

SPECvirt_sc2010 - Driving Virtualization Innovation

David L. Schmidt
Hewlett-Packard Company
d.schmidt@hp.com

Andrew Bond
Red Hat, Inc.
abond@redhat.com

Lisa Roderick
VMware, Inc.
lroderic@vmware.com

Klaus-Dieter Lange
Hewlett-Packard Company
klaus.lange@hp.com

ABSTRACT

Overview and future outlook of the #1 industry standard virtualization benchmark, SPECvirt_sc2010.

Categories and Subject Descriptors

H.3.4 [Systems and Software]: Performance evaluation (efficiency and effectiveness)

General Terms

Design, Experimentation, Measurement, Performance, Reliability, Standardization

Keywords

ACM 978-1-4503-1202-8/12/04.ACM 978-1-4503-1202-8/12/04.SPEC, Virtualization, Benchmark, Energy Efficiency, Server, Storage, Datacenter

1. SPECvirt_sc2010 OVERVIEW

SPECvirt_sc2010 [1] is designed to be a standard method for measuring a virtualization platform's ability to manage a server consolidation scenario in the datacenter and for comparing performance between virtualized environments. It is intended to measure the performance of the hardware, software, and application layers in a virtualized environment. This includes both hardware and virtualization software and is intended to be run by hardware vendors, virtualization software vendors, application software vendors, academic researchers, and datacenter managers. The benchmark is designed to scale across a wide range of systems and is comprised of a set of component workloads representing common application categories typical of virtualized environments.

Rather than offering a single benchmark workload that attempts to approximate the breadth of consolidated virtualized server characteristics found today, SPECvirt_sc2010 uses a three-workload benchmark design: a webserver, Java application server, and a mail server workload. The three workloads SPECvirt_sc2010 utilizes are derived from SPECweb2005, SPECjAppServer2004, and SPECmail2008.

All three workloads drive pre-defined loads against sets of virtualized servers. The SPECvirt_sc2010 harness running on the client side controls the workloads. Each workload must maintain a quality of service (QoS) criteria in order to be considered valid.

The benchmark utilizes the concept of a "tile" as a mechanism to increase the stress on the system under test (SUT). Each tile contains six virtual machines that are used to drive the three workloads of the benchmark (Figure 1). Two of the virtual machines are used for the web server workload, the web server VM and the infrastructure server VM which houses the data store for the webserver.

The application server workload utilizes a virtual machine that runs the application server software and a database VM. The mail server workload runs on a single virtual machine. One virtual machine is kept idle to represent VMs in a typical user environment which are not running at full capacity. Stress on the SUT is increased by adding more tiles.

All tiles must be identically configured and operate independently of each other – i.e. they each have their own unique workload dataset. Peak performance is the point at which the addition of another tile either fails the QoS criteria or fails to improve the overall metric.

The benchmark allows the freedom to select the virtualization implementation, the software applications that run within the VMs, and the hardware configuration as long as the benchmark's run rules are followed. SPECvirt_sc2010 also implements the SPECpower methodology [2] for power measurement.

The benchmarker has the option of running with power monitoring enabled and can submit results to any of three categories:

performance only – (SPECvirt_sc2010)

performance/power for the SUT – (SPECvirt_sc2010_PPW)

performance/power for the Server only – (SPECvirt_sc2010_ServerPPW)

As with all SPEC benchmarks, an extensive set of run rules govern SPECvirt_sc2010 disclosures to ensure fairness of results. SPECvirt_sc2010 results are not intended for use in sizing or capacity planning. The benchmark does not address multiple host performance or application virtualization.

2. Future Design Considerations

As hardware and software technologies continue to progress and marketplace virtualization trends evolve, The SPEC Virtualization committee continues to work on updates and future benchmarks

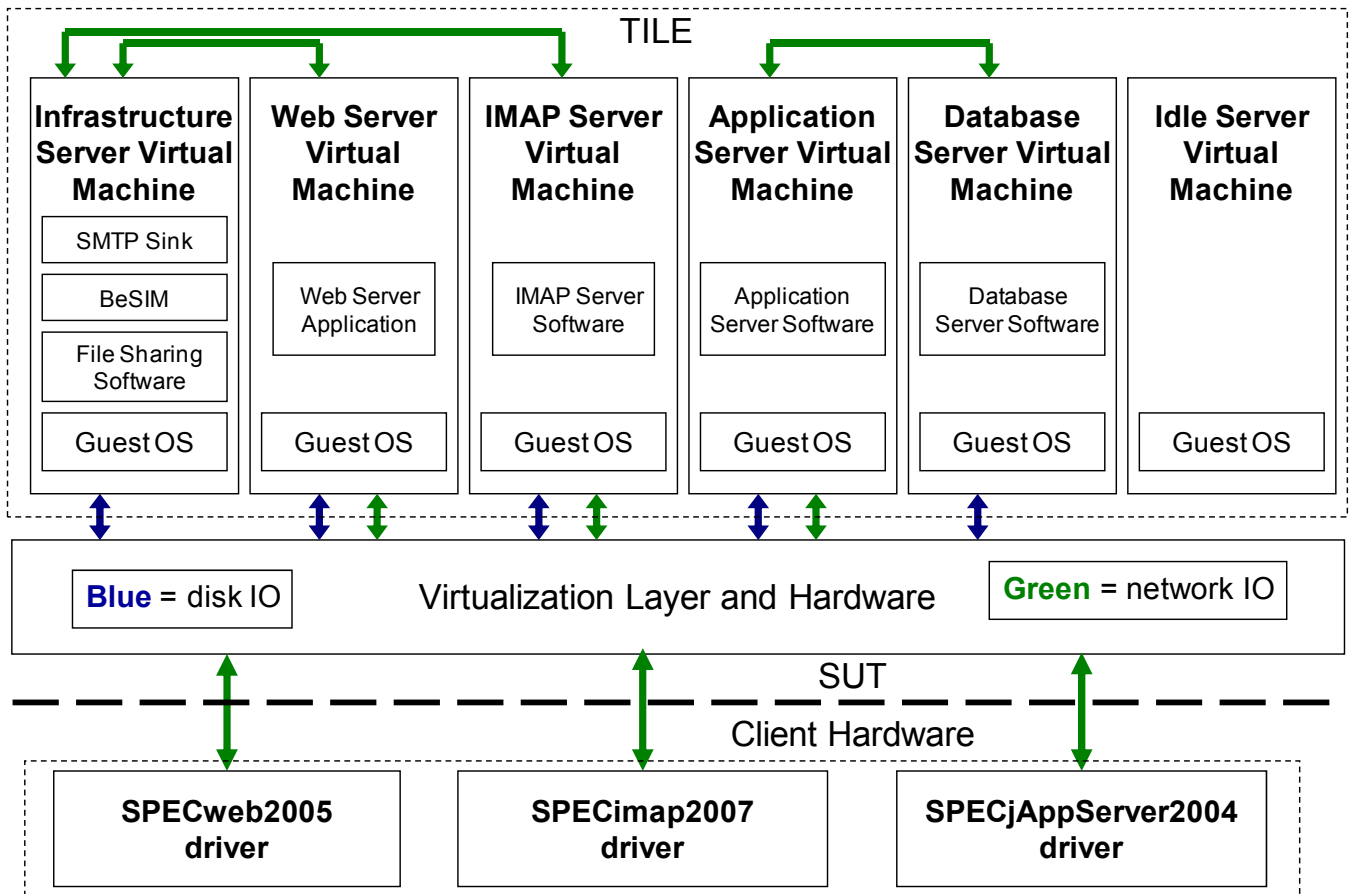


Figure 1: SPECvirt_sc2010 Benchmark Single Tile Design

to provide meaningful tools to measure these technologies. Several future design considerations for new workload models, including enterprise-class server consolidation and data center virtualization (such as VM provisioning and VM migration), are currently being discussed and developed for future benchmark releases.

3. ACKNOWLEDGMENTS

The authors want to acknowledge the additional members of the SPEC Virtualization Subcommittee who have contributed to the design, development, testing, and overall success of SPECvirt_sc2010.

The name SPEC together with its tool and benchmark names are registered trademarks of the Standard Performance Evaluation Corporation (SPEC) [3].

4. REFERENCES

- [1] SPECvirt_sc2010 home page: http://www.spec.org/virt_sc2010/
- [2] Standard Performance Evaluation Corporation home page: <http://www.spec.org>
- [3] SPEC Benchmark Methodology: http://www.spec.org/power/docs/SPEC-Power_and_Performance_Methodology.pdf