Use of virtualization and container technology for information infrastructure generation

Trofimov Valery Vladimirovich
Computer Science Department
SPbSUE
St. Petersburg, Russia
tww@mail.ru

Kiyaev Vladimir Ilich
Computer Science Department
SPbSUE
St. Petersburg, Russia
kiyaev@mail.

Gazul Stanislav Mikhailovich
Student Admission Information Support Department
SPbSUE
St. Petersburg, Russia
sgazul@gmail.com

***Abstract* —The target is to develop the approach for selecting the best location mode of virtual IT resources in generating IT infrastructure. The proposed approach allows choosing virtualization or container technology depending on the existing limitations, with accounting for the results of the author's full-scale experiments as well as the similar experiments' results of the leading software manufacturers. The important feature of the approach is the developed evaluation mechanisms. Application of the proposed methodology allows to increase the packing density of IT resources at minimum costs, or to achieve the maximum stability and efficiency of the generated IT infrastructure in relation to the set up problems.**

Keywords—component; formatting; style; styling; insert (key words)

I. INTRODUCTION

The use of virtualization technology is the key element for generating modern IT infrastructure of large hi-tech organizations. Today it is difficult to offer a technically competent and economically sound solution for building a new powerful and economical server node without using virtualization technologies [1-4, 5, 6], no matter which organization this node is designed for– be it a big production company or a large educational establishment.

At present there are several virtualization methods achieving similar result but using different abstraction levels - which, ultimately, can affect the efficiency of result and the overheads for its achieving. Due to this, with determining informatization strategy, the problem inevitably arises of choosing the most suitable technology and virtualization method with regard to their inherent benefits and drawbacks. Let's consider the most frequently used methods.

The hardware emulation method is the joint development of built-in software and hardware: a hardware virtual machine (VM) is created on the host system for emulating the hardware that is of interest for users. The method is most in-demand for the process of developing and testing software [6]. The significant drawback of this virtualization type is the slow operation of emulated applications, resulting from the need of simulating commands on basic hardware.

Complete (hardware) virtualization is used for operation management of virtualized environments with physical hardware of host system, where hypervisor is responsible for resource-sharing control of the system's physical resources between the virtualized environments. It is hardware virtualization that is used, as a rule, with building large corporate systems.

Paravirtualization is a solution for building virtual environments requiring the guest operating system (OS) to be modified for the hypervisor (as a result of modifying the operating system's source code, a special OS version, called Guest Edition, is created). Despite the need for OS source code modification being a significant drawback of the method, paravirtualization, as a result of this approach, ensures high performance of virtualized OSes and applications running under their control. The performance of such solutions sometimes approximates the performance of non-virtualized operating systems.

However, limited number of supported OS versions is a serious drawback of paravirtualization, as guest operating systems should be specially prepared for operation in a virtual environment. Additionally, the OS code modification is a difficult task requiring not only the stability but also the maximum performance of the virtualized OS. Hypervisor Xen and its modifications (Citrix XenServer, XCP) is an example of hypervisors using paravirtualization along with hardware virtualization [1, 6, 8, 10].

Operating system level virtualization - virtualized environments are allocated in the container forms located within the root operating system and using this operating system's kernel. Such method, using a single operating system, in the very general case simply isolates independent virtual containers from each other. Technically, to implement this virtualization method based on the technology sharing the resources of one server between containers, you need to introduce changes into the operating system kernel (for example, implementation of OpenVZ) [10].

Despite the technical difficulties of implementation, one should note such significant advantage of the container method as "native" performance, without "overhead" for device virtualization. Operating system level virtualization is implemented in Solaris Containers; FreeBSD jail and Virtuozzo / OpenVZ; In Linux and * BSD; In Linux Containers.

II. Performance comparison of the main virtualization technologies

We compared performances of main virtualization technologies, which, according to experts [6, 9, 10] have the maximum spread today - hypervisor virtualization and Linux containers. Docker and the kvm hypervisor were used as specific solutions for conducting tests. Docker is a platform providing convenient interface for operating with Linux containers.

Tests were conducted on the equipment frequently employed in the corporate software and hardware environment used for IT infrastructure generation (see Table 1). During the tests, the following software was used: operating systems Ubuntu Linux Server 16.4, version Docker 1.12.6, kvm.

TABLE I. Test Server hardware

Component	**Specification**
Server board	Intel S2400GP2 Granite Pass
CPU	Intel(R) Xeon(R) CPU E5-2407 0 @ 2.20GHz, 4 cores
RAM	8 Gb., PC3-12800, 1600 MHz, DDR3, ECC Reg.
HDD	Capacity: 60 Gb., Spindle speed (RPM): 5400 Cache buffer: 8 Mb.
Network	Ethernet 100 Mbit/s.

A. *Comparison of data store operation speed between Docker and VM containers*

First of all, we have tested the data store operation speed with Docker containers and the virtual machine kvm. Specialized utility fio was used for the test. The results of tests on sequential data reading and writing obtained due to the utility use are presented in Table 2.

TABLE II. Fio results: Sequential read and write (MB/s)

	Docker	**kvm**
Sequential Read	79,03	19,465
Sequential Write	78,89	38,419

The table data are presented in Figure 1 (MB / s).

TABLE III. Fio results: Random read and write (iops)

	Docker	**kvm**
Random Read	331	109
Random Write	286	679
Random Mixed	303	153

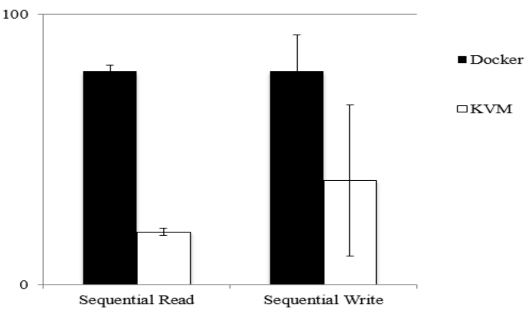

Fig 1. Sequential read and write tests results

Table 3 and Figure 2 show the data obtained during the performance test. The values are in iops units. We have compared the speed in three versions — reading, writing and reading-writing of random data.

The response time performance (ms) for random reading operation is shown in Figure 3 (in the percentile format).

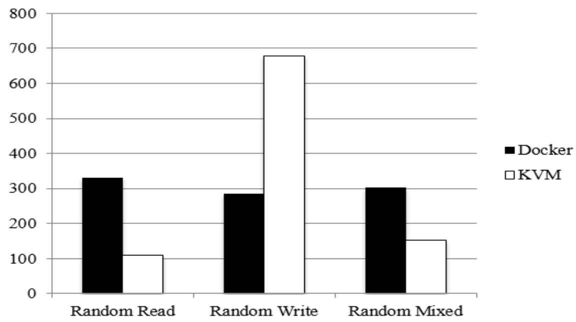

Fig. 2. Random read, Random write and Random mixed tests results

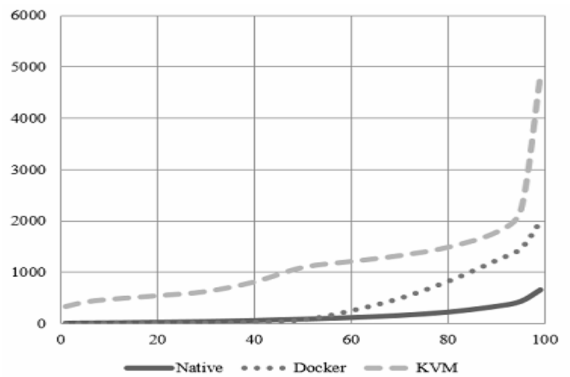

Fig. 3. Random read, Random write and Random mixed tests results

B. *Comparison of data transmission and reception speed in the network*

The results of the test conducted with the iperf utility are presented in Table 4.

TABLE IV. IPERF RESULTS (AVERAGE, MBITS/SEC)

	host	**Docker**	**kvm**
Sender	67,1	65,3	65,8
Receiver	67,1	65,3	65,8

The table data are presented graphically (see Figure 3).

Using the ping utility, we also estimated the average response time of the Docker container and of the virtual machine kvm. The average response time for the results of 10,000 measurements is shown in Table 5.

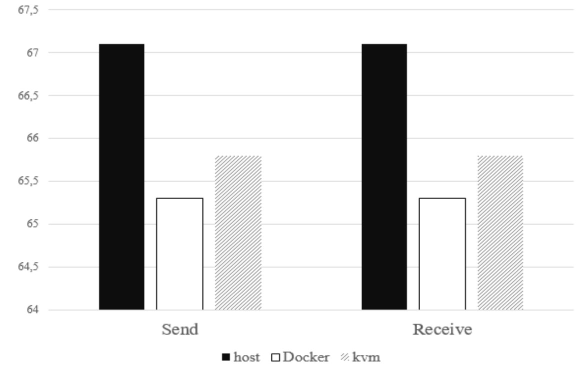

Fig. 4. Iperf results (average, Mbits/sec)

TABLE V. PING RESULTS (AVERAGE, MS)

host	**docker**	**kvm**
0,629	0,7	0,836

The Table 5 data are presented diagrammatically in Figure 5.

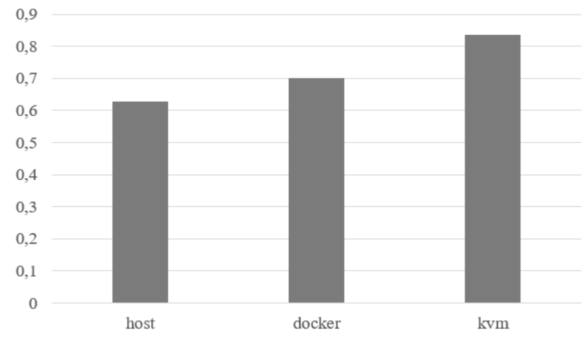

Fig. 5. Ping results (average, ms)

The conducted experiments and test results showed that with allocating large numbers of same-type virtual resources, the Docker-controlled technology of containers allows to achieve better performance - both with data store operation and operation with network interfaces.

III. SELECTION OF VIRTUALIZATION METHOD

Depending on the specific requirements for the forming infrastructure of the organization, having economic reserves and the existing material and technical base, there is a problem of choosing the optimal method of virtualization. To select one, there is an evaluation algorithm proposed, which can reasonably choose one of four possible ways of virtualization: hardware emulation, paravirtualization, full or container virtualization.

We have developed the below algorithm with regard to the earlier presented test results as well as the research of IBM corporation [1, 2, 6, 9].

Algorithms uses expert evaluation, the questions are proposed to the experts, sorted in order of importance, starting with the most significant. In a simple realization of the algorithm, on each question there is, only positive — 1 or negative — 0 answer of an expert.

A typical set of questions might be as below:

1) Is there the need for equipment, which is not available, to run virtualized systems equipment?

2) Is there a special Guest Edition version of the virtualized OS?

3) Is there the need to use different versions of the OS?

4) Is there the need to support a large number of virtual servers?

Note, that the algorithm for selecting optimization methods can be complemented by the introduction of non-integer weights used in the evaluation process; concretized position "a large number of virtual servers" considering several options, etc.

Recommended choice of answers to the questions might be:

1,0,0,0 — hardware emulation,

0,1,0,0 — paravirtualization,

0,0,1,0 or 0,0,1,1 — full virtualization,

0,0,0,1 — container virtualization.

To answer: 0,0,0,0 — the best way of virtualization - is full virtualization. However, when considering the arguments in favor of choosing the method of full virtualization, it should be noted, that this method, in the case of a large number of virtualized operating systems, requires high-performance hardware.

IV. CONCLUSION

Within the framework of this study, we examined the main virtualization technologies used today in the corporate environment. We compared performances of the main virtualization technologies, having evaluated their employment potential for solving various problems, and proposed a simple algorithm for choosing the virtualization method based on the obtained data.

REFERENCES

[1] Амелин К. С., Амелина Н. О., Граничин О. Н., Кияев В. И. Разработка приложений для мобильных интеллектуальных систем на платформе Intel Atom. — СПб: Издательство ВВМ, 2012. — 211 с.

[2] Газуль С. М. Принципы проектирования гибридной информационной системы для поддержки образовательного процесса в высшем учебном заведении // Вестник СПбГЭУ. Серия: Экономика. Выпуск 5 (72). — СПб: СПбГЭУ, 2014. — С. 58–61

[3] Газуль С. М., Бабаев Э. О., Горнов П. А. Интегральный показатель готовности информационной системы к работе в облаке // Международный научно-исследовательский журнал — Research Journal of International Studies, 2014. — № 4-2 (23). — С. 14-16.

[4] Граничин О. Н., Кияев В. И. Информационные технологии и системы в современном менеджменте. — СПб: Изд-во ВВМ, 2014. — 897 с.

[5] Minakov V.F., Ilyina O.P., Lobanov O.S. Concept of the Cloud Information Space of Regional Government//Middle-East Journal of Scientific Research, 2014. — № 21 (1). — P. 190-196

[6] Musaev A. A., Gazul S. M., Anantchenko I. V. The information infrastructure design of an educational organization using virtualization technologies // Известия Санкт-Петербург-ского государственного технологического института (технического университета), 2014. — № 27 (53). — С. 71-76

[7] Трофимов В.В. Конвергенция информационных технологий. — СПб: Изд-во С.-Петерб. гос. ун-та экономики и финансов, 2011. — 21 с.

[8] Hess K. Practical Virtualization Solutions: virtualization from the trenches, 2010. — 336 p.

[9] An Updated Perfomance Comparison of Virtual Machines and Linux Containers [Электронный ресурс] // IBM Research [сайт]. [2014]. URL: http://domino.research.ibm.com/ library/ cyberdig.nsf/papers/ 0929052195DD819C85257D2300681E7B (дата обращения 11.04.2016)

[10] Анализ современных технологий виртуализации [Электронный ресурс] // Хабрахабр [сайт]. [2017]. URL: http://habrahabr.ru/company/centosadmin/blog/212985 (дата обращения: 12.04.2017)