EMERGING TRENDS IN SOFTWARE DEFINED STORAGE

OCTOBER 2013

THE NET NET

"Software defined pretty-much-anything" is getting a lot of attention in the IT trade press. It seems new and different and promises to solve everyone's problems around prioritizing applications, sharing storage across massive distributed environments and optimizing computing resources around business data needs.

It's not quite that easy, nothing is. Software-defined technologies are in fact available today in real products. But software-defined storage, or SDS, has lagged behind developments in software-defined data center and networking domains. This is largely due to the storage process's tight confinement within the underlying physical system with its controllers and media. Virtualization first helped to loosen that hold as it logically presented available storage to create virtual storage pools.

SDS takes another step beyond that: it can abstract storage provisioning, storage management, and I/O optimization from the physical controllers to a virtualized console for unified administration. Not all SDS products do all things, but IT can choose the feature set that best benefits storage infrastructure: to unify heterogeneous storage systems, or to enable cost-effective scale-out, or to match application needs on-the-fly to their storage targets. Working products are available right now that accomplish one or more of these critical operations.

Software-Defined Storage: What Is It?

Our definition: Software-defined storage (SDS) decouples storage management programming from underlying storage hardware. The resulting control layer may be software or a virtual appliance. This layer sits on top of the physical storage stack, abstracts storage management from the physical storage layer, and converges it into a highly manageable interface. This enables IT to centrally provide application provisioning, policies, data protection, storage pooling, and reporting instead of from array-specific interfaces. Some SDS solutions also identify application I/O patterns and optimize I/O accordingly, targeting it to the optimal storage target. Some create a virtual pool of storage so that data is stored in a manner that is abstracted and free from the mechanics of underlying storage.

SDS makes application provisioning simpler by observing capacity, availability and performance needs. Storage control is usually highly automated: the more automated the layer is the better for dynamically serving applications and workloads – especially ever-changing applications and workloads in the virtual infrastructure.

In the name of easy scalability, some SDS products take a commodity storage/building block approach to scale-out. A word of caution here: no matter how virtualized or abstracted the storage system may be, at its base it remains a physical entity. Just because something is "commodity storage" does not make it an ideal choice in an SDS infrastructure. Performance does not magically increase along with capacity: storage systems are still subject to performance lags due to slow physical controllers and less-than-optimal fabric. How well the SDS layer works depends on how effectively the underlying physical storage works. Some priority applications still require high performance, highly available storage systems at the hardware layer of an SDS infrastructure.

Different Approaches

We identify three major SDS approaches in the market today: orchestration (control layer), encapsulation (virtual storage appliances), and server-side (virtual controllers).

- **Orchestration: Control layer.** This layer sits on top of physical storage systems from one or more vendors, abstracts storage management tools from the arrays, and delivers them as management tools running from a central console. IT schedules and launches storage management across heterogeneous arrays and can create virtual storage pools from attached systems. The physical data paths remain in place. Solutions are emerging from several different sources: storage virtualization vendors, large storage vendors, and startups. It is a promising approach for adding intelligence to a sprawling infrastructure, but in reality this approach takes a lot of API engineering to centrally manage heterogeneous storage systems. EMC's ViPR is making a splash and the category also includes central control products that have been around a long time such as IBM SVC and software-only DataCore.

- **Encapsulation: Virtual storage appliances (VSAs)**. VSAs typically sit on top of a physical storage infrastructure and enable various storage operations across the virtual environment. This is a very wide definition, too wide to automatically include VSAs into an SDS framework. A number of them do fit our definition if they abstract and virtualize data management from underlying storage and optimize virtual applications by matching storage availability, performance and capacity to the VMs. VSAs are here now and are a practical enablement of software-defined storage that can add agility to a data center infrastructure. VSAs are probably the largest part of the SDS market in terms of vendor count. A few examples include HP StoreVirtual VSA and VMware vSAN.

- **Server-side: Virtual controllers.** The above two approaches are common sense evolutions of virtualized storage management and virtual appliances, and make sense in an SDS definition. However, SDS is also spurring new approaches. One of these is a recently emerging architecture that uniquely distributes specific functions for a new approach to both software definition and the physical storage of data. This SDS class offers a software layer comprised of virtualized storage controllers. The virtual controllers are deployed on each virtual or physical application server. A control layer gathers information from the controllers and assigns policies to each controller, allowing IT to assign different processing policies per applications. By putting the virtual controllers at the workload layer, distributed controllers can optimize management for different needs. This class may run on commodity storage but is also well-suited to high performance computing environments.

Let's take a closer look at the last type of SDS platform. We consider the virtual controller approach to be a hybrid approach that maintains the best features of the control and data planes. It abstracts storage management tools to a control layer located next to the workloads. The virtual controllers act as multiple distributed instances to manage and optimize the I/O stream. These software controllers tune I/O performance at the server level, which minimizes latency and accelerates storage performance. A central console manages all virtual controllers and automates priority queues to optimize application performance. The underlying physical layer becomes a managed storage capacity pool matched to application I/O needs.

A leading example of this architecture that emphasizes massive grid scale-out and performance is Gridstore. Its highly scalable grid delivers SDS performance and capacity for high performance and highly available computing needs.

Benefits of Virtual Controllers

We perceive three major benefits of the Virtual Controller SDS architecture: 1) overcoming the limitations of physical storage, 2) optimizing storage performance, and 3) providing a flexible response to dynamic

application environments. Ultimately these benefits of SDS are about one overarching consideration: consistent operational and storage Quality of Service in the dynamic infrastructure.

- *Efficiency: Overcoming the limitations of physical storage.* SDS helps to overcome the inherent limitations of physical storage by adapting storage performance and capacity on demand, and by enabling better mobility and scalability. SDS deploys across a broader assortment of storage resources while greatly simplifying storage administration. When an application server is retired so is its virtual controller, and there is no need to reassign or re-provision the underlying storage resources.

- *Utilization: Aligning application I/O with storage resources.* Some SDS platforms add storage intelligence to the workloads, which optimizes I/O from the point where it enters the network. This intelligence drives higher performance and matches application needs to storage resources. SDS accomplishes this in different ways: in some cases, this might be through localized SSD and auto-tiering technologies, or it might be through storage control or QoS algorithms.

- *Quality of Service: Improving QoS throughout infrastructure.* Flexibility, storage intelligence, mobility and scalability are all big issues in achieving QoS in growing data storage environments. SDS can go a long way towards solving these problems even in its first generation. This is particularly true in SDS platforms that optimize application I/O and match them to appropriate storage resources. This ability improves QoS throughout the application/networking/storage infrastructure.

Taneja Group Opinion

Storage virtualization has always lagged behind advances in virtual servers and networking, largely because of its strong ties to physical controllers and storage media. Virtualization enabled IT to construct logical views of physical storage but most data management and provisioning is still done at the physical storage system level.

But continuing development has made software-defined storage a reality where management, provisioning and workload optimization are happening at a level above the storage. There is no SDS product that does it all since we are still in an early stage of software-orchestrated storage management. But today, right now, SDS technologies like orchestration layers, VSAs and virtual controllers are making it possible to detach data and I/O management from the siloed physical systems.

We are closely watching SDS in all its incarnations. By severing storage interaction from the physical storage location, SDS has significant potential to improve QoS throughout the storage infrastructure. Ultimately SDS is a means to provide new life and higher functionality to storage environments than was previously possible, making it a very interesting market to watch as it innovates and evolves.