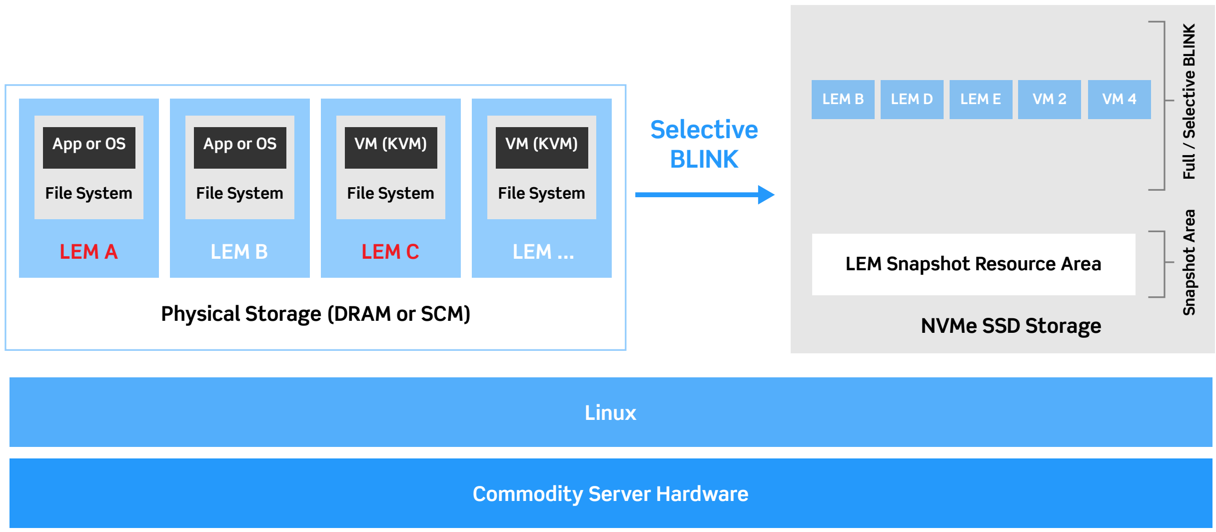

Figure 3: The Formulus Black Selective BLINK process.

But data resilience is more than backing up data; it also includes HA, and FORSA allows you to configure a LEM for HA. An HA LEM enables high-speed synchronous replication of data to a LEM residing on a second node, which can then be used if the original LEM is lost. LEM instances provisioned to FORSA KVM-based VMs allow the failover of VM instances to the HA LEMs on the second FORSA Node.

It's common to underestimate the amount of storage needed, or hesitate to commit a scarce resource, such as DRAM or SCM, to an unfamiliar technology. This isn't an issue with FORSA, as it allows expansion of LEMs as needed. Of course, a LEM expansion doesn't automatically extend the filesystem partition; you must use the standard operating system tools to take advantage of the expanded LEM.

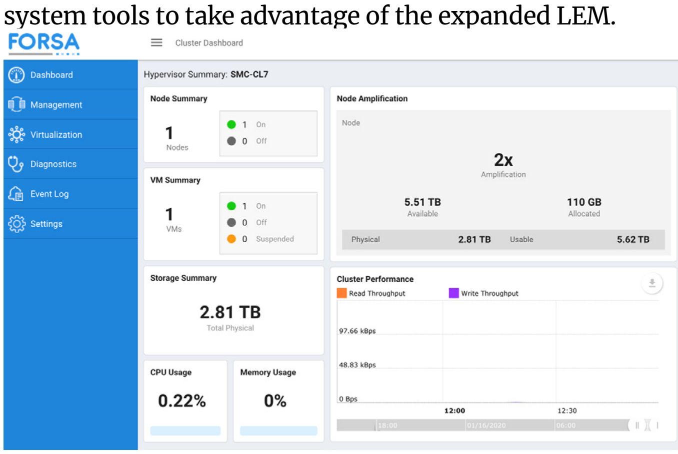

Figure 4: You can manage Forsa's in-memory storage devices (LEMs) via GUI and RESTful APIs.

In the event that the DRAM or SCM capacity of a single physical server is of insufficient size for your LEM, you can use the DRAM or SCM on a second node to create an aggregated LEM, which combines the resources of the two nodes to create a single LEM.

All of the power Formula Black delivers via FORSA would be greatly hampered if it was difficult to set up and use, fortunately, FORSA has an intuitive GUI that can be used to monitor and manage it (**Figure 4**). All FORSA functions can be accessed via APIs for those that need or want to mange FORSA functions from the command line or programmatically.

CHANGING THE GAME

Formulus Black has changed the storage game with FORSA: It's created a software stack that enables DRAM or SCM to be provisioned and managed as high-performance, low-latency in-memory storage media for the most demanding workloads.

Furthermore, the company recognized the needs of enterprise customers and have put techniques in place that allow these memory-based block devices to offer enterprise levels of data resilience, efficiency, and integrity.