Understanding All-Flash Architectures
Assessing the Right Fit for Your Data Center and Workloads

Prepared by: George Crump, Lead Analyst
Prepared On: August, 2014
All-Flash Architecture Matters

Today in the all-flash array market there are two architectures for data centers to choose from; scale out and scale up. Both have their unique strengths and weaknesses. Deciding on the right architecture for an organization’s data center first depends on identifying what its primary needs are. The problem is that needs often change but, in large part, storage architectures do not.

Most data centers purchase an all-flash array to solve a specific problem like a poor performing database application or performance issues in a virtual desktop or virtual server environment. It is when other workloads are added to the all-flash array that the weakness of the architecture and key data services can be exposed. Such workload consolidation, with its benefits in infrastructure savings and business agility, is one of the major ways that all-flash arrays can transform the data center.

What is Scale Up All-Flash

Scale up architectures are typically based on a single system with dual controllers. All of the performance capabilities are bought and paid for upfront. They are the ideal point performance solution because an organization can typically start small and add capacity via storage shelves as it is needed.

Scale up storage systems can offer excellent price-performance on a per-workload level. It's not uncommon for some scale up architectures to deliver ½ million IOPS in a single system. If an environment or application needs IOPS for a single, specific application workload, then scale up architectures with data services like data reduction can be a cost-effective way to get there.

A problem with the scale up architecture, however, is that after this initial phase of all-flash adoption is complete and the pressing performance problem is solved, IT planners often look to add additional workloads to better leverage the all-flash investment. For example, many early adopters for VDI add other virtual server applications, and many production database deployments are later consolidated with lifecycle and decision support instances. As a result, there is a need to add storage shelves to the array. As those shelves are added, the performance of the all-flash array is diluted, since the fixed processing power of the controller(s) is spread across more storage. In addition, some all-flash systems are better at certain workloads than others, with benefits from data services like deduplication, compression, space-efficient snapshots, thin provisioning, and encryption.

If more raw performance is needed or workload specific performance capabilities are needed, the only way to get it is to either split the workloads and add another array or do a controller upgrade. In most data centers, adding another array is seen as the path of least resistance. Until, of course, the storage manager is stuck managing a half dozen storage silos.
What is Scale Out All-Flash

Scale out all-flash arrays are a modular cluster of storage servers that are typically called “nodes”. As these nodes are added to the cluster, they provide greater capacity and performance-scaling to well over 1 million IOPS. In this architecture, performance gets better as more nodes are added to the system. This is ideal as more workloads can be added as nodes are added. Such additional workloads need to tap into the data services and consistent performance potential of an expanded scale out platform for IOPS and capacity and bandwidth.

As a result, scale out all-flash arrays are an ideal architecture to address the long term data center-wide performance and consolidation needs, not just the point solution performance problem that motivates most organizations to look at flash in the first place. Just as virtualization has enabled mixed server workload consolidation and agility, such scale out all-flash platforms with data services can finally enable mixed storage workload consolidation.

Most providers of scale out storage systems will report on the total performance capability of their systems, taking great pains to show linear performance scaling. But to get this scaling and reach their reported numbers requires many parallel workloads like virtual servers or SAP Landscapes. To do this, it is important that the scale out architecture can fully leverage the scale out performance so that a single volume’s performance and data services can be serviced by all the nodes in the cluster.

There are two basic challenges with scale out architectures. First, while the architecture can grow quite large, it typically can’t start very small. Some require a minimum of three nodes. Remember that most all-flash arrays are purchased to solve the performance problem caused by a specific application, and often that application’s workload is much less than the total IOPS and capacity than an initial cluster can supply. This means that deploying even a small scale out all-flash solution will typically waste some resources until additional workloads can be migrated over to it. Some vendors are now offering entry single-node options that address this issue.

The other problem is one of node sprawl. As mentioned, a node provides performance and capacity at the same time, but in many situations only one or the other resource needs expansion. These resources don’t always need to be expanded in unison. In other words, when nodes are added to meet a specific demand, performance for example, the other resource goes unutilized; in this case, expensive flash capacity. This leads to node sprawl where nodes are added but their capabilities are not fully leveraged. One way this can be addressed is by offering different node sizes to better match the expected capacity with the performance requirements of the workloads (scaling deep). It is important that the scale out architecture be able to fully tap both the capacity and performance of each node before it is expanded. It should also allow for nodes to be added that are not at their full capabilities initially.
Next Generation All-Flash Arrays Scale Deep and Out

While both architectures have their place - scale up for point specific performance problems and scale out for mixed workload consolidation and SLA performance - the data center needs an architecture that can do both, something that we call “scale smart”. A scale smart architecture should allow a data center to start with one-node, which requires it to have the redundancy of a stand-alone storage system. This initial node can be sized with the “scale-deep” capacity that is appropriate for the initial workload.

Once the first node has reached its capacity and performance limits, IT planners do not want to be stuck with the scale up problem of having to buy another system and creating another point of management across separate silos. Instead, the storage system should be able to incrementally add nodes (for performance or capacity) as needed. These additional nodes should not be separately managed systems, but be joined with the first node. The data on the first node should eventually be striped across all the nodes to present a single storage pool with a single point of management, on which all the nodes can participate equally when responding to application I/O requests and all data services globally across the cluster.

It is also important that the system be able to handle a variety of workload I/O patterns. Because the scale out storage system enables the consolidation of multiple workloads in the data center, the all flash array has to be able to handle large block, small block, random and sequential I/O patterns simultaneously. For example, many customers are mixing OLTP and OLAP workloads on a common storage cluster.

Conclusion - All-Flash the Way YOU Buy it

Such next generation all-flash architectures can finally enable broad storage workload consolidation and agility for use cases like real-time analytics. This also means flash can be purchased in a more logical way and match the scale deep and scale out capacity and performance requirements to the expanding demands. Until now, most flash purchases are in response to a specific I/O problem. Unfortunately, in many instances they are not strategic, but reactionary. The fact is most data centers are stretched so thin from a staffing perspective that they can only react to performance problems instead of proactively solving them. And that’s part of the reason that scale up all-flash arrays are finding success in the market.

A next generation architecture allows storage planners to respond to the immediate crisis with a single node approach. But once the crisis has passed and it makes sense to consider placing other workloads on the all-flash investment, then a system can expand into a scale out platform to meet growing capacity and performance needs. But it does so with extreme efficiency, making sure that each node has balanced utilization levels across both CPU and capacity.
resources, prior to expansion with all data services like deduplication and compression. This helps eliminate node sprawl. Finally, the system is able to handle the wide variety of workloads, and their associated I/O patterns with consistent low latency, that are sure to occur as the data center moves to a single flash array for all production data. across consolidated workloads like database, analytics, SAP, application-as-a-service, and other cloud services.

This independently developed document is sponsored by EMC. The document may utilize publicly available material from various vendors, including EMC, but it does not necessarily reflect the positions of such vendors on the issues addressed in this document.