

Economic Impact of Virtualization at Overstock.com



By David Merrill and Vijay Ramaswamy

August 2010

Table of Contents

Executive Summary	3
Introduction	4
Evolution of IT Infrastructure at Overstock.com	4
Storage Goals to Support a Growing Business	4
Simplify the Storage Environment through Virtualization	5
Align Storage Tiers to Business Needs	5
Simplify Data Migration	5
Reduce Operational Expenditures (OPEX)	5
Reduce Capital Expenditures (CAPEX)	5
Virtualizing the Overstock.com Storage Environment	6
Hitachi Controller-based Virtualization	6
The Virtualized Architecture	6
Multitiered Infrastructure	7
Meeting Provisioning Demands with Hitachi Dynamic Provisioning Software	7
The Economic Impact Analysis	8
Introduction	8
Changing the Rate and Cost of Growth	9
Lower Cost of Growth	10
Documenting a Lower Unit Cost	11
Qualitative Benefits at Overstock.com	13
Planning and Projecting a Lower TCO in the Future	14
Appendix A — Overstock.com Profile	15
Appendix B — Comparing the Architectures	16

Executive Summary

The lifeblood of almost every organization is data, and the requirements to protect, maintain and access the systems that store that data are absolutely necessary for supporting business operations. With the sheer amount and varying types of information that must be stored, complexity in the data center tends to grow as quickly as the amount of data itself.

Online retailer Overstock.com has been diligent in its determination to manage data growth as efficiently and cost-effectively as possible. With the expansion of its online business, Overstock.com has experienced several surges in the volume and complexity of its storage environment. As requirements for availability, performance and scalability increase so does the need to simplify administration and reduce both capital and operational expenditures.

To adeptly meet new business conditions and reduce total cost of ownership, Overstock.com implemented a virtualized, multitiered storage solution based on Hitachi Universal Storage Platform® V systems. Overstock.com makes extensive use of Hitachi data mobility and virtualization technologies to dramatically simplify its storage environment, improve storage utilization and reduce operational costs and risks.

This case study articulates the economic impact and the resulting lower unit cost of implementing a fundamentally new storage architecture. Compared to the unit costs measured in December 2007, the next year (2008) costs were reduced by 43 percent, and the subsequent year costs went down an additional 37 percent. In addition to the unit cost impact, this case study also quantifies the qualitative and operational efficiencies Overstock.com has realized.

Introduction

Overstock.com is a leading online retailer offering closeout and discount brand and nonbrand name merchandise, including bed-and-bath goods, home décor, kitchenware, watches, jewelry, electronics and computers, sporting goods, apparel and designer accessories, among other products. Overstock also sells books, magazines, CDs, DVDs and video games, and it operates an online auctions business — a marketplace for the buying and selling of goods and services — as well as online sites for listing cars and real estate for sale.

Overstock.com is based in Salt Lake City, Utah, and was founded in 1999. Overstock.com's website offers customers an opportunity to shop for bargains conveniently, while simultaneously offering their suppliers an alternative inventory liquidation and sales channel. Overstock.com continually adds new, limited inventory products to its website in order to create an atmosphere that encourages customers to visit frequently and purchase products before inventory sells out.

Evolution of IT Infrastructure at Overstock.com

The volume of business data and customer records that Overstock.com generates and retains annually is growing at an aggressive rate. This explosive growth had Overstock.com looking at other scalable and operationally efficient architectures. In April 2008, the Overstock.com data storage infrastructure was upgraded to a multitiered virtualized storage solution, based on the Hitachi Universal Storage Platform® with Hitachi Adaptable Modular Storage 2500 and 500 and third-party storage systems virtualized behind it.

Several factors have increased the organization's demand for availability, performance, scalability and simplification of the storage environment. With the increase in business operations and mandatory regulations for data protection and retention, Overstock.com data must be stored and retrievable for longer periods of time. To continue providing reliable and high-performance storage support for expanding production operations and application demands, Overstock.com also determined that the storage environment should be nearly limitless in its ability to expand with growth requirements. By optimizing data assets, promoting even greater efficiency and simplicity across the environment, and further defining storage class attributes, the IT infrastructure will remain highly aligned with business goals. Consequently, Overstock.com will be able to drive company growth and profitability while effectively managing risk.

This case study examines the latest storage architecture at Overstock.com, based on Hitachi Universal Storage Platform V systems. By making extensive use of Hitachi data mobility and virtualization technologies, including Hitachi Dynamic Provisioning software, Overstock.com expects to dramatically simplify its storage environment and reduce operational costs and risks.

Storage Goals to Support a Growing Business

Overstock.com has always focused on providing excellent coverage and customer service, while emphasizing internal efficiencies. With data volumes expected to grow greatly every year, Overstock.com developed several goals to help the company improve service levels and data access while reducing operational inefficiencies and related costs.

Simplify the Storage Environment through Virtualization

One way to effectively reduce operational costs is to reduce complexities within the storage environment by allowing data to be virtualized and managed as a single pool of storage. Through controller-based virtualization, internal and external storage systems can be brought together in a virtualized environment to increase utilization and share resources. This enables all available storage to be monitored, migrated and managed with one common interface.

Align Storage Tiers to Business Needs

All data is not equal in its importance to the organization, and the value of this data changes throughout its lifecycle. By classifying data according to specific service level requirements and then matching this data to corresponding storage tiers, organizations are better able to optimize IT infrastructures and costs. Prior to refreshing its IT environment with the Universal Storage Platform V, Overstock.com had silos of multitiered storage to support business needs. Now, Overstock.com will further define and classify its storage and distinguish the quality characteristics and performance attributes for each storage tier. In this way, the company will be positioned for sustainable data growth and mobility and ensure its ability to continuously support the ongoing alignment of data demands and resources.

Simplify Data Migration

As part of the technology refresh, Overstock.com was on the lookout for a platform that reduced risks associated with data migration. In the past, the company had been using host-based migration tools, which only move data locally between the application and the host system. This meant that data movement was somewhat cumbersome and required applications to be brought down, causing slight disruptions to operations. Being able to transparently move data was a top priority for Overstock.com, as this would allow them to avoid such disruptions to production operations. The company also wanted to standardize, automate and simplify migration procedures for all internal and external storage virtualized through a Hitachi Universal Storage Platform V.

Reduce Operational Expenditures (OPEX)

Minimizing the cost of doing business is always an asset to organizations seeking to be profitable. The IT department at Overstock.com is considered a cost center — which adds to the expense of running the company — and must operate as efficiently and cost-effectively as possible. In examining ways to reduce its total cost of ownership within the IT department, Overstock.com learned that the virtualization and tiered storage management technologies offered by Hitachi Data Systems would significantly aid in lowering the time and expense of managing the storage environment.

Reduce Capital Expenditures (CAPEX)

Managing a tight budget and supporting growing storage needs is always a challenge. Overstock.com is always looking at ways to reduce CAPEX while satisfying the demands of new business and organic growth. Overstock.com needed methods to reduce the cost of growth. With dynamic provisioning, the company is utilizing the existing disk storage more efficiently, which has fundamentally changed the raw capacity appetite curve, and thus reduced the need for additional capital outlay. Stranded capacity and allocated but not used capacity has been targeted for

reduction within this new architecture, and the first year results have shown that a new capacity demand curve could be achieved.

Additionally, when they do need to go forward for approval to acquire additional storage, it will be easier to justify the request as they can show that the existing storage is actually being used, not just allocated. This has improved the overall return on asset (ROA) measurements that are important to IT management.

Virtualizing the Overstock.com Storage Environment

“We depend heavily on Hitachi Data Systems to reliably run our business and scale as we grow. The Hitachi platform allows us to seamlessly virtualize and manage all the storage as one logical array. This architecture simplifies management, allows us to scale more cost-effectively and gives us the flexibility that we need to keep up with the demands of the business.”

*Carter Lee
VP, Technology Operations
Overstock.com*

Storage virtualization technology consolidates multiple vendor storage systems into a single pool of storage through a unified management interface. While storage virtualization is a phrase common in the storage industry, several types of virtualization solutions exist, many of which require new layers of management complexity or vendor lock-in, and may not support all server types in the heterogeneous environment.

Hitachi Controller-based Virtualization

Unlike server-based virtualization or SAN-based virtualization, for example, Hitachi platforms offer fully featured, controller-based virtualization with pervasive support for heterogeneous storage. This allows the entire data center to be managed as one virtual storage system from a single interface. The implication of that is that virtualization and tiering are handled through the Hitachi storage controller for both internal and externally attached heterogeneous storage, including mainframe storage. The Hitachi Universal Storage Platform V consolidates both NAS and SAN, and uses one set of GUI-based tools to securely migrate, configure, automate, monitor and manage that storage. Hitachi platforms scale up to 32 petabytes (PB) of managed storage capacity in a virtualized environment.

"We depend heavily on Hitachi Data Systems to reliably run our business and scale as we grow," says Carter Lee, VP, Technology Operations at Overstock.com. "The Hitachi platform allows us to seamlessly virtualize and manage all the storage as one logical array. This architecture simplifies management, allows us to scale more cost-effectively and gives us the flexibility that we need to keep up with the demands of the business."

The Virtualized Architecture

A Universal Storage Platform V forms the nucleus of the Overstock.com data center. In this case, the company is configured with internal storage and external virtualized capacity. The platforms provide external storage virtualization, performance, data mobility and (externally hosted) provisioning tools for flexible allocation, monitoring and movement of data between tiers.

The data center contains the Hitachi NAS cluster solution that functions as a highly scalable diskless file server and connects to the virtualized storage pool. (They have installed Hitachi NAS clusters at each site.) These gateways provide data accessibility to a broad range of servers and applications and are used to consolidate data behind the Universal Storage Platform V systems.

"By virtualizing external storage assets, Overstock.com was able to simplify the entire technology

refresh process. Migrations that typically took upwards of five hours took less than 30 minutes with the virtualization solutions," Lee explains.

Multitiered Infrastructure

Overstock.com implemented a virtualized multitiered storage solution based on the Hitachi Universal Storage Platform V. The Hitachi intelligent tiered storage architecture allows data to be classified according to the service level needs of business applications, and it enables seamless movement of that data to the correct storage tier. Tiering the storage architecture ensures that applications are always available and resources are appropriately managed.

Overstock has adopted a three tier storage infrastructure:

Tier 1: 15K RPM Fibre Channel Drives

Tier 2: 10K RPM Fibre Channel or SAS Drives

Tier 3: SATA

In a tiered storage environment, administrators gain the flexibility to establish guidelines for how and where the organization's information is stored. In a heterogeneous multiplatform environment, performance bottlenecks and storage utilization issues can emerge, slowing or even interrupting production. With a tiered storage architecture, Overstock.com is able to assign data according to the performance and availability needs required to meet these fluctuations, thereby eliminating barriers to sharing and improving capacity utilization.

The tiered storage architecture relies on heterogeneous virtualization technology to seamlessly and quickly move storage allocations between tiers as dictated by business demands and changes.

"The business impact of having a tool like Hitachi Tiered Storage Manager has been significant. With the implementation of Hitachi Tiered Storage Manager and Hitachi Tuning Manager, Overstock.com has been able to monitor performance of key applications and seamlessly migrate data to higher or lower tiers based upon business needs. This would have been impossible in previous architectures," adds Lee.

Meeting Provisioning Demands with Hitachi Dynamic Provisioning Software

Hitachi Dynamic Provisioning software, a thin provisioning product that provides virtual storage capacity from a storage pool, has no operational overhead and decreases the administration costs normally incurred with provisioning new storage. It also improves the availability of applications by reducing the downtime needed for the storage provisioning process. Dynamic Provisioning enables the definition of large amounts of virtual disk storage, allowing overallocation to applications and providing for planned and/or unplanned growth. When an application actually requires additional physical capacity, Overstock.com can allocate it from a storage pool nondisruptively.

With Dynamic Provisioning, Overstock.com has eliminated the complexity and costs of provisioning storage volumes on a per application basis. Without this tool, they would typically overallocate storage for each application to avoid the impact associated with incrementally allocating just enough additional storage to meet the current need. With Dynamic Provisioning they can still overallocate, but only by provisioning storage from the storage pool, as it is needed. This has reduced the need

to request and justify additional investments in storage to be allocated to applications for future needs, not immediate use.

With Dynamic Provisioning, administrators can allocate large amounts of virtual disk storage to an application but provision just enough physical storage to meet the short term requirements. When additional storage capacity is actually needed, it can be allocated from the Dynamic Provisioning storage pool with no impact to the availability of the application. Dynamic Provisioning eliminates the need to take the application down to change the storage configuration that is typically required when adding physical disk capacity. This significantly reduces the planned outages required to increase physical storage capacity and change the application server and storage system configurations. Hitachi Dynamic Provisioning significantly improves the efficiency of managing and using storage and increases the number of terabytes of storage that each storage administrator can manage. Physical disk requirements can be reduced due to more efficient utilization of the existing disk storage, resulting in environmental savings in floor space, power and cooling.

Dynamic Provisioning can also return unused storage to the Dynamic Provisioning storage pool. It does this by examining a volume's physical capacity. When it determines that no data other than zeros exist on a Dynamic Provisioning pool page, the physical storage page is unmapped and returned to the pool's free capacity and made available for re-use.

Sam Peterson, SVP, Technology at Overstock.com adds, "Managing all the storage assets as a single, virtualized, dynamically provisioned storage pool improves operational efficiencies. Provisioning tasks which typically took two to three hours the traditional way now take less than 30 minutes with Hitachi Dynamic Provisioning."

The Economic Impact Analysis

Storage Economics, from Hitachi Data Systems, is a series of methodologies, tools, solutions and planning practices that enables the application of financial and economic principles to storage technologies. The principles of economic measurement and identification of costs were introduced to Overstock.com to allow them to have quantifiable evidence of the cost impacts of the new architecture. This approach of econometrics will be employed going forward to provide continuous improvement measurements within IT management.

Introduction

The change from static tier storage architecture to one that uses virtualization of storage systems and volumes has allowed Overstock.com to significantly reduce unit costs, and reduce the raw capacity needed to maintain the current and future growth requirements. A total cost of ownership (TCO) baseline was created for the storage estate as it existed (with the old architecture) in December 2007. This baseline has been used to compare architecture improvements since then.

Detailed costs were summarized and divided by usable capacity to determine the unit costs. Overstock has defined 14 distinct storage cost categories, within three broad classifications as follows:

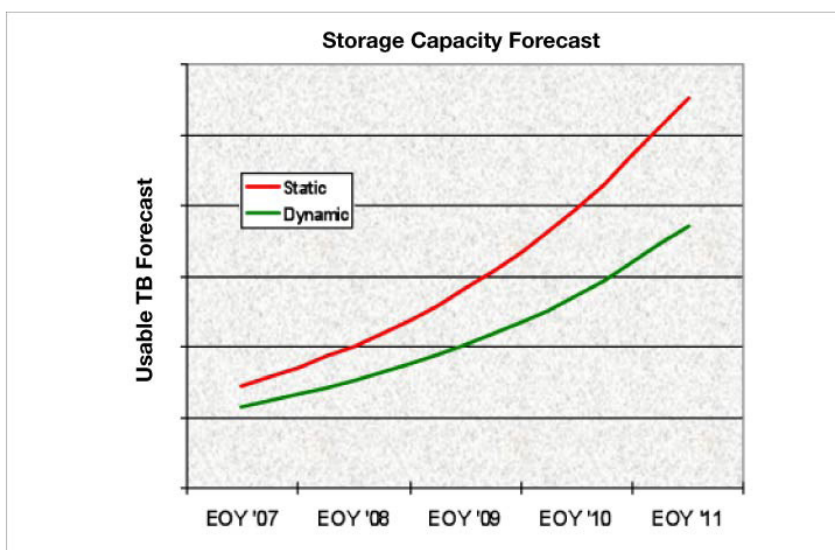
1. Capitalized Costs
 - a. Hardware depreciation
 - b. Software expense and depreciation
 - c. Hardware maintenance
 - d. Software maintenance
2. IT Infrastructure Costs
 - a. Storage management labor
 - b. Data center floor space
 - c. Power and cooling
 - d. SAN infrastructure
 - e. Disaster recovery infrastructure (circuits)
 - f. NAS infrastructure
 - g. Provisioning effort, business risk
 - h. Migration costs
3. Business Risk
 - a. Server and business risk associated with migration
 - b. Disaster recovery risk

Over time, with improvements in ROA, labor effort and effective tiering, Overstock.com has been able to see marked improvement in costs and qualitative measurements in the infrastructure.

Changing the Rate and Cost of Growth

As mentioned, Overstock.com has seen data growth rates exceed 40 percent annually. But with increasing pressure on IT budgets, a two-pronged approach was undertaken within this new architecture, as shown in Figure 1. First, reduce the overall storage "appetite" by reclaiming unused capacity and move to thin volumes. Second, reduce the cost of growth with dynamic tiers.

Figure 1. Changing the Storage Appetite with a New Architecture



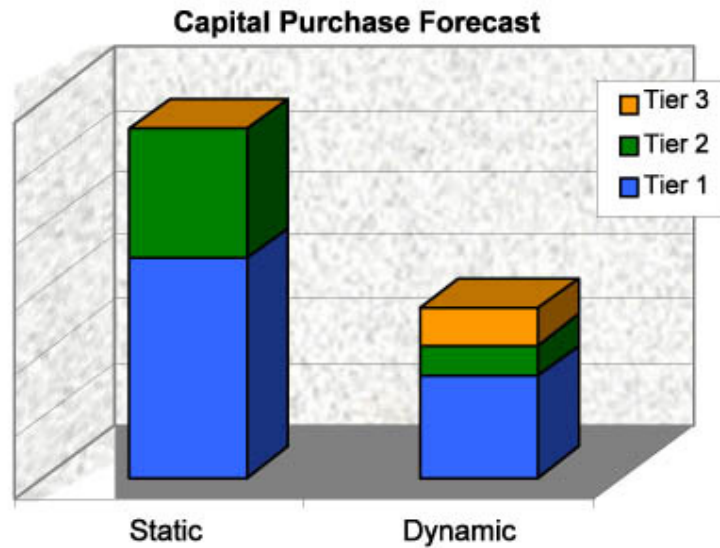
After virtualizing older, smaller storage systems, the common pool of storage presented (through the Hitachi Universal Storage Platform V) yielded more usable terabytes that were assigned to new capacity requirements. This reclaimed amount not only set a new trajectory for the above storage demand curve, but also provided real CAPEX avoidance. Figure 1 also shows a different capacity growth "slope." This is accomplished by presenting thin volumes to host.

Over the last two years, Overstock.com has been able to change the overall demand for storage (in terms of raw capacity), while still maintaining the growth rate for the business data. The difference in capacity between the red and green lines represents a real impact in terms of power, space, labor and CAPEX costs. These improvements will be quantitatively shown later in the Unit Cost econometrics.

Lower Cost of Growth

The Hitachi virtualized and tiered architecture allows data to move seamlessly up and down tiers. Overstock.com has used this function early on to migrate older data volumes to lower tiers. The rate of growing Tier 1 has dramatically changed in two years, while Tiers 2 and 3 have seen more of the capacity demand in growth. The compounded effect of growing seamlessly in three tiers and having a new appetite curve has provided Overstock.com with a much lower CAPEX commitment (see Figure 2). Comparing the two tiered static architecture in 2007 to the new architecture, the new architecture has, in the past two years, cut the CAPEX spending dramatically.

Figure 2. New Architecture Enabling Lower Cost of Growth (four years shown)

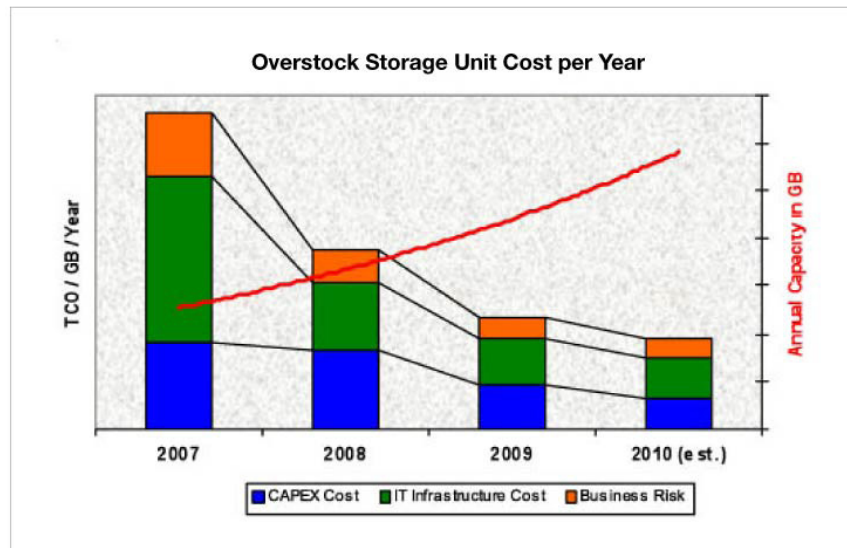


Documenting a Lower Unit Cost

The impact of a virtual, thinned and dynamically tiered storage architecture can be expressed in several impressive ways. Perhaps the management metric that matters most is that of a lower unit cost. IT departments, such as the one at Overstock.com, cannot necessarily control the demand or storage appetite, but through strategic and tactical investments they can drive down unit costs.

First year observations after the switch had been implemented include reducing the average unit cost of a GB/year of storage by 43 percent. The second year continued the unit cost improvements with a reduction of 37 percent.

Figure 3. The Rate of Unit Costs Compared to Yearly Growth



Some additional conclusions can be seen with these Unit Cost reductions (see Figure 3):

- Capacity growth (even modest) will drive down CAPEX unit costs as Moore's law is applied to the cost of disk. Some of the improvements in the CAPEX costs can be thus attributed, but infrastructure and risk costs were also reduced.
- Tiering started in 2008, but it produced more significant cost savings in 2009.
- The first year savings can be primarily attributed to:
 - Lower cost of migration of older storage systems
 - Immediate CAPEX impact of reclaiming usable terabytes
 - Simplified management with a single pool of storage for block and file access
 - Improved mean time to provision (including business risk)
- Most of capacity purchased in 2008 and 2009 has been tier 2 and tier 3.
- Cost of disaster recovery risk has not been significantly impacted (at this time) with a new architecture.
- Backup costs and processes are fundamentally unchanged; therefore, the costs of backup have not changed at the unit cost level.
- Unit cost of the SAN infrastructure rose slightly due to the higher concentration of SAN ports in front of the Universal Storage Platform V.
- NAS and SAN costs are amortized across all of the blended storage capacity.

Qualitative Benefits at Overstock.com

“Hitachi Data Systems has quickly simplified the process of nondisruptively provisioning storage. With Hitachi virtualization technologies, we've seen storage capacity savings of 50 percent on some arrays, now provision storage in 25 percent of the time, and have increased utilization rates by over 30 percent. We've reduced data migration related downtime from several hours to less than 30 minutes. Overall, by using Hitachi virtualization, dynamic provisioning and tiered storage, we've reduced our capital and operating costs for an improved return on our storage investment.”

*Carter Lee
VP, Technology Operations
Overstock.com*

Ease of Data Migration During Technology Refresh

With the implementation of Hitachi Tiered Storage Manager storage and external storage virtualization, Overstock.com has been able to substantially reduce the risk and effort required in remastering. Migration efforts that took upwards of five hours now take less than 30 minutes with fewer people.

Capacity Reclamation and Increased Utilization

Overstock.com has applied the thin provisioning capabilities of the Universal Storage Platform V on the Hitachi Adaptable Modular Storage 2500 virtualized behind the Universal Storage Platform V to improve performance and reclaim storage capacity. Storage utilization is now 80 percent or more in the virtualized thinned environment. The initial effort to virtualize older storage systems yielded additional usable capacity to be used for new projects.

Ease of Application Provisioning and Time to Market

Planning for and implementing new applications required extensive design and planning for the storage layout to ensure there was adequate storage with acceptable performance and availability. The planning included overallocation to avoid future outages due to either planned or unplanned application growth. With the implementation of Hitachi Dynamic Provisioning software, Overstock.com has been able to significantly reduce and in some cases eliminate this step, which consequently has improved time to market for the new applications.

Alignment of Storage Tiers with Business Needs

With the implementation of Hitachi Tiered Storage Manager and Hitachi Tuning Manager software, Overstock.com has been able to monitor performance of key applications and seamlessly migrate data to higher or lower tiers based upon business needs.

Improved Operational Efficiencies

Prior to implementing Hitachi Dynamic Provisioning software, Overstock.com storage administrators spent time monitoring the performance of every application, watching for hot spots. These hot spots required the movement of data to another location. With the wide striping feature of Dynamic Provisioning, they have eliminated the hot spots, freeing up the administrators for other tasks. In the past, hot spots were discovered a few times each quarter with an estimated operation impact of 40 to 50 hours per quarter. This was eliminated through the use of Dynamic Provisioning software. Also, managing the storage assets as a single pool improves operational efficiencies. Provisioning tasks that typically took two to three hours the traditional way now take less than 30 minutes with Dynamic Provisioning.

Planning and Projecting a Lower TCO in the Future

Hitachi Data Systems has projected another unit cost TCO reduction by 2010, with continuous economic improvement options beyond that time. This new architecture will enable Overstock.com, as they are ready, to implement new processes and technical features that will continue the downward trend of unit cost of TCO. Some of these options are summarized below:

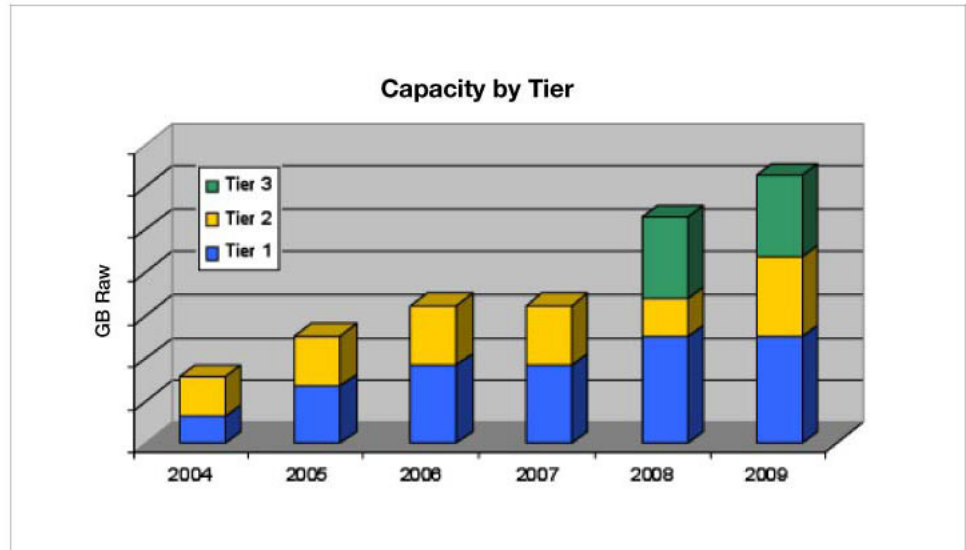
- An active archive solution can be included as a new lower cost or intelligent tier to move and index data to a very low cost tier. Federated search and data management capabilities can additionally lower discovery risk.
 - Hitachi Storage Capacity Reporter software, powered by APTARE, is a tool provided by Hitachi Data Systems that will assist Overstock.com in looking at utilization trends to assist with data allocations.
 - Policy-based management is an advanced set of tools and intelligence that can allow established policies to dictate data migration, movement, tiering and lifecycle management. This will continue to lower unit costs with policies that drive data costs over time.
 - The Overstock.com economics can be expanded to start tracking unit cost by tier, rather than the blended rate. Additional data and tier alignments will be possible with the right tier cost metrics showing patterns and trends of the users.
 - Chargeback and management reports on asset allocations mean costs can be provided to (over time) influence end user storage consumption behavior.
-

Appendix A — Overstock.com Profile

- Online closeout retailer established in 1997
 - Headquartered in Salt Lake City, Utah
 - Merchandise: includes bed-and-bath goods, home decor, kitchenware, watches, jewelry, electronics and computers, sporting goods, apparel and designer accessories
 - Web site: <http://www.Overstock.com>
-

Appendix B – Comparing the Architectures

Appendix B – Figure 1. Raw Capacity and Tiered Allocation by Year



 **Hitachi Data Systems Corporation**

Corporate Headquarters

750 Central Expressway
Santa Clara, California 95050-2627 USA
www.hds.com

Regional Contact Information

Americas: +1 408 970 1000 or info@hds.com
Europe, Middle East and Africa: +44 (0) 1753 618000 or info.emea@hds.com
Asia Pacific: +852 3189 7900 or hds.marketing.apac@hds.com

Hitachi is a registered trademark of Hitachi, Ltd., in the United States and other countries. Hitachi Data Systems is a registered trademark and service mark of Hitachi, Ltd., in the United States and other countries.

All other trademarks, service marks and company names in this document or website are properties of their respective owners.

Notice: This document is for informational purposes only, and does not set forth any warranty, expressed or implied, concerning any equipment or service offered or to be offered by Hitachi Data Systems Corporation.

© Hitachi Data Systems Corporation 2010. All Rights Reserved. WP-369-B DG August 2010