hp virtual array technology and SAN virtualization

executive summary

As storage continues to grow in both its importance and its quantity, IT managers have come to the painful realization of the impact that this growth has on their bottom line. The unfortunate paradox is that as the cost per byte of storage decreases the cost to manage the total storage has overtaken much of the savings. This, compounded with the industry's finite supply of skilled and trained IT personnel, has created a serious problem for the IT manager.

The emergence of SAN technology presents itself as the savior for storage management; however, the current realities are that SANs are more of a contributor to the current storage dilemma than a solution. Even if you overlook the differences between the vision of SANs and its current reality, the vision falls short by ignoring the details of the storage device. Most storage area management systems overlook the complexities of managing the storage device in terms of capacity and performance management. However, storage area management systems do provide an attractive management paradigm for aggregating many devices and servers.

Hewlett-Packard understands this problem – and its solution - storage virtualization. HP's Federated Storage Area Management (Fsam) strategy addresses this issue by allowing a multi-tier, hierarchal storage virtualization. HP's Fsam extends beyond other proposed SAN management solutions, where the storage network level virtualization is the key to creating and managing large storage pools. Fsam adds, with the hp SureStore Virtual Array, a second tier of virtualization within the device to address the low-level device management issues. This combination provides a powerful solution with capabilities that exceed SAN only storage management solutions.

introduction

This paper will discuss how the VA7000 series virtual array uses virtualization to allow the SAN's view of storage to be independent of physical disks. In so doing, the VA7000 series also dramatically reduces the amount of information that must be provided to configure and manage the array. This simplification makes the VA7000 series very easy to manage by administrators and SAN virtualization software alike. The VA7000 series also enforces security with customized views of storage for different hosts. Flexibility is also improved as the VA7000's internal data movement capability makes it very responsive to changes in the SAN around it. We will show how the virtualization in the VA7000 series reduces the cost of SAN virtualization, by making more efficient use of storage capacity and decreasing total data movement requirements of SAN virtualization. But, let us first lay some groundwork with definitions.

virtualization defined

There are many ways to define virtualization, SANs, and virtual arrays. The following definitions are chosen to be specific and useful in understanding multi-layer virtualization.

(Disk) Array	A group of disk drives and at least one controller that stripes data and redundancy across the drives according to RAID algorithms.

Virtualization	Presentation of a simplified view of storage that is derived from complex lower level views.
SAN (Storage Area Network)	A network of storage devices and host CPUs that allows block-level storage sharing.
Host CPU	A computer system that consumes storage resources typically by running file systems, databases, and/or applications.
Virtual Array	A disk array with virtualization capability beyond the minimum necessary to implement RAID algorithms.
Virtualization in the SAN	Virtualization capability that is neither in a host CPU nor in an array. This includes block storage servers in SANs.

The above definition of a SAN emphasizes the sharing of block-level storage. The tools of virtualization contribute primarily to the block level aspect of storage networking. The definition of a virtual array allows a wide range of virtualization features, ranging from the creation of multiple Logical Units (LUNs) in a stripe group to the ability to relocate data, based on usage.

a system of virtualization

Today's storage systems generally include two to four levels of virtualization spanning from file systems to disk drives as shown above in Figure 1. These layers were developed by various industry contributors to more efficiently manage mushrooming storage capacity. While each layer has some unique purpose and advantages, all of the layers share the purpose of making shared storage more manageable.

At the top of the virtualization stack, file systems and databases present sharable views of data to applications. The file system and database software at this level benefits from direct application coupling. The file system and database views can hide many underlying implementation differences. Although the remainder of this paper deals only with lower levels, keep in mind that the virtualization process begins at this uppermost level.

Block storage virtualization is common in host CPUs today in the form of volume managers and value added device interfaces. This software presents each host with the aggregated view of the storage that it is allowed to access. The volume manager creates logical volumes from either individual disks or the virtual disks that array controllers produce. The volume manager also controls the routing of block storage requests as they leave the host. Volume managers and drivers are responsive to changes in the host operating system (OS) and they represent the single point-of-contact for all storage clients resident on one host.

Virtualization in SANs consists of storage servers or directors combined with management software. This level extends the familiar functionality of volume management into the SAN where it is used to manage host views, share storage resources, and in some cases replicate data. This is the only level that

a system of disk storage virtualization

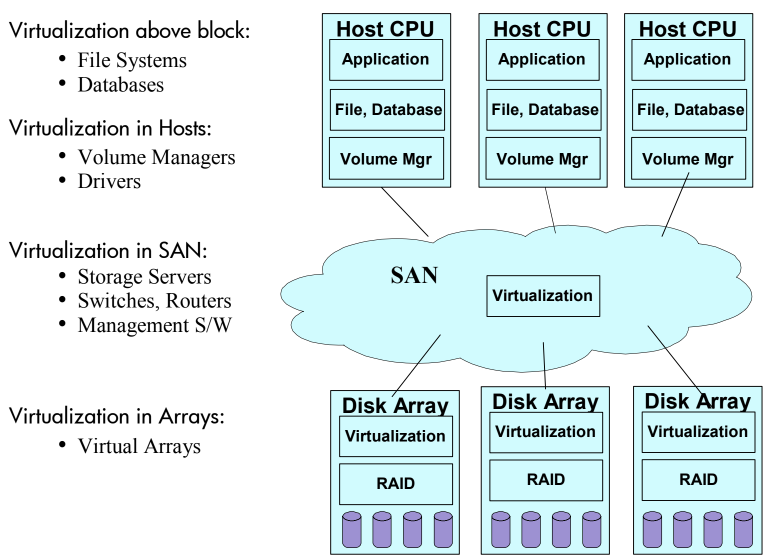


figure 1 - disk storage virtualization system

spans the entire SAN. In addition, virtualization at this level can sometimes avoid performance-critical data path intervention by managing data path elements that are already in hosts, arrays, or network infrastructure.

Finally, virtualization in arrays manages storage sharing and matches resource usage to access demand. The array level is the single point of access to any data below it. The array is extremely responsive and efficient in local management activities.

Virtualization in arrays is advantageous because it enforces security at the single point of access for each storage device. Throughout the system of hosts, fabric, and storage devices, sharing is a key goal. With this sharing, come increased security requirements. Virtualization contributes to security by allowing different views of storage to be presented to various hosts. Controlling security at the single points of access is important for managing security.

Responsiveness to change is another key goal of virtualization throughout the system. This responsiveness requires the ability to detect the need for change and to make an appropriately controlled change without disturbing the system any more than necessary. Various levels of the system respond best to different types of change depending on the detection, planning, and reconfiguration actions required. For example, the SAN has a broad view of system wide changes, whereas the array can respond to localized changes in resource usage very efficiently. Multiple layers of virtualization decompose the problem of managing SAN-wide change into bite-sized pieces.

Virtualization exists throughout modern computer systems, and will continue to exist as a layered system. Each layer, optimized for its function, works together with the other layers to simplify management, improve performance, and increase availability.

virtualization, aggregation and performance

SAN virtualization is often used to aggregate capacity and balance loads across multiple arrays. If the arrays are similar to each other in performance, then the result is a larger virtual device with uniformly upscale performance. However, if the underlying devices or storage schemas (RAID levels) have different performance characteristics, then less predictable performance will result with the aggregation. The unpredictable performance that comes from aggregating dissimilar arrays is not a function of the use or disuse of virtualization in the array. Rather the unpredictable performance is because of the different performance characteristics of the underlying arrays

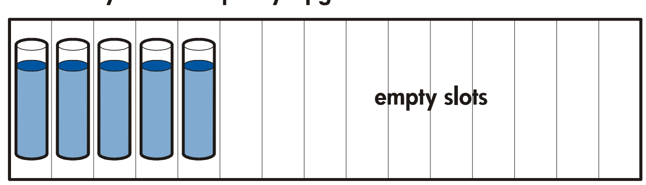
Predicting the result of most load balancing algorithms is difficult when the participants behave differently from each other. By contrast, if dissimilar storage is used in a hierarchical fashion within a SAN to reposition data from fast frequently used storage to slower, less frequently used, less expensive storage, then the performance of the SAN becomes more predictable. In such a case the storage is not really aggregated, but is used collectively to meet the overall storage requirements of the user community.

scalable performance and load balancing

Performance scalability of the SAN depends on the load balancing capability of the layered virtualization algorithms. If a load can be balanced within and across arrays, then no one component can act as a performance bottleneck. The volume manager or the SAN virtualization layers are effective at load balancing across the storage arrays, particularly where the arrays being balanced are similar in their performance as noted above. However, these layers are currently ineffective at load balancing within the array. Substantial capacity increases within a particular storage array, after the initial installation, may result in an imbalanced distribution of data among the disks in that array.

The VA 7000 series minimizes the potential imbalance of distribution of data by automatically re-striping the data across the new disks added to the array. This restriping assures a balanced distribution of data across all the disks in the array, including the recently added ones. This

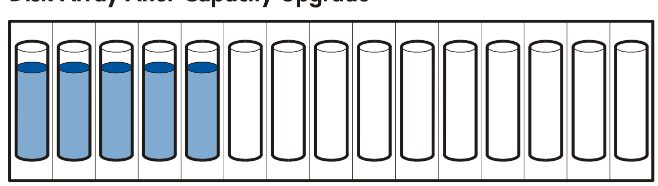
Disk Array Before Capacity Upgrade



- **Full disks.**

figure 2 - disk array before capacity upgrade

Disk Array After Capacity Upgrade



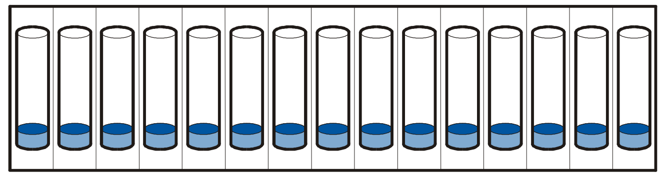
- **New disks are empty.**
- **Old disks are still full.**
- **Most of I/O load still goes to old disk because that is where the data is.**
- **Without virtualization in the array, manual intervention is required to redistribute or balance I/O load.**

figure 3 - disk array after disk upgrade

approach provides optimal data movement efficiency, since data movement does not consume SAN level bandwidth or server resources.

By using virtualization at the SAN level and off loading local array virtualization to the VA7000 series, IT managers can balance a moderate amount of optimization with high responsiveness to change at very low administrative cost. One might ask if independent optimization, at multiple levels, results in a non-optimal system. In practice, the answer to this question depends on the specifics of the virtualization algorithms, and performance characteristics of the devices and applications in play. HP's analysis and experience to date shows that in the great majority of cases the two levels of optimization complement, rather than conflict with each other. On the other hand, globally optimal solutions are very costly to achieve and maintain, even with the best tools. Significant bandwidth, compute time, and expensive administrative efforts are often required. Comparing the extreme on-going costs to achieve a hand-tuned global optimization of a storage system to the very modest cost of the separate SAN and disk optimizations, we expect many users will prefer the slightly less optimal choice at the substantially lower cost.

Disk Array After Capacity Expansion with HP VA virtualization



- **Data is automatically redistributed by the VA7000 array among new empty disks with no manual intervention.**
- **I/Os are balanced among all available disk drives.**
- **Redistribution occurs in the background so performance impact is minimal.**
- **No server, San, or administrator overhead is required to balance I/O load in the array.**

figure 4 - disk array after capacity expansion with hp VA virtualization

capacity management in virtualized storage systems

We have discussed how the VA7000 series can interoperate effectively with virtualization layers above it. The VA7000 series also has capabilities that enhance SAN virtualization. Non-uniform disk size management, RAID level determination, and workload adaptation can all be handled automatically within the VA7000 series, which significantly reduces both the management information and the data flow required to implement SAN level virtualization. The VA7000 series makes SAN management tasks much easier and more efficient regardless of whether they are automated or manual. But, how does the VA7000 reduce the amount of excess capacity and data movement required for cross-array SAN virtualization?

Storage systems need to be designed to manage and respond to change. Examples of change include capacity expansion and rebalancing of load across arrays, RAID levels, or storage types. In order to do this, SAN virtualization must associate volumes above it with storage below and frequently move data around. In addition to moving data to balance work loads across storage devices, data also must be moved from time to time to optimize the SAN for availability, to remove obsolete storage devices, to redeploy storage as applications and data sets come and go, and to perform backups.

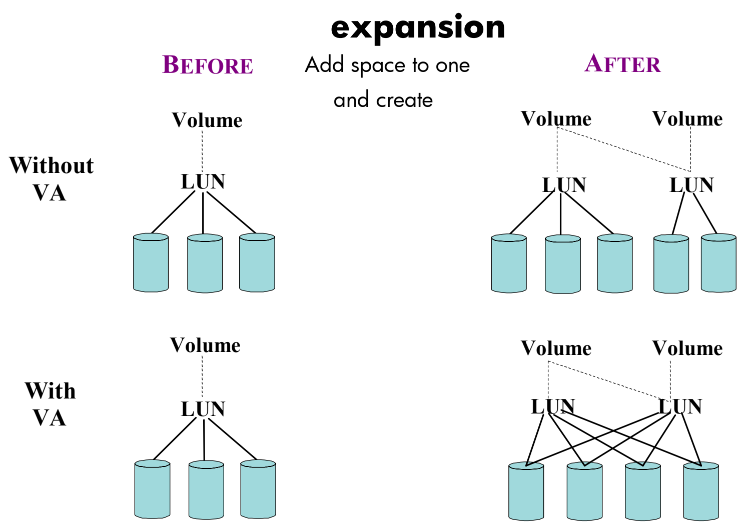


figure 5 - expansion scenario

When SANs are built on disk arrays, storage resources appear as logical units (LUNs). Whenever data is moved the destination must be *free space*. Since storage resources are represented to SANs as LUNs, there must be free space within the desired destination LUN for any data movement.

Disk arrays, with no virtualization other than basic RAID algorithms, can only create LUNs from contiguous free space, and in some implementations, only from whole disks. This means that free space of the desired type can only be made available to the SAN in large chunks that are made up of free disk drives. Therefore, SAN virtualization algorithms without array virtualization require large amounts of excess storage, or multiple movements of the same data, or both. Additional data movement and capacity requirements decrease performance and/or increase the cost of the whole storage system. What's worse, as successive changes to the SAN accumulate over time, more and more LUNs must be created. Eventually data must be moved again just to re-consolidate LUNs. This need to clean up and reuse space is a common problem with dynamic computer systems. There are many papers written to address the problems associated with free space and garbage collection.

Consider a capacity expansion scenario where the goal is to add a disk worth of usable capacity, add some of it to a LUN, and create a new LUN with the rest. With no virtualization features, adding one disk worth of space requires the creation of a new LUN that consumes at least two (2) disks for redundancy. On the other hand, the array with virtualization can add one (1) disk to an existing stripe group and create the same LUN from the aggregate. The cost savings from fewer disks and easier management shown in this example extends to more complex load balancing scenarios, where both data movement and capacity efficiencies are at stake.

Here is another example of the value gained by a layer of virtualization within the array. The VA7000 series assists SAN level virtualization by allowing LUNs to be completely independent of physical disks. The amount and type of storage resources available are constrained only by the total capacity available in the SAN, not by the groupings of physical disks within arrays. Fragmentation of physical space is managed within the VA7000's RAID blocks, not in disk-size chunks. In addition, capacity, performance, and management overheads are reduced when virtualization is coupled to RAID algorithms. The free-space management and garbage collection issues still exist. However, they are managed at a lower level in the array, not consuming SAN or server resources.

summary

We have discussed how modern computer systems consist of storage virtualization at multiple levels, and that these levels can work together to deliver simplified management and associated cost savings. While the many functions of virtualization can reside in any of the levels, some functions are best optimized in specific locations within the hierarchy. Specifically, extending the virtualization features within the storage element (array) can provide performance and management benefits beyond SAN only solutions.

In the final analysis, success is all about managing complexity in a rapidly changing world. IT managers need solutions that balance optimization, flexibility, and administrative investment. The VA7000 series virtual arrays and HP's Fsam strategy offer the winning combination of exceptional efficiency and comprehensive storage management capability for data centers worldwide.

glossary

SAN	Storage Area Network
JBOD	Just a Bunch Of Disks
FSAM	Federated Storage Area Management
IOPs	Input/Output Operations per second
RAID	Redundant Array of Independent/Inexpensive Disks
VA	Virtual Array
OS	Operating System
LUN	Logical Unit Number

for more information

For more information on HP's VA7100 or VA7400, contact any of our world wide sales offices, or visit our Web site at:

http://www.hp.com/go/storage.