DCSim: A Data Centre Simulation Tool

Gastón Keller*, Michael Tighe*, Hanan Lutfiyya and Michael Bauer
Department of Computer Science
The University of Western Ontario
London, Canada
{gkeller2|mtighe2|hanan|bauer}@csd.uwo.ca

***Abstract*—Developing algorithms to dynamically manage resources in a virtualized, multi-tenant data centre is challenging. Experimenting with such algorithms on the data centre scale is impractical due to size and complexity. Thus, there is a need for simulation tools to allow rapid development and evaluation of management techniques. We present DCSim, an extensible framework for simulating a multi-tenant, virtualized data centre.**

***Keywords*—Cloud, Data Centre, Simulator, Virtualization, Infrastructure as a Service**

I. Introduction

Algorithms for dynamic resource management in the data centre have proven difficult to evaluate due to the scale and complexity of the infrastructure on which they are intended to run. As such, simulation is becoming accepted as a means of rapidly evaluating new techniques at a speed and scale not possible with real implementations. Once a technique has been evaluated and fine tuned using a simulation, further experimentation can be performed using a real infrastructure, albeit very likely on a much smaller scale.

There is currently a lack of easily customized and extensible simulation tools that model a virtualized, multi-tenant data centre. Furthermore, such a tool must provide an application model that can simulate the interactions and dependencies between many VMs working together to provide a single service, such as in the case of a multi-tiered web application. Other features of virtualization, such as a work conserving CPU scheduler used in modern hypervisors, resource allocation and VM migration and replication must also be available. Finally, host power states (on, off, suspended) must be modelled with appropriate transition times between states. We present a new simulator, DCSim (Data Centre Simulator), designed specifically to address these requirements. DCSim is an extensible simulation framework designed to study VM management in a data centre providing an IaaS Cloud.

II. Related Work

Most data centre simulators available today do not include any modelling of virtualization. GreenCloud [1] is a packet-level data centre simulator designed to evaluate the energy costs of data centre operation. It utilizes a highly detailed simulation of network communication. MDCSim [2] is a simulation platform designed to simulate large scale, multi-tier data centres. It focuses on data centre architecture and cluster configuration, measuring both performance and power metrics. It supports three-tiered web applications, with the ability to modify and evaluate the configuration of each tier. GDCSim (Green Data Centre Simulator) [3] aims to simulate both the management and physical design of a data centre, examining the interactions and relationships between the two. The goal is to fine-tune the interactions between management algorithms and the physical layout of the data centre, such as thermal and cooling interactions with workload placement.

CloudSim [4] is a toolkit for simulating a data centre hosting a virtualized IaaS Cloud. Multiple users can create VMs within the data centre. CloudSim implements an HPC-style workload, which can be used to simulate a transactional, continuous workload such as a web server or other service [5], but it lacks a detailed model of such an application.

DCSim differs from GreenCloud, MDCSim, and GDCSim in that it is focused on a virtualized data centre providing IaaS to multiple tenants, similar to CloudSim. It differs from CloudSim in that it focuses on transactional, continuous workloads. As such, DCSim provides the additional capability of modelling replicated VMs sharing incoming workload as well as dependencies between VMs that are part of a multi-tiered application. SLA achievement can also be more directly and easily measured and available to management elements within the simulation.

III. DCSim Architecture and Features

DCSim simulates a virtualized data centre operating an Infrastructure as a Service (IaaS) cloud. In DCSim [6], [7], a data centre consists of a set of interconnected Hosts (physical machines), governed by a set of Management Policies. Each host has a set of Resource Managers that handle local resource allocation, a CPU Scheduler which decides how VMs will execute, and a Power Model which models how much power is being consumed by the host at any point in time. Each host runs a set of VMs, which in turn each run a single Application. Figure 1 outlines the general architecture of DCSim.

Virtual Machine management operations, such as VM live migration and replication, are supported within DCSim. The resource needs of each VM in DCSim are driven dynamically by an Application, which varies the level of resources required by the VM to simulate a real workload. The Application class is abstract and can be extended to implement different types of applications, but the primary application model implemented in DCSim mimics a continuous, transactional workload, such as a web server. Incoming work can be load balanced between multiple application instances running in separate VMs, and multi-tiered applications can also be modelled.

DCSim reports a number of metrics in order to help determine the behaviour of the data centre during the simulation.

* Both authors contributed equally to this work.

978-3-901882-50-0

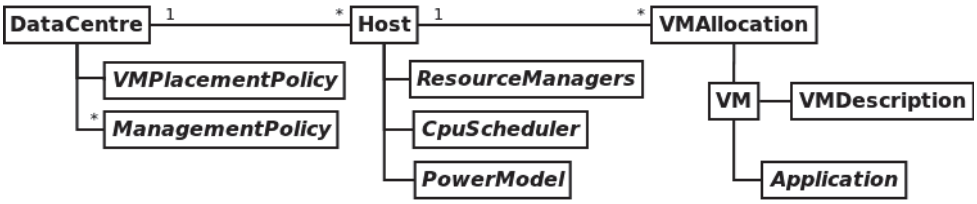

Fig. 1. DCSim Architecture

The following section outlines some of the metrics currently being computed for each simulation.

SLA Violation: When a VM requires more resources than is available to it, the VM will experience performance degradation. DCSim considers this to be an SLA (Service Level Agreement) violation, and the percentage of CPU resource required by running VMs but not available is recorded.

Data Centre Utilization: The overall utilization of the data centre is calculated as the percentage of total CPU capacity in the data centre that is currently in use.

Active Hosts: DCSim records the minimum, maximum, and average number of hosts that are *On* at any given time during the simulation.

Host-hours: Host-hours is the combined total of the active time of every host in the simulation. That is, if 10 hosts were active for 30 simulation minutes each, then 5 host-hours were used. This gives a combined measure of the number of hosts that were required to meet the workload demand throughout the entire simulation run.

Active Host Utilization: DCSim measures the CPU utilization of all hosts that are currently in the *On* state. The higher the average utilization, the more efficiently resources are being used.

Number of Migrations: The number of VM migrations triggered during the simulation is recorded.

Power Consumption: Power consumption is calculated for each host, and the total kilowatt-hours consumed during the simulation are reported.

DCSim generates a log containing details of the simulation as it progressed, including metric values at regular time intervals, host and VM states, and other events. This log can be used to generate statistics, graphs, or other visualizations in order to fully understand the behaviour of the simulated data centre and management algorithms. A visualization tool is included with DCSim which automatically generates a set of useful graphs based on the simulation log.

IV. Future Work

A number of additional features could be added to DCSim in order to improve the accuracy and usefulness of the simulation. A model of data centre networking should be added in order to make more intelligent decisions in VM placement, as well as considering bandwidth constraints on concurrent VM migrations.

Currently, DCSim simulates a data centre with a centralized management system. We intend to make it possible to simulate a distributed, decentralized management system, as such a system may provide advantages in terms of management overhead.

An interesting metric to consider when placing VMs and determining whether to power on and off individual hosts is the heat currently being generated by each physical machine. A thermal model of the data centre could be incorporated into the simulation in order to allow management algorithms to consider this factor.

Physical machine failure is a common occurrence in large data centres, and must be considered in management algorithms. The ability to simulate host failures should be included within the simulation in order to evaluate the ability of management algorithms to handle such situations.

Acknowledgements

We thank the National Sciences and Engineering Research Council of Canada (NSERC) for their support.

References

[1] D. Kliazovich, P. Bouvry, Y. Audzevich, and S. Khan, "Greencloud: A packet-level simulator of energy-aware cloud computing data centers," in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, 2010.

[2] S.-H. Lim, B. Sharma, G. Nam, E. K. Kim, and C. Das, "Mdcsim: A multi-tier data center simulation, platform," in *Cluster Computing and Workshops, 2009. CLUSTER '09. IEEE International Conference on*, 2009.

[3] S. Gupta, R. Gilbert, A. Banerjee, Z. Abbasi, T. Mukherjee, and G. Varsamopoulos, "Gdcsim: A tool for analyzing green data center design and resource management techniques," in *Green Computing Conference and Workshops (IGCC), 2011 International*, 2011.

[4] 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," *Softw. Pract. Exper.*, vol. 41, no. 1, pp. 23–50, Jan. 2011.

[5] A. Beloglazov and R. Buyya, "Energy efficient resource management in virtualized cloud data centers," in *Cluster, Cloud and Grid Computing (CCGrid), 2010 10th IEEE/ACM International Conference on*, 2010.

[6] M. Tighe, G. Keller, M. Bauer, and H. Lutfiyya, "DCSim: A data centre simulation tool for evaluating dynamic virtualized resource management," in *SVM Proceedings, 6th Int. DMTF Academic Alliance Workshop on*, Oct. 2012.

[7] DCSim project site. Distributed and Grid Systems (DiGS) Research Group, Western University. [Online]. Available: http://digs.csd.uwo.ca/