

# Using VMware vSphere 4 with Hitachi Dynamic Provisioning Software on the Hitachi Adaptable Modular Storage 2000 Family

Best Practices Guide

*By Bob Ingram*

July 2010

## Summary

This white paper describes testing that Hitachi Data Systems conducted to develop best practices for making storage provisioning and virtual disk format decisions in 2000 family environments. Hitachi Data Systems mapped Dynamic Provisioning LUs to a Virtual Machine File System datastore (VMFS datastore) and virtual disks that were created in three formats: thin, zeroedthick and eagerzeroedthick.

This testing supports the following conclusions:

- When maximum initial performance is required, use eagerzeroed thick virtual disk format to eliminate warm-up anomalies, and Dynamic Provisioning LUs for wide striping benefits.
- When maximum cost savings and low administrative overhead are required, use zeroedthick virtual disks and Dynamic Provisioning LUs.

The purpose of this white paper is to show how organizations can leverage Hitachi Dynamic Provisioning software running on the Hitachi Adaptable Modular Storage 2000 family to maximize benefits and minimize operational complexities in their vSphere 4 environments. Specifically, the paper is designed to help storage and data center administrators best provision storage within vSphere 4 environments by understanding how Hitachi Dynamic Provisioning software works and the different virtual disk format choices available to them.

**For best results use Acrobat Reader 8.0.**

## Feedback

Hitachi Data Systems welcomes your feedback. Please share your thoughts by sending an email message to [SolutionLab@hds.com](mailto:SolutionLab@hds.com). Be sure to include the title of this white paper in your email message.

# Table of Contents

<b>Test Environment .....</b>	<b>2</b>
Hitachi Adaptable Modular Storage 2000 Family .....	2
Hitachi Dynamic Provisioning Software.....	4
Servers.....	5
VMware vSphere 4.....	5
<b>Component Details.....</b>	<b>5</b>
Hitachi Dynamic Provisioning Software Operation .....	5
vSphere 4 Virtual Disk Formats.....	6
<b>Functional Lab Tests.....</b>	<b>6</b>
Deploying a VM From a Template.....	7
Storage vMotion Operations on Dynamic Provisioning LUs .....	8
Converting a Virtual Disk From Thin to Thick.....	9
VMFS Datastore Expansion .....	10
Best Practices for Dynamic Provisioning Space Savings and Virtual Disks .....	10
<b>Performance Test Analysis .....</b>	<b>11</b>
Test Methodology.....	11
Warm-up Anomaly Tests.....	11
Production Workload Simulation on Standard LUs Tests.....	14
Production Workload Simulation on Dynamically Provisioned LUs Tests .....	27
<b>Conclusion.....</b>	<b>35</b>

# Using VMware vSphere 4 with Hitachi Dynamic Provisioning Software on the Hitachi Adaptable Modular Storage 2000 Family

## Best Practices Guide

This white paper describes testing that Hitachi Data Systems conducted to develop best practices for making storage provisioning and virtual disk format decisions in 2000 family environments. Hitachi Data Systems mapped Dynamic Provisioning LUs to a Virtual Machine File System datastore (VMFS datastore) and virtual disks that were created in three formats: thin, zeroedthick and eagerzeroedthick.

This testing supports the following conclusions:

- When maximum initial performance is required, use eagerzeroed thick virtual disk format to eliminate warm-up anomalies, and Dynamic Provisioning LUs for wide striping benefits.
- When maximum cost savings and low administrative overhead are required, use zeroedthick virtual disks and Dynamic Provisioning LUs.

---

**Note:** A thin-friendly operation is one that allows cost savings through the use of over provisioning of the storage. The file system and application make use of the space efficiently, and do not claim blocks of storage that are not currently needed.

---

The purpose of this white paper is to show how organizations can leverage Hitachi Dynamic Provisioning software running on the Hitachi Adaptable Modular Storage 2000 family to maximize benefits and minimize operational complexities in their vSphere 4 environments. Specifically, the paper is designed to help storage and data center administrators best provision storage within vSphere 4 environments by understanding how Hitachi Dynamic Provisioning software works and the different virtual disk format choices available to them.

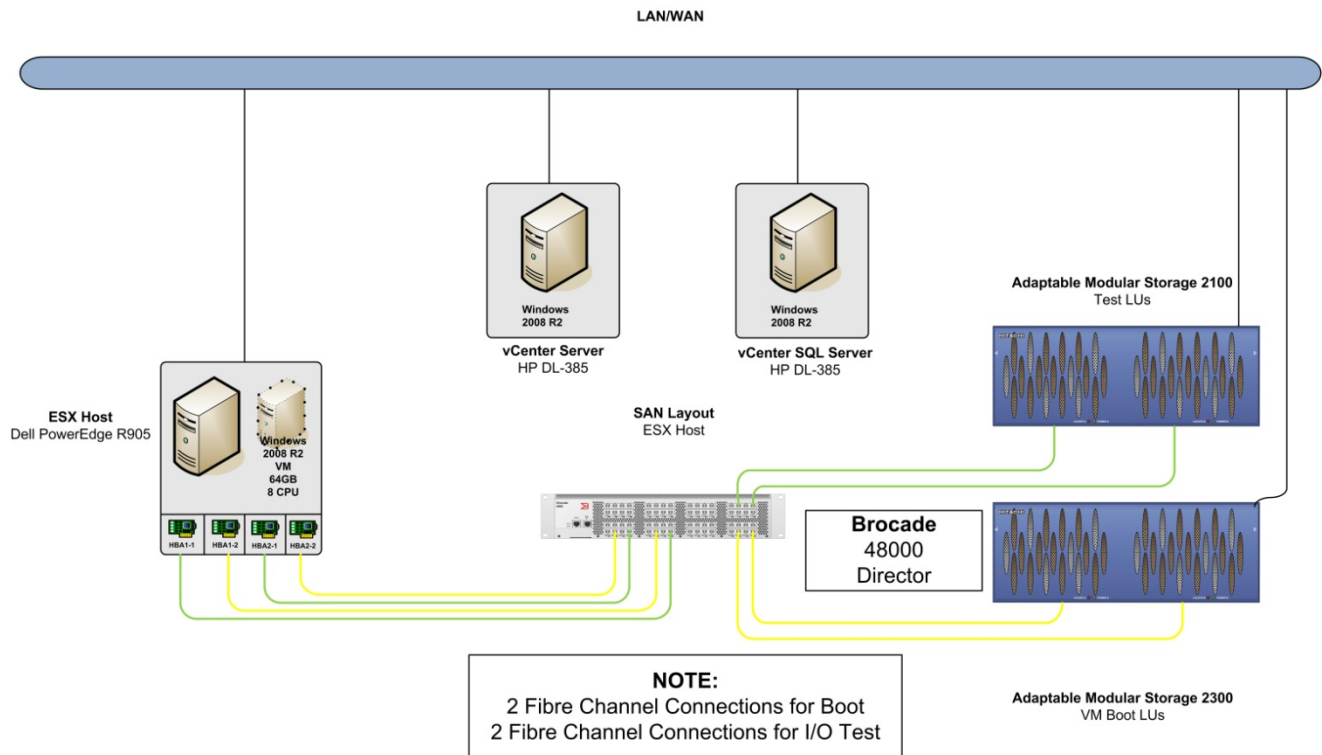
At a high level, logical units (LUs) on the Hitachi Adaptable Modular Storage 2000 family can be created as standard RAID devices or they can be provisioned as pooled LUs using Hitachi Dynamic Provisioning software running on 2000 family storage systems.

This white paper is intended for data center and storage administrators charged with planning and implementing storage in a vSphere environment. Readers need to be knowledgeable about vSphere storage administration.

# Test Environment

The following sections describe the test environment used in the Hitachi Data Systems lab. Figure 1 shows physical layout of the Hitachi Data Systems lab.

**Figure 1. Physical Layout of Test Lab**



## Hitachi Adaptable Modular Storage 2000 Family

Hitachi Data Systems used a Hitachi Adaptable Modular Storage 2100 in its laboratory testing. The 2100 is a member of the Hitachi Adaptable Modular Storage 2000 family, which are the only midrange storage systems with symmetric active-active controllers that provide integrated, automated hardware-based front-to-back-end I/O load balancing. Both controllers in a 2000 family storage system are able to dynamically and automatically assign the access paths from the back of the controller to the LU. All LUs are accessible regardless of the physical storage front-end port or the server from which the access is requested. Utilization rates of each controller are monitored so that a more even distribution of workload between the two controllers can be maintained. When coupled with VMware round-robin load balancing, a 2000 family storage system eliminates many complex and time consuming path planning tasks that storage administrators typically face.

The point-to-point back end design of the 2000 family virtually eliminates I/O transfer delays and contention associated with Fibre Channel arbitration and provides significantly higher bandwidth and I/O concurrency.

Although the Hitachi Adaptable Modular Storage 2100 was used for this solution, the information in this paper is relevant for the other 2000 family members with changes to account for capacity and performance differences. The 2000 family is an easy-to-use, scalable, cost effective storage system for mission-critical business applications. It is also a top choice for tiered and standalone storage, consolidation, business continuity, data replication, backup and archiving. The Adaptable Modular Storage 2100 offers a rich set of features in a model that scales to 159 disk drives and delivers enterprise-class performance and capabilities at

a modular price. For more information about 2000 family storage systems, see the Hitachi Adaptable Modular Storage 2000 family [web site](#).

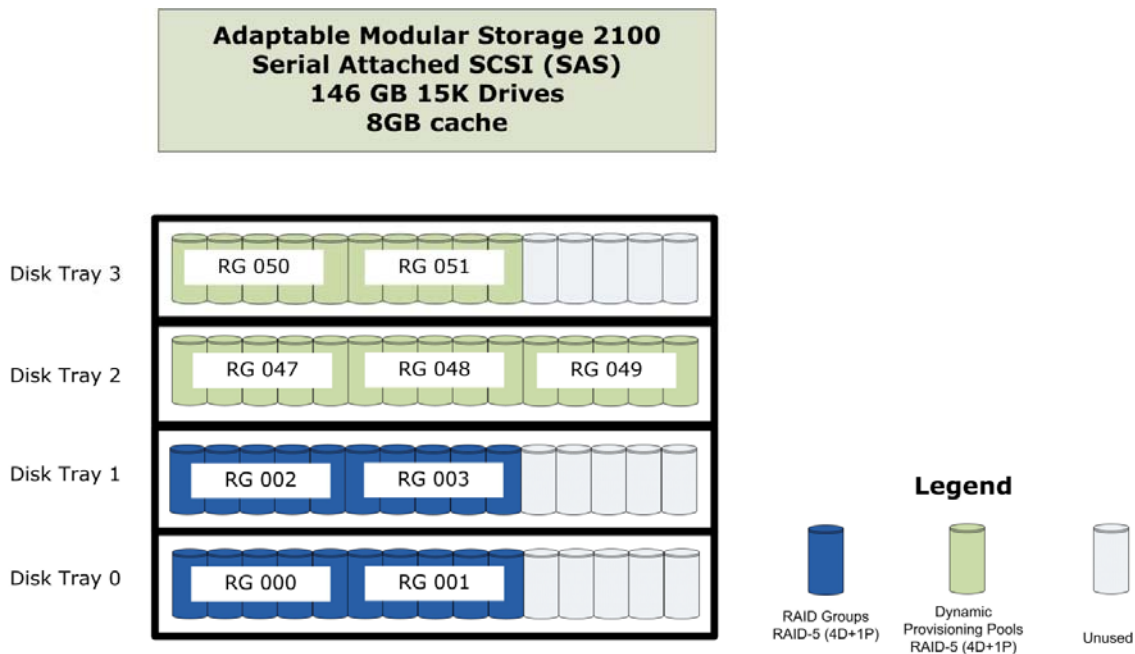
Table 1 lists the configuration for the 2100 used in Hitachi Data Systems testing.

**Table 1. Hitachi Adaptable Modular Storage 2100 Test Configuration**

<i>Item</i>	<i>Description</i>
Microcode	0883/A-S
Cache memory	4GB
Number of ports	4
Number of RAID groups	5
Size of each physical LU	531GB, (1 RAID group)
Number of RAID groups in Dynamic Provisioning pool	DP Pool -000: 2 RAID groups scaled to 4 RAID groups DP Pool -001: 1 RAID group
Number of Dynamic Provisioning pools	2
Number of LUs in Dynamic Provisioning pools	4
Size of each Virtual LU	531GB
RAID group type	RAID-5 (4D+1P)
Number of drives	60
Drive capacity	146GB
Drive type	SAS 15K RPM
Multipathing enabled	VMware native round robin

Figure 2 illustrates the RAID group configuration used on the 2100.

**Figure 2. RAID Group Configuration**



## Hitachi Dynamic Provisioning Software

On Hitachi Adaptable Modular Storage 2000 family systems, Hitachi Dynamic Provisioning software's thin provisioning and wide striping functionalities that provide virtual storage capacity to eliminate application service interruptions, reduce costs and simplify administration, as follows:

- Optimizes or "right-sizes" storage performance and capacity based on business or application requirements.
- Supports deferring storage capacity upgrades to align with actual business usage.
- Simplifies and adds agility to the storage administration process.
- Provides performance improvements through automatic optimized wide striping of data across all available disks in a storage pool.

The wide-striping technology that is fundamental to Hitachi Data Provisioning software dramatically improves performance, capacity utilization and management of your environment. By deploying your virtual disks using Dynamic Provisioning LUs from Dynamic Provisioning pools on the 2000 family, you can expect the following benefits:

- A smoothing effect to virtual disk workload that can eliminate hot spots across the different RAID groups, reducing the need for VMFS workload analysis by the VM.
- Elimination of excess, unused capacity by leveraging the combined capabilities of all disks comprising a storage pool.

For more information, see the Hitachi Dynamic Provisioning software [datasheet](#).



## Servers

Table 2 lists the servers that were used in the Hitachi Data Systems lab.

**Table 2. Tested Servers**

<i>Model</i>	<i>Configuration</i>	<i>Quantity</i>	<i>Operating System</i>
Dell PowerEdge R905	4 x Quad-Core AMD Opteron processor 1.9 GHz, 128GB RAM and equipped with 2 Emulex LPe11002 4GB host bus adapters (HBAs)	2	VMware ESX 4.0
HP DL0385	4 x Dual Core AMD Opteron Processors 2.8GHz, 8GB RAM	2	Windows Server 2008 R2 Enterprise Edition 64-bit

## VMware vSphere 4

vSphere 4 is a highly efficient virtualization platform that provides a robust, scalable and reliable infrastructure for the data center. vSphere features like the Distributed Resource Scheduler, High Availability, Fault Tolerance provide an easy to manage platform. For more information see [VMware's vSphere web site](#).

Table 3 lists the vSphere 4 software used in the Hitachi Data Systems lab.

**Table 3. Tested vSphere 4 Software**

<i>Software</i>	<i>Version</i>
VMware vCenter Server	4.0.0 build 208111
VMware Virtual Infrastructure Client	4.0.0
VMware ESX	4.0.0 build 164009

All of the virtual machines used in this lab environment ran Windows Server 2008 R2 Enterprise Edition 64-bit as their operating systems.

## Component Details

Planning an IT infrastructure that uses Hitachi Dynamic Provisioning software and vSphere virtualization software requires an in-depth understanding of how these products work. The following sections provide additional details about these components.

### Hitachi Dynamic Provisioning Software Operation

Hitachi Dynamic Provisioning software uses two primary components: a pool of physical storage capacity (Dynamic Provisioning pool) and virtual volumes (Dynamic Provisioning LUs). The Dynamic Provisioning pool is composed of standard RAID groups. Up to 64 Dynamic Provisioning Pools can be created on a single 2000 family system. A Dynamic Provisioning LU is an LU of specified size created within a Dynamic Provisioning pool. The host sees the LU as provisioned at the specified size; however, the physical space is allocated from the Dynamic Provisioning pool as needed.

As the application writes data to the Dynamic Provisioning LU, the storage processor allocates physical capacity from the Dynamic Provisioning pool while monitoring multiple thresholds to alert administrators when more physical pool capacity is needed.

The underlying storage for a Dynamic Provisioning pool can be any supported RAID level, for example RAID-5 (4D+1P) or RAID-1+0 (2D+2D), and any supported disk type, for example 15K RPM SAS, 7200RPM SATA. However, all storage in a single pool must be the same RAID level and drive type. The Dynamic Provisioning

pool presents a striped space across consecutive sets of RAID stripes using 1GB chunks. Each 1GB chunk is allocated as write I/O occurs to the Dynamic Provisioning LU.

The Dynamic Provisioning LU takes on the RAID characteristics of the underlying RAID and disk type, so plan the Dynamic Provisioning pool accordingly. A larger number of RAID groups in the Dynamic Provisioning pool increases the performance of Dynamic Provisioning LUs associated with the Dynamic Provisioning pool. Workload can be balanced across many disks by allocating a large number of disks and assigning Dynamic Provisioning LUs to hosts with a variety of workload needs.

For optimum performance on the Dynamic Provisioning pool, pre-allocate all chunks required for the application. Run a random write workload using a utility like Iometer or VDBench until all chunks are allocated in the Dynamic Provisioning pool. For more information, see the “Warm-up Anomalies” section. The latency created by adding chunks to the Dynamic Provisioning LU is small, so your environment suffers little I/O performance penalty if you do not perform this warm-up, however the wide-striping benefit of Hitachi Dynamic Provisioning software is not fully realized until all chunks are allocated.

## vSphere 4 Virtual Disk Formats

vSphere 4 uses three virtual disk formats, which are listed in Table 4.

**Table 4. vSphere 4 Virtual Disk Formats**

<i><b>Format</b></i>	<i><b>Description</b></i>
Thin	When thin format is used, the . vmdk file is allocated only to the size required by the guest OS. As write operations occur, additional space is allocated and zeroed in the . vmdk and the virtual disk grows up to the maximum allotted size.
Zeroedthick	When zeroedthick format is used, the . vmdk file is the size allocated but the space is not “pre-zeroed.” As the guest OS writes out to the . vmdk the space is zeroed as needed, but the . vmdk file does not grow in size. Zeroedthick is a Hitachi Dynamic Provisioning software-friendly format and allocates Dynamic Provisioning pool space on demand. This is the default virtual disk format.
Eagerzeroedthick	When eagerzeroedthick format is used, the . vmdk file is the size allocated, and the space is zeroed. As the guest OS writes data to the . vmdk file the VMkernel does not need to write out zeros before the OS can have access to the storage. The result is improved I/O latency but more back-end storage operations during virtual disk creation. In the case of Hitachi Dynamic Provisioning software, all storage required is pre-allocated in the Dynamic Provisioning pool. This virtual disk format is required for vSphere Fault Tolerance.

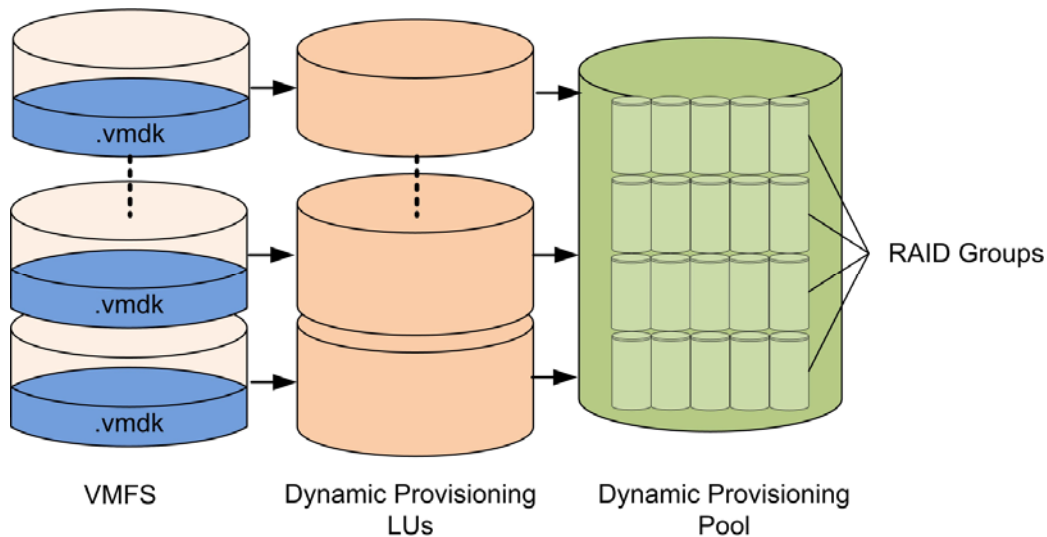
For this paper, thin, zeroedthick and eagerzeroedthick format virtual disks were tested on dynamically provisioned and standard LUs using RAID-5 (4D+1P). Additional RAID levels are supported; however, these RAID levels are out of this document’s scope.

## Functional Lab Tests

Hitachi Data Systems conducted a series of tests to determine the effect of vSphere 4 operations on Dynamic Provisioning LUs. The goal of this testing was to show which vSphere operations are Dynamic Provisioning thin friendly.

Figure 3 shows the relationship of the virtual disk (. vmdk) to VMFS, Dynamic Provisioning LU and Dynamic Provisioning pool.

**Figure 3. Logical Relationship Between VMFS, Dynamic Provisioning LUs and Dynamic Provisioning Pools**



## Deploying a VM From a Template

In these tests, all Hitachi Adaptable Modular Storage 2100 functions were performed using Hitachi Storage Navigator Modular 2 software and all vSphere 4 operations were performed using vCenter and the Virtual Infrastructure client.

### *Deploying a Virtual Machine from a Template Using Zeroedthick Format Virtual Disk*

The goal of this test was to determine how much Dynamic Provisioning pool space is consumed by the new VM and if this is an effective way to save space.

Hitachi Data Systems followed this test procedure and made the following observations:

1. Create a RAID-5 (4D+1P) group using 146GB SAS disks.
2. Add a 50GB LU to the RAID group and create a VMFS datastore called RAI D-5-source.
3. Deploy a VM using zeroedthick virtual disk format on RAI D-5-source and install Windows 2008 R2.

Windows OS uses approximately 12GB of disk space.

4. Shut down the Windows VM and convert it to a template.
5. Create a Dynamic Provisioning pool containing three RAID-5 (4D+1P) groups with 146GB SAS disks.
6. Create a Dynamic Provisioning LU of 50GB and created a VMFS datastore called RAI D-5-target.

The datastore uses 1GB of space from the Dynamic Provisioning pool.

7. Deploy a VM from the template using a zeroedthick format virtual disk on RAI D-5-target using 50GB virtual disks.

The datastore uses 14GB of space from the Dynamic Provisioning pool.

## Conclusion

The zeroedthick format virtual disk only uses the actual space needed by the guest operating system, data and applications from the storage system perspective. Deploying virtual machines from templates that have the guest OS installed on a zeroedthick format virtual disk is a Dynamic Provisioning thin-friendly operation. Hitachi Data Systems recommends using zeroedthick format virtual disks for VM templates.

### *Deploying a Virtual Machine from a Template Using Eagerzeroedthick Format Virtual Disk*

The goal of this test was to determine how much Dynamic Provisioning pool space was consumed by the new VM and if this was an effective way to save space.

Hitachi Data Systems followed this test procedure and made the following observations:

1. Using the template created in the “Deploying a Virtual Machine from a Template Using Zeroedthick Format Virtual Disk” section, deploy a VM from the template using a thick format virtual disk on RAI D-5- target using 50GB virtual disks.
2. Enable Fault Tolerance on the newly created VM.

This operation converts the virtual disk from zeroedthick to eagerzeroedthick format.

The datastore uses 50GB of space from the Dynamic Provisioning pool.

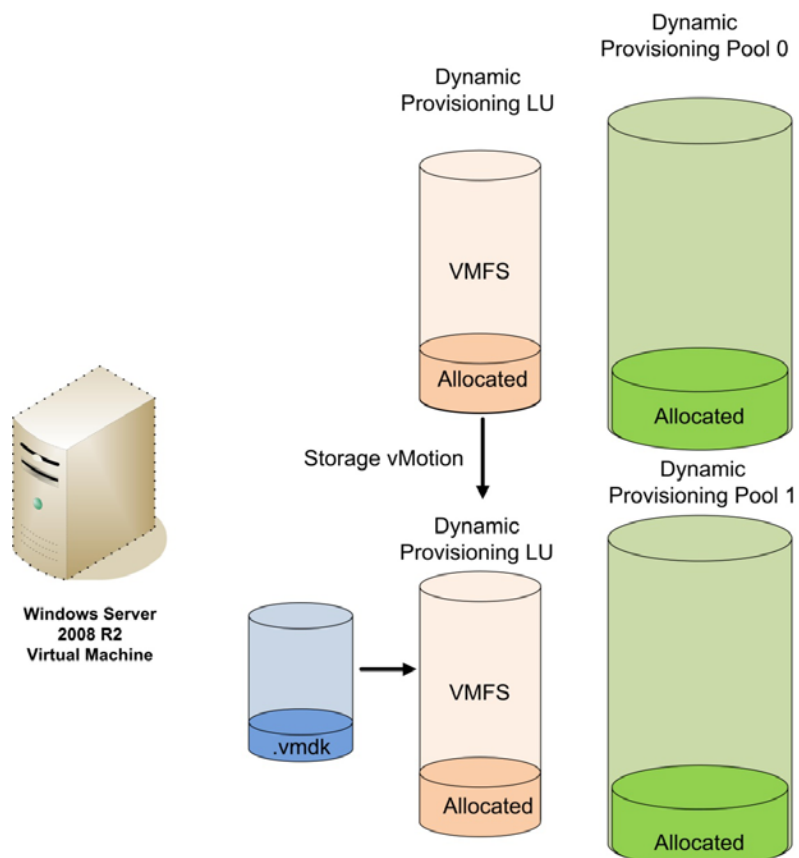
### **Conclusion**

An eagerzeroedthick format virtual disk uses the virtual disk's actual total assigned space. Deploying virtual machines from templates in eagerzeroedthick format to any virtual disk format is not a storage system thin-friendly operation.

### **Storage vMotion Operations on Dynamic Provisioning LUs**

The goal of this test was to determine how much Dynamic Provisioning pool space is consumed when a zeroedthick format virtual disk is moved from one Dynamic Provisioning pool to another. Figure 4 shows the Storage vMotion operation.

**Figure 4.Storage vMotion**



Hitachi Data Systems followed this test procedure and made the following observations:

1. Create two Dynamic Provisioning pools on the Hitachi Adaptable Modular Storage 2100, DP Pool -000 and DP Pool -001.
2. Create two 50GB Dynamic Provisioning LUs, LU-000 in DP Pool -000 and LU-001 in DP Pool -001.
3. Configure two VMFS datastores named RAID-5-vMotion-1 on LU-000, and RAID-5-vMotion-2 on LU-001 created in Step 2.
4. Create a 50GB virtual disk in zeroedthick format on the VMFS datastore RAID-5-vMotion-1 and assign it to the VM created in the deployment tests.
5. On the Windows host, run Iometer with a 100 percent write, 4KB block I/O workload to Physical Drive2 until the Dynamic Provisioning LU shows 50 percent (25GB) utilization.
6. Using the Storage vMotion GUI, migrate the virtual disk from the VMFS datastore RAID-5-vMotion-1 to the VMFS datastore RAID-5-vMotion-2, selecting the **Same format as source** radio button.

The datastore uses 25GB of space from the Dynamic Provisioning pool used in DP Pool -001.

## Conclusion

Storage vMotion is a Dynamic Provisioning thin-friendly operation.

## Converting a Virtual Disk From Thin to Thick

The goal of this test was to determine the Dynamic Provisioning pool space utilization when a thin format virtual disk is converted to a thick format virtual disk.

Hitachi Data Systems followed this test procedure:

1. Create a 100GB Dynamic Provisioning LU and assign it to the vSphere host.
2. Create a VMFS datastore on the Dynamic Provisioning LU.
3. Create a 100GB virtual disk on a VM.

The data store uses 1GB of space from the Dynamic Provisioning pool.

4. Convert the .vmdk file to thick format using the vCenter GUI

The data store uses 100GB of space from the Dynamic Provisioning pool.

## Conclusion

Converting a thin virtual disk to a thick virtual disk is not a Dynamic Provisioning thin-friendly operation because the thick provisioning in this case is eagerzeroedthick.

## VMFS Datastore Expansion

A new feature of vSphere 4 is the ability to expand VMFS datastore with no downtime. The goal of this test was to evaluate the effect of this feature on the 2100.

Hitachi Data Systems followed this test procedure and made the following observations:

1. Using the Dynamic Provisioning pool and Dynamic Provisioning LU associated with the RAI D-5-source VMFS datastore in previous tests, add a RAID group to the Dynamic Provisioning pool.
2. Increase the size of the Dynamic Provisioning LU from 50GB to 100GB.

Dynamic Provisioning pool usage does not increase.

3. After the Dynamic Provisioning LU size increases, rescan the ESX host to see the new capacity.
4. Expand the VMFS datastore.

Dynamic Provisioning pool usage does not increase.

### Conclusion

The VMFS datastore expansion feature of vSphere 4 on the Hitachi Adaptable Modular Storage 2100 is thin-friendly. Any further operations, such as expanding or adding a .vmdk file, are dependent on the virtual disk format used.

## Best Practices for Dynamic Provisioning Space Savings and Virtual Disks

Two of vSphere's virtual disk formats are thin-friendly, meaning they only allocate chunks from the Dynamic Provisioning pool as required. Thin and zeroedthick format virtual disks are thin-friendly, eagerzeroedthick format virtual disks are not. The eagerzeroedthick format virtual disk allocates 100 percent of the Dynamic Provisioning LUs space in the Dynamic Provisioning pool. While the eagerzeroedthick format virtual disk does not give the benefit of cost savings by over provisioning of storage, it can still assist in the wide striping of the Dynamic Provisioning LU across all disks in the Dynamic Provisioning pool.

When using Dynamic Provisioning LUs to overprovision storage, follow these best practices:

- Create the VM template on a zeroedthick format virtual disk, When deploying, select the **Same format as source** radio button in the vCenter GUI.
- Use the default zeroedthick format virtual disk.

Note that you can use the thin virtual disk format and achieve similar space savings; however, this creates an additional data point to monitor and additional ESX CPU overhead. Hitachi Data Systems recommends using Hitachi Dynamic Provisioning software with zeroedthick virtual disks when thin provisioning is required. This ultimately lessens management tasks for data center administrators.

- Using Storage vMotion when the source VMFS datastore is on a Dynamic Provisioning LU is a Dynamic Provisioning thin friendly operation.

Keep in mind that this operation does not zero out the VMFS datastore space that was freed by the Storage vMotion operation, meaning that Hitachi Dynamic Provisioning software cannot reclaim the free space.

- When using Dynamic Provisioning software for cost savings, do not convert a thin format virtual disk to thick.

# Performance Test Analysis

For these performance tests, Hitachi Data Systems used VDbench to drive I/O to the virtual disk. The goal was to compare virtual disk formats from both the thin-friendly aspects as well as performance distribution from the wide striping of Hitachi Dynamic Provisioning software.

## Test Methodology

The standard RAID LU tests used RAID-5 (4D+1P) groups on 146GB 15K RPM disks.

The Dynamic Provisioning tests used a Dynamic Provisioning pool containing RAID-5 (4D+1P) groups. These tests used the same number of disks and LUs to do a direct comparison.

All tests used 530GB LUs and 100GB virtual disks.

Table 5 shows the workloads used for these tests. Each workload test was run for six minutes.

In addition, Hitachi Data Systems ran all tests on thin, zeroedthick and eagerzeroedthick format virtual disks to observe any performance differences at initial creation time and after writing to all blocks on the virtual disks.

## Warm-up Anomaly Tests

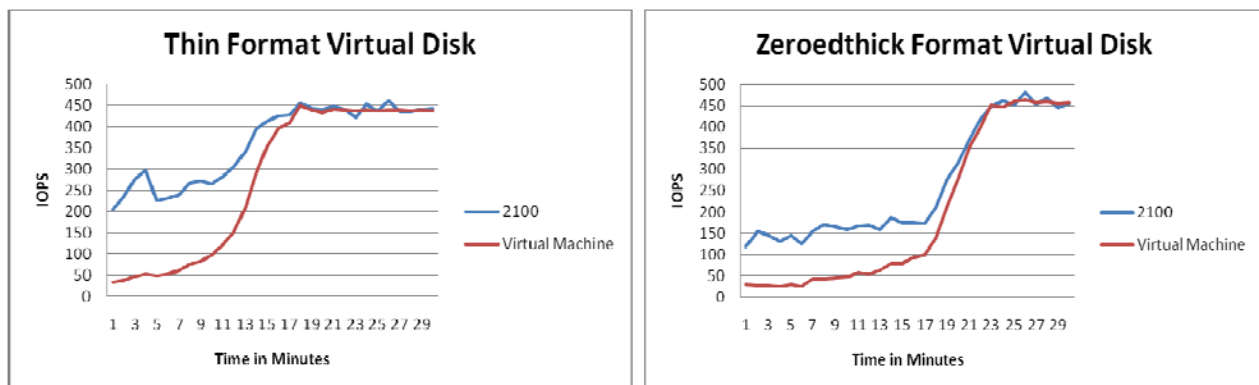
Hitachi Data Systems first performed a series of tests to examine warm-up time and compare performance on newly created thin, zeroedthick and eagerzeroedthick format virtual disks. The objective of these tests was to quantify the warm-up anomalies to assist data center administrators in deciding which virtual disk format best meets their requirements. In these tests, the zeroedthick and thin format virtual disks were warmed up and compared to the eagerzeroedthick format virtual disk. This warm-up occurs one time, when a block on the virtual disk is written to for the first time.

These tests used a Dynamic Provisioning pool on a single RAID-5 (4D+1P) group. The Dynamic Provisioning pool had a 530GB Dynamic Provisioning LU and 100GB virtual disk.

The warm-up routine ran a 100 percent write, 100 percent random, 8K block workload for 30 minutes. Hitachi Data Systems collected statistics about the 2100 storage system's Fibre Channel ports and statistics from VDbench. Note that because this was a 100 percent write workload, not a mix of read and write as seen on an OLTP workload, the warm-up routine had a higher total throughput than the OLTP workload.

For the first 15 minutes of the warm-up routine, the 2100 saw more IOPS than the VM produced. The additional IOPS resulted from the zeroing of the blocks needed before the VM was able to write to the virtual disk. After 23 minutes the IOPS were balanced and the warm-up completed. Figure 5 shows the results of these tests.

**Figure 5. Warm-up Anomaly Test Results**



Notice that the thin format virtual disk has more overhead than the zeroedthick format virtual disk, as shown by the greater difference between the VM's IOPS and the 2100 storage system's IOPS. Subsequent tests showed that a fully allocated thin format virtual disk performs the same as a zeroedthick or eagerzeroedthick; therefore the overhead only occurs during the warm-up period.

This test used an OLTP workload with 8K blocks, 75 percent read, 25 percent write, 100 percent random workload with eight threads and an I/O rate of 400 IOPS. The warm-up routine used an 8K block, 100 percent write, 100 percent random workload with eight threads run for 30 minutes. The goal was to write to every block at least once to completely zero the virtual disk.

Figure 6 shows the pre warm-up IOPS performance for these tests. At the start of the warm-up period, any write IOPS greater than 25 percent of the total IOPS are I/Os generated by vSphere for the purpose of zeroing blocks prior to writing data.

**Figure 6. Pre Warm-up IOPS Test Results**

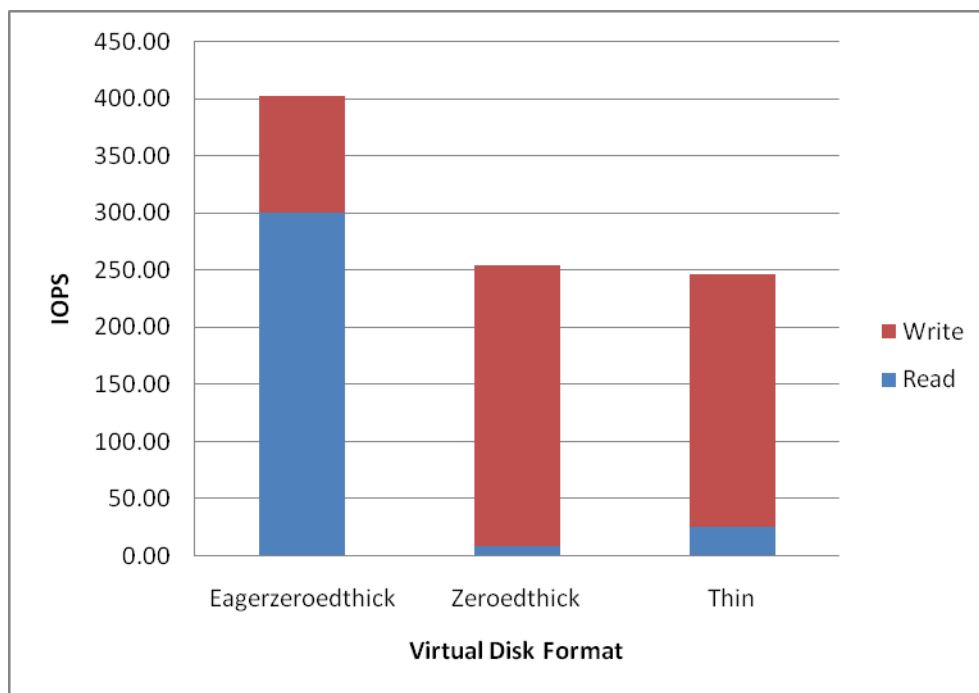
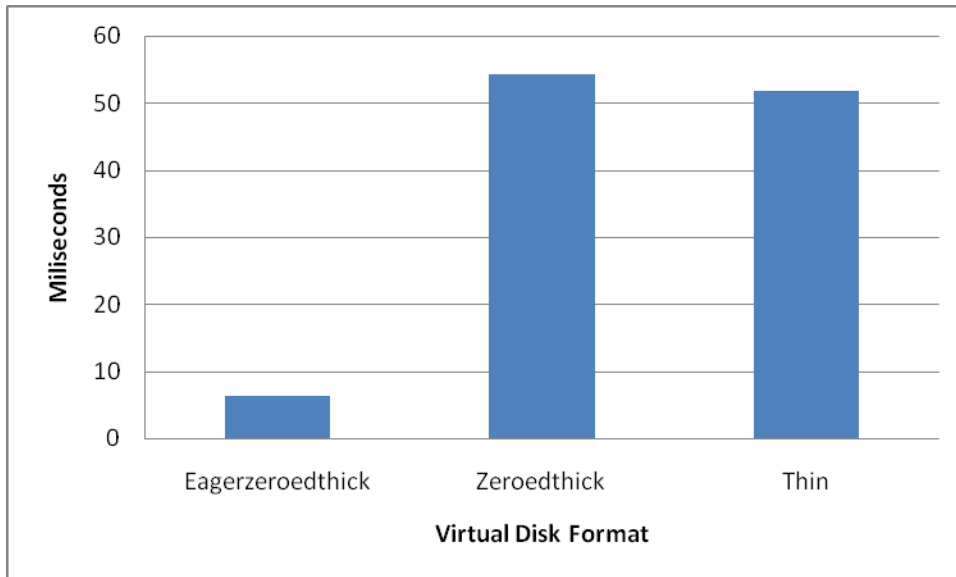




Figure 6 shows that the write IOPS for the zeroedthick format virtual disk are higher than for the eagerzeroedthick format virtual disk. This is because the VMFS issues a zero write before it can write to the blocks. The read IOPS in thin and zeroedthick are low as a result of the very high write IOPS required to zero the space before a write can occur.

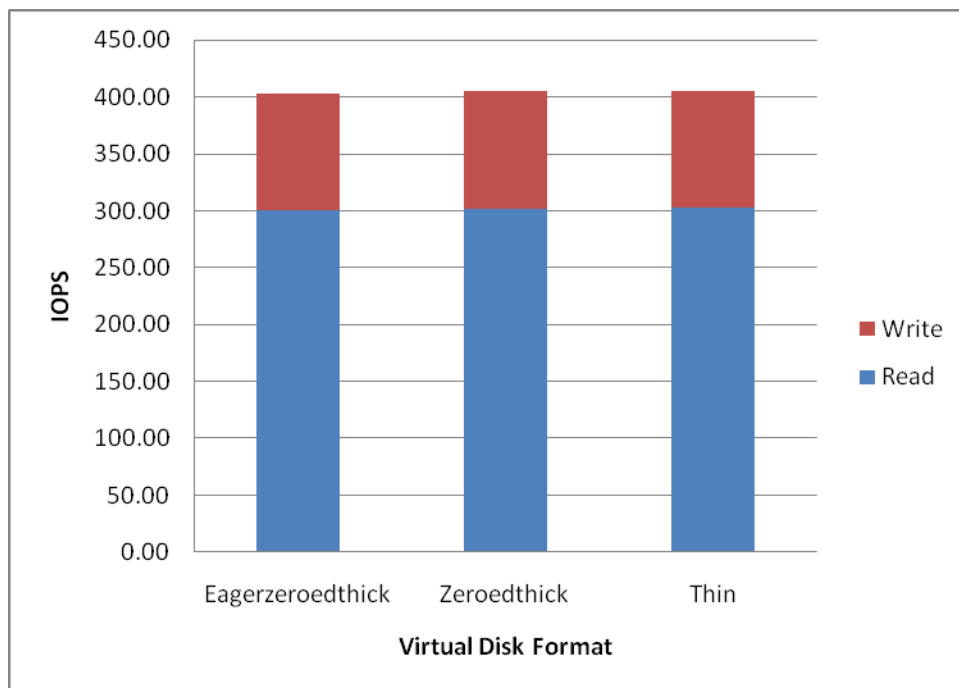
Figure 7 shows the response time results for this test.

**Figure 7. Pre Warm-up Response Time Test Results**



After a block is zeroed once, no additional zeroing is needed and all read/write IOPS for the zeroedthick format virtual disk and the eagerzeroedthick format virtual disk match, as shown in Figure 8.

**Figure 8. Post Warm-up IOPS Test Results**



These tests show that zeroedthick, thin and eagerzeroedthick format virtual disks produced the same number of IOPS after the warm-up completed. Keep in mind that after this warm-up completes, the virtual disk is no longer Dynamic Provisioning or vSphere thin friendly.

### **Conclusion**

Because of the warm-up overhead associated with thin format virtual disks, Hitachi Data Systems recommends using Hitachi Dynamic Provisioning software with 2000 family storage systems instead of vSphere 4 thin provisioning when thin provisioning is required.

## **Production Workload Simulation on Standard LUs Tests**

These tests simulated production workloads in a lab environment to determine how multiple workloads on a VMFS datastore perform. Hitachi Data Systems paid close attention to sequential versus random, block size, number of IOPS as seen by the virtual machine using the VDbench output, ESX host using ESXtop output and the Hitachi Adaptable Modular Storage 2100 using the `auperform` command.

These tests were used to examine the following:

- Best practices when creating VMFS datastores on standard RAID group LUs
- Best practices when creating VMFS datastores on Dynamic Provisioning LUs
- Best practice for balancing the I/O load across multiple VMs with low administrative overhead

In these lab tests, moving workloads between VMFS Datastores was done by creating new eagerzeroedthick virtual disks. In a production environment workloads can be migrated using Storage vMotion.

Table 5 lists the workloads tested.

**Table 5. I/O Workload**

<i>Test</i>	<i>Workload</i>	<i>Description</i>	<i>Performance Requirement</i>
OLTP	100% random 75% read 25% write 8K block	Similar to the I/O workload on an SQL, Oracle or similar database application. Workload and performance requirement were based on a Microsoft-recommended Iometer profile.	400 IOPS
Email	100% random 50% read 50% write 8K block	Similar to the I/O workload of an Exchange or other email application Workload and performance requirement were based on Microsoft's very heavy Exchange user profile 250MB mailbox on a 100GB database	230 IOPS
Streaming media	100% sequential 100% read 256K block	Similar to the workload of a streaming media application such as IPTV or Internet video streaming Workload and performance requirement were based on industry accepted requirements for streaming IPTV	100MB/s (400 IOPS)
Unstructured data serving	100% random 90% read 10% write 8k block	Similar to file serving or some SharePoint applications where unstructured data is shared Often data is written by one user and read by many Workload and performance requirement were based on a Microsoft-recommended Iometer profile	400 IOPS
Backup to disk	100% sequential 100% write 256K block	N/A	Maximum throughput

### *One Standard LU Test Configuration*

The first series of tests established a baseline and determined the degree of contention created by running multiple workloads on a single RAID-5 (4D+1P) group.

The OLTP workload was the priority workload and therefore was run at maximum I/O throughput. For the other workloads, IOPS were limited by VDbench parameters.

Figure 10 shows the test configuration with one standard LU for four workloads.

**Figure 10. One Standard LU for Four Workloads**

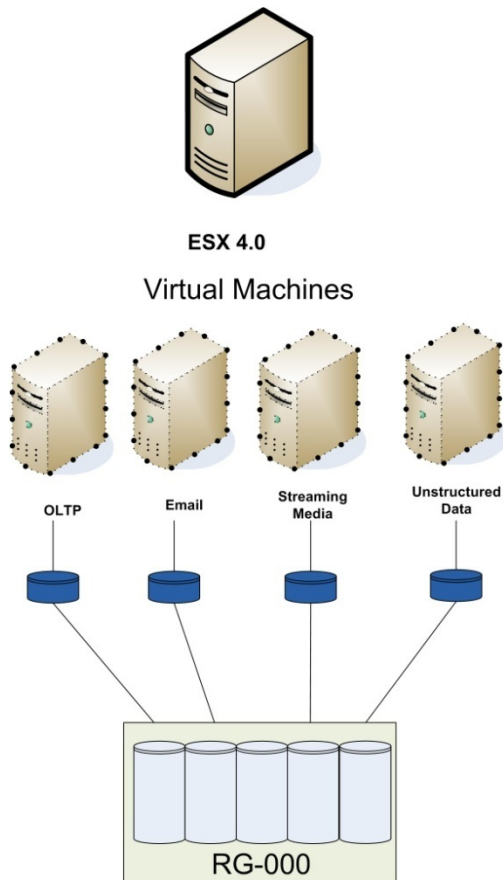
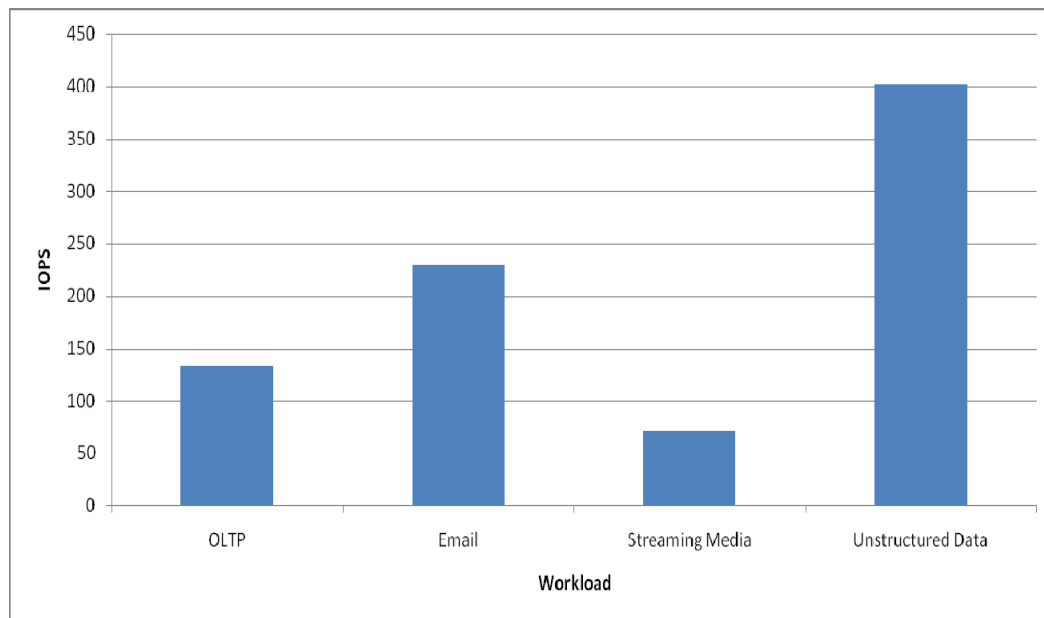


Figure 11 shows the IOPS generated by the OLTP test workload when the various workloads were added.

**Figure 11. IOPS for One Standard LU for all Workloads**



When these graphs are compared to the requirements for the applications in Table 5, it's clear that only two of the four workloads achieved the required IOPS. The disk percent busy rate for all five disks in the 2100 storage system's RAID group reached 98 percent. This means that more disks are needed to provide the IOPS needed for the OLTP and streaming media workloads.

#### *Two Standard LUs Test Configuration 1*

To relieve the IOPS constraints, an additional RAID-5 group, RG-001, was configured, and the email and unstructured data workloads were moved to RG-001. This reduced the contention for the disk resources and allowed the OLTP and streaming media access to RG-000.

Figure 12 shows the reconfigured test environment.

**Figure 12. Two Standard LUs Test Configuration 1**

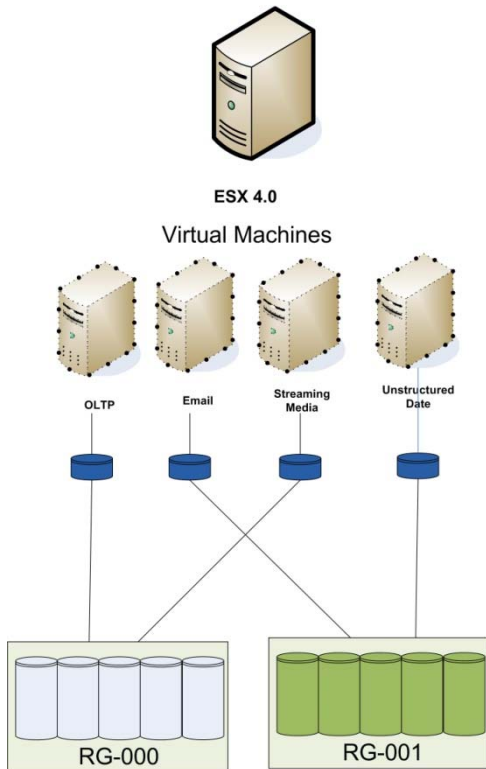


Figure 13 shows the results of this change. With this reconfiguration, three of the four workloads achieve the IOPS required. Note that streaming media has a sequential workload, all others have random workloads.

**Figure 13. IOPS for Two Standard LUs Test Configuration 1**

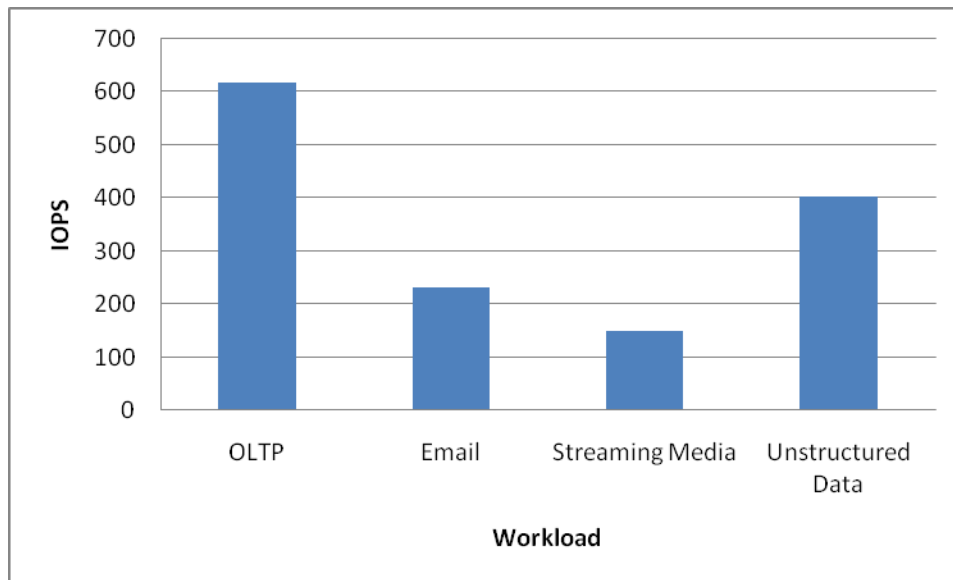
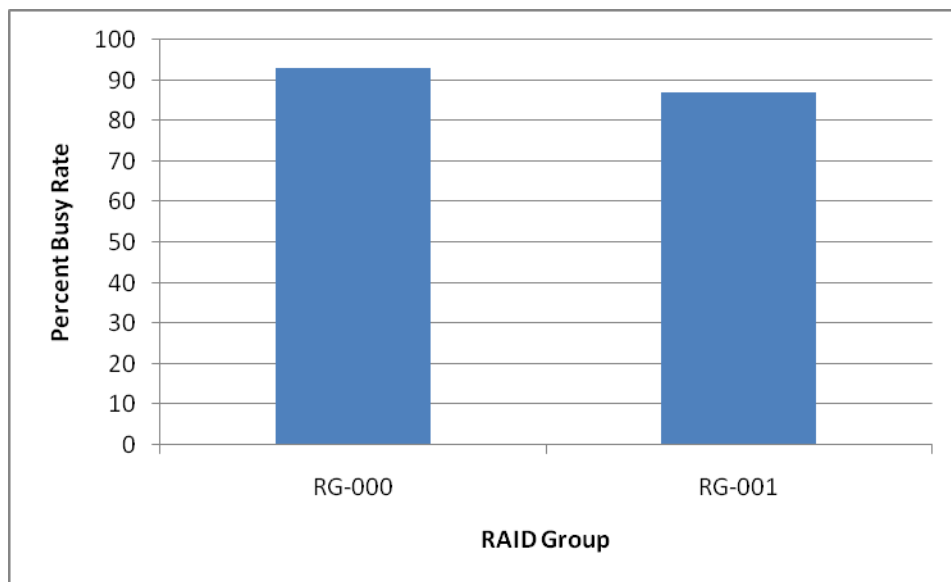


Figure 14 shows that the RAID groups are well utilized.

**Figure 14. RAID Group Utilization for Two Standard LUs Test Configuration 1**



Because the streaming media was not performing as required, the workload locations were adjusted again.

#### *Two Standard LUs Test Configuration 2*

To improve performance, the OLTP, email and unstructured data workloads were moved to RG-000, leaving the sequential streaming media to its own dedicated RAID group, RG-001.

Figure 15 shows the reconfigured test environment.

**Figure 15. Two Standard LUs Test Configuration 2**

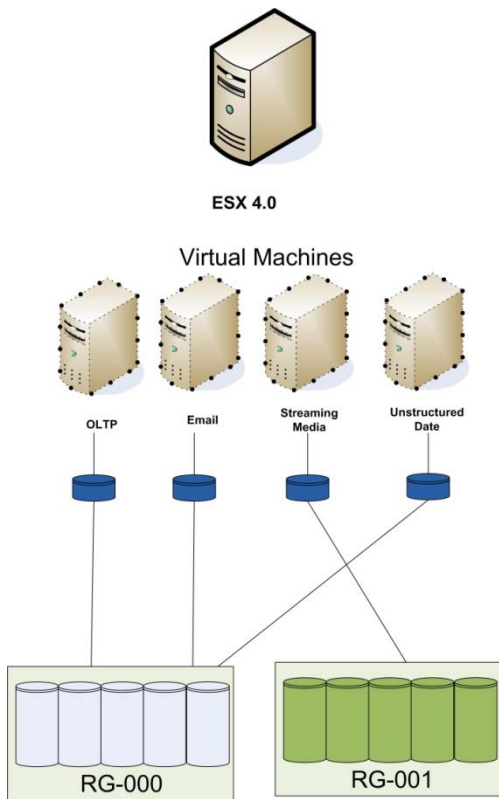


Figure 16 shows the IOPS results after the workload redistribution.

**Figure 16. IOPS for Two Standard LUs Test Configuration 2 IOPS**

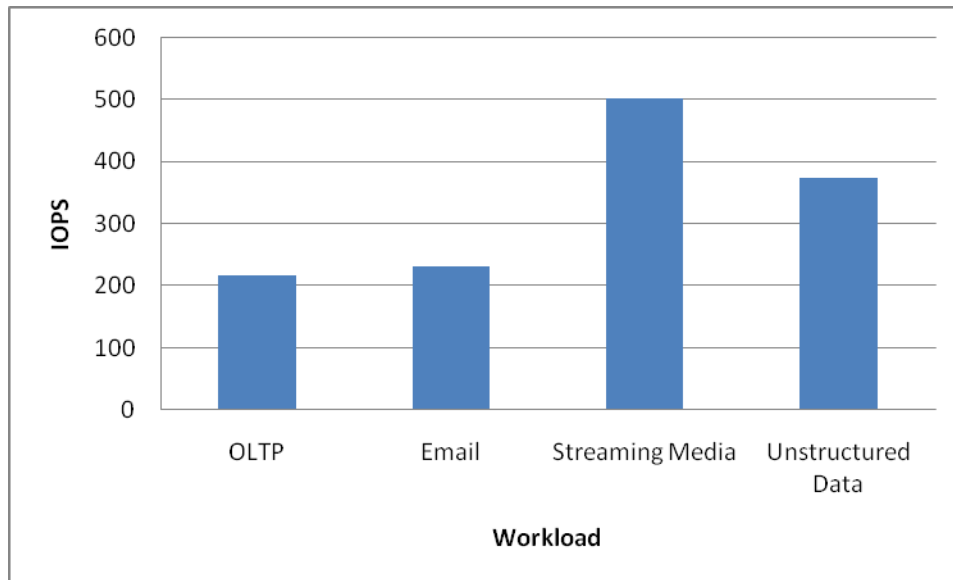
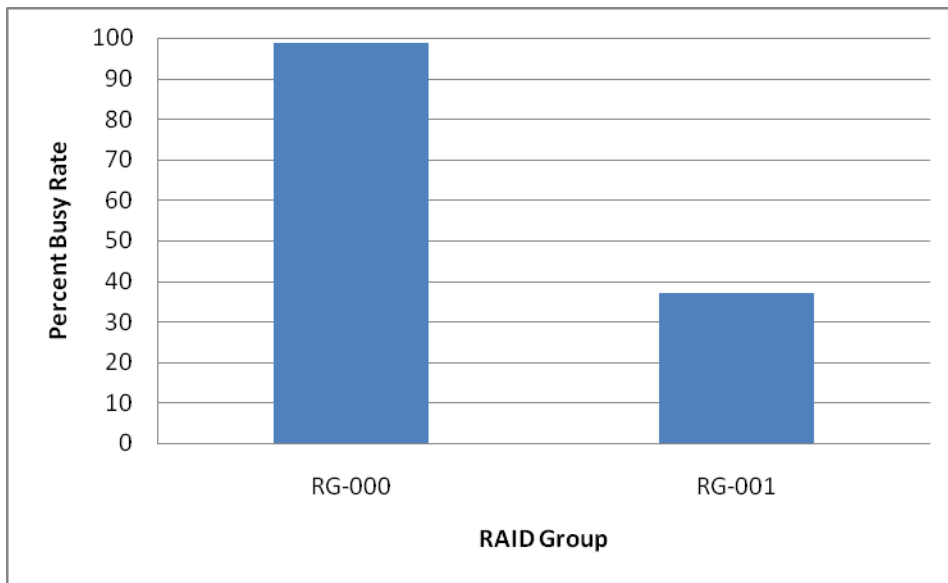




Figure 17 shows that the RAID groups were not evenly utilized.

**Figure 17. RAID Group Utilization for Two Standard LUs Test Configuration 1**



This configuration allowed the streaming media to achieve the performance requirement; however RAID Group 0 shows contention for disk resources on the OLTP, email and unstructured data workloads.

### *Three Standard LUs Test Configuration*

Because the OLTP and unstructured data were not performing as required, the workload locations were adjusted again. OLTP was given its own dedicated RG-000, email and unstructured data on RG-001. Streaming media was given a dedicated RAID group, RG-002.

Figure 18 shows the addition of a third RAID group and redistribution of the workload across 15 disks.

**Figure 18. Three Standard LUs Test Configuration**

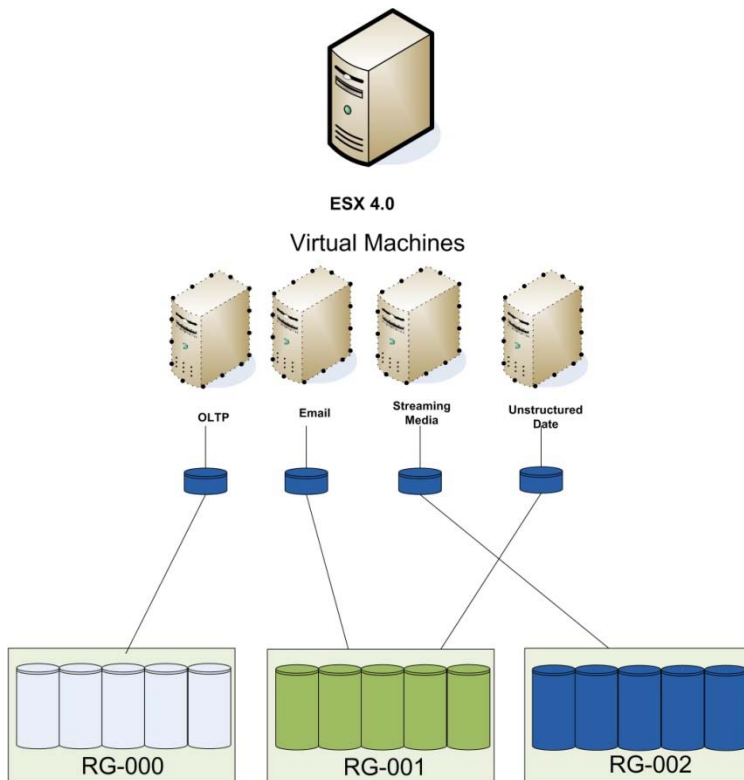


Figure 19 shows the results of reconfigured workload on the IOPs.

**Figure 19. IOPS for Three Standard LU Test**

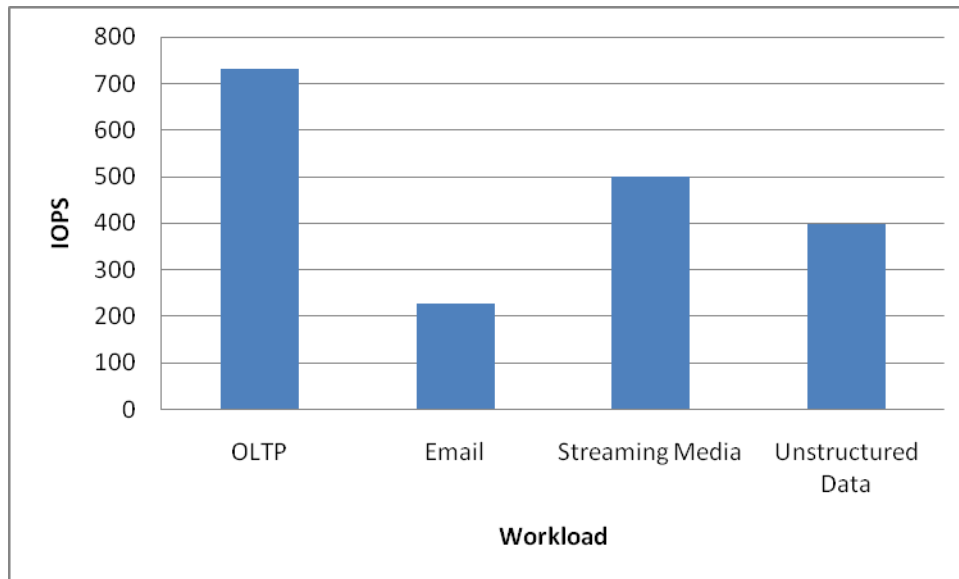
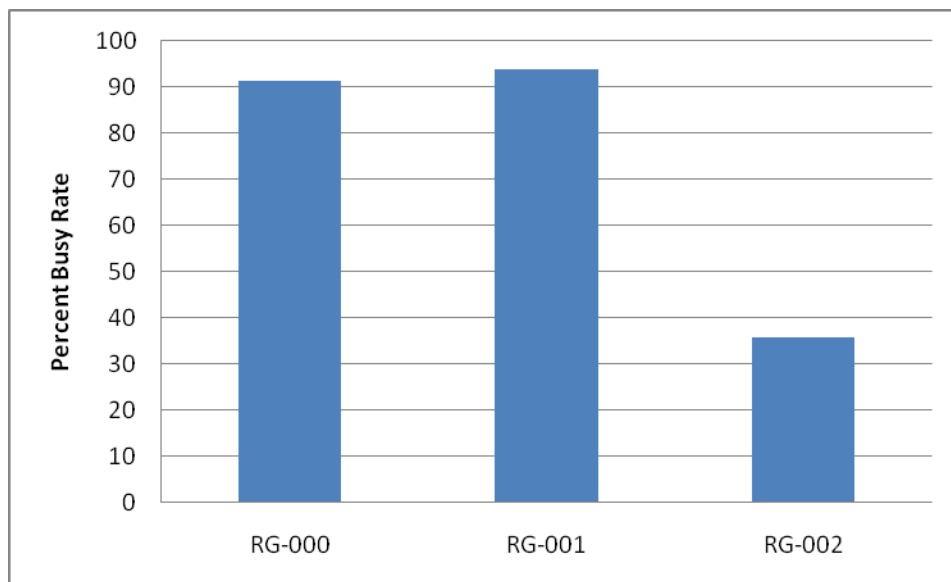


Figure 20 shows that the RAID groups were not evenly utilized.

**Figure 20. RAID Group Utilization for Three Standard LUs Test Configuration**



This configuration allowed all workloads to achieve the required IOPS performance.

#### *Standard LU Backup to Disk Test*

Using the same configuration as the three standard LUs tests, Hitachi Data Systems ran an additional test using a 100 percent sequential, 100 percent write, 256k block workload to simulate backup up to disk. Figure 21 and Figure 22 show that this workload has a large effect on the production workloads in both response time and IOPS.

**Figure 21. IOPS for Standard LU Production and Backup Workload Test**

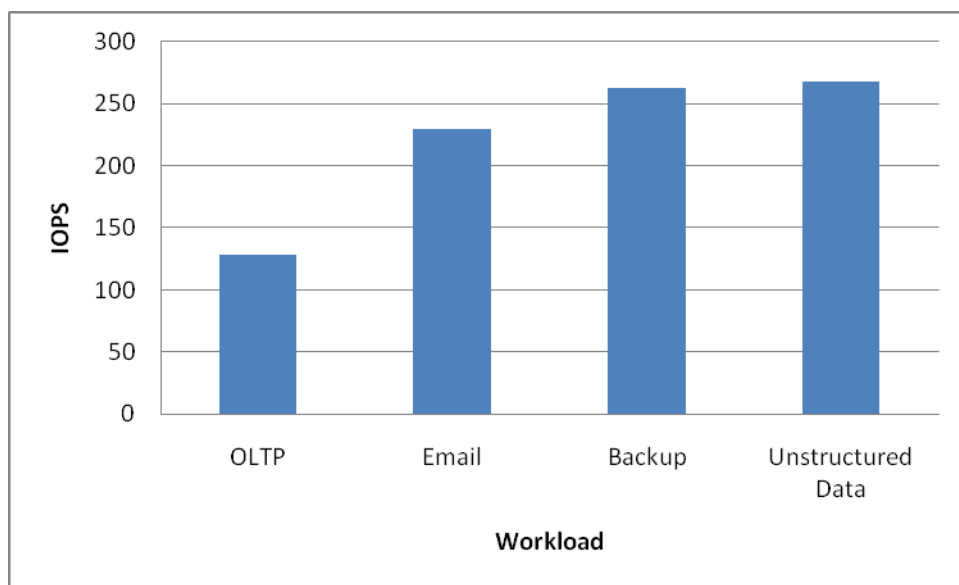
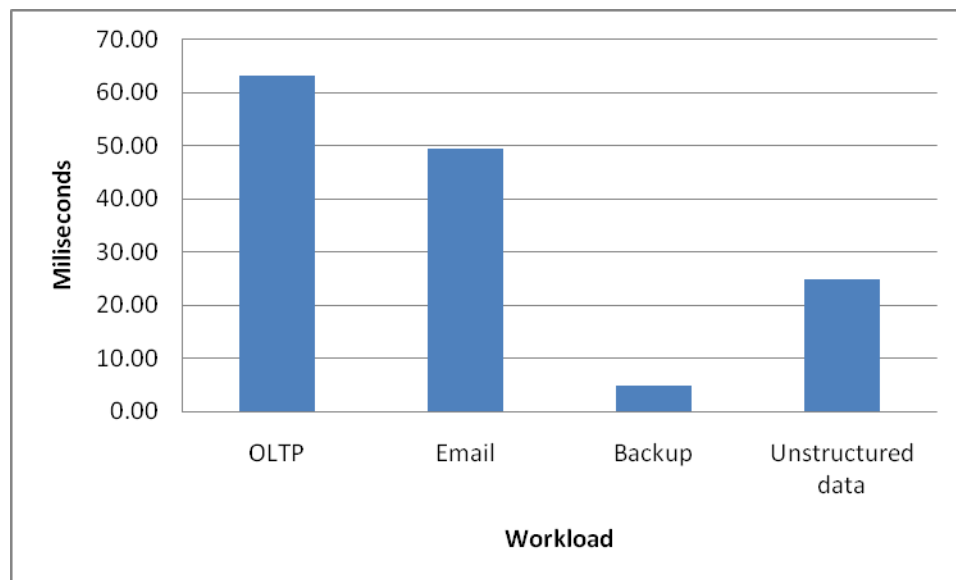


Figure 22 shows that response times were very large. Typically, response times below 10ms are required.

**Figure 22. Response Time for Standard LU Production and Backup Workload Test**



### *Standard LU Scaling Tests*

Scaling standard RAID groups can require extensive workload evaluation. Contention for disk resources can be difficult to predict and mixing workloads on RAID groups can require re-evaluation after they are put into production.

The goal of these tests was to demonstrate the administrative effort required to manually balance workloads across standard RAID LUs. Table 21 lists this test workload distribution across four RAID groups.

**Table 21. RAID Group Workload Distribution**

<i><b>RAID Group</b></i>	<i><b>Workloads</b></i>	<i><b>Notes</b></i>
RG-000	VM1 OLTP VM5 OLTP	Top priority workloads
RG-001	VM2 Email VM4 Unstructured data	Random workloads Total IOPS requirement of 630
RG-002	VM3 Streaming Media VM7 Streaming Media	Sequential workloads
RG-003	VM5 Email VM8 Unstructured data	Random workloads Total IOPS requirement of 630

Figure 23 shows the test configuration.

**Figure 23. Standard LU Scaling Test Configuration**

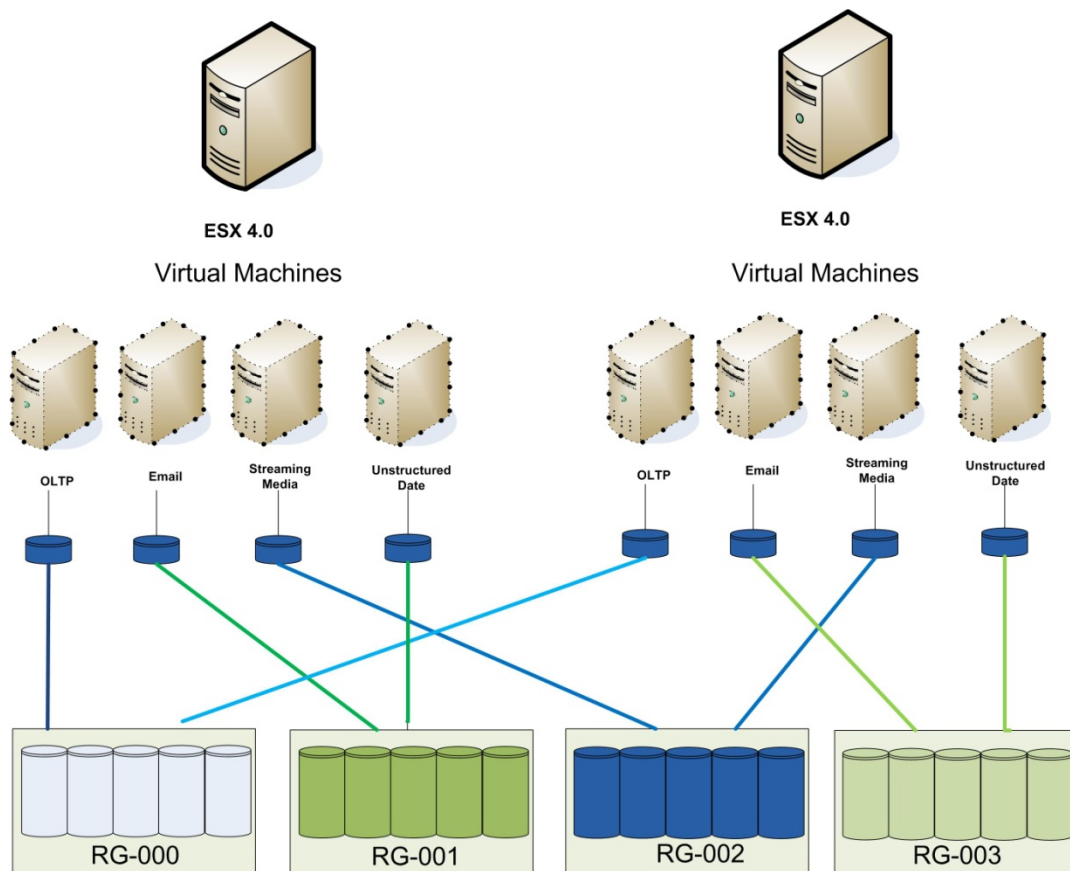


Figure 24 shows the IOPS scaling test results. These workloads are competing for disk resources that are nearly 100 percent utilized.

**Figure 24. IOPS for Standard LU Scaling Test**

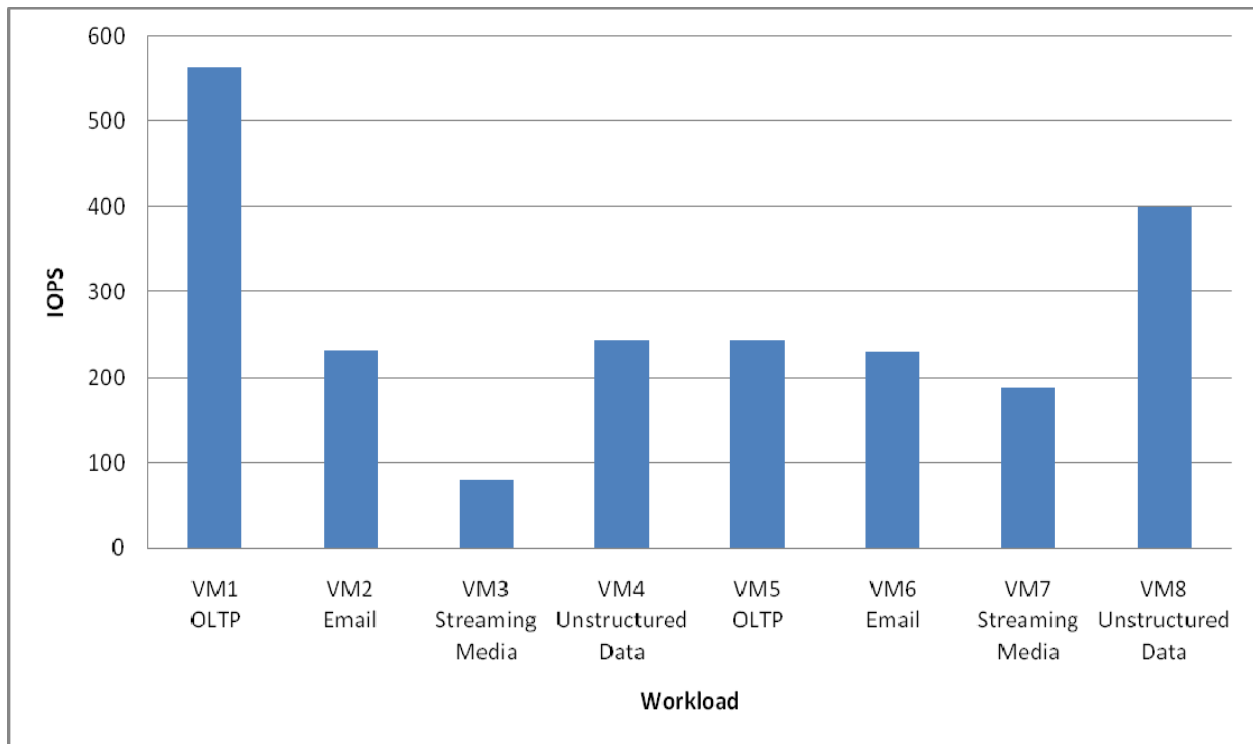
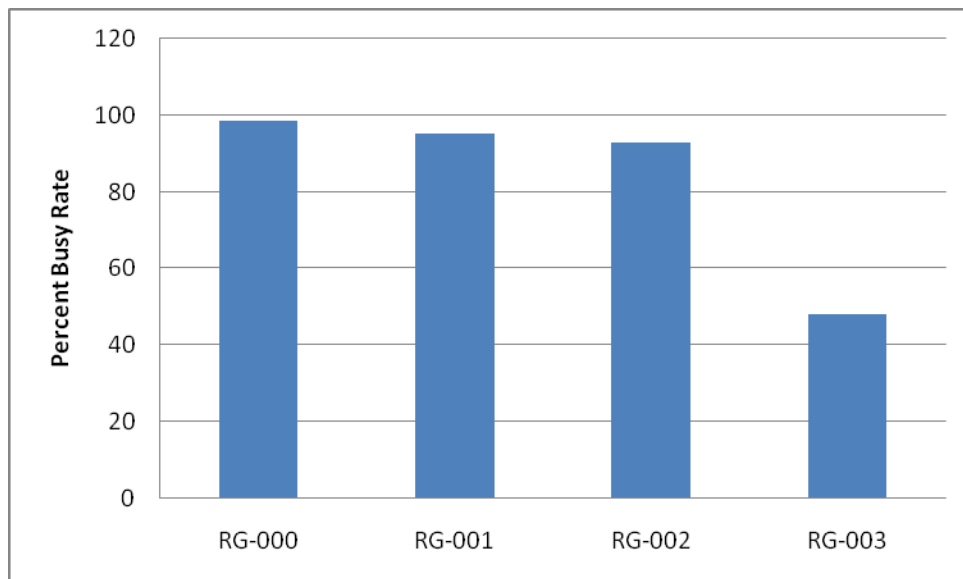


Figure 25 shows the utilization of the RAID groups.

**Figure 25. RAID Group Utilization for Standard LU Scaling Test**



### *Standard LU Considerations*

When using standard LUs, you must carefully consider workloads and be prepared to make adjustments based on these factors:

- **Sequential vs. random workload** — Sequential workloads, when mixed with other workloads (either sequential or random) on standard LUs, can become random.
- **Total IOPS available from the RAID group** — Additional physical disk reads and writes occur for every application write due to the use of mirrors or XOR parity.
- **Application requirements** — Total IOPS and response time are critical factors when designing RAID group architecture.

### *Best Practices for Virtual Disks on Standard LUs*

Zeroedthick and eagerzeroedthick format virtual disks provide similar performance after the warm-up period required by the zeroedthick virtual disk on standard LUs. Either virtual disk format provides similar throughput after all blocks are written at least one time, however zeroedthick initially shows some write latency and lower write throughput.

When deciding whether to use zeroedthick or eagerzeroedthick format virtual disks, keep the following considerations in mind:

- If you plan to use vSphere 4 Fault Tolerance on a virtual machine, you must use the eagerzeroedthick virtual disk format.
- If minimizing the time to create the virtual disk is more important than maximizing initial write performance, use the zeroedthick virtual disk format.
- If maximizing initial write performance is more important than minimizing the time required to create the virtual disk, use the eagerzeroedthick format.

### *Best Practices for Sequential and Random Workloads*

Hitachi Adaptable Modular Storage 2000 family storage systems detect sequential read workloads and speed access through aggressive cache prefetching. To take full advantage of this algorithm, follow these best practices:

- Do not mix sequential and random workloads on the same RAID group. In practice, mixing sequential with random workloads creates random workloads.
- Do not mix multiple sequential workloads on the same RAID group. In practice, adding a sequential workload to other sequential workloads creates random workloads.

## Production Workload Simulation on Dynamically Provisioned LUs Tests

Hitachi Dynamic Provisioning software can assist with balancing application workloads across many disks with Dynamic Provisioning pools and Dynamic Provisioning LUs. Even in those cases where it cannot increase the total throughput possible on disks, Hitachi Dynamic Provisioning software can reduce the amount of administrative overhead required to distribute the workload and optimize capacity.

These tests were used to demonstrate that the Dynamic Provisioning wide-striping feature dramatically reduces the administrative effort required to balance multiple workloads.

### *Dynamically Provisioned Test Configuration 1*

This test used a Dynamic Provisioning pool containing two RAID-5 (4D+1P) groups. A single LU was created and a VMFS datastore was built on the LU. Each VM contained a single virtual disk of 100GB.

Figure 26 shows the first one Dynamic Provisioning LU test configuration.

**Figure 26. One Dynamic Provisioning LU Test Configuration**

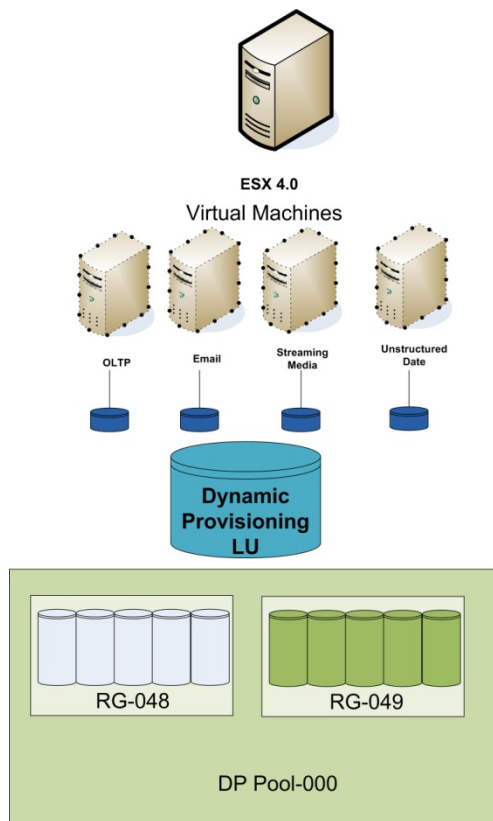




Figure 27 shows the IOPS results for the first Dynamic Provisioning LU test.

**Figure 27. IOPs for One Dynamic Provisioning LU Test Configuration**

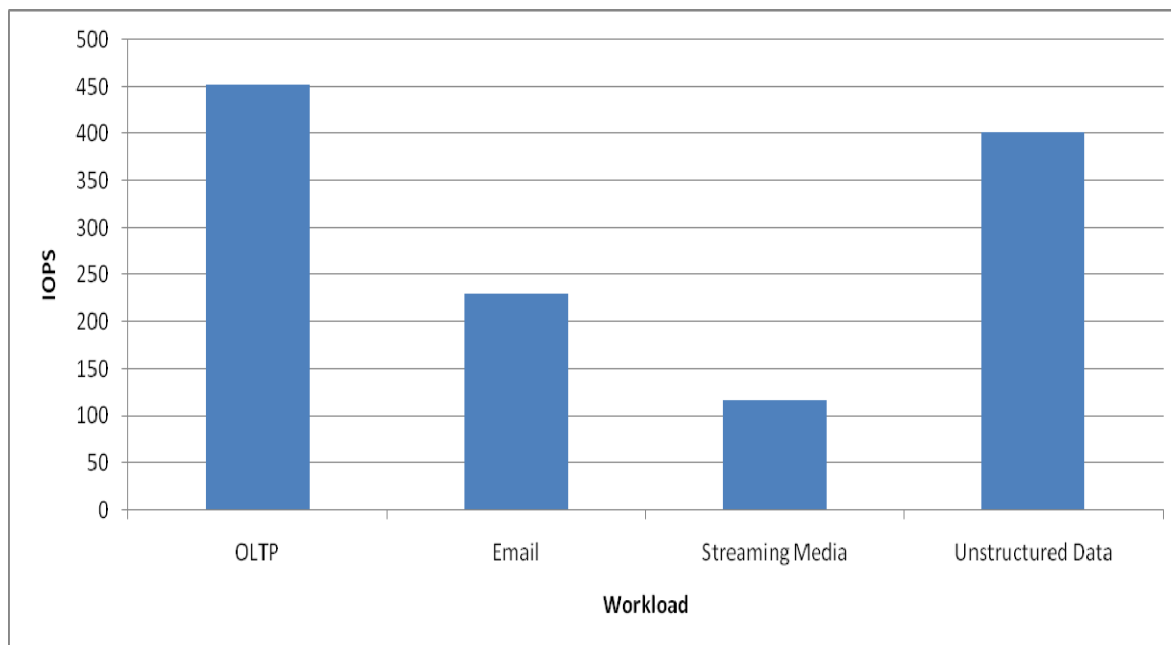
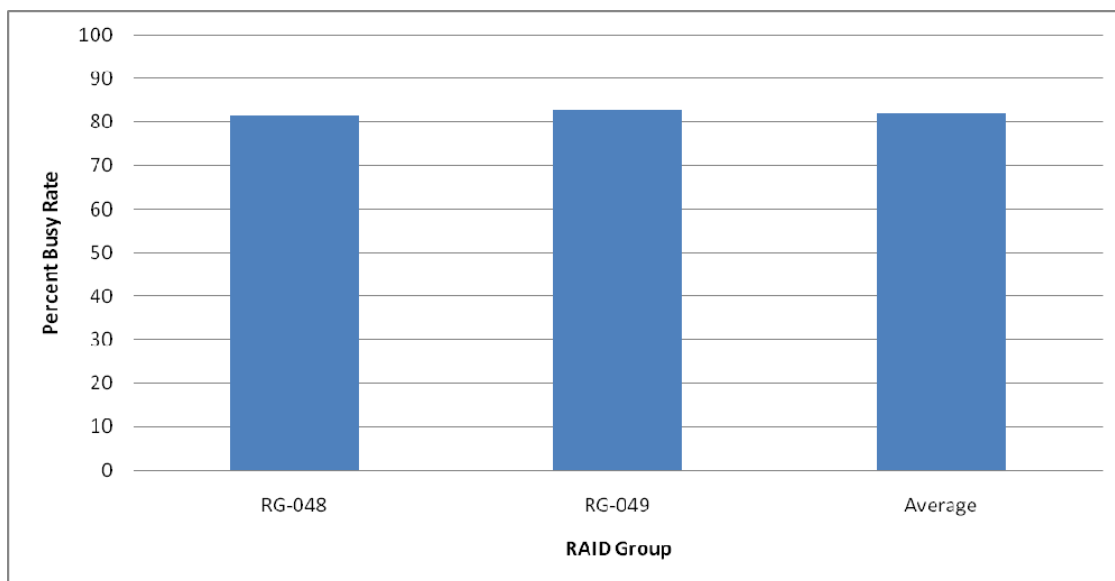


Figure 28 shows that Hitachi Dynamic Provisioning software evenly balanced the I/O workload across the two RAID groups.

**Figure 28. RAID Group Utilization for Dynamic Provisioning LUs Test Configuration**



Notice that all of the workloads except the streaming media received the IOPS required. The RAID groups in the Dynamic Provisioning pool show very similar percent busy rates, which indicates that the Dynamic Provisioning pool's wide striping feature balanced the I/O.

## Dynamic Provisioning Test Configuration 2

The goal of this test was to determine if separating the sequential workload from the random workload on different Dynamic Provisioning pools allows the 2100 to detect sequential I/O or if the virtual host bus adapter breaks up the sequential workload, making it appear as random to the 2100.

This test used two Dynamic Provisioning pools, DP Pool -000 and DP Pool -001, each containing a single LU. DP Pool -000 contained two RAID groups and was used for all random workloads. DP-Pool -001 contained a single RAID group and was used for the sequential reads on the streaming media workload.

Figure 29 shows the second test configuration using one Dynamic Provisioning LU.

**Figure 29. Two Dynamic Provisioning LU Test Configuration**

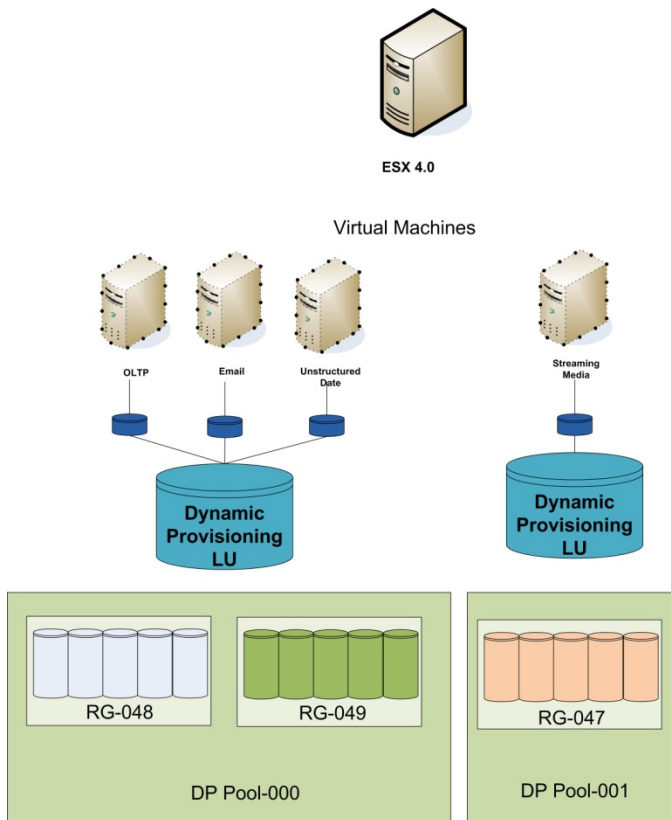


Figure 30 shows the IOPS results for this test.

**Figure 30. IOPS for One Dynamic Provisioning LU Test Configuration 2**

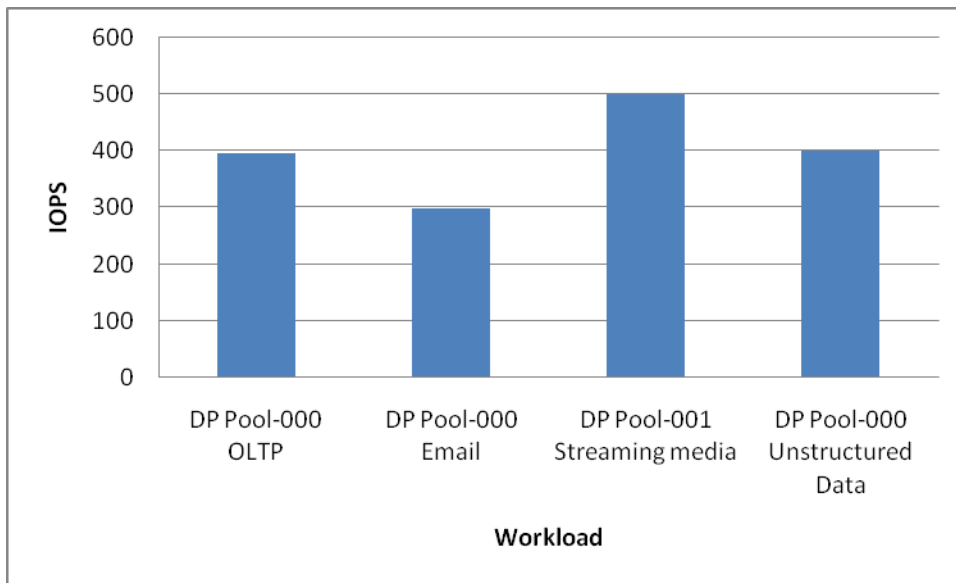
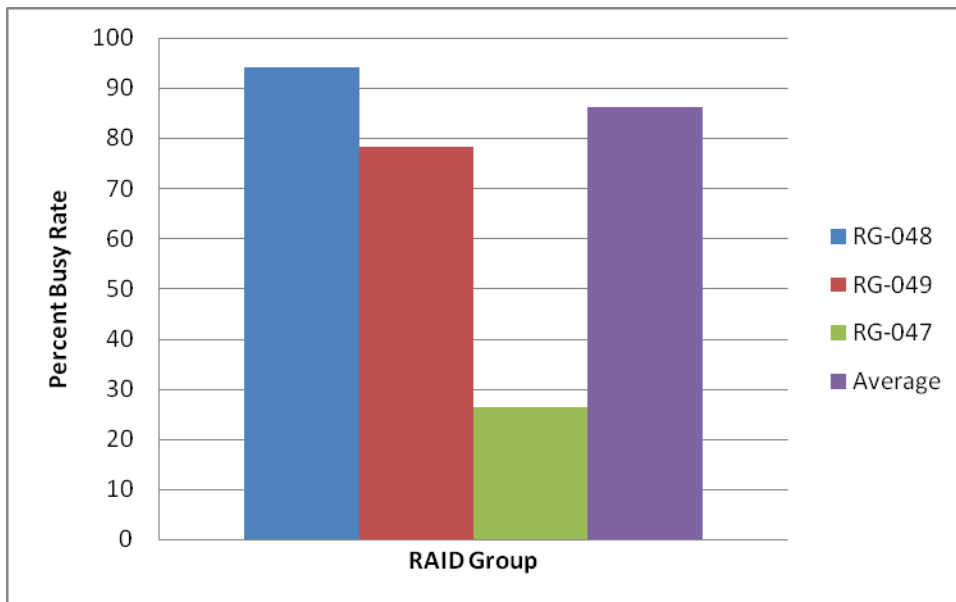


Figure 31 shows that Dynamic Provisioning software evenly balanced the I/O workload across the two RAID groups.

**Figure 31. RAID Group Utilization for Three Standard LUs Test Configuration**



After moving the sequential I/O workload from DP Pool -000, the random workloads achieved the IOPS goal. RAID group and disk load balance is good between the RAID groups in DP Pool -000.

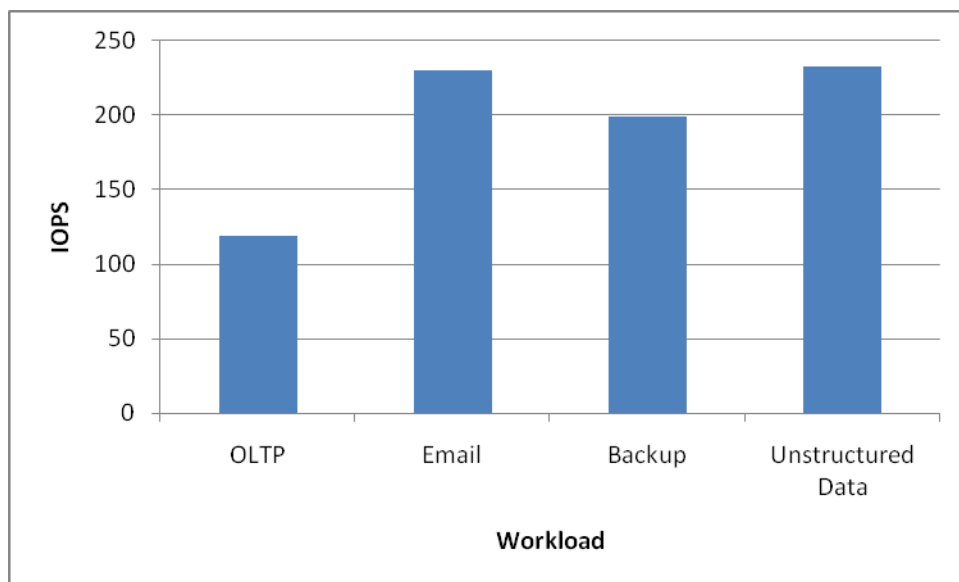
Hitachi Data Systems ran additional tests using a single Dynamic Provisioning pool with two LUs. The sequential IOPS suffered indicating that mixing sequential and random on small Dynamic Provisioning pools may degrade performance.

### *Dynamic Provisioning Backup to Disk Test*

This test used 100 percent sequential, 100 percent write workload with a 256K block I/O to simulate backing up to disk. The Dynamic Provisioning pool contained a single RAID-5 (4D+1P) group.

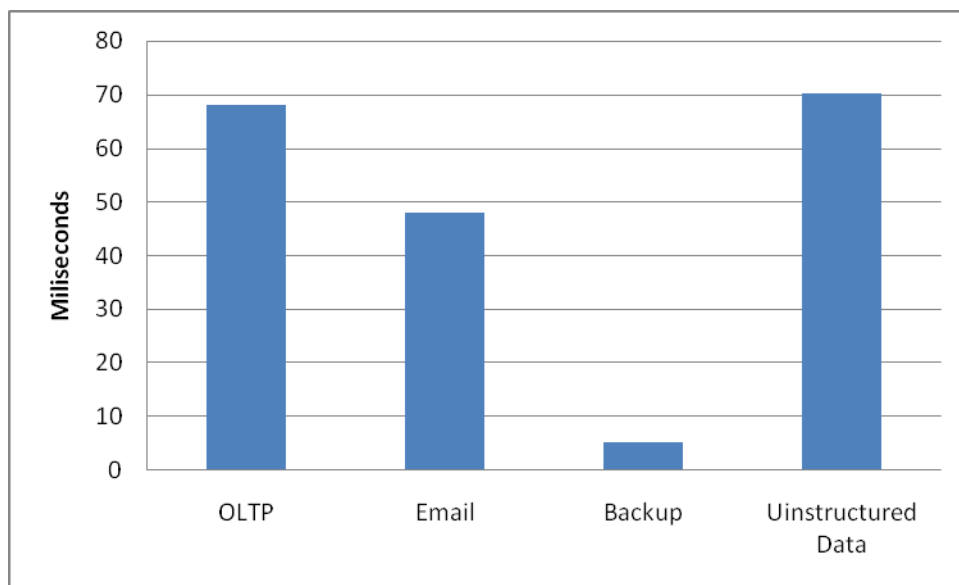
Figure 32 and Figure 33 show that this workload has an adverse effect on the production workloads in both response time and IOPS.

**Figure 32. IOPS for Dynamically Provisioned LU Production and Backup Workload Test**



As shown in Figure 33, response times for this test were very poor. For that reason, Hitachi Data Systems does not recommend mixing a sequential write workload on the same 10 disks as the production workloads.

**Figure 33. Response Time for Dynamically Provisioned LU Production and Backup Workload Test**



Testing demonstrated that streaming backup to disk has a large effect on throughput and response time for other applications. Hitachi Data Systems recommends using separate RAID groups for backup volumes to prevent contention for disk resources.

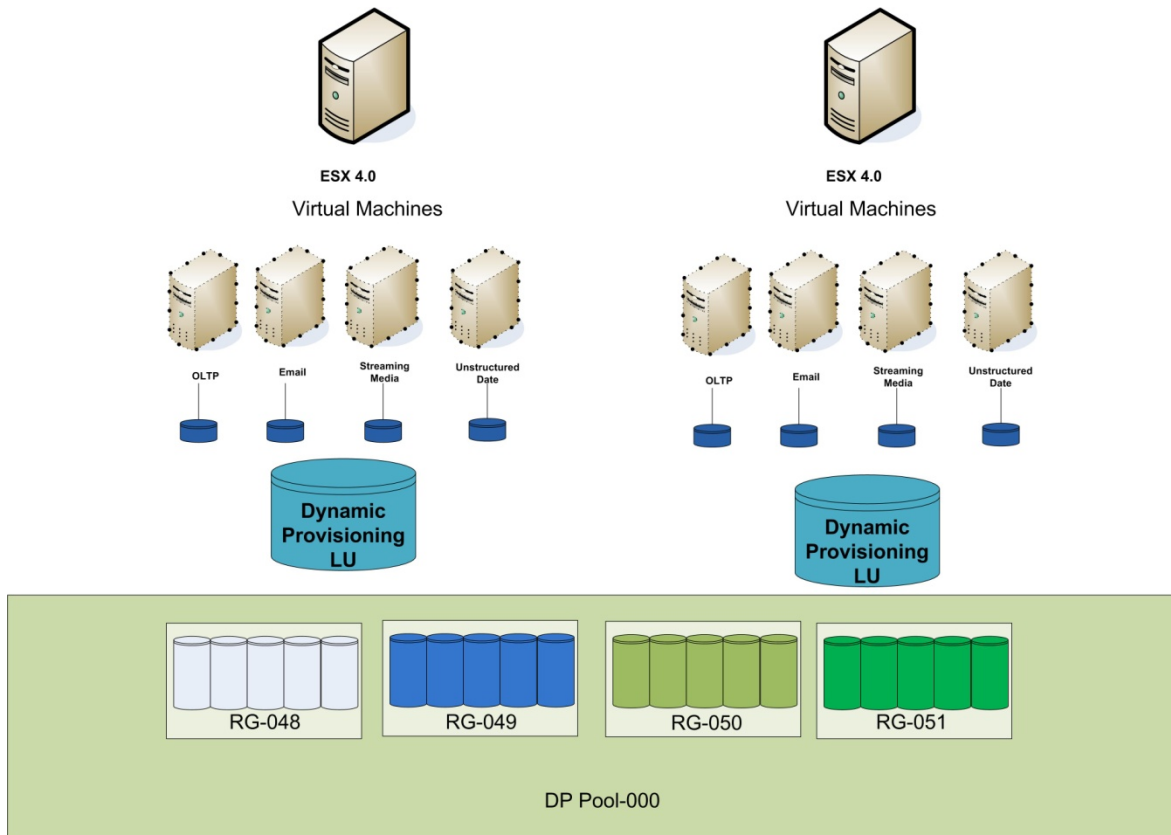
### Dynamic Provisioning Scaling Tests

Scaling standard RAID groups can require extensive workload evaluation, an administrative task that Hitachi Dynamic Provisioning software can reduce with its wide striping feature that balances workloads across many disks.

The goal of these tests was to determine if scaling to a larger Dynamic Provisioning pool assists with balancing workloads and reduces manual administrative effort.

For these tests, DP Pool -000 was extended to four RAID groups, as shown in Figure 34.

**Figure 34. Dynamic Provisioning Scaling Test Configuration**



Each of the eight VMs was assigned a 100GB virtual disk without regard to data placement. All data was wide-striped across the 20 disks in the Dynamic Provisioning pool.

Figure 35 shows the results of the IOPS test.

**Figure 35. IOPS for Scaling Dynamically Provisioned LUs Test**

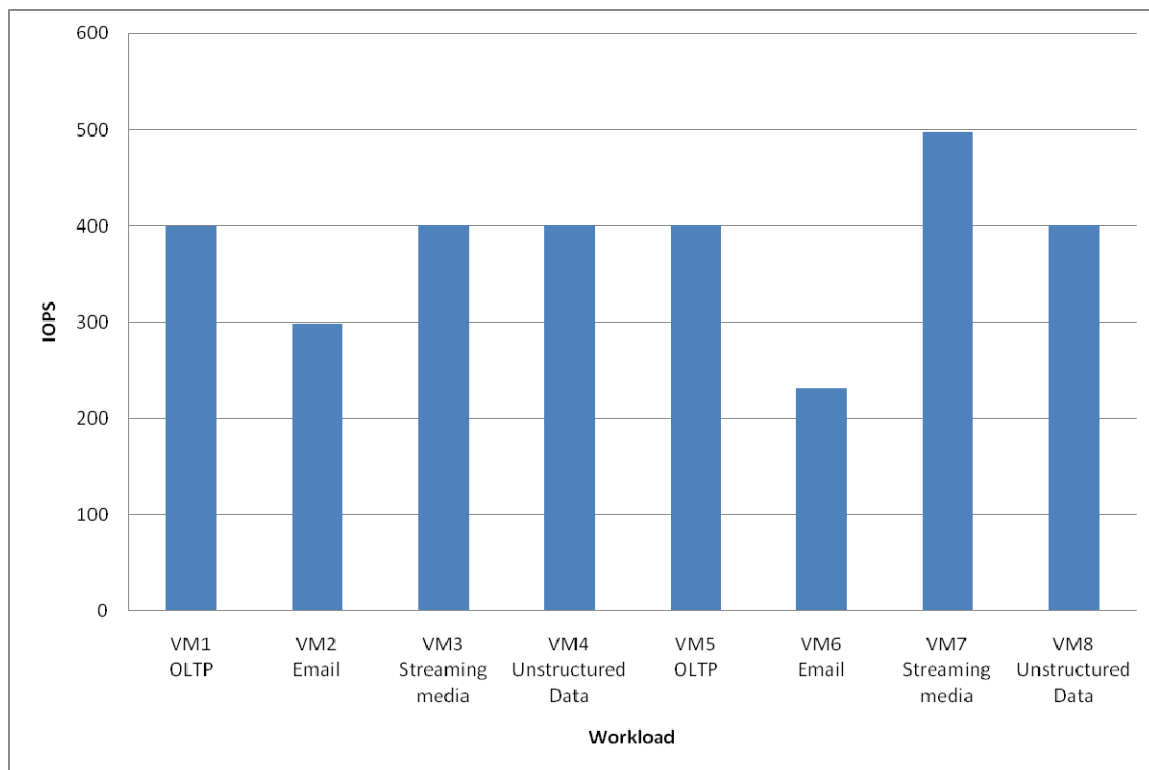
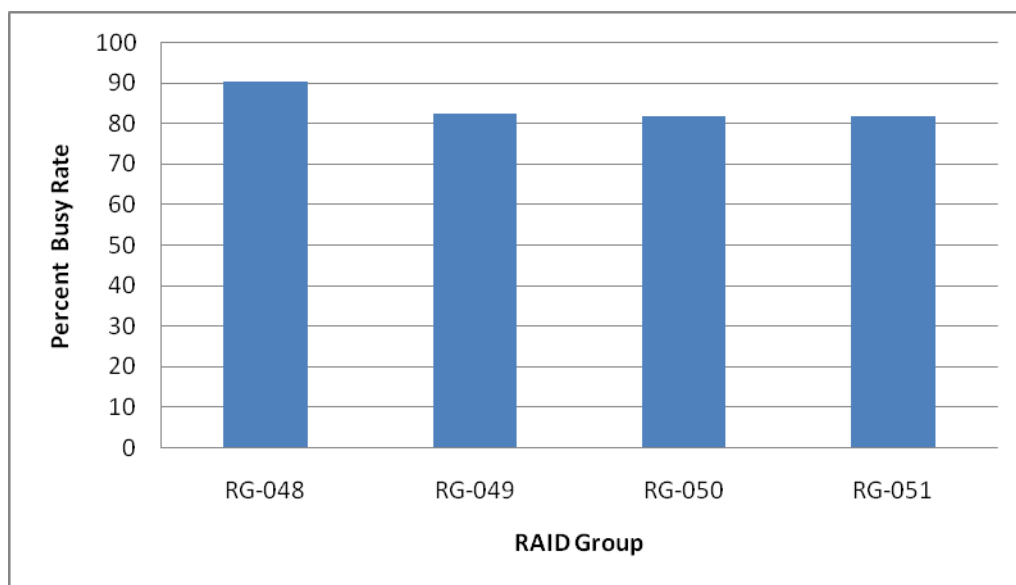


Figure 36 shows the results of the RAID group utilization test.

**Figure 36. RAID Group Utilization for Scaling Dynamically Provisioned LUs Test**



All workloads received required disk resources. No manual administrative effort was needed to balance the workload across the disks as was necessary in the standard LU tests. The sequential workloads mixed well with random workloads.

As demonstrated by the lab tests, you might realize a benefit from separating sequential from random workloads on smaller Dynamic Provisioning pools as seen in the 10-disk Dynamic Provisioning tests, but the 20-disk Dynamic Provisioning test demonstrates that Hitachi Dynamic Provisioning software can efficiently address both sequential and random workloads in a single large Dynamic Provisioning pool. When your workload is unknown, a single large Dynamic Provisioning pool can be used efficiently. When the workload is known and maximum performance is required, separating random and sequential workloads can benefit throughput.

### *Best Practices for Virtual Disk and Dynamic Provisioning Performance*

To obtain maximum storage performance for vSphere 4 when using the 2000 family storage, follow these best practices:

- Use eagerzeroedthick virtual disk format to prevent warm-up anomalies. Zeroedthick is fine for use on the guest OS boot volume where maximum write performance is not required.
- Use at least four RAID groups in the Dynamic Provisioning pool for maximum wide striping benefit.
- Size the Dynamic Provisioning pools according to the I/O requirements of the virtual disk and application.
- When larger Dynamic Provisioning pools are not possible, separate sequential and random workloads on different Dynamic Provisioning pools.
- For applications that use log recovery, separate the logs from the database on different Dynamic Provisioning pools. This separates the sequential and random workloads and can also help protect data. In the rare case where a dual hardware failure causes the corruption or loss of a Dynamic Provisioning pool, the logs are available for recovery.

## Conclusion

The Hitachi Adaptable Modular Storage 2000 family and Hitachi Dynamic Provisioning software bring significant advantages to vSphere environments. The testing described in this white paper supports the conclusions about virtual disk formats listed in Table 6.

**Table 6. Conclusions for vSphere 4 Virtual Disk Formats**

<i>Virtual Disk Format</i>	<i>Dynamic Provisioning Effect</i>	<i>Hitachi Data Systems Recommendation</i>
Thin	Dynamic Provisioning thin friendly	Not recommended due to overhead
Zeroedthick	Dynamic Provisioning thin friendly	Recommended
Eagerzeroedthick	Not storage system thin friendly	Recommended for the wide-striping benefit of Dynamic Provisioning software

The testing described in this white paper supports the conclusions about vSphere 4 operations listed in Table 7.

**Table 7. Conclusions for vSphere 4 Operations**

<i>Operation</i>	<i>Recommended Virtual Disk Format</i>	<i>Dynamic Provisioning Effect</i>
Storage of templates and .ISO images	Zeroedthick	Dynamic Provisioning thin friendly
Deploying from templates	Zeroedthick	Dynamic Provisioning thin friendly
Storage vMotion	Zeroedthick Eagerzeroedthick	Dynamic Provisioning thin friendly Not thin-friendly
Converting from thin to thick virtual disk	N/A	Not Dynamic Provisioning thin friendly
Expanding VMFS datastore	Zeroedthick Eagerzeroedthick	Dynamic Provisioning thin friendly Not thin-friendly
Performance leveling with Dynamic Provisioning software	Eagerzeroedthick	Not thin friendly Balances workload across all disks in the Dynamic Provisioning pool

To achieve Hitachi Dynamic Provisioning software's automated wide-striping of thin provisioning benefits, use vSphere and Dynamic Provisioning software as follows:

- Because of the warm-up and response time overhead associated with thin format virtual disks combined with the wide-striping of Dynamic Provisioning, Hitachi Data Systems recommends using Hitachi Dynamic Provisioning software with 2000 family storage systems instead of vSphere 4 thin provisioning when thin provisioning is required.
- Eagerzeroedthick format virtual disks eliminate warm-up anomalies and distribute the I/O across all disks in a Dynamic Provisioning pool, so Hitachi Data Systems recommends using eagerzeroedthick format virtual disks to achieve maximum performance.
- Dynamic Provisioning pools with at least four RAID groups distribute mixed workload with minimal administrative effort.
- Separating sequential and random workloads can improve performance when Dynamic Provisioning pools with fewer than four RAID groups.

For more information about the Hitachi Adaptable Modular Storage 2000 family or Hitachi Dynamic Provisioning software, contact your sales representative, channel partner or visit the Hitachi Data Systems [web site](#).





---

**Corporate Headquarters** 750 Central Expressway, Santa Clara, California 95050-2627 USA  
Contact Information: + 1 408 970 1000 [www.hds.com](http://www.hds.com) / [info@hds.com](mailto:info@hds.com)

**Asia Pacific and Americas** 750 Central Expressway, Santa Clara, California 95050-2627 USA  
Contact Information: + 1 408 970 1000 [www.hds.com](http://www.hds.com) / [info@hds.com](mailto:info@hds.com)

**Europe Headquarters** Sefton Park, Stoke Poges, Buckinghamshire SL2 4HD United Kingdom  
Contact Information: + 44 (0) 1753 618000 [www.hds.com](http://www.hds.com) / [info.uk@hds.com](mailto:info.uk@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 mentioned in this document 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. This document describes some capabilities that are conditioned on a maintenance contract with Hitachi Data Systems being in effect and that may be configuration dependent, and features that may not be currently available. Contact your local Hitachi Data Systems sales office for information on feature and product availability.

© Hitachi Data Systems Corporation 2010. All Rights Reserved.  
AS-046-00 July 2010