Will software-defined storage impact on data protection?

By Bryan Balfe   27 May 2014

"Software-defined storage" is the new buzz phrase in the cloud world and, as usual, vendor after vendor is jumping onto it. In the case of storage, you have "system" vendors on one side looking to defend their proprietary stance whilst newer, software-only companies look to turn storage hardware into a commodity. For those with an interest in data management, it begs the question: Will software-defined storage replace data management?

Before we can decide that, we need to get a definition of software-defined storage (SDS) itself and because it's a new field that's currently more of a direction than actual product, that's where it gets tricky. Analysts such as Gartner and IDC have their definitions and vendors with an interest in this space have theirs, each coming from a different direction. For the most part, it's defined as automation that can set the appropriate storage volume size, performance, redundancy and access required by workloads or applications based on policy.

The bottom line is that it enables data mobility to match the virtual server technology using it. It may even define some levels of protection, such as replication and snapshots, and can blur into wider orchestration because it still relies on networks and servers to work - doing all three adds up to defining much of datacentre with software - a panacea for many organisations that have invested heavily in cloud technologies.

Defining data management

The next challenge is defining data management and where it overlaps with the ambitions of SDS. Some hardware vendors would have you believe that snapshots and replication are as far as you need to go, and while these companies have some very capable (and important) tools in this space, snapshots and replication clearly don't answer the challenges of data growth, long-term retention, access and compliance. The current thrust of SDS development is following this same path, so in the future it may be possible for SDS technology and commodity storage to offer similar functions to proprietary set-ups, but at lower cost, of course. This presents the market with more choice and will ultimately help to lower platform costs.

We can safely say that, at least in the early days of the technology, the real overlap between SDS and data management will be firmly in the traditional hardware vendor space in most cases and, additionally, where virtualisation typically plays with the storage and protection mechanisms too. However, handing snapshot control to a data management software layer at the point that snapshots are created has a huge number of benefits, just one of which is an element of custody.

What does the issue of custody of the data mean? Custody normally refers to the owner or manager of the data in personnel or departmental terms. Here, I'm referring to data management software and I've specifically used the word custody because it alludes to knowledge of the original context of the data (what it is, its value etc.) and what happens to it

down the line, such as where it goes and the security at rest and in transit. For example, knowing a snapshot of a virtual server contains an Exchange email store means you can process it in a specific way that streamlines tasks such as indexing and deduplication and also allows for one-stage granular restore. This last point is important as many storage admins are left to recover whole volumes to scratch disk before manually sieving and plucking out the required mail or attachment if the data no longer resides on a disk snapshot.

Other functions

There are also other functions that can be carried out from snapshot copies, such as archiving. Most people associate this as a function acted on by the primary system via an agent, but there are a number of advantages of doing it in the protection tier instead; avoiding disk intensive and lengthy file scans on the primary storage being just one of them.

What happens after the snapshot is important and while no one in the SDS world is suggesting that SDS technology will go much beyond that currently, data management processes that are initiated after the snap will always have a disadvantage over data management software that looks after snapshot initiation as well.

So is control of the snapshot the battleground between SDS and data management software? It could be in many cases, but it need not be in all cases. When you analyse what's actually going on, there are a number of distinct functions taking place that have a natural division of labour. Most of the snapshot engines today create only "crash consistent" snapshots, whereas data management software combined with the same engine can manage application data to guarantee its integrity and, ultimately, its recoverability, as well as add value with functions like indexing.

There is no doubt that software-defined storage will have a growing impact on primary storage flexibility and some first-line protection processes, which will worry many array vendors. The benefit of orchestrating an abstracted heterogeneous pool of storage in an automated fashion, based on rules with one foot in the business, is an attractive one. The fact is that it's not a reality yet and many companies with big budgets are working to actively hinder its progress, or at least keep you locked in to their flavour of it.

The orchestration goals of SDS may still be some way off, but the control of heterogeneous storage snap engines and automation based on classes of service and workloads is already here. Combining that level of management with the ability to maintain the context of the original data and move it through different tiers based on policy is very powerful indeed.

ABOUT BRYAN BALFE

Channel Manager at CommVault in South Africa
▪ Understanding data sovereignty and its importance to your business - 22 Sep 2015
▪ Controlling data risk in the BYOD onslaught - 19 Aug 2015
▪ What enterprises can learn from the MSP revolution? - 13 Aug 2015
▪ Warning signs your archiving strategy is not geared for the future - 24 Jul 2015
▪ Driving forces behind cloud computing - 22 Jun 2015

View my profile and articles...

For more, visit: https://www.bizcommunity.com