Server hardware resources optimization for virtual desktop implementation

K. Makoviy[a], Yu. Khitskova[a], Ya. Metelkin[a]

[a] *Voronezh State Technical University, Voronezh, ul. 20-let Oktyabrya, 84, 394006, Russia*

Abstract

Considering the virtual desktop infrastructure implementation at the university this article discusses the problem of server hardware resources utilization. The possibility of applying the methods of linear programming to the problem of minimizing the necessary costs when selecting equipment is analyzed, linear objective function and constraints are suggested.

Keywords: virtual desktop infrastructure; linear programming; server hardware; equipment costs

1. Introduction

Virtualization is a common concept for concealing the real structure that is used widely recently for creating virtual hardware, operating system, virtual storage and network resources. Most organizations of different sizes and income implemented (has been implementing) server virtualization for the last 10 years. Server virtualizations is based on the hypervisor technology which creates a thing interlayer between hardware and guest operation system.

On the next step of developing IT infrastructure, organizations address to the technology of centralized desktop execution enhancing end user experience and IT management of desktops. While implementing desktop virtualization it is essential to understand that this serious solution requires not only adequate planning but also financial costs. Value of hardware for physical servers makes considerable contribution for the investment costs whereas optimal configuration of the servers purchased can really save considerable funds.

A mathematical model for solving the optimization problem of server resources needed for desktop virtualization implementation is described and computing results are presented.

2. The object of the study

Server virtualization is essentially a server consolidation, i.e. an approach to the efficient usage of physical servers, widely spread all over the world. This technology allows several operation systems to run on one physical server and isolate applications from each other's influence, minimize investment and operational costs, avoid overprovisioning.

Desktop virtualization or VDI (Virtual Desktop Infrastructure) uses advantages of server virtualization and cloud technologies bringing together (combining) the benefits gained from hypervisor-enabled virtualization and modern display network protocols. Desktop operating systems run on a physical server under control of host operating system i.e. 'hypervisor' whilst screen image is delivered by a network protocol to a client device which may be a PC (Personal Computer), Thin Client, laptop, tablet, etc.

One of the key perspectives of VDI implementation is a possibility to execute any application on any device for which there is a VDI client since applications are executed on the operating system running on the server, not on the device itself. Thus, desktop virtualization provides the basis for extremely promising technology allowing creation a common learning environment - BYOD (Bring Your Own Device) – a new initiative giving opportunity to use wide variety of client's personal devices in a corporate environment.

The number of client computers in a typical organization far exceeds the number of servers therefore this is so important to be able to assess server resources required to run client virtual machines. No less important is to be able to choose the optimal set of hardware servers, for example, from the rage of particular vendor. The key moment to minimize expenses of hardware procurement is a clear view of the server hardware set that needed to provide execution of required number of virtual desktops. We consider VDI implementation in a high school institute, namely the Voronezh State Technical University, which has already a centralized server infrastructure and well-designed network. Desktop infrastructure in an educational institution contains as a rule several sets of identical computers that placed in computer labs. Definitely apart from desktops in computer labs there are a large number of computers with diverse software a university that is used by staff. These computers are not the best choice to being virtualized at the first stage of the project. That is why we assume in our model that we have a number of identical in their performance requirements desktops that we have to place optimally on a purchasing hardware servers.

3. Methods

The problem of virtualized server optimization was considered previously in two aspects – static and dynamic. Static Server Allocation Problem is an approach based on a service concept, the model was introduced in [1] and designed to optimally allocate source servers to physically target servers and was proven that this model is NP-hard problem, heuristic solution based

on bin packet problem is offered. Another option of using linear programming methods for virtualized system placement representing the dynamic aspect of the problem is used for creating application placement controller pMapper [2].

There are several attempts to solve the problem of dynamic replacement of virtual machines on existed physical server infrastructure in datacenter to optimize energy consumption, minimize administrative efforts, increasing server utilization. An approach of dynamic resource allocation for large Internet –oriented data centers bases on queuing theory and Erlang's loss formula [3]. On the other hand it is proposed to use a genetic algorithm based approach, namely GABA, to adaptively self-reconfigure the VMs (Virtual Machines) in large-scale data centers [4]. All the models proposed focuses on the server virtualization not the desktop virtualization. As for desktop virtualization an allocation algorithm based on a bin-packet problem is developed [5]. It is mainly focused on achieving a balance between resource usage optimization and user satisfaction.

In this work we concentrated on the problem of server hardware assessment optimization in order to reduce financial costs while implementing desktop virtualization at the university. To achieve this goal we have to analyze resource requirements of VMs that will be used, number of VMs, and range of hardware servers of the vendor then solve optimization problem to choose a set of optimal server models and their configuration to minimize total cost.

3.1. Model description

For the model we assume that we have a need of particular number of the same virtual desktops. We plan to use them for computer labs at the university and actually we probably will have a need of several types of virtual machines for different labs but for the first approximation we will consider all virtual machines have exactly the same resources requirements.

We consider discrete set of server models actually server platforms, each of them may be supplemented by additional RAM. We can extend RAM with additional memory modules that have various amounts and prices. We assume also that performance of the server is acceptable if RAM amount is sufficient for running VMs only in virtual memory not using as a rule a paging file. In this approximation we do not consider the processor load since the main purpose is minimizing total costs at the very start of VDI implementation project. We also do not take into account the costs of network. Let's say this is a first approximation that can be refines in a future work.

For the model proposed we introduce the following variables:

$\bar{S} = \{S_1, S_2 \dots S_n\}$ – vector of server platforms that can be used for the hardware servers, where n – total number of server platform models selected for consideration;

$\bar{C} = \{C_1, C_2 \dots C_n\}$ – vector of values of server platforms $\bar{S}$, where C_i- is a value of S_i;

$\bar{N} = \{N_1, N_2 \dots N_n\}$ – number of servers of S_i -model that will be used in a final set;

$\bar{M} = \{M_1, M_2 \dots M_n\}$ – maximum amount of RAM that can be added to the server platform S_i;

$\bar{R} = \{R_1, R_2 \dots R_k\}$ – amount of memory module *j*;, where k – is the number of types of RAM modules;

$\overline{Cv} = \{Cv_1, Cv_2 \dots Cv_k\}$ – value of memory module *j*;

$\bar{P} = \{P_1, P_2 \dots P_n\}$– number of RAM slots in the server S_i;

Because our goal is to minimize costs then we have to determine an objective function reflecting the total cost of the hardware server set. The total cost of the solution consists of the value of based server platform (C_i) and the cost of additional RAM modules ($\sum_{j=1}^{k} Cv_j n_{ji}$). Thus the objective function is the following:

$$F = \sum_{i=1}^{n}(C_i + \sum_{j=1}^{k} Cv_j n_{ji})N_i \quad (1)$$

In the following we present constrains for the objective function (subject to):

1. The total amount of RAM should not exceed the one supported by this server platform:
$$\sum_{j=1}^{k} R_j * n_{ji} \le M_i, \text{ i=1..n} \quad (2)$$
where n_{ji}- number of RAM modules *j* on the server *i*.
2. The total number of RAM modules cannot exceed the number of server memory slots:
$$\sum_{j=1}^{k} n_{ji} \le M_i, \text{ i=1..n} \quad (3)$$
3. The total amount of RAM memory on all servers out of server set should provide enough memory to run necessary number of VMs:
$$\sum_{i=1}^{n}((\sum_{j=1}^{k} R_j * n_{ji})/r) \ge V, \quad (4)$$
where V – is a minimal necessary number of VMs, r – memory needed for one VM.
4. To get a solution that makes a sense we will add a constrains for numbers of servers and RAM modules to be integer:

$$N_i, n_{ji} \geq 0,\ N_i, n_{ji} - \text{integer} \quad (5)$$

3.2. Model solution

In order to obtain a solution we divided this problem into two parts:

1. Optimal filling the server slots by RAM modules, analyzing filling for 25%, 50%, 75% and 100% of the maximum amount.

$$\min \sum_{j=1}^{k} Cv_j n_{ji} \quad (6)$$

subject to:

$$\begin{cases} \sum_{j=1}^{k} R_j * n_{ji} = P_i * p \\ \sum_{i=1}^{k} n_{ji} \leq M_j \end{cases}, \text{i=1..n} \quad (7)$$

where p - filling percentage, which can be either 25%, 50%, 75% and 100%

2. Finding the right amount of instances for each server platform:

$$\min \sum_{i=1}^{n} (C_i + Cr_i) N_i, \quad (8)$$

where Cr_i is a result of (6) subject to (4).

4. Results and Discussion

For the problem solution we used the brunch and bound method [6], realized using MatLab. For example of proposed optimization model we can make a choice for implementing 40 VM's using two hardware platforms: ProLiant ML350e Gen8 and Proliant DL160e Gen9. There are five types of RAM modules available for these servers: 2Gb, 4Gb, 8Gb, 16Gb, 32Gb value 26, 136, 215, 315, 840 USD respectively. The result of optimal filling the server slots by RAM modules to minimize cost while maximizing the amount of memory is presented in Table 1,2.

Table 1. Optimal filling of server slots for hardware platform HP Proliant ML350 Gen8

Percentage of maximum supported by hardware server RAM (%)	Number of RAM modules 2Gb	Number of RAM modules 4Гб	Number of RAM modules 8Гб	Number of RAM modules 16Гб	Number of RAM modules 32Гб	Cost (USD)
25	8	0	0	2	0	836
50	6	1	0	5	0	1867
75	3	0	1	6	1	3049
100	0	0	0	10	1	3990

Table 2. Optimal filling of server slots for hardware platform HP Proliant DL160 Gen9

Percentage of maximum supported by hardware server RAM (%)	Number of RAM modules 2Gb	Number of RAM modules 4Гб	Number of RAM modules 8Гб	Number of RAM modules 16Гб	Number of RAM modules 32Гб	Cost (USD)
25	8	0	0	7	0	2413
50	2	0	0	12	2	5460
75	2	0	0	6	9	9450
100	2	0	0	0	16	13340

The obtained results are used in the second part of solution. Adding the cost of additional RAM modules to a base value of server hardware platform we receive the following objective function on the base of (8):

$$\min\{2460 * N_1 + 3491 * N_2 + 4673 * N_3 + 5614 * N_4 + 5713 * N_5 + 8960 * N_6 + 12850 * N_7 + 16740 * N_8\},$$

subject to:

$$(12 * N_1 + 24 * N_2 + 36 * N_3 + 48 * N_4 + 32 * N_5 + 64 * N_6 + 96 * N_7 + 128 * N_8) \geq 350$$

As a result for placing 350 VMs it's optimal to use 10 servers HP Proliant ML350 Gen 8 with RAM slots filled in 75%.

5. Conclusion

Desktop virtualization implementation is a next step in centralizing IT infrastructure that brings both management advantages and academic benefits creating a convenient integrated educational environment. The new model for the optimizing acquisition costs of server hardware purchased for VDI implementation is offered.

References

[1] Speitkamp, B., Bichler, M. A mathematical programming approach for server consolidation problems in virtualized data centers/ B. Speitkamp, M. Bichler // IEEE Trans. Services Comput. – 2010. – Vol. 3. No. X. – P. 266-278.

[2] Verma, A., Ahuj,a P., Neogi A. pMapper: power and migration cost aware application placement in virtualized systems/ A. Verma, P. Ahuja, A. Neogi/ Proceedings of the 9th ACM/IFIP/USENIX International Conference on Middleware. – 2008. – P. 243-264.

[3] Song, Y., Zhang, Y., Sun, Y., Shi, W. Utility Analysis for Internet-Oriented Server Consolidation in VM-Based Data Centers / Y. Song, Y. Zhang, Y. Sun, W. Shi // Proceedings of the IEEE International Conference on Services Computing. – 2009. – P. 1-10.

[4] Mi, H., Wang, H., Yin, G., Zhou, Y., Shi, D., Yuan, L. Online self-reconfiguration with performance guarantee for energy-efficient large-scale cloud computing data centers/ H. Mi, H. Wang, G. Yin, Y. Zhou, D. Shi, L. Yuan // Proceedings of the IEEE International Conference on Services Computing. – 2010. – P. 514-521.

[5] Armstrong, D., Espling, D., Tordsson, J., Djemame, K., Elmbroth, E. Conteztualization: dynamic configuration of virtual machines. / D. Armstrong, D. Espling, J. Tordsson, K. Djemame , E. Elmbroth // Journal of Cloud Computing: Advances, Systems and Applications – 2015 –P. 4-17.

[6] Taha, H. A. Operation research: An introduction, 6th edition. / Hamdy A. Taha – Moscow: Publishing House «Willams».– 2001. – 912 p.