FUJITSU

WHITE PAPER

Version 1.1
February 2011

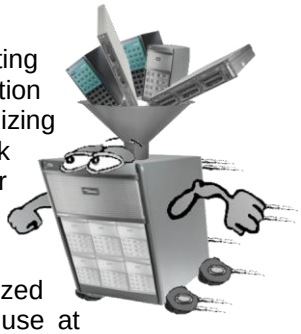
Benchmark Overview VMmark V1

Pages 6

Abstract

Conventional benchmarks are less suited for the assessment of virtualized operating systems and applications, which is why special virtualization benchmarks exist. In addition to the benchmark called "vServCon" used by Fujitsu Technology Solutions for sizing measurements which compares PRIMERGY systems to one another, the benchmark developed by VMware called "VMmark V1" enabled for a long time a multivendor comparison for highly-optimized configurations based on the hypervisor solutions from VMware, until it was replaced at the end of 2010 by “VMmark V2”.

This document describes the problems concerning benchmarks for virtualized environments as well as the fundamentals of the "VMmark V1" benchmark and its use at Fujitsu Technology Solutions.



Contents

Introduction

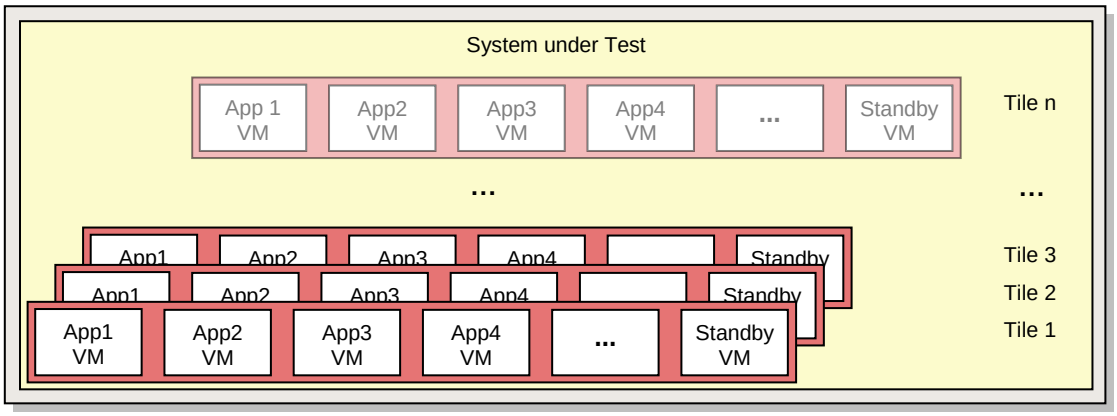
Server virtualization is an increasingly deployed scenario in the implementation of IT infrastructures. On the one hand, virtualization enables more efficient hardware to be used in an optimal way, and on the other hand the dependency on hardware specifics is being reduced. Contrary to the classic server with only one operating system and its applications, with virtualized environments, several operating systems and applications are run in parallel creating heterogeneous environments on one server. The handling of virtual machines is implemented by a virtualization layer, also called hypervisor.

Classic benchmarks are scarcely suited to measure and assess the performance capability of virtualized environments. For this purpose, it is necessary to utilize the hardware resources of a server with simultaneously working virtual machines with different workloads. The one aim of virtualization benchmarks is server consolidation. In this case, the throughput of a set collection of virtual machines is maximized on a single virtualization host by means of suitable replication. vConsolidate (Intel), VMmark V1 (VMware) and SPECvirt_sc2010 (SPEC) fall under this category. Fujitsu Technology Solutions carries out scaling measurements of virtual environments using its internal benchmark "vServCon" (based on ideas from "vConsolidate"). The other aim of virtualization benchmarks is data center operations. A server consolidation scenario for several virtualization hosts is assumed in this case. In addition to the throughputs of the virtual machines, the benchmark metric then contains ratios that reflect the efficiency of typical data center operations, such as the relocation of virtual machines. These benchmarks include VMmark V2 (VMware).

For a virtualization benchmark to fulfill its objective, it must map the real world of a data center regarding server consolidation; in other words it must consider existing servers with those application scenarios that are normally virtualized. These servers have weak utilization levels and the aim is thus to consolidate as many of them as possible as virtual machines (VMs). Therefore, such a benchmark must assess for a virtualization host both the suitably determined overall throughput across the various application VMs as well as the number of efficiently operable VMs.

The following solution concept has been established for these two objectives: a representative group of application scenarios is selected in the benchmark. They are started simultaneously as a group of VMs on a virtualization host when making a measurement. Each of these VMs is operated with a suitable load tool at a defined lower load level. All known virtualization benchmarks are thus based on a mixed approach of operating system and applications - plus usually an "idle" or "standby" VM which represents the inactive phases of a virtualization environment and simultaneously increases the number of VMs to be managed by the hypervisor. The term "tile" is the name for such a unit of virtual machines.

It must be possible to increase this well-defined load created by this group of virtual machines on a step-by-step basis until the considered system has reached its performance limit. The following illustration shows the growth of VM load on a “system under test” by operating several tiles.



An application is executed in each virtual machine where the applications are put under stress via established benchmarks. If necessary, there may also be further infrastructure components. All the individual results are then suitably summarized in one overall result. This score is an indication for the performance capability of a virtualized environment.

VMmark V1

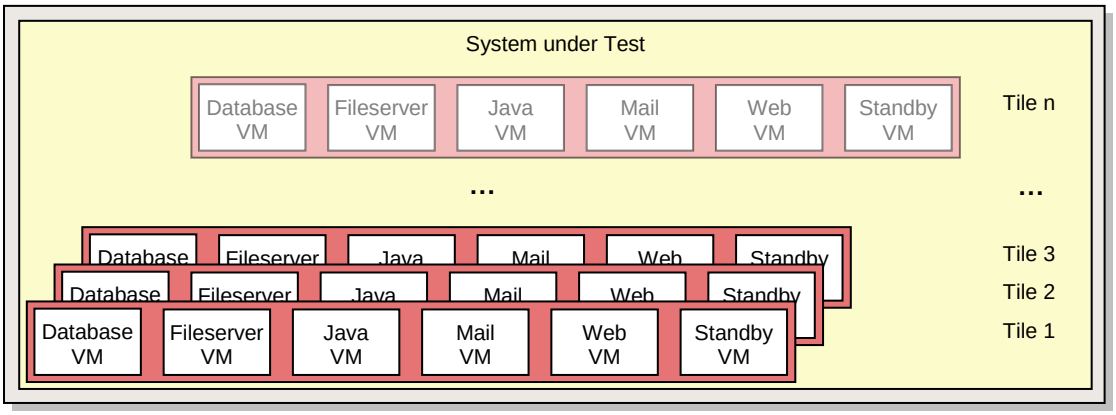
VMmark V1 is a benchmark developed by VMware to compare server configurations with hypervisor solutions from VMware regarding their suitability for server consolidation. In addition to the software for load generation, the benchmark consists of a defined load profile and binding regulations. For a long time VMmark V1 was the only established virtualization benchmark which enabled a multivendor comparison. Benchmark results achieved with VMmark V1 could be submitted to VMware and were published on their Internet site after a successful review process. Today, VMmark V1 is only available for academic use and has been replaced by VMmark V2.

VMmark V1 Benchmark

VMmark V1 is not a new benchmark in the actual sense. It is in fact a framework that consolidates already established benchmarks, as workloads, if necessary in modified form in order to simulate the load of a virtualized consolidated server environment. Five proven benchmarks, which cover the application scenarios database server, file server, application server, mail server and web server, have been adapted for use with VMmark V1.

Application scenario	Load tool
Database server	Sysbench
File server	Dbench
Java application server	SPECjbb2005
Mail server	LoadSim 2003
Web server	SPECweb2005
Standby server	-

Each of the five application scenarios is assigned to one dedicated virtual machine (VM). Then add to these a sixth VM called the "standby server". These six VMs form a "tile". Because of the performance capability of the underlying server hardware, it is usually necessary to have started several identical tiles in parallel as part of a measurement in order to achieve a maximum overall performance.

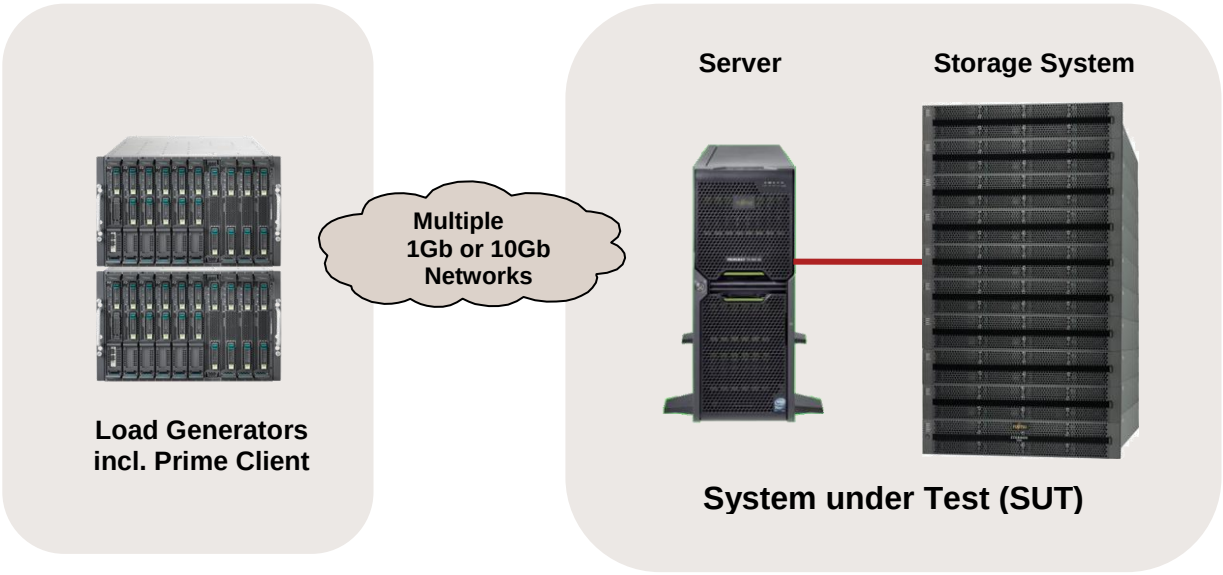


VMmark V1 requires external load generators: exactly one client system per tile. Load generators and "system under test" are connected via a suitable number of networks.

The execution of the individual load tools is controlled by one of the load generators, the so-called prime client. The prime client monitors the measurements and collects the individual performance data of the VMs.

VMmark V1 Environment

The measurement set-up is symbolically illustrated below:



VMmark V1 Score

The result of VMmark V1 is a number, known as a "score", which provides information about the performance of the measured virtualization solution. The score reflects the maximum total consolidation benefit of all VMs for a server configuration with hypervisor and is used as a comparison criterion of various hardware platforms.

This score is determined from the individual results of the VMs. Each of the five VMmark V1 application scenarios provides a specific benchmark result in the form of application-specific transaction rates for each VM. In order to derive a normalized score the individual benchmark results for one tile are put in relation to the respective results of a reference system. The resulting dimensionless performance values are then averaged geometrically and finally added up for all VMs. The outcome is the final VMmark V1 score.

In addition to the actual score, the number of VMmark V1 tiles is always specified with each VMmark V1 score. The result is thus as follows: "Score@Number of Tiles", for example "24.20@17 Tiles".

These results are listed with VMware grouped depending on the number of CPU cores of the system and sorted according to the score.

VMmark V1 scores can only be compared with each other and particularly the scores of each workload contained in VMmark V1 cannot be considered separately and interpreted or compared with the scores of the original benchmarks. A comparison of the results between VMmark V1 and VMmark V2 is also not possible; with the same performance in the virtualization environment the scores and numbers of tiles for VMmark V2 are considerably lower.

VMmark V1 Load Profile and Run Rules

The VMmark V1 rules define a standardized benchmark environment and the resources and software versions of the operating system and the application software are specified precisely. The compilation of guest operating systems and applications and their specified resource requirements are to be understood as a representative selection for a complex virtualization environment. Even if there are newer and perhaps even more high-performance versions, continuity of the load profile must be maintained for as long a time as possible in order to ensure comparability.

The following profile is used for VMmark V1.

Resource	Database Server	File Server	Java Server	Mail Server	Web Server	Standby
# vCPU	2	1	2	2	2	1
Memory	2 GB	256 MB	1 GB	1 GB	512 MB	256 MB
OS	Linux SUSE SLES 10 64-bit	Linux SUSE SLES 10 SP2 32-bit	Windows Server 2003 R2 EE SP2 64-bit	Windows Server 2003 R2 EE SP2 32-bit	Linux SUSE SLES 10 SP2 64-bit	Windows Server 2003 R2 EE SP2 32-bit
Application	MySQL 5.0	dbench (modified)	BEA JRockit 5.0	Exchange Server 2003	Apache	
Benchmark	Sysbench	dbench (modified)	SPECjbb2005 (modified)	LoadSim 2003	SPECweb2005 (modified)	
Disk Subsystem	10 GB	8 GB	8 GB	24 GB	8 GB	4 GB

The complex VMmark V1 rules are restricted to optimizations and tuning for comparison reasons; all permitted changes in the standard configuration must be documented when the score is submitted.

The disk subsystem can be individually configured and optimized according to the rules with not only the logical design regarding size and RAID level, but also the physical implementation. Since a local disk subsystem is not a sensible solution for virtualization and consolidation, a SAN-based disk subsystem is used. The performance of the disk subsystem has a direct influence on the VMmark V1 score; if there is a bottleneck an optimal score cannot be attained.

Neither should the main memory of the "System under Test" be a bottleneck for the VMmark V1 measurement. Therefore, an adequate quantity is equipped so that with the used tile number the lowest possible number of swap activities takes place at host level. CPU and network resources must also be aligned with the number of operated VMs and their load.

In general, a virtualized environment, as VMmark V1 describes, with a large number of virtual machines and different guest operating systems with high application loads, is an extremely complex system which relies on the optimal interaction of all components. Bottlenecks and also excess resources can have a negative effect on the overall score. The configuration must thus be modified and optimized for each hardware platform. Information about the individual VMmark V1 configurations is in the Performance Reports for PRIMERGY systems and in the score report for the VMmark V1 benchmark.

©Fujitsu Technology Solutions 2011

Literature

General information about products from Fujitsu Technology Solutions
http://www.ts.fujitsu.com
General information about the PRIMERGY product family
http://www.primergy.com
PRIMERGY Benchmarks – Performance Reports and Sizing Guides
http://www.ts.fujitsu.com/products/standard_servers/primergy_bov.html
VMmark V1
VMmark V1 http://www.vmware.com/products/vmmark/v1/overview.html
VMmark V1 Introduction http://www.vmware.com/pdf/vmmark_intro.pdf
VMmark V1 Results http://www.vmware.com/products/vmmark/v1/results.html
VMmark V2
VMmark V2 http://www.vmmark.com/
Benchmark Overview VMmark V2 http://docs.ts.fujitsu.com/dl.aspx?id=2b61a08f-52f4-4067-bbbf-dc0b58bee1bd
SPECjbb2005
http://www.spec.org/jbb2005/
SPECweb2005
http://www.spec.org/web2005/
SysBench
http://sysbench.sourceforge.net/
dbench
http://samba.org/ftp/tridge/dbench/README
LoadSim 2003
http://technet.microsoft.com/en-us/library/cc164328(EXCHG.65).aspx
SPECvirtualization
http://www.spec.org/specvirtualization/

Contact

PRIMERGY Hardware

PRIMERGY Product Marketing

mailto:Primergy-PM@ts.fujitsu.com

PRIMERGY Performance and Benchmarks

PRIMERGY Performance and Benchmarks

mailto:primergy.benchmark@ts.fujitsu.com

Delivery subject to availability, specifications subject to change without notice, correction of errors and omissions excepted.
All conditions quoted (TCs) are recommended cost prices in EURO excl. VAT (unless stated otherwise in the text). All hardware and software names used are brand names and/or trademarks of their respective holders.

Copyright © Fujitsu Technology Solutions GmbH 2011

Published by department:

Enterprise Products
PRIMERGY Server
PRIMERGY Performance Lab
mailto:primergy.benchmark@ts.fujitsu.com

Internet:
http://www.ts.fujitsu.com/primergy

Extranet:
http://partners.ts.fujitsu.com/com/products/servers/primergy