

MYSQL DATABASE PERFORMANCE UNDER VIRTUAL MACHINES ENVIRONMENTS

Bartłomiej Kacprzak¹, Piotr Kolasiński¹, Piotr Goetzen¹,
Agata Skowrońska-Kapusta¹, Alina Marchlewska¹

¹IT Institute, Academy of Management, Lodz, Poland
*bercik1337@gmail.com, ptkola@gmail.com, goetzen@swspiz.pl,
akapusta@swspiz.pl, amarchlewska@swspiz.pl*

Abstract

Some technical differences between virtualization platforms will be shown. Using the MySQL database we present the efficiency tests. Different hardware platforms were used: PC Intel and different mainframe computers (z/9'series). Usefulness of virtualization techniques was also shown. An example of implementation of MySQL database was used to show the selected areas of usage of virtualization environment. In this document we also show selected architectural and performance comparisons of virtual machines set up on PC (x86/x64) and Mainframe (z/9) platforms.

Key words: virtualization, database, Mainframe, Linux

1 Introduction

Today's world is facing flood of information. Systems process more and more data because of growing population and ease of access to Internet. Increasing number of customers forces service providers to add more servers and storage, build new facilities and lease greater Internet connection. All that results in greater electricity bills, more space to equip and higher maintains costs. This leads to the conclusion that thankfully to virtualization costs of implementation, maintenance and usage are less in comparison to the standalone hardware/software systems. This technology can reduce hardware needed for specific job. Because of system consolidation, servers utilize most of system resources which leads to less hardware needs.

Generally speaking virtualization is the mechanism that separates software (for instance whole operating system) from hardware. Thanks to that “encapsulation” it is possible to run multiple operating systems (OS) at this same time on one physical machine.

The very concept of virtualization is almost half century old. First working virtual machine was created in early 60's. Even back at that time system were fully functional when ran at the same time. Since that day, personal computers gained great amount of computing power. Thanks to that, anyone can run Virtual Machines (VM) on his PC. Over past few years significant growth of virtualization market can be seen. Reason for that mostly is previously mentioned computing power. On software side of things first steps were made by MIT (Massachusetts Institute of Technology) students. They created system called CTSS (Compatible Time-Sharing System) the ancestor of today's z/VM. PC's virtualization market was pushed forward by VMware company. Started in 1998, their products really popularized virtualization on small and medium sized corporations as well as home oriented solutions. Their products are also used as testbeds in professional computer systems. The university labs often use the VM to provide the laboratory sets.

2 Comparison of architectural solutions

In this section comparison of some selected software and hardware components are presented and compared [1-3]. Reason for this is because even though both systems are doing the same thing (virtualize resources and operating systems) and as for customer there is no visible difference in service itself - they use different techniques, configurations, technologies and ideas for purpose of virtualization. Before covering details it is important to note few significant differences between those systems.

Disk concept - On everyday basis we get along usually with one disk systems. In many cases, disk is used as one big place to store data. Sometimes the administrator of a system splits available space on more than one partition to help organize files and improve data security. When adding additional disk to system, the most common scenario is, that it is either a new main device or allocated as one/two partition for additional storage/backup. Since the following document is enterprise-oriented it is important to use highest I/O possible. Easiest way to do this is to use RAID arrays. Systems with multiple operational disk can significantly improve performance. As for virtualization itself, guest systems can use virtual disks (details below). From its point of view - it is a standard SATA disk drive. It can be formatted, repartitioned etc. VM user cannot notice, what type of disk he uses. It is also impossible to determine the specific location of virtual disk on physical disk without using trick like partitioning.

On the other hand, z/VM as well as whole Mainframe system uses DASDs (Direct Access Storage Device) instead of SATA drives. For systems of that scale it is natural to use multiple disk arrays. The main difference between these two types of disk management is that under z/VM it is possible to speci-

fy where, physically on a disk, data will be located. Thanks to that type of allocation important data that is needed to be quickly accessible can be located on the very beginning of a disk, and less crucial files back at the end of disk. Even though physical disks are stored in RAID arrays and grouped in logical partitions, by default system, and all users "see" this disk as physical device. However the system administrator and operators can invoke command to list *real* physical devices and manage them.

Virtual Disks (VD)

As name suggests, those are disks used by Virtual Machine. Approach to virtual disks depends on platform and software used for virtualization.

On PC systems running VMware software three main types of VDs can be marked out:

- fixed size - during VD creation system allocates whole space allowed for disk. That can prevent fragmentation, but usually results in huge space waste,
- variable size - VD size takes only as much space on host system as it needs. When more files are added, it automatically grows. Mostly common used option. When data is removed from disk, shrunk operation can be performed to trim file size back to minimal size. Huge fragmentation occurs over time, often resulting in performance drops,
- physical disk - whole physical disk, or logical partition can be "forwarded" to VM.

Both fixed and variable sized disks are seen by host computer as ordinary files.

Mainframe systems present different approach. Instead of messing around with files, system operator can allocate disk by defining where disk starts and how huge it is. Disk size is always fixed, and thanks to units used for allocation (cylinders) it is possible to physically "place" VD on specific part of physical disk making so-called mini-disk on cylinder boundary. Because I/O operations in mainframe systems are based on dedicated channel program, where cylinder number is part of disk addressing – operation for minidisk is just the translation of cylinder by adding fixed offset value. Each channel program is under control of Control Program (CP), which is a kind of supervisor, so before virtual machine read/writes data – CP translates address using arithmetic operations.

Networking

Both systems support IPv4 and IPv6 address spaces. Of course specifications are little different. VMware allows administrator to use host system network adapter in three main types of configuration:

- **Bridged** - it behaves as a ordinary NIC plugged into our existing network,

- **NAT** - creates internal network (inside VMware) with own DHCP server. Host system behaves as gateway,
- **VMnet** - similar to NAT except to the fact that there is not DHCP service, no internet connection (unless manually configured) is provided.

In Mainframe Systems there are a few different networking devices:

- **Hipersockets** – which are just shared by hardware memory pools visible in systems as networking cards – in that case communication is limited to one physical box,
- **Real Ethernet Devices – OSA (Open System Adapters)** – dedicated standalone card with own processor (currently PowerPC) and operating system (Linux). This kind of solution gives independency of main operating system and offloads TCP/IP and ARP traffic,
- **Channel To Channel Connection (CTC)** – devices based on direct wired (today using fiber link) connection between two mainframe boxes – not popular today, still exists for historical reasons.

z/VM emulates all of these devices or has possibility to connect directly such a device to one virtual machine (this kind of connection is named “attached”).

When system emulates network card (NIC), administrator has possibility to connect it to Virtual Switch fully supporting VLAN tags. This is only part of CP code, transferring from buffer of one virtual machine (Virtual NIC) to buffer of second virtual machine or to buffer of real device if destination is outside z/VM virtual machines.

z/VM forces administrator to manually configure all settings and features. Interesting feature is, that the whole TCP stack is owned and controlled by z/VM user named TCPMAINT. After logging in to the system administrator can manage almost all network-related tasks such as:

- allow/forbid specified services to be used,
- create virtual NICs,
- create virtual switches,
- attach/detach specified machines to/from switches,
- control Internet access.

Thanks to that, it is easier to control network environment. All features are located in one place.

Other devices

VMware supports many useful features like USB hub. This "forwards" USB devices from host system directly to VM. It is possible to use a flash drive/external hard drive or non disk devices like printers, scanners or even additional network adapters.

z/Series systems don't have that option, but there is really no need for it anyway. However, besides DASDs, they can share tape drives, readers, punchers etc. z/VM has also possibility to share dedicated cryptographic devices (cards with Linux operating system working on PowerPC processors).

Memory

VMware uses part of memory visible in operating system. That means, that VMware as “normal” operating system possesses maximum addressability fixed by operating system design for processes. It means also that all underground activities in operating systems (paging, process priority etc), have an influence on emulation of virtual machine,

z/VM uses memory as “low level” operating systems, but because of architecture of Mainframe has better inspection of memory access. In special cases – it is possible to cooperate between CP and second level operating system (like Linux). In that case page which was swapped out from memory in Linux – is also swapped out from Z/VM (not always true in VMware)

Processor execution

The most significant difference between the two virtualization systems exists in execution of some instructions – especially for privileged instructions. In VMware – executing privileged instruction for the first time will change its operational code in memory cells so next time virtual machine will emulate memory cell address automatically. This means each time supervisor has to trap and control all privileged instructions. In z/VM special instruction – SIE (Start Interpretative Execution) exists – which executes code directly in hardware unless special events will be trapped. That is why most of code is running efficiently.

Bus speed

In PC environment there are a few different internal buses: PCI, PCIe, FSB. In mainframe architecture there is one internal bus which uses the copper connections and a few external buses to external devices which use fiber optic technology. One – older – named ESCON (Enterprise Systems Connection) allows the transfer of data at the speed of maximum 20 MB/s (in practice – 17 MB/s). Second one – currently more popular – FICON (Fiber Connectivity) that can transfer the data at the speed of 1, 2, 4 and 8 Gb/s.

3 Efficiency tests

For purpose of tests, we created the database with 10 million clients. (10^7) Data was imported to MySQL database under z/VM Linux and Linux running in VMware on PC. Database table contains following columns:

Name, Last Name, PESEL number¹, Comment1, Comment2, Account number, Balance, Last IP address. Comment1 and Comment2 are 75 long characters fields for additional comments/notes.

Test 1

In order to perform reliable performance test, both Mainframe and PC systems were tuned, to have most matching configuration possible. Each VM had

- 1GB of RAM
- 1 CPU assigned to it

Test was performed during night hours, to ensure that there was minimal system load from other applications and services.

VMware host system was running on Intel Quad Core Q8200 (2,33GHz clock), 4GB RAM DDR2 and four 3.5" Hitachi 250 GB SATA2 disks in RAID5 setup. z/VM was hosted on IBM z/9 machine running at 1,4GHz. Both environments used Linux Debian as a operating system.

To minimize human factor that could ruin test results bash script was written. It's a configurable file, that invokes specific commands one after another. This solutions ensures that there will be no postpones between commands, so both systems will get this exact same queries at this exact same time.

The concept of the benchmark test is to send query to the database, wait for results, write them to a file and execute next query. Process Continues until query list is empty.

The list contains queries for:

- exact name,
- names starting with pattern,
- names ending with pattern,
- exact last name,
- last names starting with pattern,
- last names ending with pattern,
- exact PESEL numbers,
- PESEL patterns starting with pattern,
- PESEL patterns ending with pattern,
- exact balance,
- balance lower than,
- IP addresses starting with pattern,
- IP addresses ending with pattern,

¹ PESEL - 11 digit, equivalent of Social Security Number

- creating index for last name column,
- exact last name,
- last names starting with pattern,
- last names ending with pattern,
- creating index for name column,
- exact name,
- names starting with pattern,
- names ending with pattern.

112 queries were defined in a batch file. This allows to observe the execution time and compare results.

Majority of above lines (excluding index creating) were read only type. There was no insert or update involved, so the most important factors were:

- disk read speed,
- cache.

All other component like CPU and memory throughput had minimal impact on the performance.

After each iteration, time consumed for task was added to an end of text file. This exact same task was replicated on concurrent system. Output data was stripped out from redundant lines and processed further.

Test 2

Second test is exact copy of first one, however the difference is in VM configuration itself.

We used the following configuration:

- 1GB of RAM,
- 2 CPUs assigned to the task.

All details regarding database configuration, data within it and process of testing remains unchanged.

4 Results

After performing the tests, all output results were written to files. Graphs below (Fig. 1 and 2), help with visualizing results. Y-axis corresponds to time (minutes : seconds), X-axis corresponds to query number.

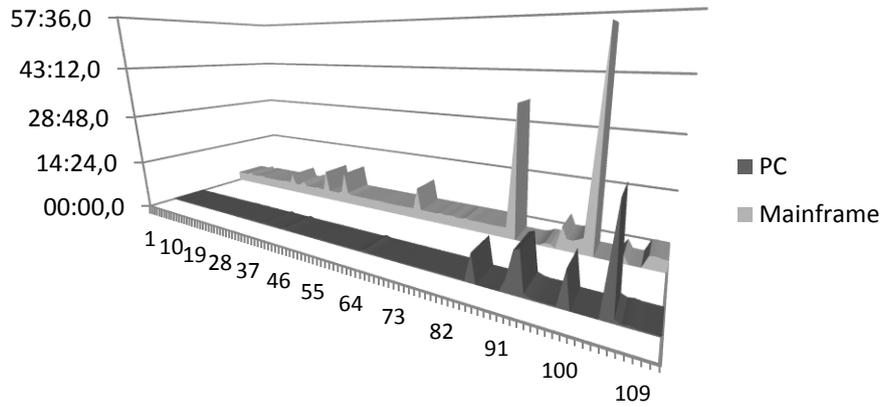


Figure 1. PC and mainframe systems results of test number 1. Less = better efficiency

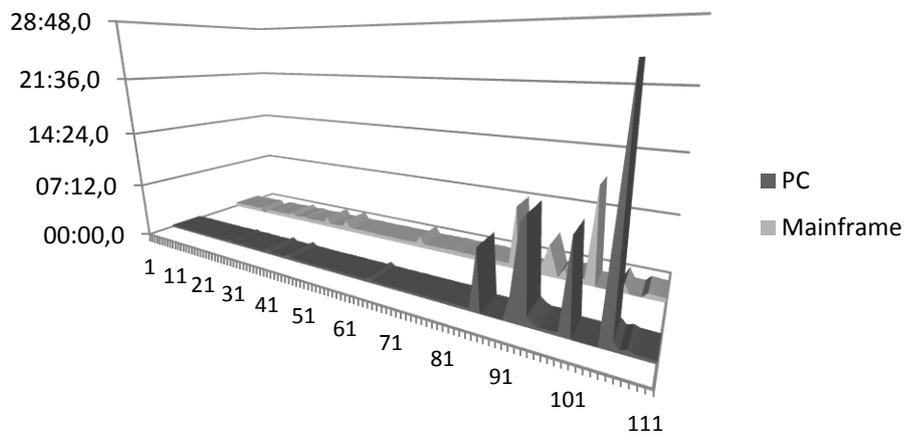


Figure 2. PC and mainframe systems results of test number 1. Less = better efficiency

All graphs show time needed to finish each query. As presented above, PC computer performs better in comparison to mainframe. The differences between the performance is smaller in the case of greater hardware configuration of Mainframe. In the case of indexed irregular expressions (last names) mainframe performs better

5 Conclusions

Beginning of document presented main differences between two virtualization systems. It is important to mention that mainframe computers are used for huge production enterprise-oriented virtual environments (and cost greater amount of money) and the PCs seems to be used for smaller VM usages

One of the most surprising things about performed test is that PC system demonstrated greater performance in most queries. This was caused by:

- Big difference in bus speed and CPU speed. Mainframe disk array used during first test was connected by slow (comparing to the PC) serial bus (20 MB/s vs 133 MB/s theoretical speed). PC CPU was clocked by 2,33 GHz clock, in comparison to mainframe CPU running at 1,3 GHz. CPU speed is rather not important, only in interrupt service time.
- It should be noted, that used Mainframe systems were true production environment. So no matter what time tests would be performed, there are some system processes going in background (transactions, systems, maintains work) that can tamper with this performance test. Also, amount of resources assigned to our z/VM machine is just a drop of what system Z can offer.
- CPU is not all. In most cases, especially under Mainframe systems we came across mass data processing. In consequence it is important to access data on disk storage and tape drives instead of doing just mathematical calculations in system memory. So it is important to configure environment with best I/O possible.
- Configuration. Crucial parameter about business and enterprise class systems is to properly configure them. Just buying product from top shelf doesn't mean it is going to be better than previous one. It needs to be customized to fit needs. Appetite for system resources should also be considered especially storage space, memory and CPU power. Just out of curiosity I repeated test with 4 cores on PC and results was clearly visible.
- Flexibility - connection of disk drive in PC is based on SATA copper wires, internal connection inside PC BOX. Adding new device is not poss-

ible in many systems (only dedicated construction has that availability). In mainframe systems connection of new device is just like to put new fiber link and also is possible to change configuration without power disruption (fully dynamic). Disks emulated in PC environment are just normal files on hosting operating system. Minidisks in z/VM are part of real disks (on cylinder boundaries) and may be shared between different virtual machines (not possible in VMware). Other devices (like crypto cards, network cards etc.) also may be shared and offloads main CPU, what is very important in compute-dedicated processes (SSL handshaking) or big volume of interruption (ARP offload).

- Power consumption – in both environments it is possible to run more virtual machines than one. In fact – during the tests, on mainframe system 6 other operating systems cooperated with the same disk array at the same time. On PC, there was only one instance of VM. If we had needed more instances – we would probably had to add more PC boxes – and power consumption in summary would be greater than for the mainframe.
- System "health" is also important. PC tests were performed when over 0.5 TB additional files were located on testing partition. After deleting them and repeating test, time required for operation drop down to half of previously required. Administrators should keep systems as fit as possible, constantly cleaning up

Although the tests showed that that for this particular configuration PC are distinguished by better performance it necessary to notice that in addition to comparing performance, one should look at the aspect of flexibility in the administration, configuration and sharing of resources. It is also worth mentioning that fine tuning of configuration of z/VM is also an important key issue in the area of performance of Linux installed in virtual machine environment [4].

In the case of z/VM, each configuration change (such as adding a new minidisc) is reflected in the running virtual machine. This fact will have great impact in systems for which it is difficult or even impossible to define the final resources at the very initial stage of its creation

It is also worth mentioning that mainframe machines posses the possibility of sharing resources among multiple virtual machines. Devices, such as discs, cryptography cards, network interfaces can be successfully shared among several hundred virtual machines. The root directory (read-only) for the dozens of virtual Linux machines located on the virtual disk is an example of practical reduction of energy consumption (less frequency of access to different areas of the physical disk).

References

1. VMware Inc., 2007, *Workstation User's Manual VMware Workstation 7.0*, VMware Inc.
2. Geiselhart G., Breneman R., Dow E., Egeler K., Gutenberger T., Hayden B, Sousa L., 2006, *Linux for IBM System z9 and IBM zSeries*, IBM Redbooks
3. Parziale L., Alves E.L., Dow E.M., Egeler K., Herne J.J., Jordan C., Naveen E.P., Pattabhiraman M.S., 2008, *Smith K. Introduction to the New Mainframe: z/VM Basics*, IBM Redbooks
4. Turk D., Bausch J., 2005, *Virtual Linux servers under z/VM: Security, performance, and administration issues*, IBM SYSTEMS JOURNAL, VOL 44, NO 2,