Backup Virtualization Brings Flexibility to Disk Backup

By Ed Ahl, Tributary Systems

Abstract: The value of Backup Virtualization, like that of server virtualization, is that it builds an abstraction layer between software and hardware. In the case of Backup Virtualization, this abstraction layer separates the backup process from the physical disk and tape backup devices. Backup Virtualization brings flexibility to disk backup, where it may be needed most.

Nowhere are there more options than in disk backup. There are devices that are just inexpensive disk systems; devices that deduplicate; devices designed for high performance, and devices designed for maximum power management through the use of MAID technology. But few disk backup systems provide all of these capabilities in one unit. Thus, disk backup systems are implemented to solve a particular objective, which changes over time. As a result, in many environments multiple disk backup targets are managed within the same backup process.

Evolving Disk Backup Complexity

For example, what could start as a need for medium-term retention (cost effectively handled by a deduplication system where performance is not a big issue), could evolve into a need for high performance backup of specific mission-critical servers. For that task, a separate, high-performance, cost-effective disk array is often added to the backup infrastructure, yet it may lack the deduplication capabilities of the former system. Then, power availability could become an issue. Instead of using tape, the decision could be to use MAID technology for long term retention of backups. Again, another system would need to be purchased to store those backups.

Not only do the capabilities of each disk backup device vary, but so does the scalability. While some systems can scale by adding nodes, these often come with a higher upfront cost. This leads to the dominant disk backup system being supported by a simple dual-controller disk array with a set capacity ceiling (which is how multiple drives can be squeezed into the case). When one of these devices fills up, the options are to lower retention times, add an entirely different system or, most commonly, add another device of the same type. The result is multiple disk backup systems all of the same type, bought multiple times to meet capacity demands, each required to be managed separately within the backup application. The management of all of these targets is further compounded by the fact that multiple backup applications exist in the environment. The whole effort to improve the backup process by implementing disk quickly becomes a nightmare.

Backup Virtualization Manages Disk Backup Complexity

Backup Virtualization helps resolve these issues by abstracting the backup application from the multitude of targets in the backup infrastructure. When Backup Virtualization is

used to manage the above process, the result is operational simplicity and better utilization of the hardware investment.

All the various disk targets, as well as any tape targets, are connected to the Backup Virtualization appliance. The backup applications all point to the single Backup Virtualization device that appears to be a tape library, albeit virtual. The Backup Virtualization appliance becomes the single point of management that controls, via pre-set policies, how and where data is stored on the available physical devices.

In the evolving infrastructure scenario above, Backup Virtualization can accommodate all of the listed devices. For example, the high-speed disk could be used as a caching area to which all backups are directed. These backups, as they complete, could be shifted to the deduplication system for medium-term storage. Then, as the backup ages it could be moved to the MAID system for better power efficiency, and then finally to an available tape system for long-term retention. Mixing disk and tape provides the added comfort of having data available on two different forms of media (disk and tape). It is important to note that the movement of data is all done without the involvement of the backup application or the resources of the server it is running on. The Backup Virtualization Appliance handles and manages all data transfers while keeping the data formats compatible with the individual applications.

In the case of dealing with multiple units of the same type that had been purchased to meet capacity demands, Backup Virtualization can automatically handle this challenge as well. The multiple disk backup systems can be connected directly to the Backup Virtualization appliance, and as one fills up, data can flow to the next available device. When there are no more devices, another unit can be added and backup can automatically start using it. Meanwhile, the backup application continues to write data to the same original backup appliance which continues to appear to be a standard tape library.

These two challenges – multiple disk backup systems of the same type and multiple backup systems of different capabilities – are often both found within the same data center. Again, Backup Virtualization can solve the combined problem as easily as it can handle the individual problem. It also addresses the final challenge of determining where tape fits in. Although many consider disk the primary backup target, tape's cost and capacity advantages cannot be ignored; tape will remain in most data centers. Plus, tape provides the additional comfort of having data protected on two entirely different forms of media.

Backup Virtualization brings more to the backup process than just herding together the various options in disk based backup and addressing the scaling requirement. It enables "non-disk friendly" applications to leverage disk to a fuller extent. Backup Virtualization also improves the tape drive backup and recovery performance issues that lead to so many disk backup projects in the first place. No matter how many devices are in the backup process, Backup Virtualization can simplify operations, increase resource efficiency and reduce hard capital costs. **<End>**