iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 4 **Issue:** III **Month of publication:** March 2016

**DOI:**

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

# A Noval Architecture to Enable Both Cloud Providers and Consumers to Monitor and Control the Cloud Resources

S.Subhashini[1], Mrs. M.Suguna[2]
*Department of CSE, Kumaraguru College of Engineering*

***Abstract-- Cloud computing resource monitoring is an important process in the cloud resource management to provide the efficient and flexible services without interruption. It would be more complex process to handle and allowing the resource monitoring services to both cloud consumers and the cloud service providers. In this work, secured cloud resource monitoring architecture is implemented for both consumer and providers namely MonPaas, which is an integrated open source software of Nagios and OpenStack. This is an adaptive distributed monitoring system which provides a logical view of the overall cloud environment to the cloud service providers and resource details to the cloud consumers. It also enables cloud consumers to find and control the other resources which can get information about their resources. The experimental tests conducted were proves that the proposed approach provides better result than the existing approach in terms of their improved monitoring service.***
***Keywords-- Cloud computing, Monitoring, Control, Resource***

## I. INTRODUCTION

Cloud computing is a favourable computing technique which recently has drawn extensive attention from both industry and academia. By merging a set of new and existing techniques from research areas such as Service-Oriented Architectures and virtualization, cloud computing is considered as such a computing paradigm in which resources in the computing infrastructure are given as services in the Internet. Some of the successful examples are S3 and Amazon's EC2, Google App Engine, and Microsoft Azure which provide users with a pay-as-you use fashion with scalable resources at relatively low prices. For example, Amazon'sS3 data storage service just charges certain cost per giga byte month. When compared to building their own infrastructures, users can save their investments a lot by migrating businesses into the cloud. With the increasing development in cloud computing technologies, it is so easy to visualize that in the near future more and more businesses will be moved into the cloud.

Number of researchers have on prior discussed access control issues for cloud computing. Daniel Nurmi and his colleagues gave an authentication system to mitigate the performance of the virtual machines (VMs) to assure that only administrators and owners could access them. Stefan Berger and his colleagues presented an authentication model based on both security labels and role-based access control (RBAC) to the control access to shared data, VMs, and network resources. Jose Alcaraz Calero and his colleagues invented a centralized authentication system that provides an organized path-based access control mechanism. It distinguishes this work is that an architecture which using an XML-based formalism is presented which can be implemented. We also address the problems of side-channel attacks and non interference in the presence of resource virtualization and multi tenancy.

## II. LITERATURE REVIEW

### *A. Underutilizing Resources for HPC on Clouds by J. Napper, P. Bientinesi, R. Iakymchuk*

The cloud computing model states the ability to scale compute resources on demand. Unlike conventional cluster systems, there are no significant upfront monetary or time investments in infrastructure or people. Instead of allocating the resources according to average or high load, the cloud user can pay costs according to current need. When resources are not used, total cost is considered as zero. Individuals can quickly scale-up and create a custom compute cluster, only paying for sporadic usage. However, there are also some disadvantages in cloud gauge services. The costs can be sub-divided into number of categories that are billed separately: for example, network, storage, and CPU usage. These two factors imply that resources have to be very conservatively scaled in current clouds, so as to reduce some benefits of scaling on demand. Eventually, in many cloud environments, physical resources are shared within the virtual nodes relying with the same or different users, which can impact performance negatively.

The article actually explores computational efficiency in a cloud environment when resources are underutilized. An observation of abnormal behaviour in the average performance on EC2: the pinnacle average performance is reached when only a part of the available resources is used, typically 25–50%. This behaviour occurs both on clusters and in single node performance of up to 32 compute cores. The solution can be obtained soon, thereby lowering the corresponding cost. Finally, it is shown that there is still available parallelism while underutilizing the resources. In some cases, adding an extra node to the cluster is absolutely free: the computation completes enough earlier to reimburse for the marginal price of the extra node.

EC2 instances on dense linear algebra algorithms. Initially, the study with the single node performance is undergone. The main goal is to analyse the firmness of achievable performance and to characterize how performance varies with the number of utilized cores. In order to compute the consistency of EC2 performance, the compute intensive applications similar to the DGEMM matrix- the matrix multiplication kernel and High-Performance Linpack (HPL) benchmark—based is executed on LU factorization—for 24 hours, repeating the experiments over different days. The results of DGEMM are presented first and then discuss the results for HPL.

*B. Resource Provisioning for Cloud Computing by J. Wong, G. Iszlai, M.l. Ye Hu,*

Static resource allocation is not cost-effective based on peak demand because of poor resource usage during off-peak periods. This work is regarded with resource allocation techniques that are relevant to autonomic resource management. The duo-stage resource management architecture provides a framework for our investigation. There are multiple application environments (AEs) at the lower level. Each AE consists of a set of computing resources those are shared by one or more than one application. At higher levels, a global arbiter undertakes resource allocation across AEs. In this paper, the processing of interactive jobs only is considered. These jobs generally have small processing necessities and good response time performance is needed. The SLAs considered are based on the probability distribution on the time of response, such as, Probability [response time $\leq x] \geq y$ where $x$ is a threshold value and $y$ is the target probability. Our approach is to use performance models to obtain results that can be used to guide resource allocation commitments. In our scrutinization, the computing resources at each AE are modelled by servers. When the global arbiter does resource allocation decisions, statistics on the number of servers that should be allocated to each AE would be very useful. This represents to significantly lower number of servers needed to meet the SLAs of all applications that are assigned to the AE.

Cloud computing infrastructure may provide service to a large number of job classes. Results on the performance difference between DA and SA for a random number of classes are difficult to obtain. This is because of the potentially huge number of possible allocation techniques that need to be evaluated. Additional complexity is introduced when the effect of scheduling discipline is incorporated in the investigation. To keep the complexity at a manageable level, the special case of two job classes is considered. In spite of this simplification, our results are directly pertained when the global arbiter, considering consideration issues such as application isolation, management and security, determines to utilize a divide-and-conquer approach where an AE contains at most two job classes. In addition, our results provide valuable perception into the performance of substitute resource allocation strategies and job scheduling disciplines, and could be made use of to spread heuristic methods for resource allocation when more than two classes are assigned to an AE.

*C. Efficient Resource Provisioning in Compute Clouds Via VM Multiplexing by Xiaoqiao Meng, Canturk Isci, Jeffrey Kephart, Li Zhang, Eric Bouillet*

VM sizing is perhaps the most vital step in both static and dynamic provisioning. VM sizing refers to the estimation of the amount of resources that should be allocated to a VM. The theme of VM sizing is to promise that VM capacity is equivalent to the workload. Whereas, over-provisioning wastes the high-cost resources, under-provisioning devalue application performance and may lead to losing customers. Actually, VM sizing is achieved through on a VM-by-VM basis, i.e., each VM has an estimated size on the basis of its pattern of workload. On a significant departure from such an individual-VM based approach, a joint-VM provisioning approach is advocated in which several VMs are consolidated and provisioned based on an estimate of their aggregate capacity requirements. According to the concept, joint-VM provisioning among the dynamic VM demand characteristics makes use of the statistical multiplexing, i.e., the valleys and peaks in a VM's demand do not coincide with the other VMs necessarily. The unused resources of a low used VM can then be directed to the other nearby located VMs when the utilization is at peak. Thus, VM multiplexing paths to a specifically significant capacity saving when compared to individual-VM based furnishings. The savings achieved by multiplexing are perceived by packing VMs densely into the hardware resources without sacrificing or losing application performance commitment. This increases the overall consolidation ratio; the virtualization overheads present in

scheduling a small amount of higher number of VMs is generally less since the VM footprints fit in the provisioned capacity.

*D. Automated and On-Demand Provisioning of Virtual Machines for Database Applications-Piyush Shivam, Azbayar, David Irwin, Laura Grit, Shivnath Babu, and Jeff Chase-2007*

Utility computing circulates compute and storage resources to applications as an 'on-demand utility' alike to an electricity grid. In this work the NIMO system is demonstrated that uses dynamic and accelerated learning mechanisms to learn achievement techniques automatically for database applications. The learned techniques are used to provision VMs for these applications in a utility setting, including (i) application performance of VM sizes based on ranking available, (ii) performance target that determines the VM sizes, and (iii) answering what-if queries. By using computational science applications as well as multi-tier web services is used for demonstration. The result of this work demonstrates System performance goals of this work based on accuracy: The NIMO system focuses of this demonstration-builds better performance models in terms of accuracy that capture the interaction among input factors, and the resulting application performance.

*E. Towards Autonomic Workload Provisioning for Enterprise Grids and Clouds-Andres Quiroz, Hyunjoo Kim, Manish Parashar, Nathan Gnanasambandam, Naveen Sharma-2009*

In this work a decentralized robust online clustering approach is presented that addresses the dispersive nature of these infrastructures; that is to identify patterns and trends, and which is most useful to hike provisioning of virtual (VM) resources. It then presents evaluating application service time using the long-term application performance monitoring; model-based approach is used. To provide feedback about the applicable requested resources the system's ability to meet QoS constraints and SLAs are established. The use of a quadratic response surface model (QRSM) is validating with tribute to traditional models, demonstrating the necessity for application-specific modelling for high-performance computing workloads. The result of this work demonstrates Average difference between requested resource (CPU speed (top) and memory (bottom)) for jobs and the corresponding analysis result in each time window for DOC, Overprov and k-means. Thus the presented approaches are evaluated using a real computing centre workload trace and the results demonstrate that the presented DOC shows better result in both effectiveness and cost-efficiency.

*F. Combinatorial Auction-Based Protocols for Resource Allocation in Grids-Anubhav Das, Daniel Grosu-2005*

Grid systems are meant for next generation computing platforms in order to providing solution for large scale problems in science and engineering. In this work the combinatorial auction model is introduced for resource management in grids. A combinatorial auction-based resource allocation protocol is presented in which a user offers a price value for each of the most feasible integration of resources desire for its tasks execution. The protocol comprises an approximation algorithm in order to resolve the combinatorial auction problem. The result of this work demonstrates Users payments, Resource profits, Payment structure for each resource and Resource utilization for each resource. The protocol is beneficial for each group of resources and the profit of all the resources as a whole is maximized and it allocates tasks constantly among resources within a family thus making use of fairly these resources.

*G. Gridecon: a Market Place for Computing Resources- Jörn Altmann, Costas Courcoubetis, George D. Stamoulis, Manos Dramitinos, Thierry Rayna[4], Marcel Risch, Chris Bannink-2008*

The Gridecon project forms the technology that is demand to create an efficient market place for trading commoditized computing resources. The market place issues that each and every owner of computing resources to offer spare computing resources as a standardized virtual machine. This work discusses the rationales for a Grid market and, in particular, the introduction of a market place for trading material computing resources. The market place presented computing resources from different providers substitutable through virtualization. The results of the present work demonstrate by using matching procedure which is subsequently delivered to the scheduler, the reservation system of the provider, and the accounting system of the market place. Thus the method results and includes the definition of a spot and future market as well as the prerequisites for a market mechanism for computing resources.

*H. Dynamic Resource Allocation For Spot Markets in Clouds-Qi Zhang, Eren Gurses, Raouf Boutaba, Jin Xiao-2011*

Cloud computing provides on-demand provisioning of resource to applications and services. In this work a cloud management framework is presented that dynamically allocates data centre resources to spot markets to maximize cloud provider's total revenue.

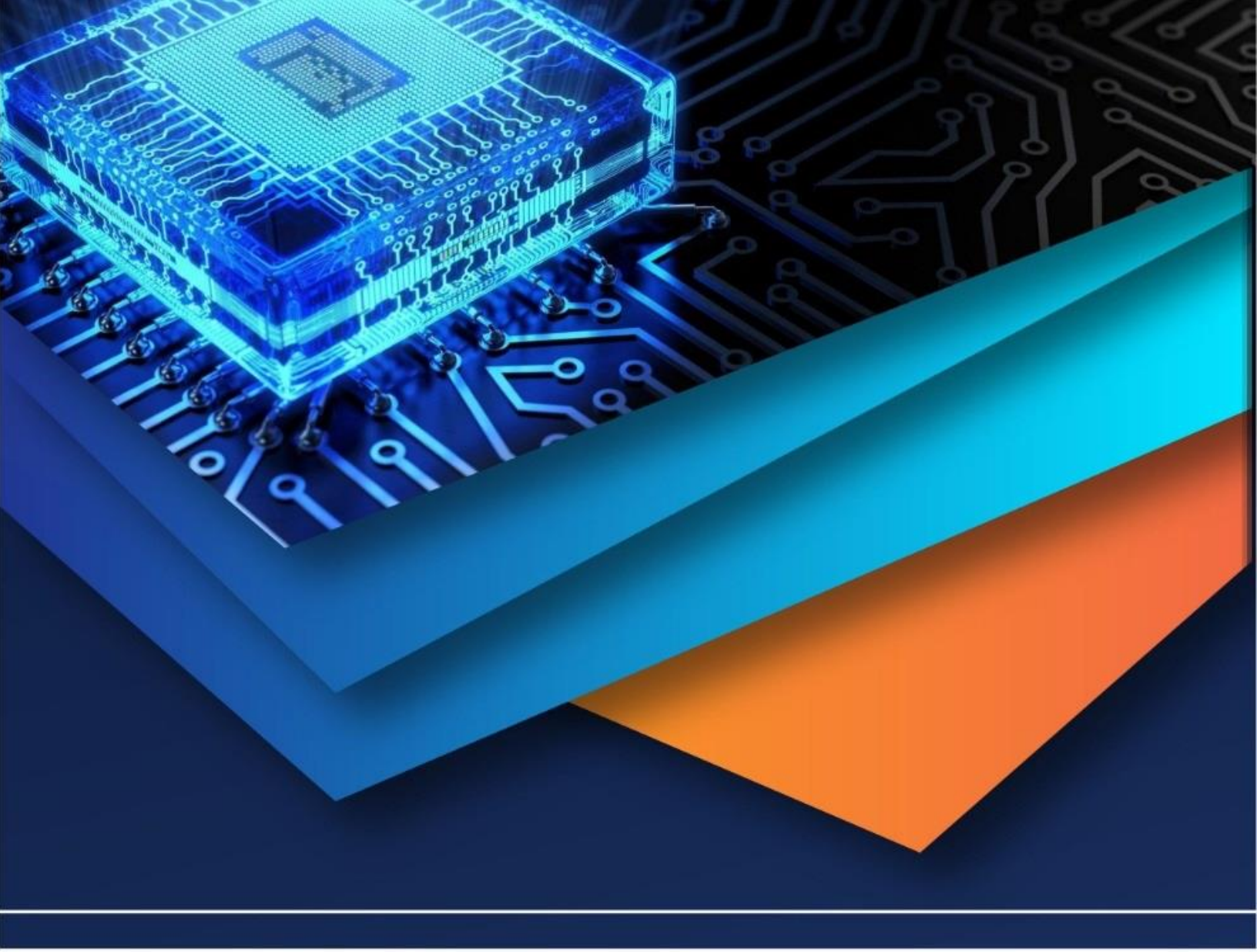
This presents a solution consists of 2 parts: (1) market analysis for forecasting the demand for each and every spot market, and (2) a dynamic scheduling and stabilized mechanism that allocate resource to each and every spot market to maximize entire revenue. The effectively allocating resources for revenue maximization is a NP-hard problem, this shows that presented algorithms can rough the optimal solutions to this complication under both fixed and variable pricing schemes. The result of this work demonstrates the Income rate with static and fluctuating demands for Static Allocation Policy and Dynamic Allocation Policy. Thus It is observed that the dynamic allocation policy causes more revenue loss, but this negative impact is outweighed by the positive revenue gain.

## III. CONCLUSION

Distributed resource auditing and prediction architecture is proposed that seamlessly combines cloud technologies, resource auditing, and machine learning-based resource state prediction. This system consists of a set of distributed services to perform all required resource auditing, data gathering, and resource state prediction functions. In this work, integrated architecture of nagios and the open stack software framework for monitoring cloud services available in the environment. The experimental tests conducted were proves that the proposed approach leads to a better result than the existing work in terms of improved result.

## REFERENCES

[1] R. Iakymchuk, J. Napper and P. Bientinesi, “Underutilizing Resources for HPC on Clouds”, roceedings of the 2011 ACM Symposium on Applied Computing, Pages 119-126, 2011.
[2] Ye Hu, Johnny Wong, Gabriel Iszlai and Marin Litoiu, “Resource Provisioning for Cloud Computing”, Proceedings of the 2009 Conference of the Center for Advanced Studies on Collaborative Research, Pages 101-111, 2009.
[3] Xiaoqiao Meng, Canturk Isci, Jeffrey Kephart, Li Zhang, Eric Bouillet, Dimitrios Pendarakis, “Efficient Resource Provisioning in Compute Clouds via VM Multiplexing”, Proceedings of the 7th international conference on Autonomic computing, Pages 11-20, 2010.
[4] Sanjay Chaudhary, “Saas (Software as A Service) On Cloud”, Wikipedia.
[5] Alexander Keller and Remi Badonnel, “Automating the Provisioning of Application Services with the BPEL4WS Workflow Language”, Proceedings of the 15th IFIP/IEEE International Workshop on Distributed Systems: Operations & Management (DSOM 2004), Davis, CA, USA, November 2004.
[6] Piyush Shivam, Azbayar Demberel, Pradeep Gunda, David Irwin, Laura Grit, Aydan Yumerefendi, Shivnath Babu, and Jeff Chase, “Automated and On-Demand Provisioning of Virtual Machines for Database Applications”, ACM 978-1-59593-686-8/07/0006, 2010.
[7] Andres Quiroz, Hyunjoo Kim, Manish Parashar, Nathan Gnanasambandam, Naveen Sharma, “Towards Autonomic Workload Provisioning for Enterprise Grids and Clouds”, 10th IEEE/ACM International Conference on Grid Computing, 2009.
[8] Anubhav Das, Daniel Grosu, “Combinatorial Auction-Based Protocols for Resource Allocation in Grids”, Parallel and Distributed Processing Symposium, 2005. Proceedings. 19th IEEE International 4-8 April 2005.
[9] Jörn Altmann, Costas Courcoubetis, George D. Stamoulis, Manos Dramitinos, Thierry Rayna, Marcel Risch, Chris Bannink, “GridEcon: A Market Place for Computing Resources”, Grid Economics and Business Models, Volume 5206 of the series Lecture Notes in Computer Science, 2013.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY