IDC
Analyze the Future

I D C V E N D O R S P O T L I G H T

Pulling the Thin Provisioning Lever to Improve Storage Management Efficiency

November 2008

Adapted from *Worldwide Storage Management Software 2008–2012 Forecast: Balanced Growth Throughout the Segment* by James Baker, IDC #213192

Sponsored by Symantec

The digital data explosion is creating more and more data, all of which has to be efficiently stored and subsequently managed if organizations hope to gain maximum value. The more data there is, the more need there is for organizations to manage it effectively. This is a major reason why IDC estimates that the worldwide storage management market will grow to $1.65 billion in 2012, up from $1.1 billion in 2007, representing a compound annual growth rate (CAGR) of 8.3%. In addition, storage utilization has long been a concern in the open systems world, where the amount of storage used versus the amount that is physically present has been as low as 30–35%. As IT managers strive for more effective storage utilization — for both efficiency and productivity reasons — thin provisioning — an approach that allows systems administrators to allocate storage capacity to an application or a host multiple times in a shared pool concept — is increasingly attractive. This paper discusses thin provisioning and how Symantec is using it in this important area of the storage management market.

The Need for More Efficient Storage Management

IT departments — and in particular storage administrators — are under increasing pressure to provide higher levels of storage utilization and efficiency. Ever-increasing volumes of digital data, including new data types such as medical images, end user–generated JPEG and music files, and Web-based transaction systems on top of traditional text-oriented data, must be handled by storage management software. In essence, the amount of data is growing and the data is being stored longer. These trends are major drivers propelling the need for more efficient storage management.

As IT departments are called to do more with existing or even fewer personnel resources, storage administrators need to simplify both the management and the provisioning of storage in an effort to be more productive and efficient. As a result, IT managers are looking to employ less storage more effectively. In addition, IT managers are aiming to deploy less storage to address power and environmental costs at the same time. By deploying less storage more effectively, IT departments are hoping to realize an energy efficiency component of "green IT." Other storage-based components of green IT include usage of more energy-efficient drives, disk spin down, and compression to reduce the raw terabytes needed by the organization.

The need for more efficient storage management, magnified by the exponential growth of digital data, is driving the development and deployment of storage management software that automates many high-level storage management tasks. One such approach is thin provisioning, which allows storage administrators to create a shared pool of storage. The physical storage capacity of the pool is provisioned to multiple applications or hosts simultaneously. In extreme cases, provisioned storage capacity combined across all hosts and applications could even exceed the amount of purchased

capacity in the system. With thin provisioning, administrators can overallocate usable capacity to an application and consume purchased physical capacity only as applications actually write data. In effect, thin provisioning provides just-in-time storage and allows administrators to simplify their storage provisioning and increase their storage utilization factor. Storage utilization also is affected by raw performance, or input/output per second (IOPS).

The Benefits of Thin Provisioning

By allowing storage administrators to overallocate storage capacity to the application or host, thin provisioning challenges the long-standing storage approach of having to dedicate capacity up front based on allocation. Consequently, thin provisioning tackles many of the issues that organizations encounter with traditional storage provisioning techniques:

- Users don't have to overrequest capacity to avoid disruption to their application. In traditional provisioning environments, users typically tended to overestimate the space they needed to account for additional "headroom" for future growth. This safety margin helped to ensure that users wouldn't have to go back to their storage administrators to ask for more storage at a later time when the need for space had become urgent. The trade-off, of course, was inefficient use of storage when comparing amount used versus amount present in the configuration.
- Storage managers can avoid the unnecessary cost and increased overhead associated with allocated capacity of a single server or application. Such costs and overhead are typical characteristics of traditional storage provisioning because traditional provisioning does not allow allocated storage to be repurposed or shared by another application.
- Storage administrators no longer have to overallocate capacity, often in order to avoid reconfiguration, downtime, or untimely interruptions. Instead, they have more time to consider larger, more complex topics such as the overall application architecture and organizational IT infrastructure.

These advantages of thin provisioning can be leveraged even further if the server is "thin provisioning aware." Thin provisioning is not just a storage administrator's decision. Rather, an additional view from the top down (server to disk), including integration with the volume manager and file system, can help to avoid compromising the benefits of thin provisioning were the user to rely solely on the storage-only view.

Technology Trends

Organizations that expect the capacity needs of a group of hosts or applications to change over time are the best candidates for thin provisioning. To avoid the possibility of an application or host saving data at a higher-than-expected rate — and therefore running out of physical capacity — organizations should use a thin provisioning solution with the following capabilities:

- Self-monitoring and alerts for the shared pool of storage (Three levels of alarm are preferable.)
- Automation of the process of configuring and dedicating capacity to thinly provisioned volumes (Methods for recapturing and/or reclaiming space once data has been deleted should be part of this strategy.)
- Capability of scaling to deliver on the overallocated promises of thin provisioned capacity

Many vendors and products support thin provisioning. While all products essentially provide the same basic functionality — they all enable capacity to be overallocated to applications and then consumed on a just-in-time basis — vendors have various approaches to their implementations. IDC recommends that organizations consider the following actions prior to a thin provisioning implementation:

- Assess the potential for increased operational risks, such as the possibility of a host or application running out of physical capacity.
- Determine an approach for balancing the use of thinly provisioned volumes against acceptable performance levels and asset utilization.
- Evaluate the way in which thin provisioning interacts with other advanced storage services.
- Investigate how thin provisioning impacts both the server components and the storage components of an infrastructure.

Early adopters of thin provisioning include relatively advanced organizations for which low storage utilization is a significant source of inefficiency and cost. While less advanced organizations share many of the same concerns in terms of their storage infrastructure, they tend to deploy new technology when it has conclusively proven itself in the market, as well as when well-established vendors begin providing solutions. This is now happening with thin provisioning, and IDC believes that thin provisioning is a valuable technology with significant benefits for users.

Considering Symantec

Symantec Corp. is a Cupertino, California–based provider of security, storage, and systems management solutions for both businesses and consumers and has operations in over 40 countries. The company's Storage and Availability Management Group provides IT organizations with a variety of products for data management, storage management, high availability, and disaster recovery.

Symantec's Veritas Storage Foundation solution provides heterogeneous volume and file system management that includes multipathing, tiered storage, and data migration and replication. The latest release of the product is thin provisioning aware and enables customers to gain the advantages of thin provisioning by supporting all thin provisioning architectures currently available, including those from 3PAR, EMC, Hewlett-Packard, Hitachi Data Systems, IBM, NetApp, and Sun. Of course, Veritas Storage Foundation is not necessarily tied to thin provisioning as its services continue to exploit traditional provisioning techniques too.

Veritas Storage Foundation is a combination of the well-known Veritas File System (VxFS) and Veritas Volume Manager (VxVM) plus Dynamic Multi-pathing (DMP) for load balancing and error recovery. Storage Foundation also includes Storage Foundation Manager for centralized management and workflow control.

According to Symantec, Storage Foundation has the ability to integrate easily into existing infrastructure across Unix, Linux, and Windows operating systems. Version 5.0 interoperates with prior versions of Storage Foundation all the way back to version 3.5. However, certain product features such as the Veritas Thin Reclamation API and SmartMove are available only with version 5.0 MP3 (Unix and Linux platforms) or version 5.1 (Windows platform).

Symantec has designed the Storage Foundation user interface with nontechnical users in mind, allowing use by end users and business data owners as well as by the more traditional storage administrators or others with specialized storage expertise. Additionally, Symantec has provided several automation techniques — such as script support, command line interpreter (CLI) support for technicians, wizards that hide multistep complexity, and the ability to control the environment through the Web — that shield the user from unwanted complexity.

Storage Foundation has a suite of applications that work together, solving the issue of too many management panes of glass. The product has been designed to be federated in that it works well with products from other vendors. Storage Foundation interoperates with APIs and management tools from other vendors.

Storage Foundation enables the following implicit advantages of thin provisioning: just-in-time capacity on demand, lowering the wasted space statistics, and spreading the I/O pattern across many disk drives to enable so-called "wide striping," or many simultaneous disk accesses.

In terms of optimizing thin provisioned environments, Symantec has added significant enhancements to Storage Foundation, including the following:

- **Getting thin and staying thin.** Storage Foundation is an advanced file system and volume manager; therefore, it can ensure that provisioning is space efficient at setup time, and as growth occurs, the expansion of volumes continues to be thin.
- **SmartMove.** This feature enables hardware-independent online migrations of application information from traditional or thick volumes to thin provisioned volumes. This capability enables organizations to move quickly into thin provisioned environments and not incur wasted storage capacity. At the application level, SmartMove ensures that the migration occurs transparently with no interruption for users.
- **Shared APIs and space reclamation of thin environments.** Symantec offers an API for online capacity reclamation. Because Storage Foundation is now "thin aware" and has been "thin optimized," it offers the Veritas Thin Reclamation API to major storage hardware vendors that integrate with VxFS and VxVM. Solving the challenge of reclaiming deleted application data, and freeing up physical storage capacity, is a benefit of the Thin Reclamation API. Once application data is deleted, Storage Foundation communicates with the storage array and the array subsequently frees up physical storage capacity that is no longer necessary. It is important to note that without this server and storage communication, storage capacity will continue to be wasted as the storage arrays have visibility only at the volume level, not at the file system and application levels. Additionally, Storage Foundation can dynamically shrink or grow both the file system and thin volumes. The value here is the reclamation of physical storage because the application no longer requires it. Freed space can be put back into the "capacity pool" and be eligible for reuse by another application. A storage-only view would have no knowledge that the space is no longer needed.

Challenges

Symantec faces a number of significant challenges in bringing this technology to the marketplace. Symantec's product announcements have been delivered, partners are aligned and have their own supporting rollout schedules, analysts are being briefed, and technical reviews of the product are favorable. However, as with all future-facing activities, Symantec's plans need to be well executed, commitments need to be honored, and the company needs to be vigilant and quick to respond to unexpected problems.

Potential customers don't yet fully appreciate the implications of thin provisioning (i.e., how to get into a thin provisioning environment and how to stay thin over time). Symantec needs to educate these customers and prospects regarding the benefits of thin provisioning as it relates to their individual environments.

Conclusion

IT departments and systems administrators are beset by storage management challenges. The tremendous growth in the amount of digital data, combined with IT staffs that are stretched thin, is driving organizations to simplify operational tasks and to increase utilization rates as measured by storage management tools. One approach that is gaining traction to achieve both of these goals is thin provisioning.

However, thin provisioning is not without its issues. Before implementing a thin provisioning storage environment, organizations need to assess the impact on both server and storage infrastructures and carefully evaluate the potential operational risks.

When implemented correctly, thin provisioning can help organizations improve their storage management efficiency by both increasing utilization and boosting staff productivity. Symantec is in a unique position to leverage thin provisioning because of its combined File System and Volume Manager capabilities that have been made "thin aware." If Symantec delivers on the product features described earlier in a timely manner and as planned, and if Symantec addresses the challenges highlighted in this paper, the company deserves to be on the short list of any organization considering a thin provisioning environment.
