Accepted Manuscript

A framework for simulating large scale cloud infrastructures

Christos K. Filelis-Papadopoulos, George A. Gravvanis, Panagiotis E. Kyziropoulos

PII: S0167-739X(17)30323-0
DOI: http://dx.doi.org/10.1016/j.future.2017.06.017
Reference: FUTURE 3515

To appear in: *Future Generation Computer Systems*

Received date : 28 February 2017
Revised date : 8 May 2017
Accepted date : 16 June 2017

Please cite this article as: C.K. Filelis-Papadopoulos, G.A. Gravvanis, P.E. Kyziropoulos, A framework for simulating large scale cloud infrastructures, *Future Generation Computer Systems* (2017), http://dx.doi.org/10.1016/j.future.2017.06.017

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

A framework for simulating large scale cloud infrastructures

Christos K. Filelis-Papadopoulos, George A. Gravvanis and Panagiotis E. Kyziropoulos
Department of Electrical and Computer Engineering, School of Engineering, Democritus University of Thrace, University Campus, Kimmeria, GR 67100 Xanthi, Greece; Email: {cpapad, ggravvan, pkyzirop}@ee.duth.gr

Abstract
Cloud infrastructures are continuously growing in size, since more cloud nodes are added to already existing hyper-scale infrastructures. These hyper-scale infrastructures are also becoming heterogeneous as different types of accelerators are added in order to increase performance per watt for certain types of applications and allow for various HPC workloads to migrate to Cloud environments. The introduction of diverse workloads that migrate in the Cloud along with increasing volume of incoming tasks results in phenomena of network congestion, underutilization and resource fragmentation. Simulators are used to analyze, study and possibly improve Cloud environments. However, existing Cloud simulation tools lack the ability to handle heterogeneous resources and tasks that span across multiple Cloud nodes. Moreover, they are mostly sequential and cannot scale to large numbers of Cloud nodes. Furthermore, they do not support over-commitment, which is a common practice in real-world Cloud environments. A framework for simulating large numbers of heterogeneous cloud nodes organized in Cells and executing large numbers of HPC tasks is proposed. The framework is inherently parallel and designed for hybrid distributed memory parallel systems, supporting CPU, memory and network over-commitment. The simulation framework is based on a time advancing loop, allowing dynamic change of the granularity of the simulator and minimizing memory requirements, since data related to the current time-step is stored. Moreover, a latency model for the currency of data in the Gateway Service and Broker is also supported. Implementation details along with discussions concerning the extensibility of the framework are given. Numerical results for simulating large number of heterogeneous resources and incoming tasks are also presented.

Keywords: cloud simulation, hyper-scale data centers, power consumption, cloud node, virtualization

1. Introduction
During the last decade, Cloud computing has become increasingly popular, due to the flexibility of allocating resources and deploying services. These advantages lead a large number of applications and services to migrate in Cloud environments. It is estimated that until 2020, 92% of the total workloads will be executed in Cloud environments [1]. The growth in diversity and volume of the migrating applications of Cloud infrastructures and end-users leads to gradually growing underlying infrastructures. These infrastructures, referred as “datacenters”, are continuously growing in complexity and size. Hyper-scale datacenters are expected to host 47% of all installed datacenter servers, by 2020. Cloud infrastructures are composed of groups of datacenters often in distinct geographical locations. The architecture of modern datacenters can be considered as a variant of the Warehouse Scale Computer (WSC) proposed by Barroso et al. [2]. The typical architecture of a WSC, or more generally a Cloud, is composed of four distinct components: a front end platform, back end platforms, a cloud based service delivery model and a network [2,3]. The front end platform is composed of client devices such as fat clients, thin clients and mobile devices. The back end platforms are composed of servers and storage. Current cloud deployments are predominantly homogeneous in terms of hardware, composed of CPU-based servers. Recently, specialized,

Download English Version:

https://daneshyari.com/en/article/6873393

Download Persian Version:

https://daneshyari.com/article/6873393

Daneshyari.com