ADVANCED TECHNOLOGIES AND APPLICATIONS SESSION
INVITED PAPER

MATHEMATICAL MODELING AND FILE SYSTEM PERFORMANCE EXAMINATION FOR TYPE 1 HYPERVISOR WITH FULL HARDWARE VIRTUALIZATION IN THE CASE OF KVM AND MS HYPER-V

Borislav Đorđević[1]*,
Nenad Kraljević[2],
Svetlana Štrbac-Savić[1],
Nenad Korolija[3]

[1]Academy of Technical and Art Applied Studies Belgrade,
Belgrade, Serbia

[2]EPS Branch TENT A - Obrenovac,
Obrenovac, Serbia

[3]School of Electrical Engineering,
University of Belgrade,
Belgrade, Serbia

Abstract:

The paper analyses the performance comparison of hypervisor type 1 in the case of KVM and MS Hyper-V platforms for virtualization. The characteristics of both tested hypervisors were examined through file system performance. The tests were performed under equal conditions and with an inimitable testing methodology, using the benchmark program - Filebench.

The performance of the tested hypervisors was compared taking into account tests performed for a system with one, two and finally three virtual machines in state of operation. Mathematical modelling was done, which is the main contribution of this work, then hypotheses about the expected behaviour were set and confirmed through the obtained results. Ubuntu 22.04 LTS, an OS (operating system) from the Debian distribution family, was execute as the guest OS.

Keywords:

MS Hyper-V, Linux KVM, Ubuntu OS, Virtual Machines, File System Performance.

INTRODUCTION

Virtualization is a technique used in information technology to create a virtual representation of computer resources. The term "virtualization" refers to the simulation of software or hardware by another software package. IBM first used virtualization in the 1960s as a technique for logically allocating mainframe computer system resources to various applications. Since there is no longer a requirement to maintain the "one server, one application" concept, today is possible to execute numerous operating systems on the identical physical platform. Utilizing such a system has numerous benefits and cost reductions, including those related to hardware, CPU, memory, and administrative personnel.

All of this is positive conclusion for virtualization in the area of reliability. Data migration from one server to another and the addition of new servers are both made simple by virtualization solutions. This adds to the technology's benefits in the area of scalability, which is considered a great improvement in hardware virtualization. Purpose of this paper is to compare two different tools of virtualization on same hardware and software resources.

Correspondence:
Borislav Đorđević

e-mail:
borislav.djordjevic@viser.edu.rs

DOI: 10.15308/Sinteza-2023-140-146

The most well-known and widely used type of virtualization is full hardware virtualization, which is the subject of this essay. VMM (Virtual Machine Monitor) is the name of the program that manages virtualization. Server virtualization is another name for the procedure of generating and managing virtual computers, and it is used most frequently in a professional IT setting.

Type 1 (native) and type 2 (hosted) hypervisors fall into these two groups [1-2]. In this case study, type 1 hypervisors were put to the test for the virtual platforms MS Hyper-V and KVM (Fig. 1).

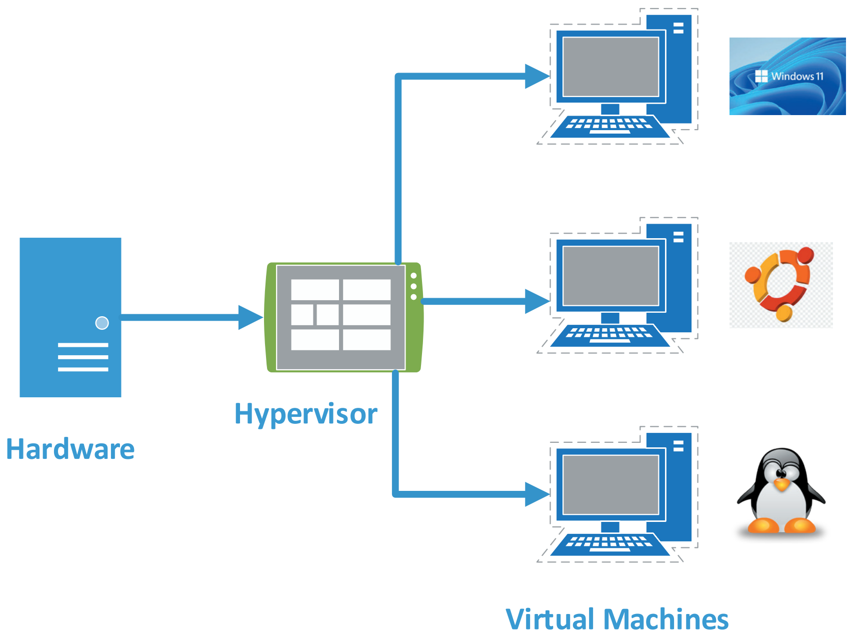

Figure 1 – Example of a type 1 hypervisor.

2. RESEARCH WORK, MOTIVATION AND GOAL

The majority of the papers in this area focuses on comparing hypervisor performance using various test methodologies and benchmarking tools. Most of these references represent quality case studies in which various system performances of hypervisors are measured and interpreted. While most of these references lack comprehensive mathematical modelling, their case studies still provide valuable scientific and practical contributions [3-8]. One of the pillars for achieving outcomes at the quality level is the employment of various tried-and-true benchmarking tools for this purpose. Because Filebench is a versatile, potent, multithreaded solution that simulates the actual application workloads, using this powerful tool is advised. The same could be advice for Fio tool, a benchmark that is comparable to Filebench, as well as certain artificial benchmarks like Bonnie++ and Postmark.

Mathematical modeling of virtualization based on hypervisors in the field of file system performance and application of the model to a case study, through the interpretation of benchmark results, are the key contributions of this research. A similar mathematical model was previously used in the case studies presented in references [9-11]. Through our core module, using mathematical modelling, we successfully interpret the file system performance in various case studies. Because there are so many variables in a complicated virtual environment, the model predicts that there won't be a single winner hypervisor and that it will depend on the case study, or the workload characteristics. In terms of competitiveness, we are mandating a mathematical model, many case studies built on the concept, and real-world performance evaluations. Reducing infrastructure costs and hardware reduction, accompanied by simpler administration, make the server variation of virtualization a fantastic answer. However, there is still much opportunity for development and many unanswered questions.

The contribution of this study is the comparison and validation of two different hypervisors, MS Hyper-V and KVM, whose quality and performance we have examined under identical circumstances. The MS Hyper-V hypervisor is also suited for the usage of paravirtualization, but both of hypervisors use complete virtualization. The well-known Linux OS distribution Ubuntu OS 22.04. LTS as the guest operating system is employed, and the Filebench benchmark tool with four different workloads was used for testing purposes. A mathematical model was constructed once the hypotheses were established and was then verified by the observed data.

2.1. KVM AND MS HYPER-V

Microsoft created MS Hyper-V, an effective hypervisor that allows for the virtualization of operating systems in a server and PC (Personal Computer) context (Fig. 2). Microsoft has integrated a Hyper-V virtualization as a role into the operating system with the release of Windows 8 for PC and Windows Server 2008 R2 for servers. When is used in PC, Hyper-V acts as a hosted hypervisor. The difference when using Hyper-V in the server or PC version is reflected in the memory usage approach. The MS Hyper-V role permits the segregation of partitions in which guest OS will run and enables administrators to establish several virtual machines [12].

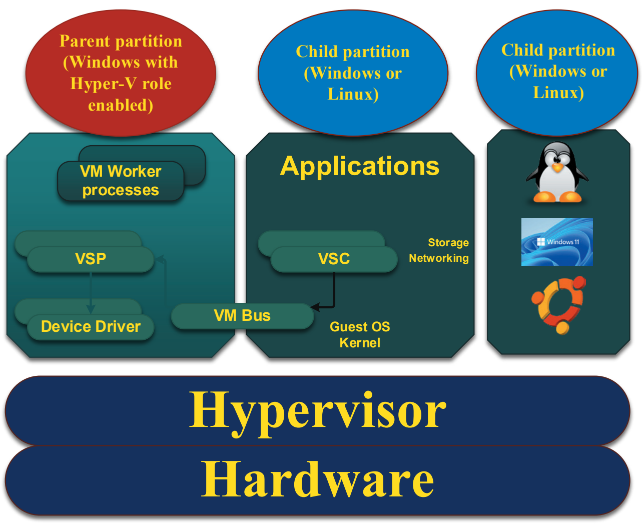

Figure 2 – Example of MS Hyper-V architecture.

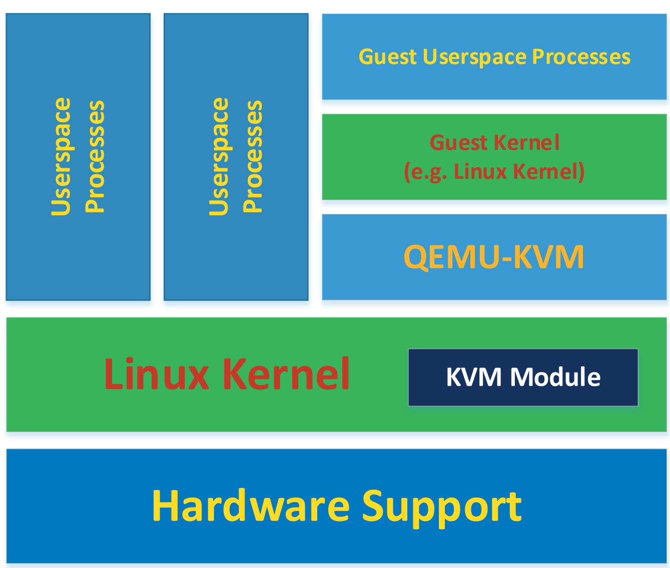

Figure 3 – Example of KVM architecture.

KVM (Kernel-based Virtual Machine) technology is practically required for virtualization under the Linux operating system. It was developed initially as a Red Hat sponsored project. KVM is a core component of the Linux kernel since version 2.6.20 and is implemented as a kernel module. It is impossible to categorize the KVM as a type 1 or type 2 hypervisor. On the one side, KVM enhances and gives the Linux kernel virtualization features, enabling Linux to be used as a bar-metal hypervisor (Fig. 3). Conversely, Linux OS is a standalone operating system that provides the underlying functionality for KVM to operate independently. As a result, it can be claimed that KVM operates above the primary OS (type 2 hypervisor), employing already developed system functionalities in the absence of its own [13-14].

2.2. MATHEMATICAL MODEL AND HYPOTHESES ABOUT EXPECTED BEHAVIOUR

Both hypervisors are implemented in the microkernel architecture. Both hypervisor are native, working directly on the hardware. The following equation (eq. 1) can be used to get the total processing time for each Time workload (Tw):

$$T_W = T_{RW} + T_{SW} + T_{RR} + T_{SR} \tag{1}$$

whereas T_{SR} and T_{RR} stand for sequential read and random read times, T_{RW} and T_{SW} stand for random write and sequential write data entering times, respectively. A file system with six input factors has the following estimated access time for each of these workloads (eq. 2):

$$T_{WL} = T_{FB} + T_{FL} + T_J + T_{HK} + T_{DIR} + T_{META} \tag{2}$$

where T_{WL} stands for the total amount of time needed to complete all operations for a specific workload, and the input factors of equation (2) stand for the amount of time needed to complete all operations involving file blocks and lists, journaling, housekeeping, metadata, and directories in the FS (File System). The following five factors can affect the workload time T_W (eq. 3):

$$T_W = f\,(\text{Bnk, GOS-FS, Hp-proc, VH-proc, HOS-FS}) \tag{3}$$

For both KVM and Hyper-V, the first and second parts, Bnk (Benchmark) and GOS- FS (guest file system), are the same. Since the same benchmark and virtual machines (with the same ext4 guest FS) were used for testing, it is reasonable to believe that these two factors will have an identical effect on the third input factor, Hp-proc (hypervisor processing). This is the period of time it takes the hypervisor to pass a request from the virtual hardware to the host drivers. QEMU full virtualization for KVM and MS full virtualization for Hyper-V make up the fourth input factor, known as VH-proc (virtual hardware processing).

Both hypervisors have their own solutions, and even while these are complete hardware emulations, their performance will undoubtedly vary. HOS-FS is the fifth input factor (host file system) of equation (3). KVM employs ext4 and the MS Hyper-V his NTFS file system. This input factor is anticipated to affect hypervisor processing times differently. Since the particular tests are designed to evaluate the performance of bar-metal virtualized guests, it is anticipated that the third, fourth, and fifth input factors of equation (3) will have the greatest bearing.

2.3. TEST CONFIGURATION AND BENCHMARK APPLICATION

It is important to use the identical hardware configuration, the identical guest OS, a high-quality benchmark test software, and the identical performance measurement methodology for testing to be appropriate and of high quality. The tests were conducted on Dell Vostro 15 3591, whose characteristics are shown in Table 1. Table 2 shows the characteristics of the hard drive used for the tests. As a guest OS, Ubuntu OS 22.04. LTS was employed.

The benchmarking software Filebench, version 1.4.9.1-3, was used for all experiments. This application may provide a huge number of workloads and is intended to assess how well file systems and storage function. Web, mail, and file server workloads are utilized to simulate settings in this article while employing services.

Table 1 – Hardware environment/Dell Vostro 15 3591.

Processor	Memory	Hard disk	Host OS
Ryzen 7 3700U, 2.3GHz, 4C/8T	16GB DDR4	Micron 2210 NVMe 512GB	Windows 11 Pro

Table 2 - Hard disk environment/Micron 2210 NVMe 512GB.

Device type	Capacity	NAND Flash memory type	Interface	Internal data rate
Solid state drive - internal	512GB	3D quad-level cell (QLC)	1xPCI express 3.0x4 (NVMe) – M.2 card	2200 Mbps (read) / 1070 Mbps (write)

3. TESTING AND RESULTS

In this essay, the performance of various virtual server platforms is compared. Data throughput and disk performance were evaluated using Filebench tool. All virtual machines were developed with identical features in order to make testing meaningful (Table 3).

We updated the base code files *varmail.f, webserver.f, fileserver.f,* and *randomfileaccess.f,* for the purposes of mail, file, web and combination of previous three server testing. By first constructing a single test virtual machine on Windows 11 Hyper-V, which had been activated as a role, was tested. The environment is tested also with two and three virtual machines using the same process. There were ten tests, each lasting 120 seconds. The utmost result is an average of all the test results that were obtained.

The Hyper-V was deactivated and its virtual machines were uninstalled to clear the environment before testing the KVM virtual platform (running Ubuntu OS 22.04. LTS with the KVM option selected). Then, the KVM virtual platform was installed and tested using the same process. The same criteria were acquired for both virtual platforms in this experiment. Figure 4 show the "Varmail" workload testing findings.

The "Varmail" workload reveals that KVM has a little better performance score than Hyper-V. Along with the random read input factors, this workload also includes synchronous random write input factors, for which the effect of FS caching is minimal. Due to the fifth input factor of equation (3), where NTFS (New Technology File System) for this workload performed better in the FS pair, Hyper-V is superior in this instance (ext4 on ext4 compared to ext4 on NTFS).

Table 3 – Virtual machine parameters/Input factors and characteristics.

Virtual processor	Virtual memory	Virtual hard disk	Guest OS
2	2GB	10GB	Ubuntu 22.04.1 LTS

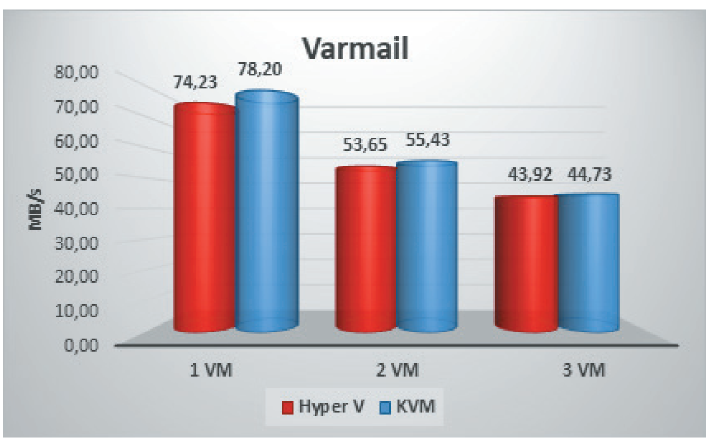

Figure 4 – Varmail test results.

Figures 5, 6 and 7, show the outcomes of evaluating other workloads.

It is obvious that Hyper-V performs quite better than KVM for the "Fileserver" workload. The FS cache effect on the guest OS and host OS is considerable in a complicated workload like Fileserver where there are random and sequential write input factors, hence KVM loses mostly due to the third and fourth input factors of equation (3). Conclusion is that that KVM has higher hypervisor latency and worse virtual hardware processing.

We can observe that Hyper-V is once more somewhat inferior to KVM for the "Webserver" task. Due mostly to the fifth input factor of equation (3), or FS pair (ext4 on ext4 in comparison to ext4 to NTFS), and the combined effect of FS caching, Hyper-V manages little bit worse in the "Webserver" workload, which contains random read input factors and very few random write input factors.

The "Randomfileaccess" workload is another one where the winner can't be concluded with certainty. Because of the significant influence of FS caching, particularly for random write, on this workload, which has many asynchronous random write and random read input factors, KVM showed a bit better performance than Hyper-V. This is mostly a result of the solid cache effect in random write and the fifth input factor of equation (3), NTFS, or the FS pair (ext4 on ext4 versus ext4 on NTFS).

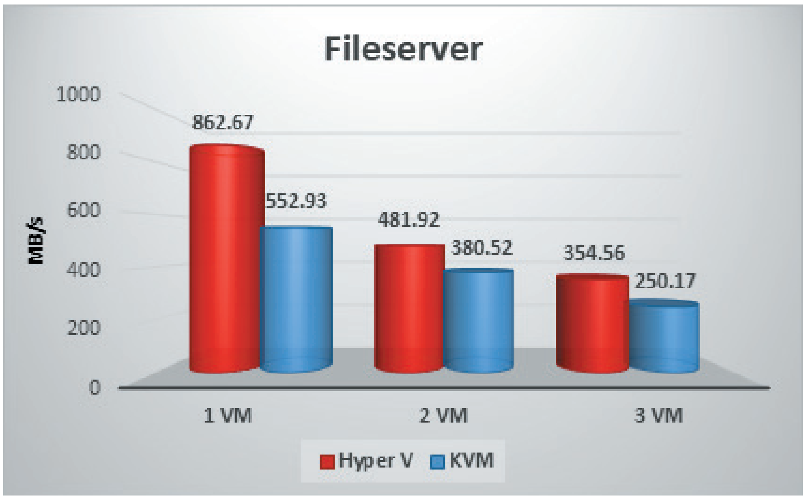

Figure 5 – Fileserver test results.

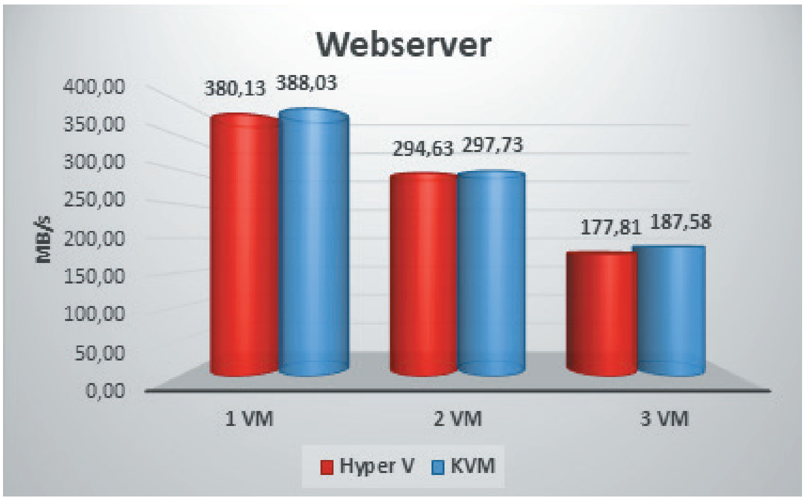

Figure 6 – Webserver test results.

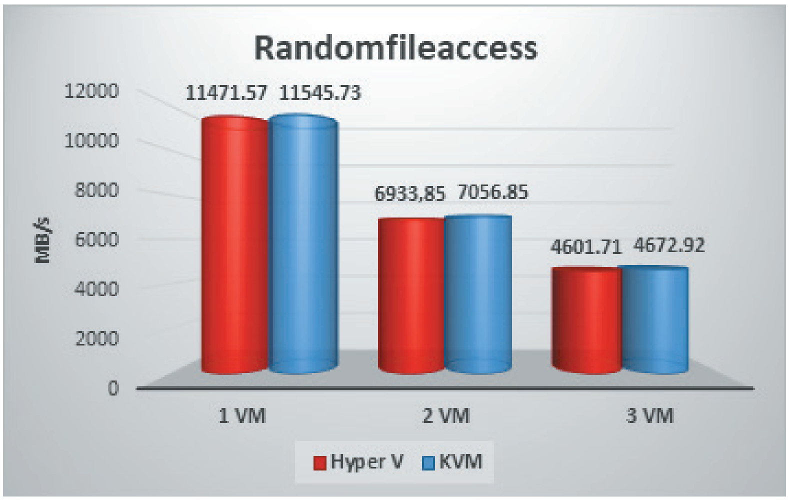

Figure 5 – Fileserver test results.

4. CONCLUSION

In the world of information technology, virtualization has already shown its worth and found a suitable place. In addition to all the advantages this technology offers, it is important to highlight its significant contribution to protecting the environment and the fact that it may be utilized to great effect in the field of green technologies. In the research described in this paper, Hyper-V dominated the most complicated workload (Fileserver), whereas KVM showed somewhat better performance in varmail, randomfileaccess and webserver testing workloads. The differences in the file system of the host OS and the FS pair played a major part in the performance disparities for this type of hardware and experiment (ext4 on NTFS vs ext4 on ext4). Additionally, virtual hardware processing and hypervisor processing, which have been shown to be the most difficult workloads, differ much on these two hypervisors for Fileserver testing results.

Future research in this area might concentrate on testing various server kinds as well as other widely used virtual platforms.

5. ACKNOWLEDGEMENTS

The work presented in this paper has partially been funded by the MESTD (Ministry of Education, Science and Technological Development) of the Republic of Serbia.

6. REFERENCES

[1] E. Bugnion, J. Nieh, D. Tsafrir "Hardware and software support for virtualization", 2017.

[2] Virtualization. Online: https://www.redhat.com/en/topics/virtualization /what-is-virtualization, 2023.

[3] What is a Hypervisor? Type 1 vs Type 2 Hypervisor, Online: https://techsupportwhale.com/what-is-a-hypervisor-type-1-vs-type-2-hypervisor/, 2023.

[4] Full Virtualization. Online: https://www.sciencedirect.com/topics/computer-science/full-virtualization, 2023.

[5] M. Portnoy: "Virtualization Essentials", 2016.

[6] J. Torbić, I. Stanković, B. Đorđević, V. Timčenko: "Hyper-V and ESXi hypervisors comparison in Windows Server 12 virtual environment", 17th International Symposium INFOTEH-JAHORINA, 2018.

[7] Microsoft, Online: https://www.microsoft.com/, 2023.

[8] Virtualization technologies. Online: https://www.microsoftpressstore.com/articles/article.aspx?p=3145750, 2023.

[9] B. Đorđević, V. Timčenko, N. Kraljević, N. Maček: "File System Performance Comparison in Full Hardware Virtualization with ESXi, KVM, Hyper-V and Xen Hypervisors", Advances in Electrical and Computer Engineering, vol.21, iss.1, 2021.

[10] Linux KVM, Online: https://www.linux-kvm.org/, 2023.

[11] KVM Internal: How a VM is created? Online: https://insujang.github.io/2017-05-02/kvm-internal-how-a-vm-is-created/, 2023.

[12] V. Dakic, "Mastering KVM virtualization" 2016.

[13] M. Polenov, V. Guzik, V. Lukyanov: "Hypervisors comparison and their performance", Computer Science Online Conference, 2018.

[14] Filebench, Online: https://github.com/filebench/, 2023.