

в периоды повышенной нагрузки и снижение в периоды ее спада.

Из этого можно сделать вывод, что использование адаптивного механизма предоставления ресурсов является эффективным при организации доступа к web-ресурсу, однако в системе распределенной обработки математических моделей наблюдается избыточное выделение ресурсов.

Наряду с проведенной работой интерес для дальнейшего исследования представляет разработка управляющего правила перерасчета количества ресурсов ВМ при решении вычислительных задач большой размерности.

Вместе с тем важно и внедрение алгоритмов приоритизации в элементе управления допуском, реализованном в адаптивном механизме.

ЛИТЕРАТУРА

1. Бородакий В. Ю., Окороченко Г. Е. Анализ средств имитационного моделирования распределенных информационных систем // Научная сессия МИФИ-2007. Сборник научных трудов. В 15 томах. — Т. 12. Компьютерные системы и технологии. — М.: МИФИ, 2007. — С. 129–130.

2. Вишневецкий В. М. Теоретические основы проектирования компьютерных сетей. — М.: Техносфера, 2003. — 506 с.

3. Пнеденко Б. В., Коваленко И. Н. Введение в теорию массового обслуживания. — М.: Наука, 1966. — 432 с.

4. A. Iosup, O. Sonmez, S. Anoop, and D. Epema, The performance of bags-of-tasks in large-scale distributed systems, in Proceedings of the 17th International Symposium on High Performance Distributed Computing (HPDC'08), 2008;

5. G. Urdaneta, G. Pierre, and M. van Steen, Wikipedia workload analysis for decentralized hosting, Computer Networks, vol. 53, no. 11, 2009. P. 1830–1845;

6. R. N. Calheiros, Rajiv Ranjany, and Rajkumar Buyya Virtual Machine Provisioning Based on Analytical Performance and QoS in Cloud Computing Environments. 2011 International Conference on Parallel Processing, 2011. P. 295–304;

7. R. N. Calheiros, R. Ranjan, A. Beloglazov, C. A. F. De Rose, and R. Buyya. CloudSim: A Toolkit for Modeling and Simulation of Cloud Computing Environments and Evaluation of Resource Provisioning Algorithms, Software: Practice and Experience, 41 (1), 2011. P. 23–50;

8. Tanveer Ahmed, Yogendra Singh Analytic Study Of Load Balancing Techniques Using Tool Cloud Analyst. International Journal of Engineering Research and Applications (IJERA), 2012. P. 1027–1030;

9. Y. C. Lee and A. Zomaya, «Rescheduling for reliable job completion with the support of clouds,» Future Generation Computer Systems, vol. 26, no. 8, 2010. P. 1192–1199;

10. Folding@home. [Электронный ресурс]. 2013. Дата обновления: 22.10.2013. URL: <http://folding.stanford.edu/home/papers> (дата обращения: 22.10.2013). ●

OPTIMIZATION OF SUPPORT OF IT-RESOURCES OF RAILWAYS

Ignatov, Nickolay A. – Ph. D. student at the department of computer systems and networks of Moscow State University of Railway Engineering (MIIT), Moscow, Russia

ABSTRACT

In this article the author focuses on peculiarities of providing virtual resources, used in cloud computing systems for guaranteed quality of service, taking into account the requirements of QoS. The article contains the description of adaptive mechanism and comparative analysis of static and adaptive mechanisms to provide resources through simulation models. Moreover it covers such computing figures for models as the average time for request processing, the level of service denial, the significance of general application of system resources with different inbound parameters.

ENGLISH SUMMARY

Background. Cloud computing is one of the most modern technology, which provides a variety of options for working with data as a network service. Users are released from worries about commissioning or systems administration. However, there are significant problems with the calculation and allocation of resources, primarily, in cloud computing systems. If they are resolved, the work of applications will be able more effective, with reduced financial costs, reallocation of unused resources and increase in productivity in moments of high load.

Objective. The author examines the mechanism of automatic adaptation to changes in workload related to the tasks that are performed by applications in the cloud environment.

Methods. The author uses specific methods of analysis of IT-systems.

Results. System based on cloud computing combines data processing centers (hereinafter – DPC) as a network of virtual IaaS (computing servers, databases, networks) and PaaS (stabilization mechanisms of load (download) and auto-scaling), so that providers are able to open applications (SaaS) from anywhere in the world with a level of costs, determined by the requirements of QoS.

Copies of the applications are examples of SaaS programs, which may belong to small or medium-sized enterprises, as well as public organizations, using applications running through the clouds. In this mechanism the application itself and platform are provided by one organization, and resources based on cloud technologies are used by different organizations.

Calculation of resources needed for the application consists of a set of independent tasks that can be modeled as service requests sent by end-users for copies of virtual applications.

An adaptive mechanism (its architecture is shown in Pic. 1) was developed in order to meet the QoS requirements of the system within the changing nature of the workload, as well as the difficulty of forecasting the volume of virtual resources specific to cloud technology.

It consists of three levels. The service provider administers various software components of the mechanism architecture. SaaS level of the mechanism



contains an admission control element, the main task of which is to admit to the further processing the requests of end users.

The received request is passed to the PaaS level, which comprises several major components.

1) Application provider receives requests of end users from admission control element and provides virtual machines and sample applications, in reliance on data from the workload analyzer and planner of productivity and system workload forecasting (hereinafter – PPSWF).

2) Workload analyzer is designed to make a forecast of the number of requests that enter the system. Such information is used for the calculation of the exact number of applications copies required for the resource allocation and purposes of QoS.

3) Planner of productivity and system workload forecasting deals with analytical modeling and represents the expected level of requests. Its function is to calculate the number of virtual applications copies, related to the tasks of QoS.

These three components of the adaptive mechanism together are able to dynamically adjust the number of virtual applications copies to maintain the desired level of QoS in situations where figures of the processing time and the level of service denial will be below the specified limit.

During the experiments the author tested two scenarios of cloud services use.

Imitation of each scenario was performed 10 times, after which an average figure of each output parameter was estimated.

The first scenario simulates workload of web-service with a lot of requests. Certain production capacity volume of application copies is required for the requests processing (for example, search for web-pages). This scenario is referred to as «Web».

Simulation was to gather statistics of requests in the data center for one week starting from 00.00 am Monday.

Key words: management, information networks, quality of service, cloud computing, virtualization, modeling, distributed systems.

REFERENCES

1. Borodakiy, V. Yu., Okorochenko, G. E. Analysis of the instruments of simulation of distributed information systems. Scientific session of MIFI, 2007. Selected scientific works in 15 volumes. Vol. 12. Computer systems and technology [Analiz sredstv imitatsionnogo modelirovaniya raspredelennykh informatsionnykh sistem. Nauchnaya sessiya MIFI-2007. Sbornik nauchnykh trudov. V 15 tomah. – T. 12. Komp'yuternye sistemy i tehnologii]. Moscow. MIFI publ., 2007, pp. 129–130.
2. Vishnevskiy, V. M. Theoretical fundamentals of designing of computer networks [Teoreticheskie osnovy proektirovaniya komp'yuternykh setey]. Moscow. Tehnosfera publ., 2003, 506 p.
3. Gnedenko, B. V., Kovalenko, I. N. Introduction in the theory of mass services [Vvedenie v teoriyu massovogo obsluzhivaniya. Moscow]. Nauka publ., 1966, 432 p.
4. A. Iosup, O. Sonmez, S. Anoep, and D. Epema, The performance of bags-of-tasks in large-scale distributed systems, in Proceedings of the 17th International Symposium on High Performance Distributed Computing (HPDC'08), 2008.
5. G. Urdaneta, G. Pierre, and M. van Steen, Wikipedia workload analysis for decentralized hosting,

The second scenario simulates workload of applications with a small number of requests which require a significant volume of production capacity of the application copy. Such a scenario is called «Scientific.» This experiment itself shows peculiarities of providing virtual resources for scientific applications (e. g., image processing and modeling of protein folding).

Results of workload simulation for model «Web» are shown in Table 1 and the results of workload simulation for «scientific model» are presented in Table 2.

Conclusions

The architecture of the described adaptive mechanism involves applications launching in the system on virtual machines, which are located on the set of consolidated servers, allowing resource sharing in a cloud infrastructure.

The simulation results show that the adaptive mechanism, that tracks changes in the load intensity for a particular time period, allocates the necessary amount of resources to achieve the desired level of quality of service and efficient allocation of resources in the cloud system.

The application of adaptive mechanism is effective in providing access to web-resource, but there is excessive allocation of resources in a system of distributed processing of mathematical models.

Along with the work undertaken, the development of control rules for resources allocation of virtual machines for computational tasks of high dimensionality is of great interest for further research of the author.

However, it is important to implement prioritizing algorithms in the admission control element, which is applied in the adaptive mechanism.

Computer Networks, vol. 53, no. 11, 2009. P.1830–1845.

6. R. N. Calheiros, Rajiv Ranjany, and Rajkumar Buyya Virtual Machine Provisioning Based on Analytical Performance and QoS in Cloud Computing Environments. 2011 International Conference on Parallel Processing, 2011. P. 295–304.

7. R. N. Calheiros, R. Ranjan, A. Beloglazov, C. A. F. De Rose, and R. Buyya. CloudSim: A Toolkit for Modeling and Simulation of Cloud Computing Environments and Evaluation of Resource Provisioning Algorithms, Software: Practice and Experience, 41 (1), 2011. P. 23–50.

8. Tanveer Ahmed, Yogendra Singh Analytic Study Of Load Balancing Techniques Using Tool Cloud Analyst. International Journal of Engineering Research and Applications (IJERA), 2012. P. 1027–1030.

9. Y. C. Lee and A. Zomaya, «Rescheduling for reliable job completion with the support of clouds,» Future Generation Computer Systems, vol. 26, no. 8, 2010. P. 1192–1199.

10. Folding@home. URL: <http://folding.stanford.edu/home/papers> (Last accessed 22.10.2013).

Координаты автора (contact information): Игнатов Н. А. (Ignatov N. A.) – proterian2000@yandex.ru
Статья поступила в редакцию / article received 30.10.2013
Принята к публикации / article accepted 21.01.2014