THE INFORMATION INFRASTRUCTURE DESIGN OF AN EDUCATIONAL ORGANIZATION USING VIRTUALIZATION TECHNOLOGIES

St. Petersburg State University of Economics, Sadovaya Str., 21, St-Petersburg, 191023, Russia
St. Petersburg State Institute of Technology (Technical University), Moskovsky Pr. 26, St. Petersburg, 190013, Russia
e-mail: sgazul@gmail.com

The main types of IT information infrastructure virtualization used to build the information system of modern educational institution are given. The prerequisites for the use of described technologies for building the information infrastructure of educational institutions are considered. The technique, which allows one to choose the most appropriate way of virtualization depending on the existing restrictions is suggested. The main advantages of using virtualization technology by educational institutions for building information infrastructure are described. These include reducing the cost of IT infrastructure of the institution, increasing the flexibility of the educational process, by allowing quick student workplace setup when computers are used as terminals to work with virtual machines, deployed in the cloud on demand, the ability to emulate the functionality of expensive physical hardware, etc. The scenarios of infrastructure network development of an educational institution are considered. The importance, economic feasibility and technical need for virtualization technology in educational institutions is stressed. The final section shows typical usage scenarios of virtualization technology in the organization of modern educational process.

Keywords: virtualization, hybrid information systems, cloud computing

Introduction

In recent years, many experts and business environment representatives mention changing the paradigm of higher education [1-4], justifying this by the emergence of new requirements for higher education, formed under the influence of global trends [6, 8]. New information technologies are implemented in the educational process in domestic and foreign educational institutions, which changes the demand for the information infrastructure resources [4, 17]. The problem of optimal use of cloud services in educational resources development becomes more urgent [8, 9, 12], in order to form a hybrid information
infrastructure for supporting the educational process [5-7]. The central element is the use of virtualization technologies in development of the IT infrastructure for a university [3, 11].

The article reviews modern IT infrastructure virtualization technologies, algorithm for choosing optimal technological solutions in accordance with specific requirements for virtualization technologies in educational institutions. Considering these requirements for virtualization technologies used to support the educational process, helps not only to determine the limits of using a particular method of virtualization, but also ensures a proper transition to implementation of cloud solutions at a minimum organizational, technical and economic costs.

**Types and basic technologies of Virtualization**

Today virtualization technology is a key element of modern IT infrastructure of large enterprises (organizations). It is difficult to find technically convenient and economically viable solution to build a new server node without using virtualization technologies [10,15]; regardless the type of organization the project is developed for – whether it is a large company or educational institution. It is possible to use one of several existing ways to implement virtualization technology, with the help of which a similar effect is achieved, but using different levels of abstraction. Each method has its sphere of application, taking into account its advantages and disadvantages. The most commonly used methods are described below.

**Hardware emulation method** - is about the simultaneous development of integrated software and hardware: the virtual machine (VM) of hardware is created in the host system to emulate the needed equipment (Table. 1). This method is mostly useful in the process of developing and testing software [15]. It allows to emulate the hardware required for debugging, which prevents the need for buying expensive physical hardware. One significant downside of this type of virtualization – is a slow work of emulated applications, arising from the need to model commands for the actual hardware.

<table>
<thead>
<tr>
<th>Applications</th>
<th>Applications</th>
<th>Applications</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guest operating system 1</td>
<td>Guest operating system 2</td>
<td>Guest operating system 3</td>
<td>Guest operating system 4</td>
</tr>
<tr>
<td>Virtual equipment 1</td>
<td>Virtual equipment 2</td>
<td>Physical equipment</td>
<td></td>
</tr>
</tbody>
</table>

An alternative option – is a full (hardware) virtualization, using Virtual Machine Manager (hypervisor). The hypervisor is used for the organization of virtualized environments with physical hardware of the host system, the hypervisor is also responsible for managing the physical system resources between virtualized environments (Table. 2). This specific method is mainly used when building large corporate systems. Such major vendors as VMware, IBM and Microsoft are developing their virtualization platforms based on Intel VT (VT-x) and AMD-V hardware virtualization technology [15].

<table>
<thead>
<tr>
<th>Applications</th>
<th>Applications</th>
<th>…</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guest operating system 1</td>
<td>Guest operating system 2</td>
<td>…</td>
<td>Guest operating system n</td>
</tr>
<tr>
<td>Physical equipment</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Another hardware virtualization type is paramervirtualization - solution for virtual environments, which requires modification of the guest operating system for the hypervisor (as a result of modifications to the source code of the operating system, there is a special version of the OS, called the Guest Edition created). Despite the fact that the need for source OS code modification is a significant drawback of the method, paravirtualization delivers high performance of virtualized operating systems and applications controlled by them, which is close in terms of performance to the performance of non-virtualized (real) OS. Paravirtualization can simultaneously support multiple operating systems (Table. 3). However, a limited number of supported OS versions is a sufficiently serious downside of paravirtualization, as guest operating systems must be specially trained to work in a virtual environment. Modification of the OS code is quite a complicated task that requires not only ensuring the stability of the virtualized OS, but also the best performance of this OS. Xen and its modifications (Citrix XenServer, XCP) – are examples of hypervisors, which use paravirtualization par with hardware virtualization [14, 20, 21].

<table>
<thead>
<tr>
<th>Applications</th>
<th>Applications</th>
<th>…</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modified Guest operating system 1</td>
<td>Modified Guest operating system 2</td>
<td>…</td>
<td>Modified Guest operating system n</td>
</tr>
<tr>
<td>Equipment quota</td>
<td>Equipment quota</td>
<td>…</td>
<td>Equipment quota</td>
</tr>
</tbody>
</table>

**Operating system-level virtualization** – is another possible solution: create a container, which helps to virtualize the environment, which are above the level of the root operating system using the OS kernel. Viewed method using a single operating system, in the most general case, just isolates independent virtual containers from each other. Technically, to implement this method of virtualization technology, based on a single server resource sharing between the containers, it is required to make changes to the operating system kernel (for example, the implementation of OpenVZ). Despite the technical complexity of implementation, significant advantage of the method should be noted, as “native” performance, without the “overhead” virtualization devices (table 4).

<table>
<thead>
<tr>
<th>Applications</th>
<th>Applications</th>
<th>…</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating system, using the kernel of host system 1</td>
<td>Operating system, using the kernel of host system 2</td>
<td>…</td>
<td>Operating system, using the kernel of host system n</td>
</tr>
<tr>
<td>Operating system</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Operating system-level virtualization is implemented in Solaris Containers, FreeBSD jail and Virtuozzo / OpenVZ; in Linux and * BSD; in Linux Containers [14, 20, 21]. Using the operating system level virtualization is particularly effective when you need to create multiple virtual servers with limited hardware resources, and each virtual server uses the same operating system. The subject method of virtualization, according to the developers of virtualization, has a good potential for development, signed a partnership agreement between Microsoft and Docker to work together in the field of technology mentioned. It is possible that in one form or another operating system-level virtualization technology will be supported by Microsoft's new version of the server operating systems Windows Server [18]. Note, that operating system-level virtualization technology is especially
effective when you want to create multiple virtual servers with limited physical resources, which may be required by many institutions, taking into an account another fact, that the funding of educational programs is poor. The structure of the educational process itself also encourages the use of this technology, especially if the training is done in the direction associated with the study of IT disciplines, when students use a large number of virtual servers to perform laboratory work (developing practical skills of administering their own servers, computer networks, etc.).

Selecting a method of virtualization

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.

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.

An important argument for the use of virtualization in the organization's infrastructure – is – the benefits acquired from the use of server consolidation: savings through greater utilization of resources used by the physical hardware, reducing the required space needed for equipment, reducing the cost of power supply equipment and the cost of cooling the work area. The costs of administering the physical hardware are also lowered due to the reduced number of physical servers [17]. Note, that the complexity of virtualized environments administration, ceteris paribus, is higher than server infrastructure administration, built without the use of virtualization technologies. This is because the security model is implemented by adding another level - the security of virtual machine. For example, the theft of physical server will be detected almost immediately, but the fact of copying the virtual machine files by an attacker may go unnoticed if the network is administered by people, poorly versed in the peculiarities of the protection of virtual environments. Confidential commercial, scientific, technical or other important information can cost significantly more than physical devices used for processing and storing it. Therefore, the damage from theft of virtual machine (or several machines) files can be significant, especially when you consider that the stolen files can be used for analysis and manipulation (eg, password guessing) as long as an attacker wants. The fact of attempts to access and manipulate with physical servers is usually detected and prevented by system administrators fast enough. Therefore, on the one hand, virtualization reduces administration costs, but on the other hand it requires administrators of a higher qualification, and as a result, higher costs of their remuneration. The important advantage of using virtualization technologies – is the possibility to abstract from the hardware. For example, in the case of failure on a physical server, it is possible to redistribute the load among the remaining equipment. Virtual machines do not bind to a computer hardware, which makes lives of IT staff of the department, responsible for the maintenance of network infrastructure, easier and reduces the risk of downtime due to physical damage of equipment, simplifies migrations, makes it easier to implement the transfer of information resources to the cloud, if necessary [5, 8].

Many experts note cost-effectiveness of the use of virtualization technology, for example, K. Hess in his book [13] results in the calculation of the cost of a number of positions for the classical physical and virtual servers (table 5).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Corporate level solution</th>
<th>Standard level solution</th>
<th>Virtual solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost</td>
<td>$18000</td>
<td>$16000 – $20000</td>
<td>$0</td>
</tr>
<tr>
<td>Power (Watts)</td>
<td>1570</td>
<td>5360</td>
<td>0</td>
</tr>
<tr>
<td>Network interfaces</td>
<td>2 + 8</td>
<td>16</td>
<td>8</td>
</tr>
</tbody>
</table>

Advantages and cost efficiency of solutions built with virtualization technologies

The most significant advantages of using virtualization methods are [14]:

1. The reduction of the cost of equipment acquisition and support;
2. The reduction of server park - the use of virtualization significantly reduces the number of physical computers, lowering the economic costs associated with the search, purchase and replacement of equipment; reduction of areas for the server database;
3. 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.
4. The increase in labor productivity by centralizing services and its relative simplicity;
5. Automation of virtual machines and user information cloning and backup processes.

Information support of the educational process of the university: from virtualization to cloud computing

Using virtualization technology in educational institutions, not only provides all the benefits from the use of virtualization itself as a technology, but also prepares the transition to hybrid systems. Using hybrid scheme allows to move part of the infrastructure of the educational institutions to the clouds of different types [3, 19], to increase flexibility and efficiency of learning process, can quickly put educational information resources into operation on demand.

Today's educational process needs technologies, which provide collaborative and distributed work. One example of practical use of such technology, is the creation of educational (corporate) portal. Portal controller based on Microsoft SharePoint or one of the popular CMS, such as Drupal, Joomla, WordPress and others allows organizing distributed user experience. The portal may be implemented using more than one server (physical or virtual). Implementation of these services using virtual machines allows making IT infrastructure of educational organization modular and flexible. For example, if, due to the failure of
server hardware, educational organization can not provide the deployment of virtual machines (virtualized applications) within its own local infrastructure, there is an alternative available for the virtual machines deployment to public cloud of IaaS type: OpenShift or Microsoft Azure, which is needed for the required modules of IT infrastructure (table 6).

Table 6. Abstraction levels of cloud services

<table>
<thead>
<tr>
<th>Abstraction level</th>
<th>Categories</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructure as a Service (IaaS)</td>
<td>Hardware abstraction, Manual administering and platform deployment</td>
<td>Hardware and software abstraction, Automated administering and platform deployment</td>
<td>Red Hat OpenShift, Microsoft Azure (MS Hyper-V)</td>
</tr>
<tr>
<td>Platform as a Service (PaaS)</td>
<td>Software as a Service (SaaS)</td>
<td>Business logic (applications, providing services to end-users)</td>
<td>Google Apps, Microsoft Office 365</td>
</tr>
</tbody>
</table>

Classical solution for reorganization of information infrastructure, providing educational portal migration, is based on the deployment of new infrastructure, followed by transfer of the data used in the new environment, which is quite time consuming and, therefore, expensive solution. An alternative solution – is modification of the infrastructure, which is based on the use of two-tier architecture, which creates the distributed services for users. For example, web services for information support of the educational process (Fig. 1) using active learning methods (ALM). “ALM” module allows using active learning methods, without specially equipped laboratories, which is important for many educational institutions [8]. The use of private cloud to build infrastructure within the organization, provides high availability and fault tolerance services used in the learning process, with the use of virtualization technologies. The resulting solution is safe, technically efficient and scalable. [9] For example, it is possible to build a hybrid solution - the portal and its integrated office applications in a private cloud, with the ability to zoom into the public cloud. Information system can allow you to connect to external services, built on the principle of public clouds, a solution used in service projects is called “Mashup”. At the user level, it is advisable to take advantage of technology Web 2.0, mash-flexible solutions.

Mashup – is a new rapidly developing area, that allows to create the necessary user resources from a variety of disparate networking sites or resources on a dynamic basis using advanced search and analytical tools. Mashup – is primarily an approach (method) to create applications. As an object, mashup – is Web-unified user resource created by the merger and subsequent configuration of available through the network third-party resources (regardless of how they are acquired) and contributing by providing their data, functionality or custom views. An example of a mashup solution is the profile page of the student in the system, which provides for the possibility of content aggregation from social networks, Twitter and other sources. Another example – is a single entry point for students and teachers, as well as a portal of educational institutions where services of distance learning and testing can be deployed. Integration of locally placed information systems of educational organizations with Google services can also be seen as an example of mashup solutions. Thus, we can talk about consolidated personal information space of working students, graduate students or teachers, as well as general co-educational space for the realization of modern active learning methods.

Let’s examine some of implementation peculiarities of virtualization technologies, linked to the specifics of modern educational institutions. At present, the study of almost any discipline, no matter whether it is technical or humanitarian, is carried out using modern computer technologies. Computing equipment allows not only to carry out complex calculations, but also to apply to a variety of information resources used by the students in the study of a particular subject through the Internet. New software application service (SW) is developed with the expectation that the software will be used on modern computers. However, in educational institutions, there usually are modern computer labs with the latest equipment installed, but classrooms with outdated computer equipment still continue to operate. Sometimes a teacher has some new software for courses, but he or she can not install the software because the installation is not possible on the old computer equipment installed in the classroom. Another common problem is connected with the fact that reading disciplines provided by the curriculum requires some new equipment, which is difficult to order, so the question of hardware emulation using computer technology becomes relevant. Using cloud technology allows us to solve some problems. Most computer education classes use a local area network, which allows you to access the corporate network of the institution (including the use of the resources of the private cloud, if there is one) and to the Internet (the use of public resources, and / or hybrid clouds). The solution provides the students the opportunity to work with virtual machines located in the clouds of various types on servers or local area network. This approach helps to abstract from the parameters of automated workstation (AWS), which is used to connect to the virtual machine. Technical implementation of solutions is provided with software support of standard AWS remote access protocols such as RDP and SSH. These protocols support a wide range of devices, including developed student AWS [2, 3, 8]. Giving students the opportunity to work with virtual machines hosted in the cloud, can successfully solve some of the problems that arise during the learning process [1, 12]:

1. Providing the access to computer resources outside normal working hours. Students enrolled in a modern educational institution, should have access to computer technologies, not only during lessons in the classroom, but also outside school hours, to prepare for classes. Almost all students today are using their own computers (PC and /
or mobile) connected to the Internet. Performance of such computers is sufficient for use as a terminal, but often not enough to install all the necessary student software used in the learning process. Note, that the installation of the software on their personal student computers, even if the specifications allow it, may not always be appropriate for several reasons. For example, the complexity of ensuring control over how students use the licensed software installed on their PCs. Another possible reason - lack of qualifications of the students (especially first-year) to install complex software.

2. Providing students access to one or more virtual machines with administrator privileges, when performing laboratory work in IT technologies. Performing laboratory work, students do not have practical experience of being the system administrator, so they may mistakenly perform actions that require serious computer reconfiguring by the administrative staff of the classroom upon completion the tasks. Using a virtual machine as a kind of “sandbox” allows you to give the students the right to administer on one hand, and on the other do not allow them to configure the operating system of the physical computer.

3. The use of cheap virtual machines for laboratory work instead of expensive physical hardware. If educational stand includes several computers (for example, laboratory work on installing of a domain controller and further configuration of the local network), then not every school can afford to provide classes with the number of computers even for small study groups for financial reasons. Using virtual cloud environments successfully solves the problem.

4. Ability to stop and save the status of virtual machines, used to perform laboratory work, allows carrying out the learning process more flexibly. When using virtual machines there is no need to plan twice as long training sessions for continual laboratory work. The process of software installing and configuring on the physical machine does not allow skipping the next lesson. Stopping and fixation of virtual machines with the ability to restart the virtual machine after a while, is supported by all developed virtualization.

5. Quick setup of computer class before classes. Educational institutions offer an extensive program for students studying computer science. The program includes the study of a variety of operating systems and application software, which can be quite resource-intensive, installing it on a physical computer in the classroom takes considerable time. Prepare the classroom for the lessons may take several hours. In addition, the need to install and uninstall the software does not allow to make quick changes in the schedule of training classes. If the computer class is actually a class of terminals for use with cloud resources, there is no such problem. However, it should be noted, that the use of virtual machines operating in the cloud, places increased demands on the quality of communication. With no connection or low-quality network connection, it is difficult to work with the cloud, and performing laboratory and practical exercises based on the use of virtual machines is impossible in this case.

The use of virtualization in the educational process should be considered comprehensively. System at the level of infrastructure support features as below:

1. The ability to provide fault tolerance;
2. The ability to provide load balancing;
3. Wide opportunities for backup.

At the user level:

1. The possibility of organizing the learning process using the AMO;
2. The modular organization of user resources, the use of mashup solutions;
3. The implementation of a single entry point;
4. The implementation of the user’s profile page;
5. Opportunities for remote access to educational resources;
6. The possibility of a distributed collaboration.

Using virtualization technologies in the implementation of the IT infrastructure of educational institutions allows to realize a modern information system of the institution. Using a hybrid approach in the implementation of infrastructure [5, 8] allows educational institutions to provide the most efficient IT component in the educational process, allows to make learning as efficient as possible, brings it to a new organizational and technical level.

References

11. Trofimov V.V. Konvergencija informacionnyx tehnologij. SPb.: Izd-vo S-Peterb. гос. un-ta’ ekonomik i finansov, 2011. 21 s.
16. Особенности разработки в облаках, Офциальный ресурс: http://www.osp.ru/os/2011/06/13009981