2012 10th IEEE International Symposium on Parallel and Distributed Processing with Applications

iCanCloud: A brief architecture overview

Gabriel G. Castañé
Dept. de Informática.
Universidad Carlos III de Madrid
Madrid, Spain
gabriel@arcos.inf.uc3m.es

Alberto Núñez
Dept. de Sistemas Informáticos
y Computación
Universidad Complutense de Madrid
Madrid, Spain
alberto.nunez@pdi.ucm.es

Jesús Carretero
Dept. de Informática.
Universidad Carlos III de Madrid
Madrid, Spain.
jcarrete@arcos.inf.uc3m.es

***Abstract*—During the last years the use of cloud computing environments for both researchers and enterprises is a major trend. On the one hand, cloud computing is a flexible and scalable paradigm that offers hardware and software services to be purchased by users. On the other hand, users have extra costs in several procedures, like calculating how many resources are needed by the applications. These tasks could be expensive and time-consuming.**

In this work we present iCanCloud, a simulation platform for modeling and simulating cloud environments. iCanCloud provides a complete set of modules for modeling with high level detail, a cloud computing environments and their underlying architecture.

I. Introduction

Currently, researchers and enterprises require powerful systems to compute huge data-sets in a feasible time-frame. The services offered by *cloud computing* environments let users to launch applications by using the resources they need without any knowledge about its internal infrastructure.

The underlying architectures of a cloud computing system can be very different depending on the owners of the cloud system. Those differences hampers to archive performance analysis of applications executed in a cloud. In addition, it would still be very difficult to reproduce performance evaluation experiments in a repeatable and controlled manner, due to the inherent variability of the cloud. At present, there are many research topics in this area: the management of data that an application can generate [1], the performance analysis of production clouds [2][3].

The inherent pay-as-you-go system for business adopted by cloud providers also hampers to obtain the right balance between costs and performance. Once an application is executed, users don't have any acknowledge about how many resources are needed by their applications. The scheduling experiments have the same issue. In order to know the behavior in management brokering policies and cloud managers, a high number of executions have to be launched in a real cloud computing environment. Thus, these experiments add extra costs.

In order to alleviate these issues in a real cloud environment, we propose the use of simulation techniques. The iCanCloud simulation platform has been designed to model the behavior of cloud computing environments and their underlying architecture. iCanCloud provides a complete set of modules: nodes, networks, virtual machines, users, applications, scheduling policies of the cloud environment and the hypervisor to provide users a high level detail without loose accuracy.

II. iCanCloud Overview

The iCanCloud simulation platform is oriented towards the simulation of a wide range of cloud computing systems and their underlying architectures. This project was initiated in 2010 and it is currently available at [4] as open source software.

iCanCloud has been designed to obtain a good trade-off between flexibility, accuracy, performance and scalability, which makes it a powerful simulation platform for designing, testing and analyzing both actual and non-existent architectures. In fact, complete high performance computing systems can be modeled using this simulation platform. The best feature of iCanCloud is its ability to model and simulate large environments (thousands of nodes) with a customizable level of detail. Distributed applications can be simulated using this framework.

III. iCanCloud Simulator Architecture

iCanCloud has been developed over the OMNeT++ platform. The network systems relies on the simulation framework simulator INET. The set of modules provided by the iCanCloud simulator kernel simulates the behavior of specific components. These modules are grouped by functionality. The modular architecture of iCanCloud simulation platform consists of three main sections, see Figure 1): user model, cloud infrastructure and cloud manager.

A. User model

The cloud job module allows users to model applications. Using this module, users can simulate applications models like map-reduce. iCanCloud offers in its repositories several applications as examples.

The cloud user module represents users who launch a set of jobs in a cloud computing environment. These jobs are executed in the virtual machines rented by users. The set of virtual machines is defined by the pair: (type, quantity), where type is the name of a virtual machine configuration allocated in the repository of iCanCloud and quantity is the number of virtual machines with the same configuration.

978-0-7695-4701-5/12 $26.00 © 2012 IEEE
DOI 10.1109/ISPA.2012.131

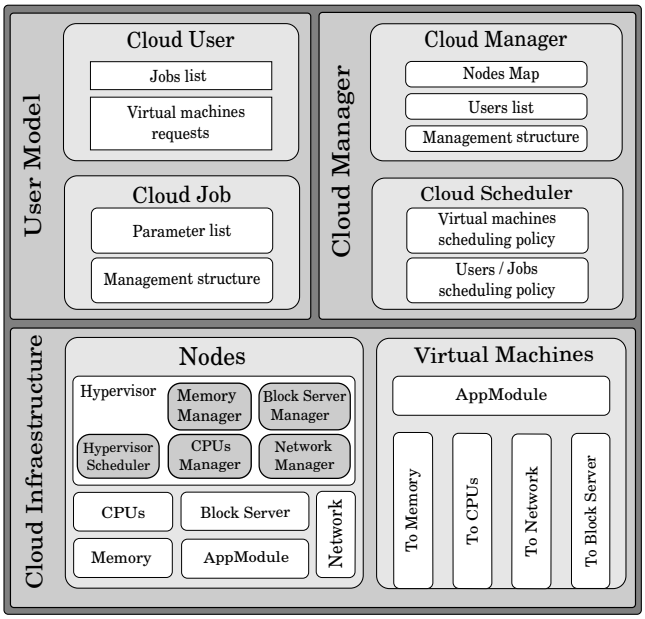

Fig. 1. Modular architecture of iCanCloud

The iCanCloud simulation platform has internal structures into user module with the purpose to link the assigned virtual machine modules. These structures provide mechanisms for searching the modules by allowing fast access to the virtual machines.

B. Cloud infrastructure model

The node module represents a physical machine. This node consists of: one memory module, one CPU module, one block server module and the network module. All of them are linked by the hypervisor scheduler module. This hypervisor is in charge of communicating and managing the access of other nodes and virtual machines to the hardware layer. In order to carry out this purpose, the hypervisor module is divided in four managers and one scheduler. Each manager (memory manager, CPU manager, block server manager and network manager) connects the applications allocated in the virtual machines with the hardware modules. Once a manager receive a request, this request is multiplexed by the managers to the hardware modules. The behavior of these managers is settled by the policies of the hypervisor scheduler module.

The structure of the virtual machine module is similar to the node module. The main difference lies in the modules that simulates the hardware components. Once an application is allocated in a virtual machine, it demands hardware resources. The request received is redirected to the hypervisor (previously selected by the cloud manager module) of a specific node. The behavior of the virtual machines is defined by user with a set of parameters. The set of Amazon EC2 [5] virtual machines types are provided by iCanCloud simulator repositories.

C. Cloud manager model

The cloud manager module is in charge of different tasks:

- The management of the cloud users by obtaining their virtual machines and cloud jobs configurations.
- The dynamic creation of the virtual machines by cloning the initial defined instances to allocate it into a node.
- The allocation of defined jobs by users into the applications list of the virtual machines.
- The results storage of a finished job.
- The destruction of a cloud job module when the job has finished.
- The destruction of the cloned virtual machines if the user doesn't need them.

The cloud scheduler module is responsible of selecting the user who is waiting to launch an application in the cloud environment. This scheduler contains the methods to select a job from a user's jobs list. These methods provide the information about the node where the manager is going to allocate the new virtual machine instance.
Users can analyze new scheduling policies by implementing this module. In addition, the behavior of the schedulers can be modified by using the flexible API offered by iCanCloud.

IV. Conclusions and Future work

In this paper we presented a simulation platform to simulate cloud computing environments. The modular design of iCanCloud provides a high level of detail without decreasing accuracy. Among others iCanCloud is flexible and scalable to let users to model different types of clouds and to embrace a wide scope of researches and studies.

References

[1] Jim Gray, David T. Liu, Maria A. Nieto-Santisteban, Alexander S. Szalay, David J. DeWitt, Gerd Heber: Scientific Data Management in the Coming Decade. SIGMOD Record, 34(4):34-41 (2005).
[2] Yigitbasi, N.; Iosup, A.; Epema, D.; Ostermann, S.; , "C-Meter: A Framework for Performance Analysis of Computing Clouds," Cluster Computing and the Grid, 2009. CCGRID '09. 9th IEEE/ACM International Symposium on , vol., no., pp.472-477, 18-21 May 2009 doi: 10.1109/CCGRID.2009.40
[3] G. Wang. A.R. Butt. P. Pandey. K. Gupta, "A simulation approach to evaluating design decisions in MapReduce setups," Modeling, Analysis & Simulation of Computer and Telecommunication Systems, 2009. MASCOTS '09. IEEE International Symposium on , vol., no., pp.1-11, 21-23 Sept. 2009 doi: 10.1109/MASCOT.2009.5366973¡¡
[4] iCanCloud simulator website. Jan 2012. http://www.icancloudsim.org/ .
[5] Amazon. Amazon Elastic Compute Cloud (Amazon EC2). http://aws.amazon.com/ ec2/.