IDC TECHNOLOGY SPOTLIGHT

Using NVMe Performance to Enable Storage Infrastructure Composability

October 2017

Adapted from *End-to-End NVMe-Based External Enterprise Storage Makes Its Debut* by Eric Burgener, IDC #US41952816

Sponsored by Kaminario

NVMe is a new storage technology that offers significant performance and efficiency advantages over the reigning SCSI protocol, but its use in external storage arrays to date has been limited. Real-time big data analytics and some high-performance database environments need NVMe performance today, but for most mainstream enterprise workloads, the need for NVMe on an individual application basis is still in the future. However, there is one use case for mainstream workloads where NVMe performance delivers significant value for workloads that do not individually require NVMe performance — the area of high-performance data mobility to enable storage infrastructure composability. In today's business environment, there is a strong need for enterprise infrastructure to be more agile, and storage agility has been limited by hardware-defined designs and network performance. The emergence of NVMe over Fabric (NVMf) provides a viable foundation for the emergence of a truly composable storage infrastructure layer. Storage infrastructure composability promises to change how IT organizations think about provisioning storage in much the same way that server virtualization changed the way in which servers are provisioned. This paper examines storage infrastructure composability as enabled by NVMf, calling out the benefits that information technology (IT) organizations can garner in the near term from this high-performance networking technology. It also looks at the role of Kaminario in enabling this new level of storage composability within enterprise infrastructure.

Introduction: The Composability Push

Today's dynamic business environment requires rapid response from supporting enterprise infrastructure and penalizes those that cannot provide it. The need for a much more agile enterprise infrastructure is a key requirement driving the digital transformation most enterprises are currently undergoing. For IT organizations supporting dynamic businesses, this means, among other things, the ability to quickly provision or deprovision servers and easily grow the resources associated with those systems as needed.

The concept of open systems composability was first introduced with server virtualization back in the early 2000s, but while this provided the agility needed for compute, supporting storage infrastructures did not provide this same level of flexibility. IDC sees a future in which all resource components of the enterprise infrastructure — compute, memory, storage, and networking — become equally composable, allowing customers to dynamically create servers that meet their requirements from logical pools of resources that are not limited by physical constraints. New technology will need to be developed across several areas to comprehensively bring composable/disaggregated infrastructures (C/DIs) to fruition, and that is easily four to five years out.

The ultimate goal for the software-defined datacenter is the composability of all resources, but it clearly acknowledges that improved composability in any one of the underlying components — compute, memory, storage, or networking — moves the customer in a good direction. IDC agrees that better composability of storage resources on their own will improve the agility of enterprise infrastructure, even if all resources do not offer that level of flexibility. To make storage composable across racks within a single datacenter, a feature that maximizes its value to the enterprise, there is a need for a low-latency interconnect that offers high bandwidth and throughput to maximize data mobility.

This is where NVMe technology — in particular NVMf technology — plays a critical role. The performance of NVMe opens up the opportunity to improve storage resource composability, but that foundation technology must be supplemented with other intelligently thought-out features that turn NVMe performance into a business solution that provides value even when most individual enterprise workloads do not specifically require that type of performance on their own. The key in this latter area is an orchestration layer.

Definitions

NVMe is more than just another I/O interface. Designed specifically for persistent storage built from nonvolatile memory media, NVMe is a storage protocol that can be used for device interfaces, array backplanes and controllers, and host connections. Today, there are NVMe solid state disks (SSDs), NVMe controllers and array backplanes, and NVMf connections utilizing Ethernet and Fibre Channel (FC). As an alternative to SCSI in enterprise storage arrays, NVMe is a much lighter-weight protocol that is specifically designed for flash rather than spinning disk media, and it supports orders of magnitude better parallelism, a critical capability for getting the full performance out of today's multicore CPUs and high-capacity SSDs. NVMe is much more efficient than SCSI, leading to much higher resource utilization in systems. IDC expects the protocol to become mainstream in enterprise arrays in 2021.

Benefits of Dynamic Storage Infrastructure Composability

Dynamic storage infrastructure composability provides the ability to deliver resource allocation and expansion nondisruptively and on demand and with zero physical manipulation of hardware components. Systems and/or additional storage resources can be immediately provisioned either manually or based on policies, ensuring that storage resources are easily and reliably allocated, managed, and protected according to governance guidelines. This provides the flexibility to increase overall system availability and minimize the performance impacts of failures, allocate resources on demand to handle daily and/or seasonal workload balancing requirements, and easily deliver different service levels, all by pulling from a centralized storage resource pool, which guarantees a very efficient use of available storage resources.

NVMe performance, particularly across the fabric, is what enables this to occur across racks within a datacenter. NVMf's ability to deliver high-performance data mobility to enable on-demand array composability provides significant flexibility benefits with today's mixed 3rd Platform computing workloads. With NVMe as the foundation, virtual arrays can be sized and configured to meet the performance, availability, security, cost, and other requirements of individual applications, all from a single storage infrastructure that is centrally managed. This simplifies the management of physical storage infrastructure, delivering benefits to customers whose applications may not necessarily require the performance of NVMe on an individual basis. It also ensures that NVMe performance is there if customers need it (for example, to add real-time big data analytics or other applications in the future that may individually require the high-end performance of NVMe).

Key Trends in Enterprise Storage

A number of macrotrends are impacting enterprise storage. The pace of business evolution is becoming increasingly dynamic, driving a strong need for agility in IT infrastructure. Physical enterprise infrastructure has largely been replaced by much more agile virtual infrastructure. Virtual infrastructure enables denser workload consolidation that utilizes IT resources much more efficiently, and the resulting I/O profiles have driven a rapid penetration of flash as a persistent storage medium for mainstream enterprise workloads. Existing SCSI protocols do not make very efficient use of flash performance, a factor that is leading to the rise of NVMe. As part of the rise of virtual computing, the hardware-defined storage architectures of the past are slowly giving way to the more flexible software-defined architectures that will form the foundation of composability. Hybrid cloud computing and composability both benefit significantly from fast, efficient data mobility, driving the need for higher-performance networks. For data mobility within a datacenter, NVMf offers a standards-driven, high-performance option that is compatible with the dominant host connection protocols in datacenters today (Ethernet and FC).

Considering Kaminario

Kaminario is an enterprise all-flash array (AFA) vendor that started shipping its flash-based primary storage platform — the K2 — in 2012. Kaminario's K2 array leverages a software-defined storage design, running on top of a scale-out storage architecture, that includes a comprehensive set of enterprise-class data services and scales to millions of IOPS and petabytes of storage capacity. K2 nodes, which include both controllers and SSDs, can be clustered to scale performance and capacity. Identified by IDC as one of the top 6 AFA vendors by revenue in 2016, Kaminario has achieved success selling to enterprises as well as cloud providers and today generates most of its revenue from software-as-a-service providers.

In August 2017, Kaminario announced its next-generation AFA for mainstream computing workloads called the K2.N, which is its NVMe-based system. The K2.N represents an evolution of the K2 hardware approach — it completely disaggregates compute nodes (i.e., storage controllers) from storage nodes (just a bunch of flash [JBOF] shelves). With the K2.N, compute nodes are called "c.nodes" and storage nodes are called "m.nodes." The K2.N offers higher storage performance density and reduced power consumption relative to older K2 models and enables customers to achieve denser workload consolidation as well as support more performance-sensitive big data analytics workloads.

Kaminario is taking an evolutionary approach with NVMe technologies, making it easy for existing customers to nondisruptively upgrade to full NVMe performance without application disruption. K2.N nodes can be added to an existing K2 cluster, and workloads can be nondisruptively migrated to the newer nodes. Older nodes can then be removed (or kept if the customer wants a cluster with tiered performance). Both system types run the same VisionOS and are administered from the same centralized management interface. Initially, K2.N nodes are connected to hosts over SCSI-based FC and/or Ethernet but ultimately will support NVMf connections. Older K2 nodes will not support an NVMf host connection capability.

The K2.N provides a strong foundation for another new capability — Kaminario Flex — that breaks important new ground in the realm of storage infrastructure composability. Kaminario Flex is a software-based storage orchestration platform for managing and orchestrating composable storage infrastructure that runs on top of K2.N. Working through the Flex dashboard, customers can dynamically compose virtual private arrays (VPAs) using a mix of available K2.N resources with no physical reconfiguration. These VPAs can pull from any mix of controller, disk, and host connection resources to support different performance tiers. VPAs can be nondisruptively scaled, in terms of both controller performance and storage capacity, and VPAs can also span physical racks given that those racks are connected over the same NVMe over Ethernet implementation. VPAs can be created on demand based on pre-established templates, data services from Kaminario's mature VisionOS operating environment (including quality of service and Kaminario Clarity cloud-based predictive

analytics features) can be applied, and VPAs can be easily integrated into datacenter workflows using Kaminario's rich set of APIs. Once VPAs are deprovisioned, all resources are returned to a common pool, and all storage is automatically reclaimed.

While the need for NVMe performance for individual workloads is still limited, it will clearly grow over time. But Kaminario's application of NVMe performance to data mobility, along with the orchestration layer to manage composability, not only enables a flexibility that is sorely needed in mainstream mixed workload environments but also is unique in the industry. Kaminario Flex supports a new availability model as hardware failures can be detected and replaced completely within the software. New tiering models supporting a dynamic allocation of NVMe disk resources on a per-array basis allow customers to more efficiently allocate performance resources exactly as needed. VPAs can very efficiently be specifically sized and dedicated to individual workloads, and additional resources can be added to meet short-term surge requirements with an ease not possible in the past.

Challenges

In 2016, close attention to the needs of mixed enterprise workload consolidation brought the top 6 AFA vendors to relative parity with each other across many metrics. In an effort to differentiate themselves, vendors made forward-looking announcements about their intended use of NVMe technology to extend performance and further improve efficiencies. Vendors are touting higher throughput and bandwidth, improved storage density, and lower latency as the key benefits of NVMe deployment. Today, many enterprises are experimenting with local NVMe SSDs in commodity servers, and IDC research indicates that the top drivers of NVMe adoption are improved scalability to meet performance requirements over time, better $/IOPS, and higher throughput or bandwidth. The use of NVMe to improve data mobility, thereby enabling composable storage infrastructure, is not even on the radar of most enterprises. While the penetration of more real-time big data analytics workloads in particular will drive an increasing need for NVMe, a relatively small percentage of customers have deployed them in production today.

Leveraging NVMe performance to support better data mobility delivers value that is much more broadly applicable even when individual applications do not require NVMe performance. Kaminario will need to evangelize the benefits that storage composability at this level brings to virtual administrators looking for improved flexibility and easier management across mainstream workloads — virtual administrators whose view of storage composability is limited by their experience with popular server hypervisors. The change in customer understanding will need to follow that which occurred with the use of snapshots in datacenter workflows as the industry moved from HDD-based snapshot implementations to flash-optimized snapshot implementations, a change that opened up significant improvements in space efficiencies and ease of use for customers willing to reexamine their stance on snapshot usage.

Conclusion

It is IDC's view that the industry is slowly evolving toward composable/disaggregated infrastructures that will eventually encompass all resource types — compute, memory, storage, and networking. Server hypervisors, by offering compute composability, have significantly improved the efficiency of enterprise computing, but composability for the other resource types has lagged. In an innovative application of NVMe technology, Kaminario has announced its intent to deliver a fully composable enterprise storage infrastructure layer, called Kaminario Flex, that promises to optimize the agility with which storage resources are allocated, expanded, and/or deallocated. This approach will improve service and availability levels, allow storage to be more closely tailored to the exact needs of individual applications, and allow for much more efficient use of storage resources unfettered by the limits of physical configuration.

To fully utilize a highly composable storage infrastructure layer, customers will need to evolve their mindset around what is possible with storage. In the past, centralized storage was a shared resource that was difficult to tune for optimal performance in mixed workload environments. A storage infrastructure layer flexible enough to enable the on-demand allocation of storage resources in the form of virtual private arrays on an individual application basis will meet required service levels while using those resources much more efficiently. What is interesting about Kaminario's application of NVMe technology here is that it leverages NVMe performance in a way that is meaningful for all applications, not just those that require NVMe performance and/or throughput on an individual basis. Kaminario customers can implement NVMe technology when the capabilities start to ship in early 2018 and garner immediate benefits while positioning themselves well for a future that will require NVMe performance for an increasing number of application workloads.
