Comparative Analysis of Performance on Cloud Computing Environments: a Brazilian Public Sector Perspective

Breno G. S. Costa, Fernando Escobar, Priscila Solis, Aletéia Araújo, Jacir Bordim

[1]Department of Computer Science – University of Brasília (UnB) – Brasília – DF – Brazil

brenogscosta, fernando.escobar.br@gmail.com, pris, aleteia@cic.unb.br, bordim@unb.br

Abstract. *The use of public cloud services is growing at a rapid pace and many organizations either have started using these services or plan to do so in a short time. This is due to the need to collect the benefits of the platform, including cost reduction and agility in providing new information technology services. Public cloud providers have similar offers of infrastructure as a service, but with different characteristics and costs, and it is necessary to evaluate how the resources of an organization's internal infrastructure compare to the resources offered by those providers, creating a mapping between them. This takes advantage of the experience and knowledge that the organization's technical team has over the internal infrastructure and reduces the risk of cost increase, time increase, or even to offer a cloud service with lower performance than it is already provided using the organization's own resources. This study compares the performance of a computational unit of an organization with a similar unit in two public cloud providers, using metrics to measure processing, storage and network performance, and comparing the costs charged by these providers. The results show that there is a significant variation in performance: in network performance, one provider outperformed the other by approximately one order of magnitude.*

1. Introduction

Industry experts see Cloud Computing (CC) as something that can revolutionize information technology (IT), since it changes the way IT is provided and consumed. It also changes the current landscape in which organizations manage their own IT infrastructure to another where IT is consumed as a service, bringing along its benefits [ISACA 2012]. According to Gartner [Gartner 2017], the use of public cloud services is growing at a rapid pace and this can be seen by the significant growth rates of large providers such as Amazon Web Services (AWS). In a survey cited in Gartner's study, with almost 3000 participants, 21% of them already was using public cloud services, while 56% was planning to implement the cloud by the end of 2017. For many of these organizations, IT modernization is almost a synonym of increasing use of CC.

The Brazilian Federal Court of Accounts (TCU) published research providing a detailed analysis of public CC offerings and presenting the risks and benefits of their use by government bodies [Tribunal de Contas da União 2015]. The Ministry of Planning published a standard [MPOG 2016] that prohibited dozens of Brazilian government agencies from building or expanding data centers, stating that they must hire public CC providers in their place. It is therefore expected that, in the next years, there will be several public CC contracts in the Brazilian government with an initial focus on infrastructure as a service (IaaS), as defined by the National Institute of Standards and Technology (NIST) [Mell et al. 2011].

But, in the Brazilian Government, an agency can not hire suppliers at will. In the opposite sense, it must schedule a public reverse auction and the suppliers that offer the smaller price and meet the stated requirements will be hired. It means the agency will know its suppliers only after auction procedure is completed. In this context, the agency will begin to provide virtual machines (VMs) in the cloud and not just in the local infrastructure. Agencies have knowledge and experience related to capacity planning in local infrastructure. A new system has the technical, functional, and non-functional requirements mapped to virtual processing, memory and disk power in accordance with that experience. With the advent of public cloud providers, a mapping between a locally provided VM and its equivalent in each hired provider has to be created.

To address this, the key contribution of our work is a method to create a mapping between the processing power of local infrastructure and the processing power of hired cloud providers. This mapping is based on performance measurements, on the cost of each provider and on two common scenarios of VM use. This study compares a reference VM of a government agency with a similar VM in two public clouds. A reference VM is one that has the configuration of virtual processors (vCPUs), memory, and disks more prevalent in the local infrastructure. The goal is to compare the performance between them, so that the technical team can more accurately suggest equivalent capacity across different cloud providers evaluated. Although public cloud environments allow vertical growth and decrease (capacity variation) more easily than traditional environments [Marston et al. 2011], even automatically, when using elasticity characteristics, this characteristic is usually offered at higher prices. When there is a time commitment, for example a full year, there is a considerable cost reduction and some of the workloads will fit into this usage model and may benefit from the cost reduction.

Having a precise estimate of the size of new VMs to be created on the Cloud can lead to cost and time savings. In a large volume of new VMs, starting the service with inadequate capacity, monitoring performance, and making adaptations can affect service delivery if it decides to go into production first to evaluate later, or can cost additional effort and time until the ideal size is found. Thus, knowing the mapping between the local environment and the new environments will save time and potentially cost.

The rest of the paper is divided into the following sections: Section 2 analyzes related works, Section 3 describes the proposed method, Section 4 describes the testbed of the experiments, Section 5 shows the results, Section 6 presents a discussion about the results, and Section 7 concludes and describes future works.

2. Related Work

O'Loughlin and Gillam [O'Loughlin and Gillam 2013] extensively evaluated the vCPU performance variability of a particular VM type on AWS and showed that the performance is a function of the underlying hardware, more specifically of physical CPU model in which the vCPU is running. Other studies [Farley et al. 2012] [Ou et al. 2012] corroborated this finding.

Several other studies have proposed strategies to allocate the best public cloud VM in order to run a specific workload and meet user cost and performance goals. Chiang et al. [Chiang et al. 2014] have proposed Matrix, a resource management and performance system, that uses machine learning models to predict cloud performance of new workloads from previous benchmarking on physical and virtual machines. Matrix, in addition to predicting performance, selects the most appropriate VM for execution and automatically adjusts the configuration of VM resources. [Grozev and Buyya 2017] uses a combination of heuristics and machine learning approaches to learn the application's performance characteristics and to adapt to workload changes in real time. It is a fully automated approach. [Boza et al. 2017] proposed a tool based on M(t)/M/* queuing theory models and a process that can be used by tenants to properly plan and budget their cloud computing costs. [Yadwadkar et al. 2017] has presented PARIS, a data-driven system that uses a hybrid off-line and on-line data collection and modeling framework to provide accurate performance estimates with alleged minimal data collection.

[Entrialgo et al. 2017] proposed Multi-Application Load Level based OptimizatiOn for VIrtual machine Allocation (MALLOOVIA). It is an allocation strategy that combines a long-term and a short-term strategy, organized in two phases, to take advantage of both reserved and on demand VMs. It uses integer linear programming to find the best VM set, composed mainly by reserved VMs (cheaper) but complemented with on-demand VMs, that meets performance goals and minimizes VM allocation cost.

Although these studies have published strategies and tools to find an optimal VM set to run a specific workload on the cloud, the complexity and effort to collect the required data and to execute the machine learning models for performance prediction may be an obstacle to their use. Brazilian agencies are on the beginning of cloud adoption process and lack skilled personnel to deal with advanced strategies and with the intricacies of VM families and pricing models. This restriction applies also to small and medium sized organizations, according to [Boza et al. 2017]. The work described here differs from the others by the simplicity of its proposed method, offering a low effort way of selecting a proper VM configuration on the cloud provider.

3. Method

To create a computational mapping between the local infrastructure and a public cloud provider infrastructure, we suggested the following steps:

1. Identify what is the reference VM of your virtual infrastructure.
2. Create a similar VM on the cloud provider you want to compare to.
3. Run performance benchmarks on each VM and collect the results.
4. Calculate the cost of running the VM on the agency's usage scenarios.
5. Compare performance and cost results and create the mapping. This mapping will guide the decision about VM capacity on the new cloud provider using the knowledge and experience with the local infrastructure.
6. Repeat the previous steps on a regular basis to update the mapping.

On the next sections, we describe a proof of concept of the proposed method.

4. Testbed for Experiments

The performance comparison will consider four metrics: the throughput that each vCPU supports, I/O throughput, network throughput, and cost of running VMs on each cloud provider. The chosen providers were AWS and Google Cloud Platform (GCP). Both are well-known providers and they are among the top three providers in Gartner's magic quadrant for CC [Gartner 2016]. Moreover, both have representatives in Brazil and meet the regulatory, legal and compliance requirements of Brazilian government related to cloud.

Inside public cloud providers, different availability zones mean different datacenters, and each datacenter has a specific support infrastructure. In this study, the lowest cost zone was selected in each provider since this is the main criterion for contracting used by the Brazilian government. When allowed by the cloud provider, the most recent CPU micro-architecture generation was chosen. The goal was to run and compare, on each cloud provider, a similar VM that has the lowest cost and the best performance.

The agency currently has a virtualized infrastructure which is the primary way of provisioning computing services. There are 1200 active VMs and about 20% of them have little or no dependency on other corporate services. That is, these VMs need only the computing and storage resources to provide the service that they were created for and they can be migrated more easily to the cloud.

Reference VM - The most prevalent general-purpose VM in the organization, hereafter referred to as the reference VM, is a VM with 4 vCPUs 2.1 GHz Intel Xeon E7 (Haswell) and 8 GB RAM, running Windows 2008 server R2. The hypervisor used by the organization is VSphere[1].

A similar VM was created in the GCP cloud, which allows arbitrary choice of vCPUs and RAM size. The hypervisor used by GCP is KVM [Kivity et al. 2007]. The zone chosen for creating the VMs was us-central1-a. On AWS, a M5.xlarge VM was created. It is the 5th generation of M type (more details on section 5.1) and it has 4 vCPUs, but 16GB RAM. As the measurement is based on vCPU, I/O and network throughput, this memory difference has no impact on the results, according to preliminary tests performed. The zone in which the VM was created was us-east-1c.

To measure whether the workload performance is proportional to the number of vCPUs, we created two additional VMs in the organization, and in each cloud provider. The additional VMs had the number of vCPUs representing half of the reference VM and a fourth, namely two vCPUs and one vCPU, with the same proportional decrease in memory. In total, 9 VMs were created. On the agency side, the reference VM was labeled as Org.Big and the others as Org.Medium and Org.Small. The same notation was used to name AWS and GCP VMs.

In public cloud providers, general purpose VMs were used. Their configuration is balanced between processing and memory, similar to the reference VM. There are other pre-defined categories, such as VMs with a focus on computing, memory, or network, with an imbalance between these functions. For storage, Hard Disk Drive (HDD) based services were selected, since it is a cheaper option and it is the option used on reference VM. Both providers offer Solid State Drive (SSD) based storage, with higher I/O per-second (IOPS) and throughput, but at a higher cost per GB.

5. Results

According to [Hennessy and Patterson 2011], performance metrics that are most useful and informative to a user are execution time and throughput. Execution time is defined as being the time for an application to execute. Throughput can be thought of as the work done per unit of time for applications that are doing some specified unit of work, for example a file compression.

[1]VMWare VSphere. Link: http://www.vmware.com/products/vsphere.html/

Current best practices for measuring computer performance have led to a preference for benchmark programs drawn from actual end-user applications, as opposed to synthetic benchmarks [O'Loughlin and Gillam 2017]. These benchmarks are implemented by the Standard Performance Evaluation Corporation (SPEC)[2] benchmark suite. Following this approach, we have used the IntRate, a suite composed by 10 integer SPEC benchmarks. In particular, we have used SPEC CPU 2017[3], the most recent version of SPEC CPU benchmark software. It does a comparative measure of integer computationally intensive performance.

In relation to I/O, the benchmark chosen was IOZone[4], which is also widely used in measurements of this type in the literature [Xavier et al. 2013] [Lee et al. 2013]. IPerf[5] was used to evaluate the performance of the network. It is a widely used tool to provide a measurement of the maximum throughput of a network. The choice of benchmarking tools was based on the tool's reputation, evaluated by checking its use in academic articles, and also its availability for Windows Server 2008 R2, the operating system used in the organization's VMs. The primary objective of the experiment is to compare performance between a local VM and a similar one running on public cloud environments. The secondary objective is to compare each of these VMs with others from the same provider, proportionally varying the number of vCPUs and memory and verifying whether there is proportional performance.

For IOzone, which measures the throughput in various I/O operations, the default parameters were used. They vary the size of the file, as well as the data record inside it, in 8 operations. For the purpose of this comparison, we will show the results only for write and read operations. For iPerf, which measures network throughput, standard TCP protocol parameters were used.

5.1. Experiment 1 - vCPU

This experiment will measure vCPU performance (throughput) of each VM and compare them to reference VM. In addition, it will verify whether computing performance varies proportionally to the number of vCPUs.

On GCP, there are some options for frequency and generation of physical processor micro-architecture (inside parentheses, below) that can be chosen when creating a VM. In the chosen zone, us-central-1, the options were 2.6 GHz Intel Xeon E5 (Sandy Bridge), 2.2 GHz Intel Xeon E5 v4 (Broadwell) and 2.0 GHz Intel Xeon (Skylake). Although the processor frequency is listed in a decreasing way, the performance between them, on that specific case, has the opposite behavior, and this can lead to a bad choice. Tests showed that Sandy Bridge had a 92% lower average performance than Broadwell, which is three generations ahead of Sandy Bridge; and Broadwell performed 42% less than Skylake, a generation ahead of Broadwell [Doweck et al. 2017]. Skylake was the option chosen for the GCP provider in this study.

On AWS the frequency and the generation of the processor is determined by the instance type. Type M is the AWS option for general purpose VMs. M4 instances (4th

[2]S.P.E.C. Link: http://www.spec.org/
[3]SPEC 2017. Link: http://www.spec.org/cpu2017/Docs/index.html
[4]IOZone - Filesystem Benchmark. Link: http://www.iozone.org/
[5]IPerf - Network Performance Benchmark. Link: http://iperf.sourceforge.net/

Table 1. VM CPU throughput Using SPEC CPU2017 Benchmarks.

CPU2017	ORG BIG	ORG MED	ORG SML	AWS BIG	AWS MED	AWS SML	GCP BIG	GCP MED	GCP SML
Throughput	9.30	4.93	2.53	10.40	5.52	2.27	6.66	2.77	1.96

generation) use Intel Xeon E5 v4 (Broadwell) 2.3 GHz processor. Recently, AWS has announced M5, a new generation, which is based on Skylake micro-architecture and which performs 14% faster than M4. The M5 Family was available in the AWS zone where the experiments were run and it was therefore the choice for this study. On the Organization side, the processors available to support vCPU on local VMs are Intel Xeon E7 v3 (Haswell).

Table 1 shows the measurements for each VM using benchmarks from SPEC CPU 2017. Although they use different hypervisors, the variation in performance between them is due primarily to the micro-architecture generation of physical processor in which the VM executes [Ou et al. 2012][Liu 2015][O'Loughlin and Gillam 2017].

For each SPEC benchmark, a performance ratio is calculated using the time needed to run the benchmark on a SPEC reference machine to the time needed to run it on the System Under Test (SUT). In our study, each VM is a SUT. Higher scores mean that more work is done per unit of time. All ratios calculated are averaged using a geometric mean, which is reported as the overall metric by SPEC. Running the entire SPEC CPU 2017 Intrate suite took approximately ten hours per VM. Within the organization, and on each cloud provider, the measurements varied proportionally to the number of vCPUS, as shown in Table 1.

5.2. Experiment 2 – I/O

This experiment will measure the I/O throughput of all VMs and compare them. IOZone is the file system benchmark that was used in this experiment. It generates and measures a variety of read and write operations on files. The operations analyzed in this experiment were File Read, which measures performance of reading a file that already exists in the file system, and File Write, which measures performance of writing a new file in the file system. IOZone creates temporary test files with sizes ranging from 64KB to 512MB. The size of the records varies from 4KB to 16MB. All results are in MB/s.

Tables 2 and 3 show throughput measurements on each VM for write and read operations, respectively. By analyzing their data, it can be seen that AWS VMs have a higher I/O throughput, GCP VMs are in second place and organization VMs, third. For writing operations, AWS BIG performed 108% higher than the reference VM, while GCP BIG performed 40% higher than it. For reading operations, the order was the same and the percentages were 61% and 16% for AWS BIG and GCP BIG, respectively, when compared to the reference VM. It is important to note that both providers offer a proportional increase of IOPS (IO per second) when increasing the number of gigabytes (GB) of storage, with different upper and lower limits. Scenarios with volumes much larger than the ones analyzed in this study, 500GB for instance, can give different results when comparing the same providers.

Table 2. I/O Throughput (MB/s) for Write Operations.

	ORG BIG	ORG MED	ORG SML	AWS BIG	AWS MED	AWS SML	GCP BIG	GCP MED	GCP SML
Mean	1323	1443	1350	2751	2456	2177	1847	1552	1578
Minimum	245	254	183	552	458	475	303	399	376
Maximum	2460	2528	2371	3868	3425	2993	3215	2859	2466

Table 3. I/O Throughput (MB/s) for Read Operations.

	ORG BIG	ORG MED	ORG SML	AWS BIG	AWS MED	AWS SML	GCP BIG	GCP MED	GCP SML
Mean	3270	3319	3177	5265	4519	4248	3787	3500	3574
Minimum	1160	1124	1062	2128	1882	1743	1525	1437	1356
Maximum	8407	8185	7447	11185	10220	10119	9665	10114	9808

5.3. Experiment 3 - network

This experiment will measure network throughput between different VMs in the same zone in each cloud provider and compare them to reference VM results. IPerf benchmark was used. The SML VM was used as the server, while the other two VMs were used as client machines within the same zone.

The experiment consisted of generating on each client machine a flow of TCP data for one minute and measuring it second by second. This flow was executed 30 times at different times of the day, in order to record the throughput variation. As there were close values between the measurements of both VMs that served as clients in the organization and in each provider, the values were consolidated and are presented in Table 4 with reference only to the BIG VMs. As was to be expected, there was great variability in the measurements, which is reflected in the ratio from the deviations and the mean, as well as in the interval between the maximum and minimum values.

Table 4. Network Throughput (Mbps).

	ORG BIG	AWS BIG	GCP BIG
Mean	18	434	3592
Deviation	4	242	543
Minimum	5	124	430
Maximum	44	986	5100

5.4. Cost Comparison

It is not possible to estimate the cost of the reference VM in the organization due the lack of tools to do this measurement. However, it is possible to calculate the cost of VMs on cloud providers. For this, two usage scenarios were defined.

In the first scenario, VMs run 12 hours a day, 5 days a week for 4 weeks, consuming a total of 240 hours per month. This scenario represents the applications that are used only during extended business hours, from 8:00 AM to 8:00 PM, and from Monday through Friday. The monthly costs of chosen VMs on providers in this scenario can be seen in Table 5. This scenario is based on on-demand prices, which are proportionally higher than prices of VMs used for larger periods. The values are in US dollars.

In the second scenario, VMs run uninterruptedly 24 hours a day, 7 days a week, consuming a total of 730 hours per month. In this scenario, it is possible to take advantage of loyalty programs which establish a consumption commitment for up to three years. This is the option with the highest discount rate and because of that it was the choice for this scenario, whose costs are presented in Table 6. The costs related to AWS in this scenario is the amount corresponding to a month of VM use, however AWS require prepayment of three years of use to reach this level of discount rate. GCP does not have the prepayment method and the costs presented are monthly payments, considering the loyalty use for three years. The values did not include operating system license costs, since the goal is to compare providers' exclusive offerings.

Table 5. Monthly Cost(US$)–First Scenario

	AWS BIG	AWS MED	AWS SML	GCP BIG	GCP MED	GCP SML
vCPU	46.08	23.04	16.08	38.36	19.18	9.59
Storage	2.25	2.25	2.25	1.30	1.30	1.30
Total	**48.33**	**25.29**	**18.33**	**39.66**	**20.48**	**10.89**

Table 6. Monthly Costs(US$)–Second Scenario

	AWS BIG	AWS MED	AWS SML	GCP BIG	GCP MED	GCP SML
vCPU	54.02	27.01	18.98	52.67	26.34	13.17
Storage	2.25	2.25	2.25	1.30	1.30	1.30
Total	**56.27**	**29.26**	**21.23**	**53.97**	**27.64**	**14.47**

6. Discussion

The experiments showed that, in relation to vCPU performance, AWS had a 12% higher throughput when comparing to the reference VM, while GCP performed 28% lower than the reference VM. Although GCP presented 36% less performance when compared directly to AWS, it has a lower cost on both scenarios. In the first scenario, GCP had a price 18% lower than AWS. In the second scenario, the difference was tiny, only 4%. These differences must be considered when migrating a CPU-bound workload to the cloud and the performance proportion should be considered when estimating the required computational capacity at the provider.

In relation to I/O throughput, AWS had outperformed the reference VM by 108% and 61% for File Write and File Read operations, respectively. GCP had performed at 40% and 16% above the reference VM, respectively. Since the cost difference for VMs in scenario 2 is small, AWS should be the choice for workloads that are I/O-bound and that fit scenario 2 of use (Section 5.4). For I/O-bound workloads that fit scenario 1, GCP should be the choice because it performs better than the reference VM and has a lower cost than the AWS, and the lowest cost is the main criterion of provider choice in the Brazilian government.

In terms of network throughput, AWS outperformed the reference VM by 2200%, while GCP outperformed by 19600%. It is expected that cloud providers have a reliable and fast network infrastructure, since it is one of the main CC characteristics, according to NIST[Mell et al. 2011]. However, GCP had an outstanding performance even when compared to AWS, the world biggest cloud provider, according to Gartner[Gartner 2016]. It was published in the media that GCP made a huge multi-billion dollar investment in infrastructure over the last three years [IEEE 2018], part of it improving its network and maybe this is the reason why it outperformed AWS by such a large margin. GCP should be the choice for workloads that are network-bound.

For other workloads, in which there is not a clear characterization of more intensive use of a specific resource, GCP should be the choice, since it has lower cost of provision and superior performance (when compared to the reference VM) in two of the three experiments. In summary, the mapping created for provisioning VMs in cloud providers in this organization is:

1. If the workload is CPU-bound or I/O-bound and fits into the second usage scenario, create the VMs on AWS with 11% less vCPUs than the VMs that would be created on the local infrastructure.
2. Otherwise, create the VMs on GCP with 40% more vCPUs than the VMs that would be created on the local infrastructure.

7. Conclusion

This work compared a computing unit from a virtual infrastructure of a public agency with a similar unit on two public cloud providers: GCP and AWS. Metrics related to CPU, I/O and network were used, as well as calculating the execution costs in two usage scenarios.

The measurements and cost estimates made in this study provided a more accurate assessment of the need for computational infrastructure of workloads that are migrated from the organization to these providers, as well as propose a criteria-based choice model (in this case, lower cost and superior performance), which can be used in other providers in a similar way. The results indicated significant performance differences between public cloud providers across the metrics used, as well as differences in cost. They confirmed the importance of making objective comparisons between the organization's current infrastructure and those provided by the cloud, as part of the plan for CC adoption.

Each service in the cloud has many parameters and features, and they differ from provider to provider, as seen in Sections 5 and 6. This makes the comparison challenging. The choice of representative workloads is important because it may reduce the number of test cases when evaluating a category of systems, minimizing the decisions in terms of the balance between the computing, storage and network resources. As future work, there is the intention to evaluate other cloud providers, evaluate workloads with different characteristics, as well as to measure the performance and cost of other components, such as GPUs (Graphical Processing Units). Furthermore, considering the multi-CPU environments used nowadays, analyzing parallel computation could be of great value.

References

Boza, E. F., Abad, C. L., Villavicencio, M., Quimba, S., and Plaza, J. A. (2017). Reserved, on demand or serverless: Model-based simulations for cloud budget planning. In *Ecuador Technical Chapters Meeting (ETCM), 2017 IEEE*, pages 1–6. IEEE.

Chiang, R. C.-L., Hwang, J., Huang, H. H., and Wood, T. (2014). Matrix: Achieving Predictable Virtual Machine Performance in the Clouds. In *ICAC*, pages 45–56.

Doweck, J., Kao, W.-F., Lu, A. K.-y., Mandelblat, J., Rahatekar, A., Rappoport, L., Rotem, E., Yasin, A., and Yoaz, A. (2017). Inside 6th-Generation Intel Core: New Microarchitecture Code-Named Skylake. *IEEE Micro*, 37(2):52–62.

Entrialgo, J., Díaz, J. L., García, J., García, M., and García, D. F. (2017). Cost Minimization of Virtual Machine Allocation in Public Clouds Considering Multiple Applications. In *International Conference on the Economics of Grids, Clouds, Systems, and Services*, pages 147–161. Springer.

Farley, B., Juels, A., Varadarajan, V., Ristenpart, T., Bowers, K. D., and Swift, M. M. (2012). More for your money: exploiting performance heterogeneity in public clouds. In *Proceedings of the Third ACM Symposium on Cloud Computing*, page 20. ACM.

Gartner (2016). Magic Quadrant for Cloud Infrastructure as a Service, Worldwide.

Gartner (2017). Developing a Public Cloud IaaS Adoption and Migration Framework.

Grozev, N. and Buyya, R. (2017). Dynamic Selection of Virtual Machines for Application Servers in Cloud Environments. In *Research Advances in Cloud Computing,187-210.*

Hennessy, J. L. and Patterson, D. A. (2011). *Computer architecture: a quantitative approach*. Elsevier.

IEEE (2018). http://techblog.comsoc.org/2018/01/18/ google-expands-cloud-network-infrastructure-via-3-new-undersea-cables-5-new-regions/.

ISACA (2012). Guiding principles for cloud computing adoption and use.

Kivity, A., Kamay, Y., Laor, D., Lublin, U., and Liguori, A. (2007). kvm: the Linux virtual machine monitor. In *Proceedings of the Linux symposium Vol.1*, pages 225–230.

Lee, E., Bahn, H., and Noh, S. H. (2013). Unioning of the buffer cache and journaling layers with non-volatile memory. In *FAST*, volume 13.

Liu, R. (2015). *Performance Analysis and Configuration Selection for Applications in the Cloud*. PhD thesis, Rice University.

Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J., and Ghalsasi, A. (2011). Cloud computing—The business perspective. *Decision support systems*, 51(1):176–189.

Mell, P., Grance, T., and others (2011). The NIST definition of cloud computing.

MPOG (2016). Boas práticas, orientações e vedações para contratação de Serviços de Computação em Nuvem.

O'Loughlin, J. and Gillam, L. (2013). Towards performance prediction for Public Infrastructure Clouds: an EC2 case study. In *Cloud Technology and Science (CloudCom), 2013 IEEE 5th International Conference on*, volume 1, pages 475–480. IEEE.

Ou, Z., Zhuang, H., Nurminen, J. K., Ylä-Jääski, A., and Hui, P. (2012). Exploiting Hardware Heterogeneity within the Same Instance Type of Amazon EC2. In *HotCloud*.

O'Loughlin, J. and Gillam, L. (2017). A performance brokerage for heterogeneous clouds. *Future Generation Computer Systems*.

Tribunal de Contas da União (2015). Acórdão 1739/2015-Plenário.

Xavier, M. G., Neves, M. V., Rossi, F. D., Ferreto, T. C., Lange, T., and De Rose, C. A. (2013). Performance evaluation of container-based virtualization for high performance computing environments. In *21st Euromicro International*, pages 233–240. IEEE.

Yadwadkar, N. J., Hariharan, B., Gonzalez, J. E., Smith, B., and Katz, R. H. (2017). Selecting the best vm across multiple public clouds: A data-driven performance modeling approach. In *Proceedings of the 2017 Symposium on Cloud Computing*, pages 452–465. ACM.