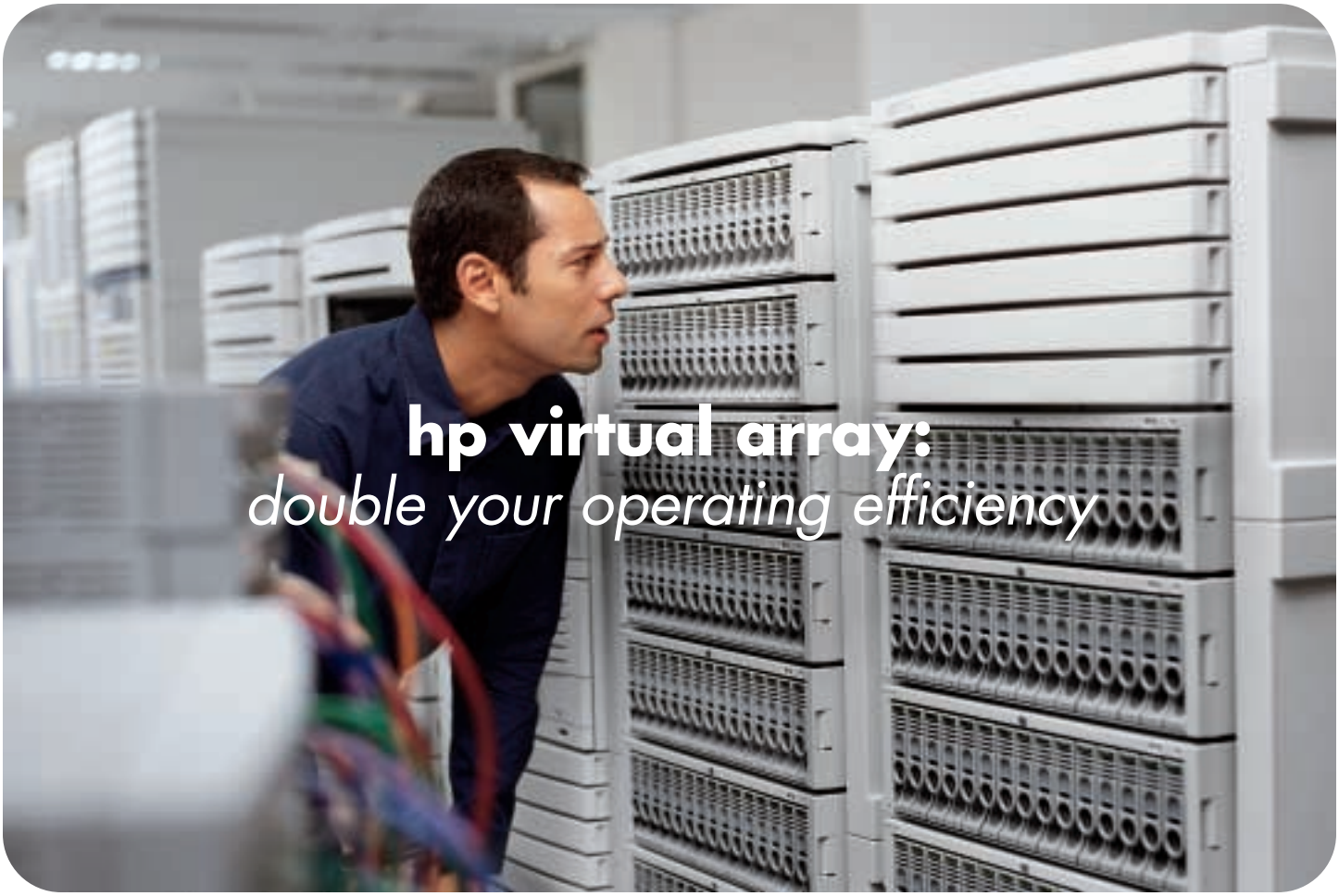

hp virtual array:
double your operating efficiency

invent

In today's market environment, you are constantly challenged to find ways to both grow your business and save costs. And with the current market growth uncertainties, saving costs is getting a lot of attention by everyone from Main Street to Wall Street to inside the Beltway. HP has invented a solution that will get you on the road to saving money and set you up to grow revenue streams once the economy turns around. That solution is HP's Virtual Arrays.

It may sound complicated at first, but HP's Virtual Arrays are simply just another automated advancement. They free your organization from the many manual functions required by the current vintage of arrays on the market today. Just as automated printing machines revolutionized the publishing industry in the last century, HP revolutionized the personal and corporate printing industry with its industry-leading LaserJet printers. Now, HP's Virtual Arrays are revolutionizing the storage industry, creating industry-leading storage solutions that save your organization 2x in operating costs over the current set of arrays on today's market.

HP's Virtual Arrays are automated to maximize your operating efficiencies and maximize your return on investment. They are automated to save you money.

The following customer quotes and two articles explain how HP's Virtual Arrays deliver on this money-saving promise. Read them for the details or have your trusted IT professionals do this for you. Just remember that HP's Virtual Arrays are the only automated arrays on the market today that offer you these cost-saving features at a mid-range price.

Can you think of even one profitable publishing house that has not fully adopted automated publishing equipment? Can you think of any successful organization that has not adopted laser printing? The same will be said for HP's Virtual Arrays. **Capture these savings today.**

customer quotes

section 1: hp storage white paper

November 2001

section 2: "hp extends virtualization to the array"

—by D.H. Brown Associates, Inc. November 15, 2001

After seven years and 20,000 sales, we hardly consider virtualization new. Many companies have optimized their performance with HP, including:

Amazon.com

"We look at three things when selecting data warehouse disks. Price per terabyte, throughput capabilities, and availability features. With the HP Virtual Array 7100, the price per terabyte was extremely competitive and the throughput was phenomenal. Our current Superdome VA7100 configuration is capable of driving three to four gigabytes of I/O per second, and the autoraid features of the VA7100 exceed our availability requirements."

Mark Dunlap
Director of Data Warehousing
Amazon.com

Room & Board

"After reviewing various competitive scenarios for our SAP environment, we selected HP as our single vendor with rp5450 servers and a VA7100 for our SAN. We found the set-up to be fast and easy, and the performance excellent. It was an easy decision to add a second virtual array, a VA7400, to accommodate our growth and position us for the future."

John Focht
Systems Administrator
Room & Board

State of New Mexico

"In consolidating from eight controllers to the HP Surestore VA7100 we expected to see some performance degradation in running our Children, Youth, and Family Department databases. We haven't seen any degradation and are pleased with the performance. We've found it very easy to allocate space with the VA7100 and have saved time with reduced administrative overhead. With the AutoRAID feature, we've eliminated worrying about mirroring. We let AutoRAID go to work and don't even have to think about it."

Tom Elder
DBA 3 Children, Youth and Family Department
State of New Mexico

Santa Clara

"We needed to move to the next generation of technology as we merged data centers. We had outgrown the large server farm scenario and wanted to focus on consolidation. In discussing our current and future needs for our NT environment, the VA7400 was the best solution. We have transformed to an enterprise environment with great expandability for the future. And as a hospital environment with needs for an always-on, 24x7 operation, we've found a reliable solution with HP."

Gary Davis
Hospital Information Systems Manager II
Santa Clara

ViaSat

"We selected the HP Surestore Virtual Array 7400 storage based on the number of Oracle ERP instances that would be needed. We found the Virtual Array has many more features and capabilities than the Sun solution that was being proposed. We felt it would scale the way we needed, provided strong price/performance, and offered excellent operating efficiencies. It has saved many hours and long nights for our DBA's and system administrators. Plus, we're using the HP Surestore tape library for efficient tape backup data protection."

Mike Johnson
IT Manager
ViaSat

virtualization, simplification and storage

In 2001, HP announced two new disk arrays—the HP Surestore Virtual Array 7400 and the HP Surestore Virtual Array 7100. This paper will focus entirely on the virtualization technology that makes these the easiest arrays to manage and the most intelligently simple arrays on the market. This paper answers the question, "What is the value of HP's Array Virtualization?"

For those familiar with data center environments, virtualization is not a new concept. Virtualization already exists to some extent at every point in the solution stack. For example, servers, operating systems, databases, file systems, volume managers, drivers, switches, and storage devices all require virtualization to achieve their purposes.

At the array level, HP's Virtual Architecture simply expands on already familiar ground, and it cleanly fulfills the promise of virtualization—it hides complexity from the administrator and can have a dramatic and positive effect on real-world performance.

This reduction in complexity greatly simplifies and streamlines the data center environment.

hp's array virtualization:

- reduces the time spent managing individual arrays
- allows storage administrators to manage more storage with less effort
- reduces the opportunities for human error
- frees up precious IT resources to work on revenue-generating projects
- self-manages the RAID configuration for optimum performance

These are not trivial benefits. Data centers are already complex, and with the continued explosion in storage capacity they will only become more so. HP's Array Virtualization is evolutionary, not revolutionary. It is a logical progression in array technology. It is proven. It is necessary. It is the perfect storage area network (SAN) technology. And it is offered only by Hewlett-Packard!

array virtualization defined

The purpose of virtualization in any technology is to hide complexity from the user, or in the case of disk arrays, to hide complexity from the storage administrator and provide a standard environment for application development and increased price/performance.

Disk arrays are complex devices designed for complex tasks. A disk array with 50 disk drives is more complex to manage than a disk array with one drive. Armed with virtualization, an array could potentially allow the 50 drives to be perceived and managed as one big drive or as one big pool of storage. The power of virtualization is the power of simplification.

Now for the definition: Virtualization in arrays is about creating and managing virtual storage devices. It is about taking blocks of storage on the disk drives and presenting them as LUNs (logical units of storage). What system administrators see, then, are not the actual physical disk drives but rather a created, simplified "virtual" view of the actual physical storage, i.e., the LUNs.

HP's Virtual Array works with LUNs just as traditional arrays do. However, the virtual array doesn't stop there. The virtual array actually manages the disks down to the level of the smallest available "cluster" or "chunk." Further, while traditional arrays utilize static address-translation algorithms for managing the chunks, the virtual array uses a dynamic mapping system.

This allows the array to dynamically allocate and de-allocate clusters of any RAID type without affecting the logical view as seen from the server. In other words, clusters can be moved and new RAID 5 and RAID 1+0 stripes can be created or extended dynamically without the host view being changed and without system administrator intervention.

Just as a file system on the host presents a virtualized view of the storage to the application or user, the virtual array presents a virtualized view of the storage to the file system. For example, just as you can create and delete files in a file system, you can create and delete LUNs within the virtual array. This degree of virtualization is what gives the virtual array its unique ability to relieve the system administrator of many of his/her mundane storage management responsibilities.

the reasons you need more virtualization in the array

- capacity requirements will always grow
- IT departments will always have a limited budget and limited human resources for managing storage
- virtualization improves the efficiency of storage administrators
- virtualization reduces the training required for people to manage storage
- data always grows, but human head count doesn't—virtualization reduces the number of people required to manage storage

key features of hp's patented virtual technology

- automatically stripes every LUN across all disks in a very large redundancy group
  Benefit: simplifies overall management; reduces the number of LUNs required to achieve a balanced workload and eliminates "hot spots"

- automatically adds new disk drives into existing RAID groups
  Benefit: capacity upgrades require no IT training, are faster, and hot spots are virtually eliminated

- dynamically, and without any human intervention, optimizes the RAID level to the application workload
  Benefit: performance tunes the array 24 hours per day and eliminates downtime for reconfigurations

- allows LUNs to be created in seconds
  Benefit: reduces setup times

- allows immediate use of the array after LUN configuration
  Benefit: speeds implementation time

- allows hundreds of LUNs to be created of virtually any size—from a single megabyte to over 2TB
  Benefit: simplifies configurations, improves application performance, and greatly decreases management complexity and potential for security errors

- allows any combination of disk drive capacities and speeds within the array
  Benefit: simplifies disk drive inventory control and capacity management

- allows disk drives and controllers to be moved to any slot in the array
  Benefit: reduces the chance for human error

why now is the time for array virtualization

Data centers have become enormously complex. Interestingly, new technologies that initially held the promise of simplifying the data center have, in fact, added to the complexity—at least for now. Fibre Channel, SANs, and the low cost and ease of adding storage have enabled data centers to expand both in size and complexity.

The world is moving from SCSI to Fibre Channel. Fibre Channel and SANs allow for more devices to be connected together, with greater performance, and at far greater distances. All of these are wonderful benefits, but along with those benefits comes an ability to now create configurations that are larger, more complex, further apart, and harder to manage than anything envisioned even a few years ago. Couple this new reality with the exponential increase in storage, fueled by its low cost, and you have a recipe for losing control of the data center.

Consider this: while human resources remain static, environments grow larger and more complex. Simplification is the only realistic answer. Adding newer and slicker management software oftentimes can help, but in the long term, adding software tends to have the opposite effect of reducing complexity. Software may give you a central place from which to manage your hardware and it may simplify processes, but it doesn't necessarily eliminate those processes. Armed with better management software, environments are encouraged to add more hardware and eventually the environment becomes as complex as it was before, but for different reasons. And when the human resources are already stretched to the breaking point, this is a recipe for complexity, stress, long hours, and human error. Virtualization is the answer. It solves the basic underlying problem. It permanently simplifies the environment for the system administrator.

managing traditional storage

"The mistakes are all there waiting to be made."
—Chessmaster Savielly Grigorievitch Tartakower on the game's opening position

Configuring and managing traditional storage is time-consuming and challenging even for experienced system administrators. When configuring storage, a wide variety of factors must be accounted for, judgments must be made, and steps repeated over and over again.

The potential for error is high. The cost of configuring an array improperly is also high. The administrator must consider the following factors:

- capacity, cost, performance and availability requirements
- requirements for future additional capacity and performance
- number of disk drives and their capacity
- performance characteristics of the disks
- which RAID level will meet desired capacity, cost, and availability needs
- number and type of RAID groups
- number of LUNs based on application, performance, and array configuration needs
- size of LUNs
- configuration of the server volume manager
- cache configuration options
- stripe depth configuration
- implementation plan: who, what, and when
- time to bind LUNs

In addition, the same processes must be followed whenever disk drives are added or the environment changes. Plus, these additional factors must be considered:

- current configuration
- desired additional capacity, performance, and availability
- whether the new disks will be stripe extensions of existing disks or be independent groups

here is a typical process for setting up a traditional array:

1. Determine number of disks, number of RAID groups, disks and disk type per RAID group, RAID level of each group, total LUNs, LUNs per RAID group, stripe depth.
2. Determine volume manager configuration, stripe size and depth, LUNs per logical volume.
3. Using the command station, set up the LUNs and their RAID levels and assign them to particular disks.
4. Set up the cache page size depending on the size of the I/Os coming in from the host.
5. Finally, before the new LUNs can be used, disks must be formatted, which can take many hours per array.

configuring an array for a database

Properly configuring an array for a database typically involves a large problem set with many variables. Many database administrators have been taught to isolate different pieces of the database in an attempt to optimize performance, availability, and recovery. This process, although based on sound objectives, is far too error-prone. This typically involves a large problem set with many, many variables. Unfortunately, database administrators oftentimes don't have all the critical information. They don't know the precise database performance requirements for each of the pieces, and they don't know the performance behavior of the array in its multitude of configurations.

In these real-world environments, it is typically far too time-consuming to try a number of different storage configurations, so database administrators typically apply rules from previous installations. The changing characteristics of newer versions of the database typically result in an unbalanced configuration that has "hot spots" that limit the performance of the system.

This entire process can take from a few hours to several days, depending on the skill of the administrator and the number and size of the LUNs. During much of this time the array is either unusable or must operate in a degraded performance mode. In other words, LUNs cannot be utilized until they have been formatted. This formatting takes up a lot of the array's internal resources and bandwidth. After a LUN has been formatted, it can be used; but as long as other LUNs in the array are also going through their format process, the entire array will suffer from degraded performance.

Now, a short word about human error. Every step of this process has the potential for human error. Except in the grossest cases, errors would probably not result in data loss, but every miscalculation in this process would easily result in a decline in performance. Some of these declines could be huge. For example, miscalculating the RAID levels or the cache page size could severely degrade the array's performance.

the system administrator's dilemma

Change is the issue. Many environments change over time and this makes their initial configurations progressively less and less optimal. The dilemma is that an optimum configuration today eventually becomes out-of-date, and typically over time, performance degrades in traditional arrays, but the administrator usually has neither the time nor the window of opportunity to bring the system down and perform the necessary reconfiguration.

Access patterns change. Different areas of the database become more highly used. Sometimes at the end of the month certain data that normally lies unused now becomes highly used and requires higher performance. All of this poses problems for the conscientious system administrator.

At first, the administrator notices a change in performance, but it is not enough to justify bringing the system down. However, over time, the performance continues to degrade until eventually the problem is so severe that the administrator is forced to bring the system down over the protests of the users. He must then go through all 13 configuration steps, including binding all the LUNs and waiting for the reformat to complete. After that the cycle begins again: Performance initially is great, but over time it degrades until finally the pain increases to the point that a reconfiguration again becomes justifiable.

managing the hp virtual array

Configuring an HP Virtual Array is much simpler than the process for configuring a traditional array. Remember the complex steps involved in configuring a traditional array? The steps involved in configuring the HP Virtual Array consist of:

- determining the total capacity and performance requirements
- for each application, determining the number and size of LUNs
- determining the number of disks required for the necessary performance
- creating the LUNs—note: creating LUNs takes only a few seconds per LUN

Step away from the array; the configuration is now complete. Every other step is automatic. RAID levels are automatic. The different capacity, number, and speeds of disk drives are automatically accounted for. The cache page size is automatically set. Even the disk formatting is performed automatically. Moreover, after the LUNs are set, the array is immediately available to accept data. The array does the work, not the administrator.

Not only is this the initial configuration process, but the process is just as simple for any subsequent reconfigurations. In other words, if LUNs have to be deleted and new ones created, the process is just as simple. Note: As with any array, if you wanted to delete the LUNs but save the data, you would have to do a backup and restore.

adding capacity with hp virtual arrays

HP's Virtual Array Architecture also simplifies the process of adding capacity to an array. Today, many traditional arrays allow the administrator to add disks on-the-fly—in other words, to add a disk drive when the array is up and running and accepting I/Os.

However, when the disk drive is added to a traditional array, it is not part of any LUN. It is not formatted, and it is not able to accept data. An administrator must go in and manually perform those functions. If the disk drive is to be added to an already existing RAID group, then the data in that group must first be backed-up and later restored to the newly created LUN that now contains the new disk drive. And except for the backup, the steps are the same as when an array is first configured.

The HP Virtual Array accepts new disks while the array is up and running and accepting I/Os as with some higher-end traditional arrays. However, the HP Virtual Array takes it one step further. Once the disk is inserted, the array automatically includes that disk into the existing disk space and stripes all LUNs across that disk. This means that even without the creation of any additional LUNs, the array performance will improve because of the additional available spindle. Only the HP Virtual Array automatically adds the new disks to existing LUNs. Further, any newly created LUNs are also automatically spread across all the disks in the array, including the additional disk.

time to implementation: formatting the array

As mentioned earlier, after new disks are added to a traditional array, it then takes several hours to complete the formatting of the RAID group. During this format phase, no data can be written to the new LUNs. With some implementations, the array is offline until all the LUNs have been formatted. In other implementations, I/Os can be written to already formatted LUNs even while other LUNs are going through the format process, although performance is very slow.

Because executing the disk format command uses up so much of the array's internal bandwidth, array performance is greatly reduced until all of the disk formatting has been completed. With HP's Virtual Array Technology, the array is immediately available as soon as the LUNs have been configured. The disk formatting is done as the writes are done. In other words, as writes are sent to disk, the formatting is accomplished for only those blocks being written to. This means that while there is a small hit to performance for that individual write, there is very, very little impact on overall array performance.

automating the cache parameters

Configuring a traditional array typically requires setting the cache parameters such as the percentage of read and write cache, the size of the cache pages, and, in some cases, the allocation of cache to specific LUNs. In making these determinations, there is ample opportunity for error.

With HP's Virtual Arrays, all of this is preset and automatic. And this means that all the parameters within the array are tuned to work in unison with the stripe size and the array hardware. First, the cache is set at 80% read and 20% write, is shared between controllers, and is treated as a "pool." Second, the cache page size is set at 64K and is set to automatically destage to disk every 4 seconds whether the page is full or not. The 64K size minimizes the number of I/Os to the back-end in sequential environments and provides a carefully calculated balance within the array between the number of cache pages and the speed of the back-end in random environments.

performance

Traditional arrays are susceptible to "hot spots" and to changes in the environment that make the initial configuration obsolete. The HP Virtual Array virtually eliminates these critical performance issues.

First, the HP Virtual Array is far less likely to experience a hot spot—in other words, it will almost never experience a condition where a few disk drives become a performance bottleneck in the array. Here's why: the virtual array always (and automatically) stripes all of the LUNs across all of the disks in the RAID group. For example: assume a virtual array loaded with a total of 60 disks had 30 disks in each of its two RAID/redundancy groups. Every LUN in that group would be spread across all 30 disks.

Every LUN would have 30 spindles at its disposal. And don't forget, in the virtual array the spare disk capacity is also spread across all of the spindles, i.e., there are no unused spindles in an HP Virtual Array.

Second, the virtual array automatically performance tunes the array 24 hours per day, 365 days per year. The RAID level is matched with the workload. Data with access patterns that would benefit from RAID 1+0 storage are automatically directed to a RAID 1+0 section of the array. Infrequently used data, or data whose access patterns match RAID 5DP performance characteristics, are directed to the RAID 5DP section of the array. These are the same rules that a storage expert would use to optimize an application's performance. In addition, like an expert storage administrator, any changes to the configuration are made only during low-usage periods or when disks are added. Note: When disks are added, the existing LUNs are automatically extended across the new disks. This normally requires data movement within the array. However, administrators have the option to postpone this data movement by simply turning off the array's "Auto-Include" feature.

faqs

Question: Isn't this virtualization technology new and doesn't that make it risky and untried?

Answer: It would be new for HP's competitors, but HP has been virtualizing arrays since 1995, and since then HP has sold over 20,000 virtualized arrays. The technology is neither new nor risky and it offers a simple and compelling value proposition, i.e., great ease of management and great data protection. For arrays, management costs are far more costly than the initial purchase price. The HP virtual array is the easiest array in the world to configure, add capacity, and manage—the administration savings are significant. It also has the best availability of any mid-range array on the market. See the white paper titled "VA7100 Hardware High Availability Features" for more details.

Question: Shouldn't a good system administrator know which RAID level the data is in? With the virtual array, I have no idea if it is in RAID 5 or RAID 1+0.

Short answer: Both the HP Surestore Virtual Array 7100 and the HP Surestore Virtual Array 7400 can be configured to run in either fixed RAID 1+0 or AutoRAID modes. And to help with performance analysis, the controller can provide data on the actual usage of each RAID level.

Long answer: With the HP Virtual Array, if your data is frequently used, it WILL be in RAID 1+0 and will have the best performance. The virtual array RAID level policies were developed after researching the decision patterns of experienced system and database administrators. Almost always, technology progresses from highly manual to highly automated operations. In almost every case, the developers of automation simply replicate the best of the already developed manual processes. This is exactly what HP did with the HP Virtual Array. Also please note: The Virtual Array policies are improved over those of the Model 12H. The virtual array really does strive in almost every instance to do any background data movement during periods of low array activity.

Question: On which disks is my data kept? I suppose the virtual array can be trusted to handle the RAID level decisions, but I also need to know on which disks my indexes and redo logs are kept. The way the virtual array moves data and stripes across all the disks means I have no idea where my data is kept.

Answer: First, if you are worried about the integrity of your data, the HP Virtual Array does a combination of things to protect your data that no other array does. First, the HP Virtual Array offers end-to-end checksum, ECC protected memory, parity coherence, disk scrubbing, and RAID 5DP which, in a typical configuration,

gives 100X the data protection of traditional RAID 5 and 10X the data protection of RAID 1. **Second:** The HP Virtual Array stripes across all the disks in the redundancy group. This involves more disks but is not so different from any other RAID 1+0 implementation. Third: array striping is the way of the future. It's faster and safer.

Question: Doesn't all this "behind the scenes" movement of data require a huge performance hit? Surely, the trade-off for virtual technology is slow performance.

Answer: Not true. For years, HP's original virtual array, the AutoRAID Model 12H, was used for HP's V-class TPC-C benchmark tests. And today, HP Virtual Array products have been used for the new rp8400 benchmark testing and are scheduled to be used for the Superdome TPC-C tests. Remember, the HP Virtual Array mimics the policies of experienced system administrators. Would an experienced system administrator do a reconfiguration of the array during a period of high workloads? No! And neither would the HP Virtual Array. The HP Virtual Array policy is to NOT perform the background tuning operations when the array is under a greater than 60% workload.

Question: Isn't manual always better than automatic? In cars, manual transmissions give better performance than automatic transmissions because they give the driver more control over performance; likewise wouldn't manual RAID configurations be better than HP's Virtual Array Technology and its automatic RAID configurations?

Answer: Manual transmissions in cars would not give better performance if drivers were only allowed to shift gears once and could never change them after that. In essence, this is what you have with traditional disk arrays. You are stuck with the initial configuration unless you bring the array down and go through a time-consuming and complex reconfiguration every time the environment changes. HP's Virtual Array Architecture tunes the array automatically, 24 hours per day. It is the hands-down winner in real-world performance. A better analogy would be to compare the multiple manual processes required to set the type, load the paper, and actually print books on the old-style printing presses with those of the automated printing systems of our generation. Obviously, automation in printing presses adds to greater performance. It works the same for arrays.

Question: Wouldn't striping the data across such large RAID groups make the disk rebuild times very long and take up a high percentage of the array resources?

Answer: Because of HP's RAID 5DP this is practically a non-issue for HP Virtual Array products. The HP Virtual Array's RAID 5DP gives each redundancy group 10x the protection of RAID 1. RAID 5DP requires that 3 drives would have to fail before there would be data loss. Thus, even AFTER a drive fails, the data is STILL protected with the same degree of protection as standard RAID 5. Finally, this means that a single drive failure does not put the data at risk and therefore does not require an emergency rebuild. The HP Virtual Array can take the time and do the rebuild in the background without impacting incoming I/Os. Also, the HP Virtual Array will rebuild the RAID 1+0 data first since that is the most vulnerable after a failure. In all cases, data integrity is ensured and performance is preserved. No other array can make this claim.

Question: If your environment is totally stable, wouldn't a manual configuration by an experienced system administrator result in better performance than one derived from the virtual array's policies?

Answer: Let's first admit it: Totally stable environments are rare. But the answer is that if the environment were totally stable, and if the administrator configured the array absolutely

accurately, the traditional array and the HP Virtual Array would have similar performance. However, the HP Virtual Array would still have the performance advantage because of its ability to efficiently load balance across all the disks in the array both at the initial configuration and after capacity growth. In both stable and changing environments, the HP Virtual Array is the right choice. Now, if the environment is not totally stable, then the HP virtual array is the hands-down winner. When both arrays are initially configured, the traditional array should be faster at least for a day or so, but after that the HP Virtual Array will catch up and continue to operate at peak performance for as long as the array is plugged in, while the traditional array will get slower over time. This is how it works. When the HP Virtual Array is first turned on, it doesn't know which data needs to be in RAID 1+0 and which data needs to be in RAID 5. After a day or so of reading the access patterns, it will figure out which RAID level is best suited for which data. Once that happens, the performance will be as good as in any manually configured array. And, don't forget, if the environment should change after that, even slightly, the HP Virtual Array will adjust while a manually configured array can do nothing but keep plugging along in a degraded mode.

Question: Are you saying that this technology is for everyone? I'm not looking for anything new.

Answer: If it's new to you, then it is understandable that HP's virtual technology might seem unnecessary. After all, you've succeeded without it up till now. But sometimes improvements do come along that really do hold out the promise of a widespread advantage. The dilemma of our rapidly advancing technological age is to know which new technologies should be ignored and which are the ones that need to be grasped.

Consider: It is always easier to manage a smaller number of things than a larger number of things. At first, bits were grouped into bytes. When there were too many bytes to keep track of, they had to be grouped into blocks. When there were too many blocks to manage, someone invented volumes and volume managers. What do you think the next logical step is when there are too many volumes?

Answer: HP's Virtual Array.

summary

HP's Virtual Array with HP's patented Virtual Technology is the industry's most intelligent disk array. Because of its unique ability to greatly simplify storage management, it significantly reduces overall IT management costs and practically eliminates mistakes caused by human error. HP's Virtual Technology also has a significant positive impact on real-world performance by automatically eliminating "hot spots," and by performance tuning the array 24 hours per day, 365 days per year.

for more information

For additional information on HP Virtual Arrays and other HP storage products and solutions, please call your local HP sales representative or visit our Web site at www.hp.com/go/storage.

virtual value in hp's va7000 series

Virtualization is a key to solving complexities associated with storage configurations. As Information Technology continues its non-stop evolution, tools and devices that improve manageability while increasing business efficiency will command a premium in the marketplace. HP has taken a significant step forward in delivering its Virtual Array (VA) 7000 series with a full complement of virtualization functions. Enterprises that value a return on investment in areas such as data-sharing, optimized self-tuning performance, and data-access flexibility must evaluate HP's Virtual Array offerings.

SAN virtualization is in effect the aggregation of multiple storage devices in a centralized configuration, a storage pool. This simplifies the storage administrator's role by enabling all resources to be seen as a single entity. SAN virtualization by itself, however, offers little assistance in efficiently managing requirements of single applications. It pays scant attention to the capacity and performance dynamics of individual, heterogeneous devices within the storage pool. HP, in its Federated Storage Area Management (FSAM) strategy, acknowledges the importance of SAN virtualization. In its VA7000 Series, HP extends these concepts and benefits to the device-array level. Together, they demonstrate the full advantage of interdependencies between managing SAN virtual storage pools and managing virtual storage at the device level.

HP's VA7000 Series Characteristics

Feature	VA7100	VA7400
Capacity	1 TB (raw) in 3U enclosure; 14 TB in 2M rack	7.6 TB (raw) using the Disk System (DS) 2400 (up to six DS2400s chained); 14 TB in a 2M rack of 73 GB disks
LUNs (max)	128	1,024
Disk Drives Supported	15 per 3U enclosure; same disk support as 7400	Up to 105 drives with DS2400; 18 GB 15 K rpm; 36 GB 10 K and 15 K rpm and 73 GB 10 K rpm
I/O ports	1 Gb/sec. FC host and disk ports	Two 2 Gb/sec. FC host ports; two 1 Gb/sec. FC disk ports
Cache	256/512/1024/ MB mirrored	512/1024 MB mirrored per controller
Transfer Rate	90 MB/sec. 12,000 I/Os per sec. from cache; 3000 I/Os per sec. back-end	160 MB/sec.; 28,000 I/Os per sec. from cache; 7500 I/Os per sec. back-end

In the HP Surestore VA7000 family, "virtual array" refers to treating the disks within the array as a pool of storage blocks rather than as physical disks. With this technology, the VA7000 family emerges as a second tier of storage virtualization. Extending virtualization to the device level, the SAN's view of storage becomes independent of physical disk attributes. This characteristic reduces the cost associated with SAN virtualization by simplifying storage management, enabling more efficient use of available device capacity, and decreasing the amount of required data movement at the device level.

VA7000 virtualization, managed at the array level, offers the advantage of providing greater granularity when managing a pool of logical storage. Logical Unit Name (LUN) capacity is definable as a logical set of storage blocks rather than as a specific set of predefined physical disks. Each LUN becomes scalable from a few megabytes to several terabytes by adding capacity without operator intervention (the array automatically controls the addition of new capacity to the storage pool).
This "self-tuning" capability automatically restripes data across new disks added to the array, minimizing any potential imbalances of data distribution within the VA7000 series. Redistribution occurs in the background with no impact on SAN bandwidth or server performance. This function significantly reduces the efforts required by the system administrator to keep storage resources in balance.

Also with VA7000 virtualization, the number of LUNs is no longer restricted by the amount of physical storage. Administrators can configure LUNs up to the limit permitted by the architecture without concern for physical disk groupings within arrays.

Disk arrays without virtualization support LUNs only from contiguous free space, requiring allocation of large storage chunks even if only a small portion is necessary. As changes occur over time, data must be moved or storage reconfigured to recover unused space.

Within the VA7000, fragmentation of physical space is managed at the RAID (Redundant Arrays of Inexpensive Disks) block level rather than in disk-sized chunks. The impact of free-space management and "garbage collection" is absorbed at the array level, without consuming SAN or server resources.

super redundancy enhances RAID

As alternatives to the "self-tuning" LUN-management capability described above, the VA7000 series also enhances LUN performance through its use of RAID configurations "redundancy groups."

Many variations to RAID configurations have surfaced over the past 10 years. Most provide moderate variations on the standard RAID 1 (mirroring) and RAID 5 (striping) techniques.
In its VA7000 Series, HP delivers a variation that substantially adds to its availability characteristics. RAID 5DP (Double Parity) approaches a RAID 6 solution in that it enables recovery from simultaneous failures in two disks without loss of any data.

While this might seem superfluous in RAID arrays of only a few disks, it becomes critical in configurations supporting a large number of disks in a virtual array. The potential for disruption due to disk failure grows with the number of disks involved in an array. The goal of virtualization is to remove the physical constraints, which enables dozens of disks to be used in a single virtual array. HP has addressed this concern with its RAID 5DP solution, increasing data availability by two orders of magnitude over traditional RAID 5 implementations. Additionally, use of storage is made more efficient as very large RAID groups (up to dozens of disks) are now manageable using only two parity disks rather than a parity disk for each group of five disks.

reliability characteristics of hp's va7000 series

- redundant, hot swappable controllers, fans, power supplies, and an internal fibre-channel hub eliminates planned downtime
- online firmware and capacity upgrade using mixed capacity disk drives
- Checksum algorithm used for end-to-end data and data-path protection against data corruption
- double parity supports recovery from two simultaneous disk failures (RAID 5DP)
- error Correction Code (ECC) protects mirrored cache from corruption before writing to disk
- three-day battery backup protects user data in cache from power outages and disasters
- phone-home capability is provided through hp's Predictive Support programs and Event Monitoring Services (EMS)

"Redundancy groups" are physical disks that contain redundant (copies of) data, defined by the RAID type assigned to the data. Each redundancy group is divided into LUNs, addressable by a host.

The VA7100 supports only one "redundancy group" with each of the 15 disk drives in the enclosure addressable by both controllers with no physical restrictions. The two fibre-channel loops (one to each controller) include a transmit line and a receive line for the transfer of data to and from the controller. Controller-to-controller communication and loop failover occur via an internal N-way bus.

The VA7400 supports two redundancy groups in which one controller manages access to one redundancy group (consisting of all the odd-numbered disk slots) and the other controller to the second group (consisting of all the even-numbered disk slots). Data movement and failover occur in the same manner as in the VA7100.

The variation in data management at the array level provides great flexibility in matching performance to application need. Self-tuning enables immediate response to data and application variables ensuring more work completed per transaction and maintaining more consistent service-level performance. RAID 1+0 uses a large number of disks in a single configuration to deliver strong performance to selected LUNs. It does so, however, at the expense of additional space for mirroring the data. RAID 5DP incurs a performance penalty for write operations and can affect system performance if frequent updates to large volumes of data are required. Its advantage is in the data redundancy provided and the improved cost efficiency vs. mirroring. Dual-redundancy groups in the VA7400 enhance the efficiency of data movement through the controllers, improving system performance.

software extends virtual functions

While hardware design and functionality are critical to enabling performance, software function can maximize the efficiency and flexibility of the device. HP's software packages extend the management and flexibility of the VA7000 family while exploiting the virtualization capability.

HP's Software for VA7000 Series

Feature	VA7100	VA7400
CommandView SDM	HP-UX with all PA-RISC computers; Windows and Red Hat Linux with Intel Pentium III/500 MHz computers or above	HP-UX with all PA-RISC computers; Windows and Red Hat Linux with Intel Pentium III/500 MHz computers or above
Enterprise Management Smart Plug-ins	HP-UX with all PA-RISC computers; Windows and Red Hat Linux with Intel Pentium III/500 MHz computers or above	HP-UX with all PA-RISC computers; Windows and Red Hat Linux with Intel Pentium III/500 MHz computers or above
Secure Manager VA	HP-UX with all PA-RISC computers; Windows and Linux with Intel Pentium III/500 MHz computers or above; supports 128 World-Wide Names and 128 secure LUNs	HP-UX with all PA-RISC computers; Windows and Linux with Intel Pentium III/500 MHz computers or above; Managed by Command View SDM; supports 128 World-Wide Names and 1,024 secure LUNs
Business Copy VA	HP-UX with all PA-RISC computers; Windows and Linux with Intel Pentium III/500 MHz computers or above; supports 127 Business Copy LUNs	HP-UX with all PA-RISC computers; Windows and Linux with Intel Pentium III/500 MHz computers or above; Managed by Command View SDM; supports 1,023 Business Copy LUNs
Auto Path VA	Windows 2000 with Intel Pentium III/500 MHz computers or above	HP-UX with all PA-RISC computers; Windows and Red Hat Linux
Network Management Tools	Not Applicable	HP OpenView NNM, HP Toptolls, CA Unicenter, Tivoli NetView, BMC Patrol
Operating Systems	HP-UX 11.0; Windows NT/2000; Red Hat Linux 6.2	HP-UX 10.2, 11.0, 11i; Windows NT/2000; Solaris 2.6, 7.0, 8.0; AIX 4.3.3; NetWare 5.0/5.1; MPE/iX 6.5, 7.0; Red Hat Linux 6.2, 7.0

The major HP Surestore software packages[1] (identified on page 2.2) include,

- *Command View SDM* speeds and eases local and remote management, including install, configure, monitor, and control via a web browser, direct host attach, or SAN. HP Predictive Support programs and Event Monitoring Service (EMS) provide phone-home capability.
- Smart Plug-ins allow *HP Command View SDM* to integrate and process storage events via SNMP with *HP OpenView Network Node Manager (NNM)*.
- *Secure Manager Virtual Array (VA)* controls server access to data on a LUN-by-LUN basis, even in simultaneous heterogeneous array sharing environments.
- *Business Copy Virtual Array (VA)* supports LUN replication for development, testing, or backup situations. Data replication takes place at a specified point in time while the system is operating and entirely within the storage array, reducing the potential for I/O bottlenecks.
- *Auto Path Virtual Array (VA)* routes data around any path's failure to maximize data availability. It provides the ability to self-configure and to automatically manage multiple I/O paths. It also provides dynamic load balancing to ensure peak performance.

[1] CA Unicenter TNG and BMC Patrol are supported in addition to the HP OpenView Storage Node Manager, OpenView Allocator, Builder, Optimizer, and Accountant.

06/02
5981-2067ENUC