Towards Understanding the Workload of an Infrastructure as a Service Platform

Loïc Perennou, PhD student, Outscale & ISEP
Sylvain Lefebvre, PhD, ISEP
Mar Callau-Zori, PhD, Outscale

Outscale is a public Infrastructure as a Service (IaaS) provider established in 2010. IaaS helps organizations achieve business agility by offering on-demand computing resources [1]. Outscale and its mother company Dassault Systemes team up to provide collaborative solutions for industries and smart cities, through the 3DExperience platform running in the cloud. Outscale has 12 points of presence in 3 regions: North America, Europe, and Asia. It invests 15% of its revenues back into R&D to develop its own API-driven and Amazon EC2-compatible cloud orchestrator, called TINA OS [2].

Being an operating system, TINA OS must optimize resource allocation. Its objectives are both 1) to maximize the number of admissible virtual machines (VMs), and 2) to minimize interference between them. Interference between VMs may arise when servers' resources are overcommitted. This optimization problem is hard because throughout the partnership with Dassault Systemes and other clients, the IaaS platform is large-scale and shared by users who request resources for various and temporary use cases.

In order to develop better resource allocation algorithms that now heavily use machine learning [3, 4, 5], it is necessary to understand the nature of cloud workloads. The workload of Google's general-purpose cloud backend is published in [6], and its heterogeneity (in terms of machine characteristics, job requirements, job runtime) and dynamicity (in terms of machine state, job resource consumption, job arrival rate) are characterized in [7]. In [8], the characteristics of Google's workload is compared with Grids. A description of VMs' size, runtime and arrival rate in Microsoft Azure is given in [9] along with the dataset.

We observe that 1) studies focus on hyper-scale clouds, and 2) due to a lack of data, interference between VMs and user actions related to some virtual resources (management of volumes, snapshots, images) have not been studied yet.

In the context of my PhD, I recorded data from Outscale's platform to understand its workload and help improve TINA OS. I acquired data from one region during 3 months. The acquisition involved two systems : TINA OS, and a probe that I developped.

On one hand, the logs emitted by TINA OS contained information about the management of virtual resources. Timestamped user actions (e.g, starting a VM) are described in Table 1. Virtual resources' attributes (e.g, the number of cores of a VM) are described in Table 2.

On another hand, I developed a software probe that was deployed by operation teams on servers in production. Written in *python* and based on the *psutil* library, the probe reads and averages VMs' hardware resource consumption periodically (5 minutes). The different metrics (e.g, cpu usage) are described in Table 3.

Dependent resource	Actions	Dependency
volume	create from / delete	snapshot
snapshot	create from / delete	volume
image	create from / delete	VM
VM	run from / start / stop / terminate	image
volume	attach to / detach from	VM
security group	create / delete	
VM	put in / remove from	security group

Table 1: Timestamped user actions and virtual resource dependencies.

I am now planing to analyze Outscale's workload and compare it with other providers when data is available. The goal is to identify patterns. For instance, I will study the periodicity of VMs starts and stops and their lagged correlation. I expect bursts of VMs starts to be followed by symmetrical bursts of stops, which would suggest that VMs start and stop in groups. This could allow me to describe the size of groups, to be compared with the number of VMs per security

Name	Description	Owned By
core	number of virtual cores	VM
ram	amount of RAM	VM
family	ex: general-purpose, compute-optimized, ...	VM
generation	processor generation	VM
platform	linux or windows	VM
size	storage space	volume, snapshot
type	general-purpose, cold, hot storage	volume
iops	input output per second	volume
tag	freely typed text describing a resource	all resources (optional)

Table 2: Virtual resources' attributes.

Name	Description
cpu	5-minute avg of usage in (% of a core). May be > 100 for multi-core VMs
ctx switches	5-minute avg of the number of involuntary context switches per second
read	5-minute avg of sum of the number of bytes read on all attached volumes (B/s)
write	5-minute avg of the sum of the number of bytes written on all attached voumes (B/s)

Table 3: Timestamped measurements of physical resource usage of VMs.

group. A second kind of analysis will quantify the distribution and periodicity of VMs' hardware resource usage depending on some attribute (e.g, VM type, runtime, or tag). I hope the results will help optimize Outscale's server overcommit policies.

References

[1] Harvard Business Review Analytics Service, "Business agility in the cloud," https://hbr.org/resources/pdfs/tools/Verizon_Report_June2014.pdf, August 2014.

[2] "TINA OS," posted at https://en.outscale.com/tina-os-cloud-orchestrator/.

[3] Y. Jiang, C. s. Perng, T. Li, and R. Chang, "Self-adaptive cloud capacity planning," in *2012 IEEE Ninth International Conference on Services Computing*, June 2012, pp. 73–80.

[4] M. Carvalho, W. Cirne, F. Brasileiro, and J. Wilkes, "Long-term slos for reclaimed cloud computing resources," in *Proceedings of the ACM Symposium on Cloud Computing*, ser. SOCC '14. New York, NY, USA: ACM, 2014, pp. 20:1–20:13. [Online]. Available: http://doi.acm.org/10.1145/2670979.2670999

[5] J. Xue, F. Yan, R. Birke, L. Y. Chen, T. Scherer, and E. Smirni, "Practise: Robust prediction of data center time series," in *2015 11th International Conference on Network and Service Management (CNSM)*, Nov 2015, pp. 126–134.

[6] J. Wilkes, "More Google cluster data," Google research blog, Nov. 2011, posted at http://googleresearch.blogspot.com/2011/11/more-google-cluster-data.html.

[7] C. Reiss, A. Tumanov, G. R. Ganger, R. H. Katz, and M. A. Kozuch, "Heterogeneity and dynamicity of clouds at scale: Google trace analysis," in *Proceedings of the Third ACM Symposium on Cloud Computing*, ser. SoCC '12. New York, NY, USA: ACM, 2012, pp. 7:1–7:13. [Online]. Available: http://doi.acm.org/10.1145/2391229.2391236

[8] S. Di, D. Kondo, and W. Cirne, "Characterization and comparison of cloud versus grid workloads," in *2012 IEEE International Conference on Cluster Computing*, Sept 2012, pp. 230–238.

[9] E. Cortez, A. Bonde, A. Muzio, M. Russinovich, M. Fontoura, and R. Bianchini, "Resource central: Understanding and predicting workloads for improved resource management in large cloud platforms," in *Proceedings of the 26th Symposium on Operating Systems Principles*, ser. SOSP '17. New York, NY, USA: ACM, 2017, pp. 153–167. [Online]. Available: http://doi.acm.org/10.1145/3132747.3132772