Chapter 12
Virtualization Benchmarks

Klaus-Dieter Lange, David Schmidt, and David Morse

The concept of partitioning a computer's physical resources to create virtualized user environments has been around for decades. From the mainframes of the 1960s and 1970s through the current era of x86 servers, the goal of maximizing the utilization and efficiency of business IT resources has driven the ongoing development of server virtualization technologies. The fundamental component of server virtualization is the hypervisor, which is computer software, firmware, or hardware that creates and runs virtual machines. Hypervisors can run directly on a host's hardware, or can run as an application within a traditional OS, to manage and control virtual machines (VMs) on the host.

At the end of the twentieth century and the beginning of the twenty-first century, a new wave of hypervisors, led by VMware's VMware Workstation and VMware ESX Server, ushered in a renaissance of server virtualization in the computer industry. These new hypervisors ran on newer x86 servers which were increasing in popularity with businesses. In less than 10 years, many hardware and software vendors such as HP, Sun, Microsoft, Citrix, Oracle, and Red Hat introduced competing virtualization solutions. Additionally, the open-source community developed hypervisors like Xen.

One of the early uses of these hypervisors was to take advantage of improved performance of newer generations of server hardware products to consolidate the business applications running on older servers onto fewer new ones. This consolidation increased the utilization of servers in a data center and because fewer servers were needed it led to reduced space, cooling, and energy costs.

Further development of hypervisors introduced redundancy capabilities allowing VMs to be moved between physical servers and storage pools within a cluster of hypervisors while still being fully active. This capability allowed IT departments to perform maintenance on the hardware, add new hardware to a given cluster, and balance the resource utilization across multiple servers, all with no downtime for the end users. New VMs can be provisioned from predefined templates for rapid deployment. These hypervisor capabilities provide the infrastructure for what is now known as "The Cloud."

With the introduction and rise in popularity of new server virtualization products in the early 2000s, the inevitable question arose: Which solution performs better? The question is not a simple one, as virtualization environments consist of hypervisors, server platforms, storage, and networking. No common workload was devel-

S. Kounev et al., *Systems Benchmarking*, https://doi.org/10.1007/978-3-030-41705-5_12

oped to measure head-to-head performance of different hypervisors, as traditional benchmarks were intended to run on bare-metal servers, not within a virtualized environment. The traditional benchmark would be constrained by the resources provided to the VM rather than the entire virtualization solution. Likewise, running a traditional benchmark on the bare-metal server with the hypervisor present, even if possible, would not provide a meaningful measure of the hypervisor's performance.

In this chapter, we provide an overview of established benchmarks for evaluating the performance of virtualization platforms. We focus on the SPEC VIRT series of industry-standard benchmarks released by SPEC (SPEC VIRT_SC 2010, SPEC VIRT_SC 2013, and SPECvirt Datacenter 2020) while also considering the VMmark benchmark released by VMware.

12.1 SPEC Virtualization Benchmarks

In October 2006, SPEC established the OSG Virtualization Working Group to explore the possibility of developing an industry-standard benchmark to measure virtualization performance. The initial members of the working group were representatives from AMD, Dell, Fujitsu Siemens, Hewlett-Packard, Intel, IBM, Sun Microsystems, and VMware. In March 2007, the working group became a full subcommittee with its initial charter: *"The goal of the subcommittee is to develop a standard method for comparing virtualization performance of data center servers. The subcommittee's deliverable will be a benchmark that will model server consolidation of commonly virtualized systems including mail servers, database servers, application servers, web servers, and file servers. The benchmark will support hardware virtualization, operating system virtualization, and hardware partitioning schemes for server consolidation scenarios."*

12.1.1 SPEC VIRT_SC 2010

The first SPEC virtualization benchmark was SPEC VIRT_SC 2010 (Lange et al., 2012), released on July 14, 2010. It was designed to be a standard method for measuring a virtualization platform's ability to manage a server consolidation scenario and for comparing performance between virtualized environments. SPEC VIRT_SC 2010 measures the performance of the hardware, software, and application layers within a virtualized environment with a single hypervisor host. It uses three modified SPEC benchmarks as a workload to stress the system under test (SUT). Each of these three applications, SPECweb2005, SPECjAppServer2004, and SPECmail2008, drives predefined loads against the SUT. The benchmark requires the use of a set of clients to support the benchmark harness and drive the workloads on the SUT. SPEC VIRT_SC 2010 also supports the use of the SPECpower methodology to measure power usage during the benchmark. Results can be submitted in three categories: performance only (SPECvirt_sc2010), performance/power for the SUT (SPECvirt_sc2010_PPW), and performance/power for server only (SPECvirt_sc2010_ServerPPW).

12.1.1.1 Design

The benchmark suite consists of several SPEC workloads that represent applications that, at the time, industry surveys reported to be common targets of virtualization and server consolidation. The workloads were modified to match a typical server consolidation scenario's resource requirements for CPU, memory, disk I/O, and network utilization for each workload. The SPEC workloads used were:

- SPECweb2005: This workload represents a web server, a file server, and an infrastructure server. The SPECweb workload is partitioned into two virtual machines (VMs): a web server and a combined file server and backend server (BeSim). SPEC VIRT_SC 2010 uses the support workload from the original benchmark with a modified dataset.
- SPECjAppserver2004: This workload represents an application server and backend database server. Specifically, SPECjAppServer2004 was modified such that it created a dynamic load, increased the database scale, and decreased the session lengths. Additionally, the injection rate for queries varied significantly during the course of the benchmark. A sequence of different injection rates are cycled through during the course of a run, where each tile starts at a different injection rate in this sequence in order to create a "bursty" utilization pattern for the workload.
- SPECmail2008: This workload represents a mail server. Specifically, the harness employs the SPECmail IMAP component with new transactions.

SPEC VIRT_SC 2010 employs a fourth workload called SPECpoll developed explicitly for the benchmark. SPECpoll serves two functions: It sends and acknowledges network pings against an idle server VM in the 100% load phase to measure its responsiveness, and to all VMs in the 0% load phase (active idle) during power-enabled runs. SPECpoll ensures that sufficient resources are allocated to the idle server to function during the benchmark.

The four workloads described above run across 6 VMs in a set known as a "tile." Figure 12.1 shows the structure of the tile and its interaction with the SUT and client harness. A tile will deliver a specific amount of stress to the SUT and each workload must achieve a minimum level of Quality-of-Service (QoS). Scaling the benchmark on the SUT entails running an increasing number of tiles. Peak performance is reached at the point in which the addition of another tile fails to achieve the QoS criteria. The final benchmark result is the sum of the score achieved for each tile. The VMs of the same type were required to be configured identically across all tiles; only items like VM, IP, and NFS share names could be unique.

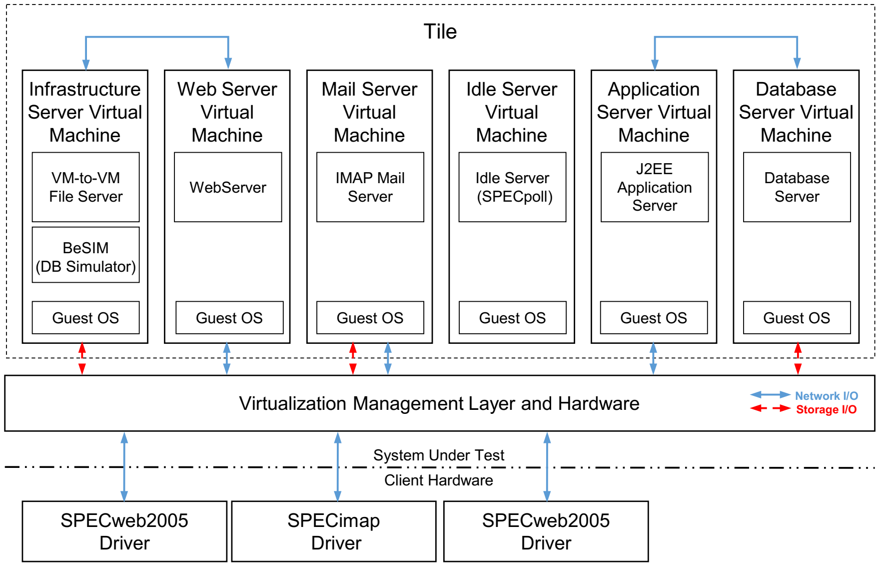

Fig. 12.1: SPEC VIRT_SC 2010 single-tile layout

To allow for fine tuning of the workload scaling, the final tile employed can be a "partial tile." This partial tile throttles down the workload drivers so that less stress is delivered to the SUT. The partial tile's score is proportionally scaled down.

12.1.1.2 The Need for a New Benchmark

SPEC VIRT_SC 2010 was released in July 2010 and several vendors published results on server configurations with 2–16 CPUs. Within a year of its release, a few trends became clear:

- As the number of tiles used by the benchmark increased, each injection rate used by the application server workload was exercised simultaneously by one of the tiles. In this case, the overall utilization across all tiles became more constant, removing the desired variability (i.e., burstiness) during the course of a run. As a result, the overall SUT CPU utilization could be driven to near 100%, which was not representative of real-world use cases.
- The workload levels for the tiles were too low. The initial utilization levels for the workloads within a tile were intended to be representative. However, as time went on, the amount of virtual resources needed for each VM decreased significantly. Within a year of the release, no VM needed more than a single vCPU, and the memory footprints for several VMs were less than 1 GB. This was significantly less than intended for the benchmark.

- A result of the reduced virtual resources needed per tile led to an increase in the number of tiles a SUT could support. Results that supported more than 17 tiles (102 VMs) on a 2P server became quite common. Feedback from customers expressed the large number of VMs reported in the results were unrealistic. Additionally, the benchmark harness struggled with the number of tiles being run on 8P and 16P configurations, topping over 100 tiles in some cases.

These concerns drove the SPEC Virtualization Subcommittee to develop a replacement benchmark.

12.1.2 SPEC VIRT_SC 2013

SPEC VIRT_SC 2013, released in May 2013, represents a significant update to its predecessor SPEC VIRT_SC 2010 which retired in February 2014. While still employing the concept of a tile for its basic unit of work, the design of the tile itself changed, with a set of tiles sharing a single database VM. Each workload was overhauled to increase its stress level and the idle server VM was replaced with a batch server VM with a new workload. Workload injection rates were made variable on the mail server in addition to the application server. Lastly, the web server introduced encryption in its web requests. All of the above updates are made for a much more robust tile.

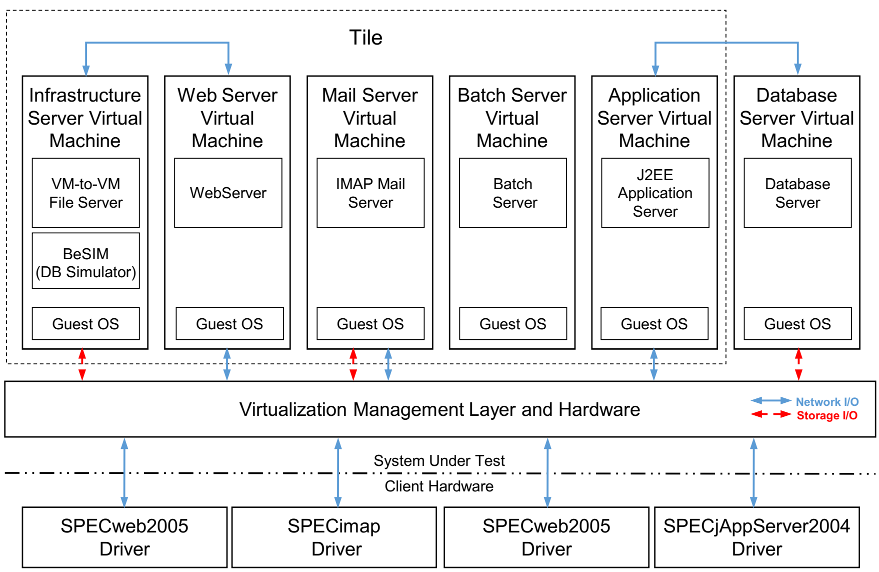

Fig. 12.2: SPEC VIRT_SC 2013 single-tile layout

Figure 12.2 shows the updated tile layout for SPEC VIRT_SC 2013. Notable enhancements for the new benchmark over its predecessor are as follows:

- **Shared database VM.** To ensure the presence of multi-vCPU VMs in the benchmark, the application server workload's database VM was pulled out of the tile. With SPEC VIRT_SC 2013, the application server VMs for every four tiles share the same database server VM, each with its own data within the database. This configuration requires the database VM to consume more resources to handle the increased database activity. Figure 12.3 shows a multi-tile configuration with the shared database VMs.
- **Web-server workload implemented SSL encryption.** To increase the utilization on the web-server VM, SSL encryption was introduced into the web-server workload. With the latest version of SPEC VIRT_SC 2013, SSLv3 and TLS 1.x encryptions are supported.
- **New batch workload.** To introduce more burstiness in the benchmark's workload profile, the new batch server VM replaces the idle server VM in the tile. The batch workload is based on one of the SPEC CPU 2006 training workloads, which runs 10 copies of the workload every hour and is idle for the rest of the hour. The 10 "jobs" must complete within 15 min, necessitating resource allocation from the SUT sufficient to satisfy this requirement. The batch jobs are staggered from one tile to another to avoid an unreasonable spike in server utilization at the beginning of the benchmark.
- **Mail server workload profile is now bursty.** Again, to add workload variation, the mail server workload now has a bursty profile akin to the application server's workload profile.

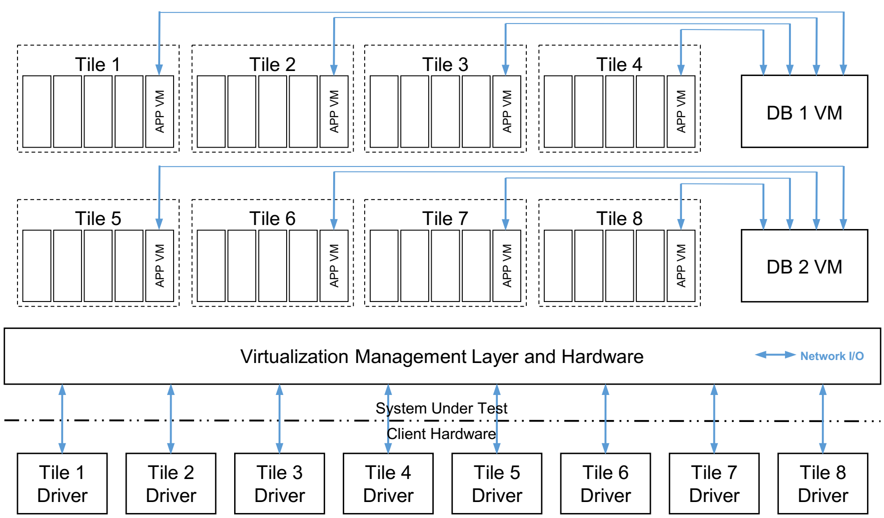

Fig. 12.3: SPEC VIRT_SC 2013 multi-tile layout

The operation of SPEC VIRT_SC 2013 is the same as SPEC VIRT_SC 2010. Scaling the workload is identical with the exception that only one database VM is added for every four tiles. All VMs of the same type must be identical across tiles. Partial tiles are allowed in SPEC VIRT_SC 2013. As with SPEC VIRT_SC 2010, SPEC VIRT_SC 2013 also supports measuring power performance.

SPEC VIRT_SC 2013 achieved its goal of a more intensive tile. At the time of the release, a 2P server that would require 28 tiles (168 VMs) to saturate it with SPEC VIRT_SC 2010 would need only 6.6 tiles (37 VMs) with SPEC VIRT_SC 2013. The benchmark continues to be active with new publications on the latest generations of hardware.

12.1.3 SPECvirt Datacenter 2020: The Next Generation

As the virtualization industry evolved, more complex environments became common. Configurations with multiple hypervisor hosts, shared networking, and common storage, all controlled by a central management application, are prevalent in modern data centers. Simple server consolidation is no longer the most interesting use case for businesses. The need for an industry-standard, multi-host virtualization benchmark became more urgent. Such a benchmark is needed to factor in common data center operations such as dynamic provisioning of VMs, automatically balancing resource utilization across multiple hosts, and introducing new physical resources into an environment. The subcommittee's charter was expanded to reflect this goal: *"The goal of the subcommittee is to develop standard methods for comparing virtualization performance of data centers. The subcommittee will develop and maintain benchmarks that represent typical virtualized infrastructure for various enterprise customer scenarios, such as server consolidation and multi-host virtualized environments."*

With this goal in mind, the first SPEC virtualization multi-host benchmark, SPECvirt Datacenter 2020, is planned to be released in 2020. It is a completely new virtualization benchmark designed to measure the performance of a different use case than SPEC VIRT_SC 2013, a multi-host virtualized data center. In addition to measuring traditional host capacity performance like SPEC VIRT_SC 2013, SPECvirt Datacenter 2020 also measures the virtual data center's ability to dynamically deploy VMs, balance workload levels across a cluster of hosts, and utilize new host resources that come online during run time. SPECvirt Datacenter 2020 also introduces preconfigured template VMs to simplify its setup and use.

12.1.3.1 Design

SPECvirt Datacenter 2020 uses five workloads contained within a 12-VM tile as its unit of work; see Figure 12.4 for the tile layout. Some of the VMs within a tile are deployed from a template during the course of the benchmark, while others

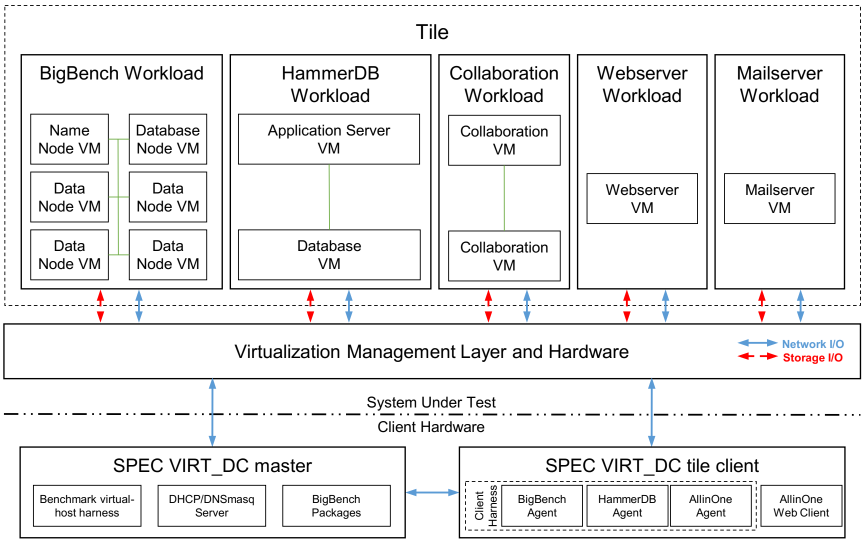

Fig. 12.4: SPECvirt Datacenter 2020 single-tile layout

are brought online from a powered off state. This behavior models the dynamic nature of a real-world virtualized data center environment. The workloads used in SPECvirt Datacenter 2020 are:

- a synthetic workload that simulates the stress of a pair of collaboration servers interacting with each other, modeled from real-world data. Two Collaboration Server VMs interact with each other to run this workload.
- a synthetic workload that simulates the stress of a web-server environment based on the SPEC VIRT_SC 2013 web-server workload. A Web Server VM runs on the SUT and interacts with a remote process that runs on the client.
- a synthetic workload that simulates the stress of an IMAP mail server application environment based on the SPEC VIRT_SC 2013 mail server workload. A standalone Mail Server VM runs this workload.
- a transactional database workload based on the HammerDB load testing and benchmarking tool.[1] The workload utilizes two VMs running on the SUT: an Application Server VM and a Database Server VM.
- a big data workload based on a modified version of BigBench that utilizes an Apache/Hadoop environment to execute complex database queries. The workload runs across six VMs on the SUT: a Name Node VM, a Database VM, and four Data Node VMs.

Unlike SPEC VIRT_SC 2013 where all of the tiles were started at the beginning of the benchmark and began their measurement intervals at the same time,

[1] https://www.hammerdb.com

SPECvirt Datacenter 2020 employs a more complex run profile as shown in Figure 12.5. During the first phase of the measurement interval, 1/4 of the hosts within the SUT are in maintenance mode. The benchmark will then bring the tiles' workloads online, starting or deploying the VMs used for each tile as needed. The ability of the SUT environment to deploy more rapidly and start a tile's workloads will be reflected as a longer active duration for that tile's measurement intervals. Once the target number of tiles for Phase 1 have been deployed, the SUT will remain at steady state until the end of the phase. Phase 2 begins with the activation of all of the hosts that were in maintenance mode during Phase 1. The SUT then will be able to take advantage of the newly available resources to balance the load on the systems. Phase 3 then sees the deployment of additional tiles to fully saturate the entire SUT environment.

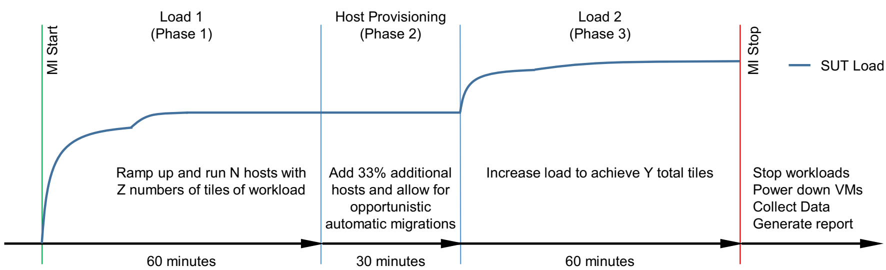

Fig. 12.5: SPECvirt Datacenter 2020 measurement interval (MI) profile

The performance score for each tile uses the aggregate throughput for each workload across all phases during which it was active. These throughput scores are normalized to reference values and then combined using a weighted geometric mean. The final score is the sum of the weighted means for all tiles. SPECvirt Datacenter 2020 also permits the use of a partial tile to allow for finer tuning of the saturation of a SUT. However, unlike SPEC VIRT_SC 2013, a partial tile for SPECvirt Datacenter 2020 consists of a subset of the workloads within a single tile; the partial tile will contain only the VMs needed to run the subset of workloads. The order of adding the workloads to a partial tile is fixed to ensure reproducibility of the benchmark results.

12.1.3.2 SPECvirt Datacenter 2020 Template VMs

SPECvirt Datacenter 2020 is a highly complex benchmark, even more so than the SPEC VIRT_SC benchmarks. In an effort to focus more attention on tuning the virtualization solution rather than tuning the application stacks within the tiles' VMs, SPECvirt Datacenter 2020 utilizes pre-built template VMs provided with the benchmark kit to create and deploy all of the VMs needed to build the benchmark harness (master controller and clients) and all of the tile's workloads. No modifications are needed—or allowed—within the VMs beyond the provided configuration and control scripts. This frees up the focus of performance tuning to be solely at the hypervisor and host level. At the time of the initial release, the template provides scripts for VMware's vSphere and Red Hat's RHV virtualization products. Other toolkits for different or newer versions of hypervisors are allowed but must be reviewed and approved by the SPEC Virtualization Subcommittee.

12.2 VMware's Virtualization Benchmarks

VMmark is a virtualization benchmark developed and maintained by VMware. It is intended for hardware vendors aiming to showcase the performance of their products using the VMware ESXi hypervisor. The first VMmark multi-workload server consolidation benchmark was released in August 2007 and measured the single-host performance in virtualized environments. Its successor, VMmark 2, was enhanced with multi-host virtual machine capabilities that addressed the increasing virtualization of bursty and heavy workloads, dynamic virtual machine and data store relocation, and the automation of many provisioning and administrative tasks across large-scale multi-host environments. In this new paradigm, some of the stress on the CPU, network, disk, and memory subsystems is generated by the underlying infrastructure operations. While still focusing on user-centric application performance, this benchmark also accounted for the effects of infrastructure activities on the overall platform performance.

12.2.1 The VMmark 3 Benchmark

Over the years, virtualization has become more common and end users are now considering highly scalable workloads and more complex online transaction processing (OLTP) workloads. VMmark 3 was developed to address this evolution as well as the additional challenges resulting from the increased load, frequency, and complexity of infrastructure operations.

The unit of work for a benchmark targeted at evaluating virtualized consolidation environments is generally defined as a collection of virtual machines executing a set of diverse workloads and the VMmark 3 benchmark follows the convention of its predecessor and refers to it as a tile. The total number of VMmark tiles (see Figure 12.6) a multi-host platform can accommodate provides a coarse-grained measure of that platform's consolidation capacity. This concept is similar to some server benchmarks, such as TPC-C (see Chapter 9, Section 9.3), that scale the workload in a stepwise fashion to increase the system load.

Tiles are relatively heavyweight objects that cannot capture small variations in platform performance. To address this, both the number of tiles and the performance of each individual workload determine the overall benchmark score. Each workload

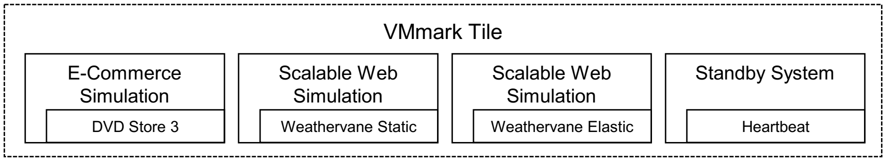

Fig. 12.6: VMmark tile

within a tile is constrained to execute at less than full utilization of its virtual machine. However, the performance of each workload can vary to a degree with the speed and capabilities of the underlying platform, for example, the addition of a fast disk array might result in disk-centric workloads producing a more favorable score. These variations can capture system improvements that do not warrant the addition of another tile. However, the workload throttling forces the use of additional tiles for large jumps in platform performance.

When a tile is added, the performance of the workloads in existing tiles might decrease. However, the aggregate score should increase if the system has not been overcommitted and the minimum Quality-of-Service (QoS) requirements are met. This results in a flexible benchmark metric that provides a measure of the total number of workloads that can be supported by a particular multi-host platform as well as the overall performance level within the workload virtual machines.

12.2.1.1 Workloads

A meaningful consolidation benchmark should be based on a set of relevant data center workloads. A survey of data center applications led to the inclusion of the workloads shown in Table 12.1 representing popular applications commonly run by VMware customers and a series of common infrastructure activities described later in this section. Rather than developing workloads from scratch, existing workloads and benchmarks were used where possible in order to reduce the implementation effort and to provide a well-understood foundation upon which to build.

Table 12.1: VMmark 3 workloads

Workload	Application	Virtual Machine Platform (CentOS 7.2 64-bit-—16 GB disk)	vCPU	Memory [GB]	Data Drive(s) [GB]
Scalable Web Simulation	Weathervane Auction Static	Auction Web (2)	2	8	–
		Auction App (2)	4	14	–
		Auction LB	2	4	–
		Auction MSQ	2	4	–
		Auction DB	2	8	20
		Auction NoSQL	2	16	100
Scalable Web Simulation	Weathervane Auction Elastic	Elastic Web (2)	2	4	–
		Elastic App (2)	2	8	–
		Elastic LB	1	4	–
		Elastic DB	2	8	–
E-Commerce Simulation	DVD Store 3	Database	8	32	100 + 250
		Web (3)	1	4	–
Standby System	None	Idle	1	2	–
Deploy	None	Infrastructure	4	8	100

Scalable Web Simulation Scalable web applications are used to provide a wide variety of services such as social networking and online auction websites. These applications typically will have a core application that implements the business logic, surrounded by a variety of support services such as load balancers, web servers, message servers, and databases. The application logic may be distributed among multiple independently deployed services. Each tier in a scalable web application deployment might stress different infrastructure resources. For example, the application servers might have high CPU demands, while the data services might place high demands on storage or memory.

Weathervane[2] is an application-level benchmark for virtual infrastructure and cloud performance tests. The Weathervane application, named Auction, is a scalable web application that implements a website for hosting real-time auctions. The Auction application uses a scalable architecture that allows deployments to be easily sized for a large range of user loads. A deployment of the application involves a wide variety of support services such as caching, messaging, NoSQL data store, and relational database tiers.

In VMmark 3, the Weathervane workload uses two independent instances of the Weathervane Auction application, a static instance and an elastic instance. The virtual machine configuration used in these application instances are shown in Table 12.1. Each instance includes:

- a load balancer running HAproxy 1.5.18,
- web servers running Nginx 1.12.0,
- application servers running Tomcat 8.5.13 and Java 1.8.0.121,
- a message server running RabbitMQ 3.5.3,
- a database running PostgreSQL 9.3, and
- a NoSQL data service running MongoDB 3.0.14.

In the static application instance, all of these services run on their own virtual machines. In the elastic application instance, the message server, load balancer, and NoSQL data service share a single virtual machine.

The static application instance, as its name implies, injects a relatively consistent load on the SUT. The elastic application instance, on the other hand, is both elastic and bursty. As in today's data centers it is increasingly common to have self-scaling applications that dynamically add and remove resources to meet demands, VMmark 3 takes advantage of Weathervane's elasticity-related capabilities to add and remove an application server and a web server throughout the benchmark run. This elastic component (along with the cyclical application profile generated by DVD Store 3 described below) allows VMmark 3 to represent more accurately today's bursty environments. The load for Weathervane is generated by a workload driver that simulates users interacting with the Weathervane Auction application. The load generated by each user is constant as long as the application can satisfy its quality-of-service (QoS) requirements. These QoS requirements specify the 99th-percentile response time for each operation as well as the required mix of operations performed

[2] Weathervane: https://github.com/vmware/weathervane

by all users. The performance metrics from Weathervane include the operation throughput, the average response time for each operation, and the percentage of each operation that completes within the response-time limits.

E-Commerce Simulation Databases running transactional workloads support a wide array of applications, typically as part of a multi-tier architecture. Databases tend to be resource-intensive and exercise most server and infrastructure components. In many cases, database systems also face strict response-time demands. Transaction processing often exhibits bursty behavior, resulting in widely varying resource demands over time. The ability of the underlying platform to support usage spikes is critical to maintaining acceptable performance.

DVD Store Version 3 (DS3)[3] is a complete online e-commerce test application with a back-end database component, a web application layer, and driver programs. The DS3 driver simulates users logging into a web server and browsing a catalog of products using basic queries. Users may select items for purchase, and then proceed to check out or continue shopping. Each web server communicates with a database server that maintains user accounts and inventory data.

The DS3 workload used in VMmark 3 utilizes four virtual machines in each tile, three web servers and one database server. The three virtual machines in the DS3 web tier (DS3WebA, DS3WebB, and DS3WebC) each run the Apache 2.4.6 web server, and the DS3 database tier runs the MySQL database. One of the web servers delivers a constant load to the database throughout each benchmark interval. The other two web servers deliver periodic load to the database during the benchmark interval to create a bursty overall load profile and varying resource demands. For VMmark 3, each web server is driven by 24 driver threads when active. The performance metric for this workload is the total number of transactions per minute. Minimum QoS metrics must also be met.

Virtual Machine Cloning and Deployment Creating a new virtual machine and installing a guest operating system and applications can be time-consuming. Using virtual machine cloning technology, administrators can make many copies of a virtual machine using a single installation and configuration process. Cloning, configuration, and deployment operations create bursty loads on platform resources, particularly the storage subsystem as the virtual machine files are copied.

The infrastructure workload: (1) clones the VMmark template virtual machine, (2) powers-on and pings the clone, (3) takes a snapshot, (4) performs a hot add of CPU and memory, (5) takes another snapshot, (6) creates a small MySQL database, (7) then reverts the snapshots, (8) pings the clone again, and (9) finally deletes the clone.

The benchmark then waits 40 s and repeats this process, continuing for the duration of the benchmark period. The number of concurrent clone and deploy operations increases with the number of tiles and the number of hosts in the benchmark cluster. The performance metric used is the number of clone and deploy operations per hour.

[3] DVD Store Version 3: http://github.com/dvdstore/ds3

Dynamic Virtual Machine Relocation Between Servers Live migration technology such as VMware vMotion leverages the complete virtualization of servers, storage, and networking to move an entire running virtual machine seamlessly from one server to another. During a vMotion operation, the active memory and precise execution state of a virtual machine are rapidly transmitted over a high-speed network from one physical server to another and access to the virtual machine's disk storage is instantly switched to the new physical host. This transition can result in bursty loads on platform resources, particularly the networking subsystem. VMmark mimics the manual relocation of a virtual machine, which can be a common task performed by an administrator.

This infrastructure workload acts on one of the AuctionMSQ virtual machines selected in a round-robin fashion from among all the tiles. A destination host is selected at random from among all hosts in the benchmark cluster (other than the virtual machine's current host). The virtual machine is moved to the destination host, left there for 2 min, and then returned to its original host. VMmark then waits another 2 min and repeats this process, continuing for the duration of the benchmark period. The number of concurrent relocation operations increases with the number of tiles and the number of hosts in the benchmark cluster. The performance metric used is the number of relocations per hour.

Dynamic Virtual Machine Relocation Across Storage Live migration of virtual machine disk files across or within storage arrays enables enormous flexibility for storage maintenance, upgrades, and load balancing. Storage relocations can create bursty loads on platform resources, particularly the storage subsystem.

In this infrastructure workload, VMmark relocates a virtual machine's disk files to a maintenance partition, then returns them to their original location. This round-trip approach models an administrator temporarily evacuating a disk partition, performing maintenance on the storage system, and then returning the system to its initial state.

This infrastructure workload acts on one of the standby server virtual machines selected in a round-robin fashion from among all the tiles. The virtual machine's files are moved to the maintenance partition, left there for 2 min, and then moved back to their original location. VMmark then waits another 2 min and repeats this process, continuing for the duration of the benchmark period. The number of concurrent storage relocation operations increases with the number of tiles and the number of hosts in the benchmark cluster. The performance metric used is the number of relocations per hour.

Simultaneous Server and Storage Virtual Machine Relocation The live migration of virtual machines simultaneously across both servers and storage (vMotion without shared storage) allows even more flexibility than either capability alone. This infrastructure workload produces a combination of the infrastructure loads created by the individual operations.

In this infrastructure workload, VMmark uses vMotion to relocate a virtual machine while simultaneously invoking the storage relocation of the same virtual machine's disk files to a maintenance partition. After two and a half minutes, the virtual machine is returned to its original host and the files are returned to their original location. VMmark then waits another two and a half minutes and repeats the process. This workload models an administrator temporarily evacuating a host and disk partition, performing maintenance on the host and/or storage system, and then returning the system to its initial state.

This infrastructure workload acts on one of the DS3WebA virtual machines selected in a round-robin fashion from among all the tiles. The number of concurrent relocation operations increases with the number of tiles and the number of hosts in the benchmark cluster. The performance metric used is the number of relocations per hour.

Automated Load Balancing Automatically balancing resource demands among multiple physical servers using technology such as VMware's Distributed Resource Scheduler (DRS) has become a fundamental part of modern virtualized data centers. Intelligently allocating and balancing resources allow the underlying platform to respond effectively to bursty-load conditions even when utilizations are high.

VMmark requires DRS to be enabled and running at (or above) a specific level to ensure that rebalancing occurs in a timely manner when utilizations are high. This should improve overall performance by addressing load imbalances occurring during the benchmark interval.

12.2.1.2 Scoring Methodology

VMmark 3 aggregates the throughput metrics of all application and infrastructure workloads to create a single overall benchmark metric that can be used to quickly compare different platform configurations. If any of the workloads within any tile fails to run, produces errors during a run, or fails its minimum QoS requirement, the entire VMmark run is considered to be incompliant. After the completion of a compliant VMmark benchmark run, each individual application and infrastructure workload reports its relevant performance score (see Table 12.2). These scores were collected every 60 s during the standard 3 h run resulting in a series of meaningful numbers for each of the workloads. VMmark 3 automatically generates graphs of key performance metrics for each workload as shown in Figure 12.7.

The scores of the application and infrastructure workloads are computed and aggregated separately based on the geometric mean, and the final benchmark metric is the weighted arithmetic mean of the scores (geometric means) for the application-workload component (80%) and the infrastructure-workload component (20%). These weights were chosen to reflect the relative contribution of infrastructure and application workloads to overall resource demands.

The VMmark 3 metric shows the virtualization overheads of the individual workloads as well as the scalability of the entire system. Therefore, results for multi-tile runs are reported as the aggregate score for all tiles, the individual scores for each of the tiles, and the scores for the workloads within the tiles as well as the individual scores for each infrastructure workload. If two different virtualization platforms

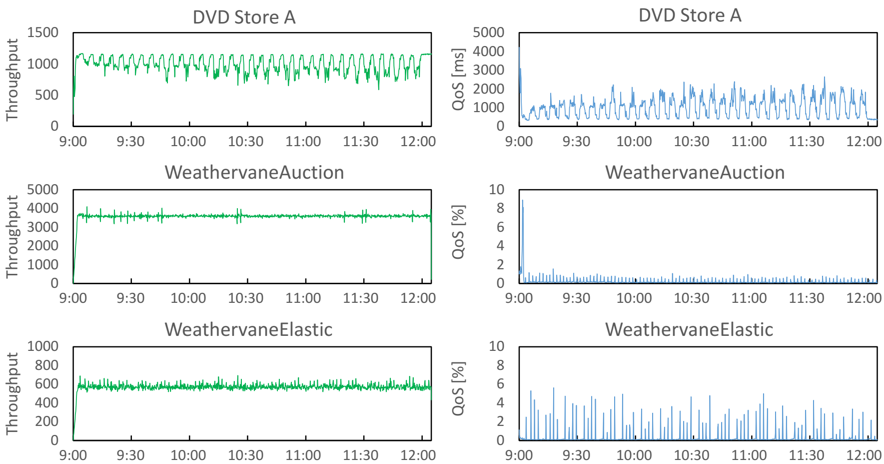

Fig. 12.7: Throughput and Quality-of-Service (QoS)

Table 12.2: Individual VMmark workload scores

Workload name	Applications(s)	Scores
Weathervane static	Auction	Operations/s
Weathervane elastic	Auction	Operations/s
DS3WebA	Apache, MySQL	Transactions/min
DS3WebB	Apache, MySQL	Transactions/min
DS3WebC	Apache, MySQL	Transactions/min
Standby server	None	None
Clone and deploy	Infrastructure	Deployed VMs/h
vMotion	Infrastructure	VM migrations/h
Storage vMotion	Infrastructure	VM migrations/h
XvMotion	Infrastructure	VM migrations/h
Distributed Resource Scheduler (DSR)	Infrastructure	None

achieve similar VMmark scores with a different number of tiles, the score with the lower tile count is generally preferred. The higher tile count could be a sign that the underlying hardware resources were not properly balanced. Studying the individual workload metrics is suggested in these cases.

12.3 Concluding Remarks

We provided an overview of established benchmarks for evaluating the performance of virtualization platforms. We focused on the SPEC VIRT series of industry-standard benchmarks (SPEC VIRT_SC 2010, SPEC VIRT_SC 2013, and SPECvirt Datacenter 2020) while also considering the VMmark benchmark by VMware. The discussed benchmarks provide users with the capability of measuring different virtualization solutions on either single-host or multi-host platforms, using workloads and methodologies that are designed for fair comparisons. Great effort was taken to ensure a wide range of virtualization solutions can utilize the benchmarks and they have been used by hardware and software vendors to showcase, analyze, and design the latest generations of virtualization products.

References

Lange, K.-D., Schmidt, D., Bond, A., and Roderick, L. (2012). "SPECvirt_sc2010—Driving Virtualization Innovation". In: *Proceedings of the 3rd ACM/SPEC International Conference on Performance Engineering (ICPE 2012)*. (Boston, MA, USA). ACM: New York, NY, USA, pp. 251–252 (cited on p. 268).