

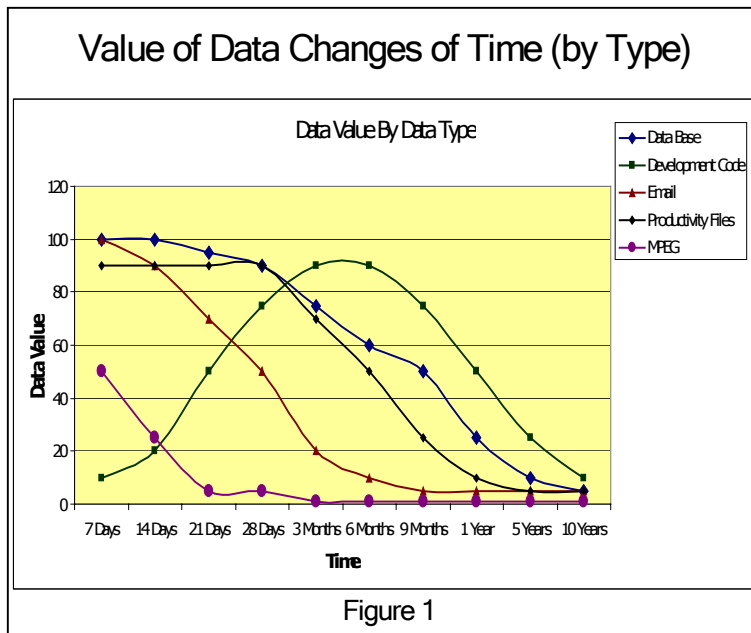
VERITAS Software: The Foundation for Tiered Storage

Date: April, 2005
Author: Nancy Hurley
Title: Senior Analyst

Abstract: Implementing a tiered storage infrastructure enables organizations to better utilize existing resources, reduce management complexities and reduce overall costs. VERITAS offers organizations a full portfolio of data management, migration, and protection solutions to enable an automated tiered storage environment. VERITAS' provides solutions that give users an overall understanding of the entire storage infrastructure, managing the file system and provisioning of volumes, processes which are critical to effectively tiering storage. VERITAS Storage Foundations provides all of the functionality for automating storage tiering and provides the basis for effective ILM implementations.

The Value of Tiering Storage

Today, more than ever before, IT organizations are being asked to do more with less. They are being asked not only to store and protect growing stockpiles of data for extended periods of time but also to make sure that data is readily available – and easily accessible – for a variety of business (e.g., service-level agreements) and regulatory purposes. And they are being asked to do all this with the same, and often even fewer, IT dollars.



Today the tendency among many IT organizations is either to throw more disk (usually of the high-end variety) at the current “hot spot” or to divert existing resources for a temporary “quick fix”. However, the Enterprise Strategy Group (ESG) believes there is a better option – one with significant and potentially long-lasting benefits – that is to implement a “tiered” storage environment as part of an over-arching information life-cycle management (ILM) strategy.

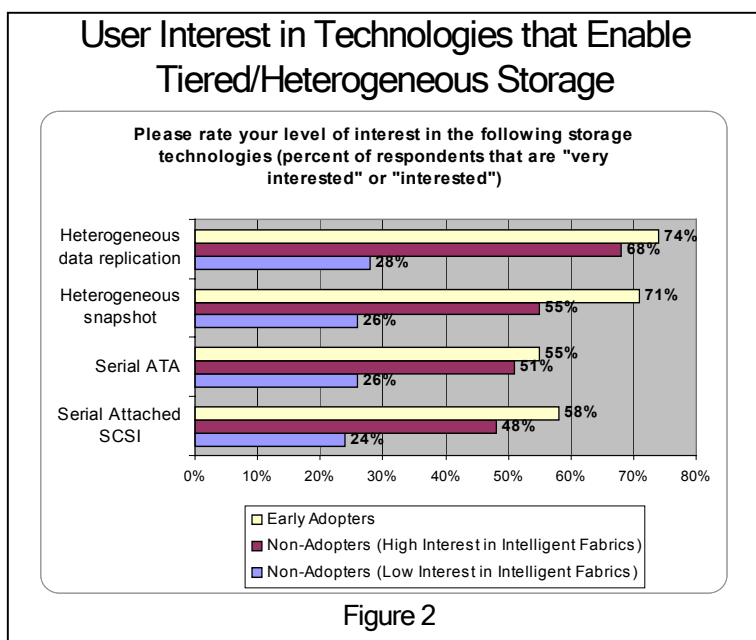
The concept is simple: Match storage resources to the changing business value of the data being stored (Figure 1). Mission- or business-critical data (i.e., data that is frequently accessed and leveraged for business-critical applications such as OLTP) would be placed on higher-performance, higher-capacity Fibre Channel (FC)-based disk arrays, while data requiring less stringent performance and availability would be stored on lower-performance, lower-capacity Serial ATA (SATA) disk systems, for example. As the value of the data to the

organization changes over time, the data is moved, or migrated, automatically among available storage resources to the appropriate tier, or level, of storage according to pre-determined policies.

Administrators, in conjunction with business-line managers or executives, would establish these tiers of storage based on the performance, availability, disk type (FC, SCSI, ATA, SATA), retention requirements (e.g., WORM), and retention needs (e.g., lifespan), to name a few, of the various data types to meet established SLAs or regulatory guidelines. While the type and number of storage tiers may vary from organization to organization, generally speaking organizations would create several tiers of storage (e.g., silver, gold, and platinum storage resources) from, most likely, multiple storage hardware vendors, that have storage service attributes associated with them (e.g., gold and platinum have local copy services capabilities for disk-based application recovery), which would then be doled out according to application requirements.

While the concept of tiered storage is being touted by most array vendors (who recognize the days of all data being stored on high end, high cost arrays are over), what is often overlooked is the data protection component of a tiered storage environment. The same type of principles should apply to the data protection side of the equation. Data needs to be protected according to its value at any point during its life-cycle. In other words, data that is constantly changing requires a different level of protection than data that is static, or unchanging, but is subject to regulatory scrutiny. Similarly, some data, though static in nature, may have more stringent recovery requirements than other “more-active” data types.

Along these same lines, data that is mission-critical today may not be in a week’s, months, or a year’s time. The goal, again, is to match the data protection criteria (e.g., mirroring, replication, backup to tape or disk), the backup methodology (e.g., incremental, snap, etc.), the backup frequency, and recovery requirements to the data *throughout* its life-cycle and apply disk, tape, and content-addressable storage (CAS) resources accordingly.



However, the tendency among many IT organizations (of all types and sizes) is to establish “blanket” primary and secondary data protection strategies. In fact, it is not uncommon for organizations today to store all data – and many terabytes of it – on high-end primary storage arrays and to protect all data in the same way, regardless of its current – and changing -- value to the organization. As ESG research consistently indicates that the top storage management problem end users face is meeting the backup window, it should be obvious that this practice needs to change. In today’s business and economic climate, clearly the model “one-size-fits-all” no longer applies. An effective tiered storage environment not only recognizes the changing behavior of data over time but also acts on these changes, based on policies, automatically migrating data back-and-forth among storage resources with the knowledge of the application.

ESG has confirmed with multiple end users that have implemented tiered storage environments that by placing, protecting, and migrating data in this fashion, organizations can expect to see significantly lower primary and secondary storage resource costs as well as improved utilization (data that is not deemed mission-critical is migrated off high-end arrays, freeing up costly disk space for future applications). Early adopters of tiered storage networks and those interested in tiering storage show significant interest in the technologies that enable tiering storage (Figure 2). Additionally, tiering storage and protecting it according to its value should help organizations reduce backup window issues and better meet regulatory compliance laws (e.g., Sarbanes-Oxley, SEC17a-4, HIPAA, etc.) as they evolve.

The benefits of changing data migratory patterns to meet evolving data requirements can be huge – and, contrary to what many vendors tell you, don’t require users to “rip-and-replace” established storage and backup infrastructures. Instead, organizations can leverage what they have and supplement it with complementary hardware and software

technologies where needed. However, the challenge is to implement a tiered storage environment that meets the specific needs of your organization. ESG recognizes that this type of ILM implementation can be difficult to coordinate, and can vary significantly from one organization to the next one of the most confusing aspects is the myriad of “ILM” enabling solutions that are being marketed by storage vendors.

While a number of software solutions are available today to enable management, migration and protection, users should keep an eye toward the future and look for solutions that provide integration and coordination between these various data management activities. VERITAS is one of the few vendors in the market that not only has a full portfolio of software solutions to enable an automated tiered storage environment, but also has a vision as to how to address the more complex issue of accessing and recovering the data in an appropriate manner depending on the needs of the organization.

The VERITAS advantage

As its name states, VERITAS Storage Foundation, which includes file system, volume management, policy based data migration, and granular protection is the cornerstone of VERITAS’ ILM/tiered storage capability. VERITAS Storage Foundation allows users to create volumes, intelligently place data on storage resources, migrate data automatically and non-disruptively among tiers of both heterogeneous storage resources and heterogeneous operating systems, and protect data according to policy via FlashSnap snapshot technology. The components of Storage Foundation that are key to enabling tiered storage environments are:

- **VERITAS Volume Manager (VxVM):** Creates and manages virtual volumes, which can be set up, expanded in real time with no application impact or reallocated based on capacity demands. Volumes can be created across multiple disks, and multiple heterogeneous storage subsystems. Each volume can have specific performance and availability characteristics. In a tiered storage environment, this enables more effective utilization of storage resources and provides higher performance for defined tiers of storage.
- **VERITAS File System (VxFS):** Enables administrators to set policies regarding the placement of data across virtual volumes. VxFS enables organizations to implement multi-volume file systems providing the flexibility to place information across disparate and heterogeneous storage hardware, if the organization so chooses. The policy based placement automatically enables tiering by directing data to volumes that are best matched (performance, availability, capacity, etc) to the needs of the information. Additionally, the metadata can automatically be separated to streamline I/O access. VxFS also automates file and database migration based on pre-defined criteria such as time since last modified, time since last accessed, application type, and type of file. Unlike static HSM migration, VxFS can migrate data to multiple heterogeneous resources over the course of its lifetime, from Tier 1, to Tier 2 to archive resources for instance.
- **Quality of Storage Service (QoS):** QoS is enabled via the capabilities of VxVM and VxFS. QoS enables administrators to move files and databases among storage resources based on pre-defined policies. QoS policies also enforce various protection schemas, where FlashSnap can be implemented at different intervals depending on the protection requirements. In addition, the target for the snapshots can vary (FC vs SATA arrays) and support heterogeneous storage subsystems based on QoS policies.
- **Portable Data Containers (PDCs):** PDCs enable customers to perform rapid operating systems data migrations. For example, PDCs enable an organization to have a single copy of production data, “export” the information from a Solaris environment and “import” it into a Linux environment with minimal downtime (particularly compared to NFS file transfers or tape-based recoveries) and without the requirement of additional storage capacities. PDCs enable another dimension of tiered storage; use the most appropriate server operating system for the application and business while still maintaining a consistent storage infrastructure and operations.

Other notable features of the Storage Foundation include:

- **Multiple Disk Spanning:** Allows data to be grouped together into logical disks that can be dynamically resized according to the requirements of the application. Moreover, the multi-volume file systems enable volumes to span multiple, and heterogeneous storage controllers. This removes the “per subsystem” thinking and restrictions when allocating LUNs to applications.
- **Dynamic Multipathing (DMP):** DMP provides load balancing and availability (in the event of an interconnect failure) across multiple I/O paths (FC or IP) between the server and storage array for optimum performance and availability.

- **Automated Performance Tuning:** Tunes each write automatically for performance purposes.

The above features ensure maximum storage utilization and optimum administrator productivity; and a variety of other capabilities, including online administration, hot relocation, panic prevention, RAID support, VERITAS Cluster Server, and Storage CheckPoints are designed to virtually eliminate any downtime.

ESG believes that one of the key advantages VERITAS brings to the table is its ability to provide organizations with an end-to-end ILM/tiered storage software capability – one that extends well beyond VERITAS Storage Foundation. VERITAS has a comprehensive portfolio including email archiving, application performance monitoring and a variety of remote and local copy (e.g., VERITAS Volume Replicator and NetBackup Advanced Client), data analysis, modeling, heterogeneous storage management, centralized workflow, (e.g., CommandCentral Storage), and policy management, and reporting (e.g., CommandCentral Service) tools.

In particular, integration between VERITAS Storage Foundation and VERITAS CommandCentral makes it easy for organizations to manage their various storage resources. VERITAS CommandCentral Storage and CommandCentral Service serves as a single, integrated portal into the storage environment, displaying information about capacity, usage, costs and service levels for online storage, data protection, servers and clusters. This information is required in order to effectively implement a tiered storage environment and ensure resources are being utilized effectively and applications are adequately served.

VERITAS is also known for its backup and recovery solutions, and the company continues to innovate on the NetBackup and BackupExec platforms in order to provide granular, application aware data protection. The latest version of NetBackup will also enable users to recover data in a variety of ways, either performing full recoveries in case of failure or granular searches when specific files need to be accessed from archive. NetBackup has moved beyond being a backup and recovery solution to being the overarching manager to access data that is protected in a myriad of ways with various products in the portfolio, including snapshot technology in Storage Foundations. This is extremely powerful functionality and illustrates the value of integration. In the future, VERITAS will enable a single view and means to recover data from any “data protection system” (disk, tape, CAS etc), regardless of the methodology of placement (via backup application, snapshot, migration or copy on write).

VERITAS End User- Results of Using Storage Foundations to Tier Storage

Results:

70% increase in storage utilization

Reducing backup windows (now only 1.5TB backed up each day)

Recaptured 30TB of tier one storage (by moving data to tier 2 and 3 storage)

Tier 2 50% less than Tier 1, Tier 3 66% less than Tier 2
75% reduction in TCO equates to 140% ROI in one year

Figure 3

ESG believes the next big opportunity for VERITAS, as it pursues its utility computing vision, is to make Storage Foundation more application-aware. Doing so will provide administrators with more granular control of the data and how this data is eventually migrated among storage resources. VERITAS has taken some initial steps in this direction with product introductions such as VERITAS Storage Foundation for Oracle RAC, etc.

End user Results

No two organizations are alike, nor is the data they generate. Therefore, it follows, that no two tiered storage environments will be – or should be – alike either. However, what ESG has learned from discussions with customers is that despite these differences, organizations tend to share a common view of ILM – more specifically, of tiered storage. Organizations don’t necessarily think about tiered

storage as three rigid buckets (i.e., primary, backup, and archival) into which all data must be arbitrarily dumped. Instead, they view it as process with both hardware and software components that is first and foremost about using primary data storage more effectively and more efficiently.

Tiering storage is about recapturing high-rent primary disk space and re-directing data that doesn’t belong on this class of storage to secondary and tertiary targets (e.g., SATA, CAS, tape, or any combination thereof). For some organizations, this process involves three tiers (e.g., primary, secondary, and archive); for others it may mean two

(e.g., backup and archive). VERITAS Storage Foundation can facilitate the movement of data among these varied resources; it is agnostic to both disk and OS platforms.

One VERITAS customer ESG spoke with implemented three storage tiers to address a growing data management problem it was experiencing with its single-tier primary-disk storage environment. Data volumes for the firm were increasing 386% year over year while the size of its IT staff remained constant. By migrating some of its data off primary storage (Hitachi Lightning) and onto secondary (Hitachi Thunder) and tertiary (Xyratex SATA disk) storage tiers, the organization was able to gain control of its storage environment and cut costs significantly. The firm relied on VERITAS Storage Foundation to manage and move data among these tiers.

The company was able to boost storage utilization by 70%, reduce its backup window drastically by backing up only 1.5TB (versus 58TB) of data daily, and recapture 30TB of tier-one storage by moving data to tier two and two three storage – and it did all this using VERITAS Storage Foundation and backup products. Overall, the organization says it saw a 75% reduction in TCO, or about a 140% ROI in one year (Figure 3).

However, other end users we have talked with have implemented tiered storage in a variety of ways and for many different reasons. An industrial firm that generated very large files but needed to access those files frequently choose to establish two tiers: one (e.g., high-end disk) for primary storage and a second (e.g., CAS or NAS) for archives. While a large bank dealing with a very large Exchange environment (e.g., > 26,000 Exchange users) choose to implement a two-tiered storage environment that directs all email directly to a CAS-based archive, eliminating the use of primary disk for any email storage.

The point here is that ILM doesn't dictate a specific type of tiered storage infrastructure. But to establish an effective tiered infrastructure, organizations do need to follow a few guidelines. First, they need to assess their storage assets (i.e., determine how much storage they've got and what type of data is stored on them). Second, they need to define and implement appropriate tiers of storage, and match data to the appropriate tiers. And, lastly, they need to move, or migrate, data among the defined tiers as needed. VERITAS is one of the few vendors in the industry that can provide the portfolio of software solutions that address each phase of the process.

The Bottom Line

There is no question that end users will benefit from tiering storage and automating data migration between tiers. VERITAS Storage Foundation is one of the industry's most-comprehensive tiered storage-enablers. The software provides all the tools necessary for users to turn existing resources into efficient tiers of storage based on policies dictated by changing business requirements. ESG believes in the value of integrated software solutions; the true power of the portfolio is realized when solutions that monitor and analyze can trigger actions such as the provisioning of capacity, migration of data and protection of data regardless of where it is in its lifecycle. VERITAS is on its way to providing a comprehensive, integrated solution that will enable much more than an automated tiered storage environment- but for now that's a good place for users to start and realize tangible cost savings.

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. and is intended only for use by Subscribers or by persons who have purchased it directly from ESG. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of the Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at (508) 482-0188.