Dynamic VM Migration: Assessing Its Risks & Rewards Using a Benchmark

Krishnamurthy Srinivasan
Intel Corporation
5000 W Chandler Blvd
Chandler, AZ
krishnamurthy.srinivasan@intel.com

Sterlan Yuuw
Intel Corporation
5000 W Chandler Blvd
Chandler, AZ
sterlan.yuuw@intel.com

Tom J. Adelmeyer
Intel Corporation
5000 W Chandler Blvd
Chandler, AZ
tom.j.adelmeyer@intel.com

ABSTRACT

Dynamic migration of virtual machines (VMs) across physical servers has the potential to increase the utilization of the servers and hence drive down the data center costs. However, IT practitioners are leery of using this capability for increasing resource utilization due to concerns about the impact of such migrations on the performance of the applications, particularly the response times seen by the users of the applications. The relative newness to the industry of many of the tools used to automate VM migrations for resource utilization; data from researchers; as well as the recommendations from some analysts justify such caution and warrant quantifying the risks as well as potential rewards before deciding how aggressively this capability should be adopted. This paper will discuss the requirements for a benchmark to be used for such quantification. We will also discuss adaptations to SPECvirt_sc2010*, originally developed as a single server benchmark, to meet these requirements[1]. We will also present risk-reward quantifications obtained using this benchmark for a simple case and the broader use of the benchmark for other cases.

Categories and Subject Descriptors

C.4 [Computer Systems Organization]: Performance of systems (Design studies; Measurement techniques; Modeling techniques; Performance attributes) I.0 [Computing Methodologies]: General

General Terms

Management, Measurement, Performance, Design, Experimentation.

[1] The modifications we have made are non-compliant with the SPEC* run-rules. The data presented here are only to illustrate the points discussed in this paper and cannot be compared with any other SPECvirt_sc2010* results.

ICPE'11, March 14–16, 2011, Karlsruhe, Germany.

Keywords

Virtualization, Server, Data Center, Virtualization 2.0, Cloud, Migration, Live Migration, Benchmark.

1. INTRODUCTION

In the recent economic recession, most enterprise IT shops, faced with the universal pressure to cut costs, have adopted virtualization aggressively to reduce the number of servers deployed through server consolidation. According to the International Data Corporation (IDC)*, "customers came out of the economic downturn with a 'virtualize first' mentality" when it comes to deploying new servers [4].

The cost reduction achievable is directly proportional to the degree of consolidation or the number of VMs that are run on a single physical server. In an informal survey, we found that typical IT shops run 10-15 VMs on a single server with an average CPU utilization of only about 25-35%.

Why are the IT shops not more aggressive with server consolidation given the motivation to reduce costs? To answer this question, we need to understand the nature of the workload patterns of the applications run by them. Most of these workloads vary over time. Speitkamp and Bichler present an analysis of load changes in a large commercial data center and have also made the data publicly available [7]. They have identified load changes that could be explained with diurnal or other seasonal variations or just plain random variations. They also present groups of applications with either congruent or opposing patterns of load level variations.

Given the variations in the loads on individual applications, the data center manager faces multiple choices in determining the number of applications to consolidate on a given server:

1. Be conservative and consolidate only as many application as possible allowing for all the application loads to peak simultaneously;
2. Be aggressive and allow for only the average load-level of the applications at any instance; or
3. Make a choice somewhere in between the above two.

Armbrust, et al., [1] have presented the higher cost incurred due to choice 1 above vs. the risk of losing customers due to poor response time with choices 2 and 3.

1.1 (Non-)Use of Dynamic VM Migration to Optimize Utilization

If applications can migrate from a server with high resource utilization to one with low resource utilization, we can achieve higher resource utilizations at a lower risk of running into performance problems as the probability of the loads on all the applications across a number of servers peaking simultaneously reduces exponentially with the number of servers and applications.

However, we found that even IT shops experienced with live migration of VMs for purposes such as hardware maintenance did not use VM migrations for optimizing resource utilization often. Our findings are validated by a formal survey conducted by Forrester Consulting* [3].

IT managers were apprehensive about the impact of migrations on the performance of applications and specifically the response times seen by the users of the applications. Several researchers have underscored the importance of considering an applications' use of platform resources such as the CPU, memory, and network/disk input/output (I/O) in picking the right applications to deploy side-by-side on a single server. Pu, et al., found that running CPU- and network-intensive loads in conjunction increases the performance [5]. Armbrust, et al., have found that in the Amazon EC2* environment, the consolidation of CPU- or memory-intensive benchmarks together on a single server causes lesser performance deterioration compared to consolidation of I/O-intensive benchmarks [1].

Virtualization software vendors allow IT personnel to perform VM migration for resource optimizations (i) manually; (ii) using scripts; or (iii) high-level automation tools for policy-based migration decisions [9, 10]. Some of them consider only the CPU and memory requirements of the applications.

Burton Group* has summarized some concerns regarding dynamic load balancing across servers [2]:

"Load balancing allows the IT organization to more efficiently use existing physical infrastructure, but can also come with risks if left unchecked. For example, not all VM load-balancing services count I/O when determining where to place a VM. Instead, the service relies on free CPU and memory capacity. Without proper I/O accounting, a VM could become I/O bound if placed on a physical host that is using most of its I/O channels. Administrators would have to react to the problem (likely discovered by a call to the help desk) by manually moving the VM to another host."

1.2 Why Use a Benchmark?

It appears the risks and potential rewards of resource optimization through VM migrations depend on the performance characteristics of the applications, the variations in application load levels and the ability of the resource optimization tools to look at the resource needs of the applications comprehensively. Additionally, the resources that need to be considered could also be specific to the virtualization environment. For example, the VMware ESX* environment allows setting the depth of the various queues in the I/O path and these settings could have a significant impact on the I/O latency and hence the response of times of certain applications [8].

One option is to use mathematical models to predict the utilization improvements and performance degradations associated with greater degrees of server consolidation. As we discussed in the previous paragraph, the model will highly depend on the virtualization environment, the logic used by the automation tools and so on. Construction of models with reasonable levels of accuracy and the validation of the models will require considerable data based on the specific user's applications and infrastructure. Such models will be valuable once they mature. The insights that could be gained from benchmark characterization are needed to develop the models, and could address the concerns of the IT personnel meantime.

1.3 Rest of the paper

We will first discuss the requirements for a benchmark to be able to quantify the risks and benefits of VM migration. Next, we will describe SPECvirt_sc2010*, a benchmark released recently by SPEC* for performance characterization of consolidation on single servers [6], and how it meets the requirements for a VM migrations benchmark. Subsequently, we will discuss the modifications we made to SPECvirt_sc2010* to use it for our current purpose. Finally, we will use a simple, but reasonably realistic scenario to characterize the resource utilization, aggregate throughput and response time with and without VM migrations.

Our goals in this paper are (i) to highlight the current state of dynamic resource optimization through VM migrations; (ii) to demonstrate the use of SPECvirt_sc2010* to characterize VM migrations in a particular scenario; and (iii) to point out the parameters that need to be changed for broader use of this benchmark in other scenarios. Our goals do NOT include developing better algorithms for automating VM migration decisions or to provide data on the impact of VM migrations that are broadly applicable to all scenarios.

2. Benchmark Requirements

We will list the requirements a benchmark should meet for it to be useful as a tool to understand the impact of VM migrations on the resource utilization and the application performance:

1. The benchmark should comprise a heterogeneous mix of component benchmarks that represent typical enterprise applications;
2. The component benchmarks should exhibit different CPU, memory and I/O usages;
3. These component benchmarks should have load levels that change dynamically during the run of the benchmark;
4. The variations in the load level should be customizable to mimic the seasonal and random changes seen in a particular user's applications;
5. It should be possible to obtain throughput as well as response times as performance measures of the benchmarks;
6. The benchmark should provide metrics to show power consumption relative to its performance (e.g., using a performance-per-watt or PPW metric);
7. It should be possible to run a mix of lightly- and heavily-loaded component benchmarks and be able to

fine-tune these load mixes to reflect sizes of a specific user's application; and

8. It should be possible to scale the benchmark from a single server to several servers (most IT shops migrate VMs within clusters of 10-30 servers; however, some 3rd party service providers may have significantly larger clusters).

3. SPECvirt_sc2010*

SPECvirt_sc2010* has been developed by SPEC* to model server consolidation and comprises modified versions of three standard benchmarks from SPEC*, representing three commonly consolidated applications [6] (Figure 1):

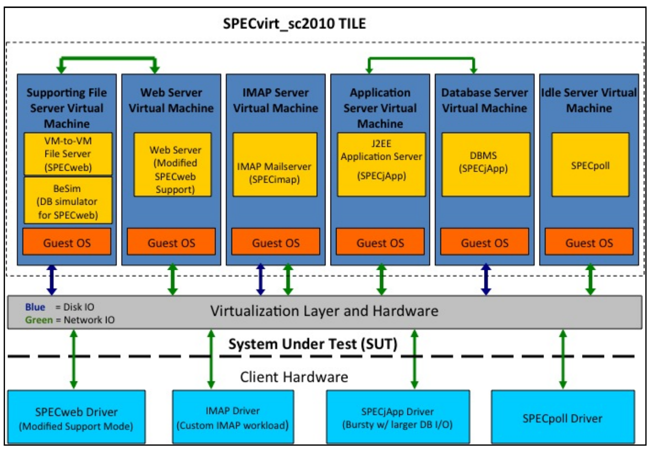

Figure 1. One Tile of SPECvirt_sc2010

1. *SPECweb2005 Support** representing web servers; it has been partitioned into two VMs – one running a web server and another running the Backend Simulator and a file server.
2. *SPECjAppServer2004** representing application and database tiers of a multi-tier application. The application server and database server are run in two separate VMs that communicate with each other.
3. SPECmail2008* representing mail servers. This runs in a VM by itself.

In addition to the above three applications, SPECvirt_sc2010* also includes an *Idle Server* VM to represent very infrequently used applications commonly found running on real-world servers.

A set of the above VMs is called a "tile." Each tile has a fixed amount of throughput rate by default (please see feature 3 below for the exception to this default). A virtualization platform is assessed by determining the maximum throughput it can handle by adding more tiles as long as some quality of service criteria based on response times are met by all the applications.

In the next section we will describe some highlights of SPECvirt_sc2010* that are relevant for our current topic.

3.1 Key Features of SPECvirt_sc2010*

SPECvirt_sc2010* meets almost all the requirements listed earlier in Section 2. With respect to the last requirement in that list, viz., scalability to multiple nodes, the benchmark was not developed with that requirement as a goal and the benchmark has not been tested with a very large number (several 100s) of VMs yet to the authors' knowledge. Here are some of the key features.

1. The three main component benchmarks have performance scores based on their throughput.
2. They are also expected to meet Quality of Service requirements based on response times.
3. The load level on the individual tiles can be controlled to be 0.1X to 0.9X of the default load levels in steps of 0.1X. For keeping the all the published results comparable, the load level can be altered at most for only one tile. However, for experimental purposes, we can alter the load on any or all the tiles.
4. The benchmark also supports gathering power consumption data during the runtime and reports out the performance-per-watt for either just the server or for the server and storage.
5. The benchmark has been designed to use non-trivial amounts of network and disk I/O. In addition to using the Support component of SPECweb2005* which is I/O intensive, the benchmarks were also modified to increase the I/O. Each benchmark is also distinct in its use of CPU, memory and I/O resources. The Mail Server and Database Server are disk-write intensive; the Supporting File Server VM is disk-read intensive (the amount of disk reads depends on the memory available to cache the web content); the Web Server VM has a significant network transmit volume; the Web Server and Supporting File Server VMs have high inter-VM network traffic; and the Web Server and the Application Server VMs are the most CPU- and memory-intensive.
6. A feature that makes SPECvirt_sc2010* a good candidate to study dynamic resource management is the dynamic variations to the load level of SPECjAppServer2004* benchmark. The injection rate (IR) of transactions for this benchmark is changed along a curve with 30 points over the run-time of the benchmark as shown in Figure 2. The IR is kept constant for 40 seconds at each point on the curve. The starting IR is offset by 7 points along the curve shown in the figure for each tile (the first tile starts at IR point 1, the second tile at 8, and so on) to minimize the chances of two or more tiles having the same load levels throughout the run. The IR curve, the amount of time spent at each point and the offset value are all configurable. Changes to these values render the results non-compliant according to the SPEC* rules. However, these changes are allowed for experimental purposes.

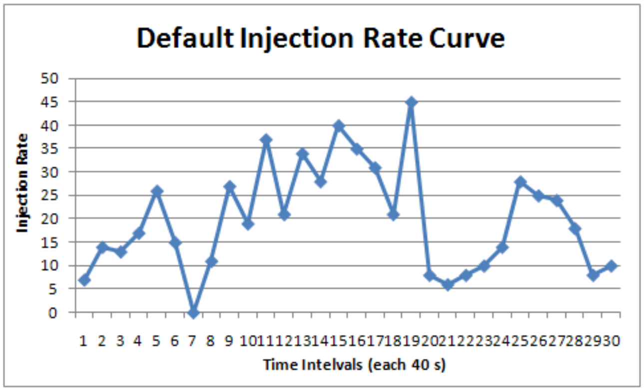

Figure 2. Default injection rate variation in SPECjAppServe2004*.

4. Experimental Scenario: Scope and Assumptions

Our goal in conducting these experiments is to show that SPECvirt_sc2010* could be used with just some configuration changes (no changes to the code) to assess the benefits and pitfalls of VM migrations for dynamic resource management. We will do it with a simple scenario while pointing out the changes needed to assess other scenarios.

- We will run the VMs on two servers and migrate the VMs between them. Changing the number of servers is straight forward as we will see in the next section.
- Half of the VMs have identical or synchronized variations in the load levels; and the load levels of other half complement the first (i.e., the peaks and troughs of one half are synchronized while the peaks of the first half coincide with the troughs of the second half and vice versa). It is common to find several applications with such synchronized or complementary load patterns [7]. This pattern of load variations increases both the need for and the benefit from migrations. A different pattern may reduce both, but a larger number of servers will offset the reduction. SPECvirt_sc2010* allows changing these assumptions easily.
- A simple script is created to monitor the CPU utilization and reactively move VMs conservatively one at a time. Agents that monitor other resources, use their a priori knowledge of application load trends or move VMs more aggressively (e.g., move multiple VMs between the same nodes simultaneously as allowed by some virtualization environments now) could show greater benefit due to VM migration.
- Our watermark CPU utilization for each server at which we look for opportunities to migrate is 85%. We chose this to assess the impact of migrations at a CPU utilization level that is higher than what is encountered in most IT shops, but still leaves some headroom.

5. Modifications to SPECvirt_sc2010*

SPECvirt_sc2010* was developed to be a single server benchmark. We will describe the modifications we made to its configurations to use it in the experimental scenario we discussed in Section 4. There are two factors that made it possible to use this benchmark with minimal changes for studying VM migrations between multiple hosts:

1. The Internet Protocol (IP) addresses of the VMs do not change when the VMs migrate from one host to another. Hence, benchmark clients are oblivious to the number of physical servers that host the VMs or the migration of the VMs across the hosts; and
2. The benchmark exposes several parameters to easily modify its behavior without having to change the code.

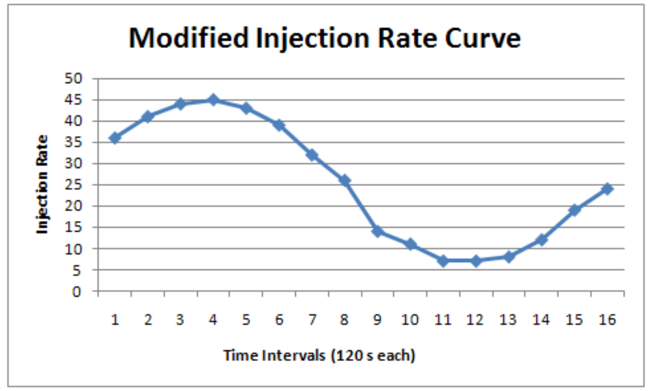

Figure 3. New injection rate variation in SPECjAppServe2004*.

The following are the key modifications made were to SPECvirt_sc2010*:

1. Run half of the tiles on one server and the other half on a second server.
2. The default time spent at each IR point, viz., 40 seconds, and the upward and downward swings between successive IR points made it hard for the benchmark to benefit from VM migrations. Also, in the real-world applications, while short bursts in loads are also encountered, many increases in loads last longer (e.g., logins in the morning). Some of these periodic changes also show a gradual increase in load level, staying for some time at the peak level, followed by a gradual reduction in the load level. We changed the IR curve to have only 16 points and to spend 120 seconds at each point (Figure 3).
3. We changed the offset for the starting IR point for the consecutive tiles to be 8. Since the IR curve has a total of 16 points, alternate tiles have the same starting point and hence identical behavior. Here are the changes made to the run.properties file of SPECjAppServer2004*:

```
stepRate =120000
startPointMultiplier = 8
burstyCurve=36,41,44,45,43,39,32,26,14,11,
7,7,8,12,19,24
```

4. The software-only virtual NICs for inter-VM traffic were replaced by virtual NICs with physical NICs as the communicating VMs could end up being on two different hosts.

6. Experiments and Results

The experiments were conducted using two servers, each running 4 tiles of SPECvirt_sc2010* modified as described in the previous section. The first three tiles on each server ran the default load levels for the benchmark and the fourth tiles were run only at 30% of the default load. This load level was chosen since prior experiments indicated that this load could be run at 80% CPU utilization with no VM migration. The runs were repeated with and without VM migration to compare the throughput, response times and platform resource utilizations.

6.1 HW Configuration

The two servers each had two Intel Xeon® X5570 (2.93 GHz) processors and 48 GB memory. Each server had six Gigabit Ethernet ports, four of which assigned to the VMs, one for VM migration and one for management. In the beginning, one network port was assigned to all the VMs on a tile. After migrations, some ports were supporting 2 full tiles. The server was connected to a fiber channel storage area network using a 4 GB per second adapter. Enough disks were allotted so that the average disk latency was around 2 ms.

6.2 VM Migration Triggering Condition

The average CPU utilizations on both the servers were measured every 40 seconds. If the utilization on one of the servers exceeded 85% for three consecutive measurements, it was checked to see if the other server's CPU utilization is 10% lower than the first. If it was, one VM at a time was migrated in the order of "Mail Server, Supporting File Server, Web Server, Database Server and Application Server." Mail Server was migrated first as it did not depend on any other server on the same host. The other VMs with mutual dependencies were migrated one after the other. A second VM was not migrated until the first one's migration is completed and it had started running on the target server.

6.3 Results

In this section, we will look at (i) CPU utilization; (ii) throughput; (iii) QoS/response times; and (iv) consolidation density for two cases – one with VM migrations and one without.

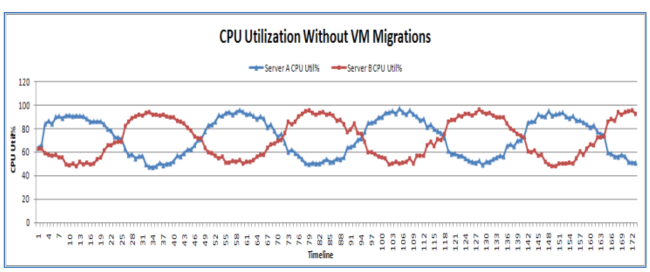

Figure 4. CPU Utilization without migrations

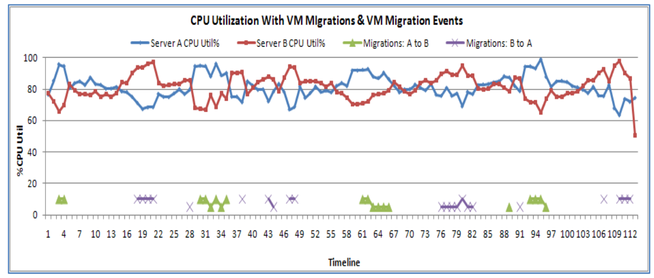

Figure 5. VM migrations and CPU utilization

6.3.1 CPU Utilization Impact

The graphs in Figure 4 show that as the load starts to increase on the VMs in server A, the load starts dropping in host B (after an initial warm-up period when both the hosts show the same load levels). In the case where VMs are migrated, when the migration triggering condition described in section 6.2 is met, the VMs are migrated to server B. In that case, the CPU utilization is higher than the target maximum rate of 85% for a much shorter duration than the case where the VMs are not migrated. Figure 5 shows when VM migrations happen and how it impacts the CPU utilization on both the servers.

As mentioned earlier, our script does not migrate multiple VMs at a time. VM migration could be even more effective in preventing prolonged spikes in resource utilizations if such capabilities are leveraged.

6.3.2 Throughput Impact

VM migrations helped to increase the aggregate throughput by 9.4% compared to the case when VMs were not migrated. This was due to VM migrations keeping the periods of very high CPU utilization much shorter.

6.3.3 QoS and Response Time Impact

When VMs were not migrated, all the instances of SPECjAppServer2004* running the default load failed to pass the QoS requirements whereas once the VMs were allowed to migrate all the instances passed. It was noticed that VM migrations helped to reduce the response times for the three Dealer transactions of SPECjAppServer2004* 4-8X as shown in Table 1 (except the two tiles running only 30% load):

Table 1. Impact of VM Migrations on SPECjAppServer2004 Dealer Transactions 90th Percentile Response Times

	90% Response Time (s)	
Tile	No Migration	With Migration
1	5.70/4.70/7.30	0.90/1.20/1.10
2	6.00/5.20/7.20	1.40/1.60/1.50
3	6.90/5.80/7.80	0.70/0.90/0.90
4	7.10/5.90/7.80	1.00/1.20/1.10
5	5.10/4.60/5.80	0.90/1.10/1.10
6	3.50/3.50/4.30	1.30/1.40/1.40
7	0.10/0.20/0.20	0.10/0.20/0.20
8	0.10/0.20/0.20	0.10/0.20/0.20

6.3.4 *Consolidation Density Impact*

As started earlier, the runs without VM migrations failed to pass QoS requirements when running eight tiles (or 48 VMs) on the two servers. We were able to run only six tiles (or 36 VMs) while meeting the QoS requirements. On the other hand, when VM migrations were allowed, we were able to run 8 tiles (though the last two tiles – one on each server -- were able to handle only 30% of the default full load of the other tiles.

7. Conclusions

In this paper, we saw that while IT managers are trying to reduce data center costs through server consolidation, many of them are not thus far using dynamic VM migrations to maximize consolidation densities. Analysts and researchers encourage such a cautious approach at least for certain application classes and with most resource management tools. A flexible benchmark (i) with a mix of CPU-, memory-, and I/O-intensive workloads; (ii) that could be configured easily to model the dynamic changes in demands for these resources; and (iii) having the ability to measure throughput, response times and power, will help answer the questions IT managers have about the risks and rewards associated with the use of VM migrations for resource optimization. SPECvirt_sc2010* meets many of the requirements. Though it was designed as a single-server benchmark, we showed it could be used for characterizing two servers with VM migrations across them with just some configuration changes. We discussed a simple, but realistic enough scenario and using SPECvirt_sc2010* quantified the benefits of VM migration including avoiding spikes in resource utilization; improving the ability to meet QoS criteria; and increasing consolidation densities.

We intend to extend these studies in a few directions:

1. Alter the virtual machine or platform configurations (e.g., memory size, number of NICs, and number of disks) to study the impacts of constraints in resources other than the CPU;
2. Conduct the study with automated resource managers that are bundled with some of the virtualization environments; and
3. Alter our script for triggering migrations to monitor additional resources and response times and use this additional data in making migration decisions.

We hope that IT users, vendors, researchers and others would conduct similar studies for other application types, load patterns, larger number of servers and so on, and create a critical mass of public knowledge needed. In addition to VM migrations, automated VM provisioning and other areas also need more studies using realistic heterogeneous workloads and employing real-world characteristics such as varying load levels so that virtualization technologies could be leveraged to their full potential to minimize data center costs.

8. REFERENCES

[1] Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., and Zaharia, M., A View of Cloud Computing, *Communications of the ACM*, April 2010, Vol. 53, No. 4.

[2] Burton Group*: Stuck Between Stations: From Traditional Data Center to Internal Cloud, (Author: Wolf, C.), Version: 1.0, Sep 22, 2010.

[3] Forrester Consulting*: Virtualization Management And Trends, January 28, 2010, http://www.ca.com/files/IndustryAnalystReports/virtual_mgmt_trends_jan2010_227748.pdf .

[4] IDC* - Press Release: Virtualization Market Accelerates Out of the Recession as Users Adopt "Virtualize First" Mentality, 28 April, 2010, http://www.idc.com/about/viewpressrelease.jsp?containerId=prUS22316610&sectionId=null&elementId=null&pageType=SYNOPSIS.

[5] Pu, X., Liu, L., Mei, Y., Sivathanu, S., Koh, Y., Pu, C., Understanding Performance Interference of I/O Workload in Virtualized Cloud Environments, *2010 IEEE 3rd International Conference on Cloud Computing (CLOUD)*, Miami, FL, 5-10 July 2010, 51-58.

[6] SPEC*: SPECvirt_sc2010 Design Overview, 06/29/2010, http://www.spec.org/virt_sc2010/docs/SPECvirt_Design_Overview.html.

[7] Speitkamp, B., Bichler, M., A Mathematical Programming Approach for Server Consolidation Problems in Virtualized Data Centers, IEEE *Transactions on Services Computing*, Vol. 3, No. X, 2010.

[8] VMware* Communities: Storage Queues and Performance, June 1, 2010, http://communities.vmware.com/docs/DOC-6490.

[9] vSphere*2 Resource Management Guide, http://www.vmware.com/pdf/vsphere4/r41/vsp_41_resource_mgmt.pdf.

[10] Windows Server 2008 R2 Hyper-V™ Live Migration http://download.microsoft.com/download/5/B/D/5BD5C253-4259-428B-A3E4-1F9C3D803074/LiveMigrationWhitepaper_Final.docx.
