The Benefits of Server-Side Flash with VMware vSphere Flash Read Cache on Hitachi Compute Blade 500

Lab Validation Report

By Jeff Chen

Contributor from Hitachi, Ltd.: Tsuyoshi Inoue

July 16, 2014
Feedback

Hitachi Data Systems welcomes your feedback. Please share your thoughts by sending an email message to SolutionLab@hds.com. To assist the routing of this message, use the paper number in the subject and the title of this white paper in the text.
Table of Contents

Product Features........................................................................................................................................2
  Hitachi Virtual Storage Platform G1000..................................................................................................2
  Hitachi Compute Blade 500....................................................................................................................2
  Brocade Storage Area Network Switches................................................................................................2
  VMware vSphere 5................................................................................................................................3

Test Environment Configuration..............................................................................................................4
  Software Components..........................................................................................................................5
  Network Infrastructure..........................................................................................................................5
  SAN Infrastructure..................................................................................................................................5
  vSphere Flash Read Cache – Flash Devices Installation......................................................................7

Test Methodology......................................................................................................................................9
  Test Cases...........................................................................................................................................11

Analysis..................................................................................................................................................12
  Benefits of Using vFRC.......................................................................................................................15
  Considerations....................................................................................................................................16
The Benefits of Server-Side Flash with VMware vSphere Flash Read Cache on Hitachi Compute Blade 500

Lab Validation Report

VMware vSphere 5.5 introduces new functionality to leverage flash storage devices on a VMware ESXi host. The vSphere Flash Infrastructure layer is part of the ESXi storage stack for managing flash storage devices that are locally connected to the server. These devices can be of multiple types (primarily PCIe flash cards and SAS/SATA SSD drives) and the vSphere Flash Infrastructure layer is used to aggregate these flash devices into a unified flash resource.

The flash resource created by the vSphere Flash Infrastructure layer can be used for two purposes:

1. Read caching of virtual machine I/O requests (vSphere Flash Read Cache)
2. Storing the host swap file.

The goal of vSphere Flash Read Cache (vFRC) feature is to enhance performance of certain I/O workloads that exhibit characteristics suitable for caching. Please note that the vFRC feature requires a VMware Enterprise Plus license.

This lab validation report shows the benefit of vFRC and compares the performance between SAS attached SSDs and PCIe flash cards. It also provides guidance for the following questions:

- What type of workload will benefit from vFRC?
- PCIe flash devices cost much more than SAS attached SSDs. Which flash device to use?
- How to cut cost by vFRC?

This paper is intended for storage or data center administrators implementing storage within a VMware vSphere environment.

Note — Testing of this configuration was in a lab environment. Many things affect production environments beyond prediction or duplication in a lab environment. Follow the recommended practice of conducting proof-of-concept testing for acceptable results in a non-production, isolated test environment that otherwise matches your production environment before your production implementation of this solution.
Product Features

The following are the hardware and the software products tested for this report.

Hitachi Virtual Storage Platform G1000

Hitachi Virtual Storage Platform G1000 provides an always-available, agile, and automated foundation that you need for a continuous infrastructure cloud. This delivers enterprise-ready software-defined storage, advanced global storage virtualization, and powerful storage.

Supporting always-on operations, Virtual Storage Platform G1000 includes self-service, non-disruptive migration and active-active storage clustering for zero recovery time objectives. Automate your operations with self-optimizing, policy-driven management.

Virtual Storage Platform G1000 supports Oracle RAC and VMware Metro Storage Cluster.

Hitachi Compute Blade 500

Hitachi Compute Blade 500 combines the high-end features with the high compute density and adaptable architecture you need to lower costs and protect investment. Safely mix a wide variety of application workloads on a highly reliable, scalable, and flexible platform. Add server management and system monitoring at no cost with Hitachi Compute Systems Manager, which can seamlessly integrate with Hitachi Command Suite in IT environments using Hitachi storage.

The Hitachi Compute Blade 500 chassis contains internal Fibre Channel and network switches for the high availability requirements of Hitachi Unified Compute Platform for VMware vSphere.

Brocade Storage Area Network Switches

Brocade and Hitachi Data Systems have collaborated to deliver storage networking and data center solutions. These solutions reduce complexity and cost, as well as enable virtualization and cloud computing to increase business agility.

This lab validation report uses the following Brocade products:

- Brocade 6510 Switch
- Brocade VDX 6740 Data Center Switch
VMware vSphere 5

VMware vSphere 5 is a virtualization platform that provides a datacenter infrastructure. It features vSphere Distributed Resource Scheduler (DRS), High Availability, and Fault Tolerance.

VMware vSphere 5 has the following components:

- **ESXi 5** — This is a hypervisor that loads directly on a physical server. It partitions one physical machine into many virtual machines that share hardware resources.

- **vCenter Server** — This allows management of the vSphere environment through a single user interface. With vCenter, there are features available such as vMotion, Storage vMotion, Storage Distributed Resource Scheduler, High Availability, and Fault Tolerance.
Test Environment Configuration

The testing of vFRC took place in the Hitachi Data Systems laboratory using Hitachi Compute Blade 500 and Hitachi Virtual Storage Platform G1000.

Table 1 describes the details of the components used.

<table>
<thead>
<tr>
<th>Hardware</th>
<th>Description</th>
<th>Version</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hitachi Virtual Storage Platform G1000</td>
<td>4 × 8 Gb/sec Fibre Channel ports used&lt;br&gt;8 GB cache memory allocated by CLPR&lt;br&gt;8 × 900 GB 10k RPM SAS disks, 2.5 inch SFF used</td>
<td>80-01-22-00/02</td>
<td>1</td>
</tr>
<tr>
<td>Hitachi Compute Blade 500 chassis</td>
<td>8-blade chassis&lt;br&gt;2 Brocade 5460 Fibre Channel switch modules, each with 6 × 8 Gb/sec uplink ports&lt;br&gt;2 Brocade VDX 6746 Ethernet switch modules, each with 8 × 10 Gb/sec uplink ports&lt;br&gt;2 management modules&lt;br&gt;6 cooling fan modules&lt;br&gt;4 power supply modules</td>
<td>SVP: A0170-D-8920&lt;br&gt;5460: FOS 7.0.2c&lt;br&gt;VDX6746: NOS 4.1.2</td>
<td>1</td>
</tr>
<tr>
<td>520BH2 server blade</td>
<td>Half blade&lt;br&gt;2 × 12-core Intel Xeon E5-2697 processor, 2.70 GHz&lt;br&gt;96 GB RAM&lt;br&gt;6 × 16 DIMMs&lt;br&gt;PCIe Expansion Module&lt;br&gt;Fusion-io 1.2 TB ioDrive2</td>
<td>BMC/EFI: 01-27</td>
<td>2</td>
</tr>
<tr>
<td>Brocade 6510</td>
<td>SAN switch with 48 × 8 Gb Fibre Channel ports</td>
<td>FOS 7.0.1a</td>
<td>2</td>
</tr>
<tr>
<td>Brocade VDX 6740</td>
<td>Ethernet switch with 40 × 10 Gb/sec ports</td>
<td>NOS 4.1.2</td>
<td>2</td>
</tr>
</tbody>
</table>
Software Components

Table 2 describes the software used in this testing.

Table 2. Software Components

<table>
<thead>
<tr>
<th>Software</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hitachi Storage Navigator</td>
<td>Microcode Dependent</td>
</tr>
<tr>
<td>Hitachi Dynamic Provisioning</td>
<td>Microcode Dependent</td>
</tr>
<tr>
<td>VMware vCenter server</td>
<td>5.5.0</td>
</tr>
<tr>
<td>VMware ESXi with Enterprise Plus license</td>
<td>5.5.0</td>
</tr>
<tr>
<td>Microsoft® Windows Server® 2008</td>
<td>Enterprise edition, R2</td>
</tr>
</tbody>
</table>

Network Infrastructure

The network design used in this environment provided ample bandwidth and redundancy based on Hitachi Unified Compute Platform for VMware vSphere.

SAN Infrastructure

The Hitachi Virtual Storage Platform G1000 controller used in this environment has 32 ports for connections to the Brocade 6510 enterprise fabric switches.

The Hitachi Compute Blade 500 HBAs were zoned to four ports on the Hitachi Virtual Storage Platform G1000 controller.

Figure 1 illustrates the physical SAN connections. Following best practice, this environment used two Brocade 6510 switches to create two separate fabrics.
Figure 1
Hitachi Virtual Storage Platform G1000 is built with multiple controllers to provide input-output hardware-based load balancing. For the best performance and reliability, this environment used at least two ports from each controller. At least two unique paths exist from the ESXi host to the storage system to maximize availability. The multipathing policy was set to **round robin** in ESXi.

Table 3 shows the zone configuration used for this testing.

**Table 3. Zone Configuration**

<table>
<thead>
<tr>
<th>Hitachi Compute Blade 500 Blade/Host</th>
<th>Host HBA Number</th>
<th>Fibre Channel Fabric</th>
<th>Zone Name</th>
<th>Storage Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blade 0/ESX 0</td>
<td>HBA1_1</td>
<td>Fabric 1</td>
<td>CB500_ESX0_HBA1_1_VSP_1A_2C</td>
<td>1A, 2C</td>
</tr>
<tr>
<td></td>
<td>HBA1_2</td>
<td>Fabric 2</td>
<td>CB500_ESX0_HBA1_2_VSP_1C_2A</td>
<td>1C, 2A</td>
</tr>
<tr>
<td>Blade 1/ESX 1</td>
<td>HBA1_1</td>
<td>Fabric 1</td>
<td>CB500_ESX1_HBA1_1_VSP_1A_2C</td>
<td>1A, 2C</td>
</tr>
<tr>
<td></td>
<td>HBA1_2</td>
<td>Fabric 2</td>
<td>CB500_ESX1_HBA1_2_VSP_1C_2A</td>
<td>1C, 2A</td>
</tr>
</tbody>
</table>

**vSphere Flash Read Cache – Flash Devices Installation**

Two Hitachi Compute Blade 500 server blades were used in this test environment.

Figure 2 shows the flash devices tested with vFRC.
There were two options to install flash devices on the Compute Blade 500 server blades:

- Install two SAS attached SSD drives in the existing 2.5 inch slots. In this case, no Compute Blade PCIe Expansion Module is needed.
- Install one or two PCIe Fusion-io ioDrive2 flash cards in the Compute Blade PCIe Expansion Module.
- The ESXi hosts were deployed by VMware Auto Deploy, so no internal storage was used.
Test Methodology

This section describes the test methodology used. The purpose of the tests was to do the following:

- Show the benefit of using vFRC.
- Compare the I/O performance between vFRC with SAS and PCIe flash devices.

A large number of IOPS need to be generated in order to measure the possible maximum number of IOPS that vFRC flash devices can accommodate. To better manage the large workloads, a benchmark tool called Vdbench was used. This tool provides the following advantages for this test:

- Easily define the workloads in detail
- Control of IO rate and run behavior
- Capable of running multiple instances and threads to simulate multiple applications

Table 4 shows the seven workload definitions used for this test.

<table>
<thead>
<tr>
<th>Workload</th>
<th>Block Size (KB)</th>
<th>Read %</th>
<th>Random %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max Throughput 50%</td>
<td>32</td>
<td>50%</td>
<td>0%</td>
</tr>
<tr>
<td>Real Life</td>
<td>8</td>
<td>65%</td>
<td>60%</td>
</tr>
<tr>
<td>Random</td>
<td>8</td>
<td>70%</td>
<td>100%</td>
</tr>
<tr>
<td>SQL Server®</td>
<td>64</td>
<td>66%</td>
<td>100%</td>
</tr>
<tr>
<td>Web Server</td>
<td>8</td>
<td>95%</td>
<td>75%</td>
</tr>
<tr>
<td>Exchange 2007</td>
<td>8</td>
<td>55%</td>
<td>80%</td>
</tr>
<tr>
<td>OLTP</td>
<td>2</td>
<td>67%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Storage Configuration

To show the maximum benefit of vFRC, only one parity group of SAS with minimum size of 8 GB of Cache Logical Partition (CLPR) was used on the Virtual Storage Platform G1000.
Figure 3 shows the storage configuration and VMDK creation.

Only one parity group of RAID-6 (6D+2D) 900 GB 10K SAS was placed in Hitachi Dynamic Provisioning Pool 0. Since the maximum size of a LDEV is 3 TB, to fully utilize the capacity of the 4.7 TB parity group, two 2.35 TB LDEVs were created as pool volumes. One virtual volume was created from Pool 0 and presented as a datastore to ESXi hosts. 40 GB VMDKs were created from the datastore and attached to a Vdbench VM. For each VMDK, 8 GB of vFRC were allocated as VMware’s recommendation of 20% of the VMDK capacity.
Test Cases

Table 5 lists the three main test cases. All seven workloads were executed for each test case. For each workload, the I/O were issued to the target VMDK for 10 minutes, after a 30 minute warm-up run. The data was cached into vFRC during this warm-up run.

<table>
<thead>
<tr>
<th>Test Case</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>Run all seven workloads without enabling vFRC</td>
</tr>
<tr>
<td>vFRC with SAS SSD</td>
<td>Run all seven workloads with vFRC using internal SAS attached SSDs</td>
</tr>
<tr>
<td>vFRC with PCIe Flash</td>
<td>Run all seven workloads with vFRC using internal PCIe Fusion-io ioDrive2</td>
</tr>
</tbody>
</table>
Analysis

This is an analysis from the results of the vSphere Flash Read Cache testing.

Figure 4 shows the maximum IOPS achieved on each test case. A higher IOPS number indicates higher performance.

![IOPS Comparison](image)

**Figure 4**
Figure 5 shows the I/O response time comparison between each test cases. Lower response times are better, in terms of storage and cache device performance.

![Response Time Comparison](image)
Over all, much higher IOPS with better response time was reached when vFRC was used. A summary and analysis of the results are listed below:

- With vFRC, Web Server workload reached more than four times of IOPS compared to the baseline.
  - The vFRC is most suitable for high-read random workload like web servers.
- With vFRC, SQL Server workload reached almost twice of the IOPS than the baseline.
  - SQL Server workload is also a relatively high-read and random workload, and also a good candidate for the vFRC use case.
- There was no difference for OLTP workload.
  - Workload with high random and small block size, like OLTP with 2 KB block size, will not benefit much from vFRC.
- Except for Max Throughput workload, the PCIe flash device only performed slightly better than SSD.
  - For the most of the case, SSD seems like a more cost effective choice for the vFRC.
- For the Max Throughput 50% workload, vFRC with SSD performed almost same as baseline.
  - There were several factors that were considered contributors to this result.
    - SAS 2.0 has the lowest bandwidth of 750 MB/sec whereas FC 8 Gbps is 1 GB/sec and PCIe 2.0 is 8 GB/sec.
    - Bandwidth for SAS attached SSD is between 200 MB/sec and 500 MB/sec depending on the device compared to ioDrive2 which is 1.3 GB/sec.
    - Local SSDs are managed by a local RAID controller. Even though no RAID was configured for this test, there might be an overhead from this local RAID controller.
    - Virtual Storage Platform G1000 was used in the back end baseline test. Virtual Storage Platform G1000 is an enterprise grade storage with features like Read Ahead during the sequential reads. Although Cache Logical Partition was configured to minimize storage array caching benefits as much as possible, the baseline IOPS may have been set higher compared to running on basic storage.
Benefits of Using vFRC

On certain workloads, such as high-read random workload, vFRC could provide tremendous I/O performance improvements. With this I/O performance improvement, physical server nodes can be consolidated.

Figure 6 shows an example of how cost could be reduced with vFRC if vFRC with SSD provides twice the I/O performance improvement. With twice the I/O improvement, the server nodes could be reduced to half, and VMware License and support fees could be reduced as well. Please note that this example is purely from an I/O performance perspective.

<table>
<thead>
<tr>
<th>Without vFRC</th>
<th>Unit Price($)</th>
<th>Qty</th>
<th>Subtotal (K$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Server X – 64 Nodes</td>
<td>$ 5,000.00</td>
<td>64</td>
<td>$ 320.00</td>
</tr>
<tr>
<td>VMware License</td>
<td>$ 3,500.00</td>
<td>128</td>
<td>$ 448.00</td>
</tr>
<tr>
<td>VMware Support 3yr</td>
<td>$ 2,500.00</td>
<td>128</td>
<td>$ 320.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td><strong>$ 1,088.00</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>With vFRC and SSDs</th>
<th>Unit Price($)</th>
<th>Qty</th>
<th>Subtotal (K$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Server X – 32 Nodes</td>
<td>$ 5,000.00</td>
<td>32</td>
<td>$ 160.00</td>
</tr>
<tr>
<td>SAS SSD 600GB</td>
<td>$ 2,500.00</td>
<td>32</td>
<td>$ 80.00</td>
</tr>
<tr>
<td>VMware License</td>
<td>$ 3,500.00</td>
<td>64</td>
<td>$ 224.00</td>
</tr>
<tr>
<td>VMware Support 3yr</td>
<td>$ 2,500.00</td>
<td>64</td>
<td>$ 160.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td><strong>$ 624.00</strong></td>
</tr>
</tbody>
</table>

Figure 6  

42% Savings!
Considerations

The following needs to be considered before deploying VMware vFRC to your environment.

- The vFRC provides more benefit if your environment is I/O bounded. This means your applications perform better with faster I/O response. If your environment is bounded by other resources, such as CPU, memory, or network, the benefit from vFRC may be limited.

- This test was performed as a small scale test without using the full capacity of the SSD or PCIe flash devices. The result might be different in larger scale environments.

- In many cases, the performance delta between SAS, SSDs, and PCIe Flash devices is negligible, and the cost savings of SAS SSDs are significant. This makes SAS SSD a great match for vFRC.

- If your application saturates the SSDs bandwidth, then PCIe flash devices are a good alternative for these specific workloads.
For More Information

Hitachi Data Systems Global Services offers experienced storage consultants, proven methodologies and a comprehensive services portfolio to assist you in implementing Hitachi products and solutions in your environment. For more information, see the Hitachi Data Systems Global Services website.

Live and recorded product demonstrations are available for many Hitachi products. To schedule a live demonstration, contact a sales representative. To view a recorded demonstration, see the Hitachi Data Systems Corporate Resources website. Click the Product Demos tab for a list of available recorded demonstrations.

Hitachi Data Systems Academy provides best-in-class training on Hitachi products, technology, solutions and certifications. Hitachi Data Systems Academy delivers on-demand web-based training (WBT), classroom-based instructor-led training (ILT) and virtual instructor-led training (vILT) courses. For more information, see the Hitachi Data Systems Services Education website.

For more information about Hitachi products and services, contact your sales representative or channel partner or visit the Hitachi Data Systems website.