HAL
archives-ouvertes.fr

Cost in the Cloud: Rationalization and Research Trails

Kahina Hamadache, Viliam Simko, Rustem Dautov, Fotis Gonidis, Paraskevi Zerva, Ivan Dario Paez Anaya, Ariana Polyviou

▶ **To cite this version:**

Kahina Hamadache, Viliam Simko, Rustem Dautov, Fotis Gonidis, Paraskevi Zerva, et al.. Cost in the Cloud: Rationalization and Research Trails. International Conference on Advanced Cloud and Big Data (CBD 2014), Nov 2014, Huangshan, Anhui, China. International Conference on Advanced Cloud and Big Data (CBD 2014), 2014. <hal-01158925>

HAL Id: hal-01158925
https://hal.inria.fr/hal-01158925

Submitted on 2 Jun 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Cost in the Cloud: Rationalization and Research Trails

Kahina Hamadache[1], Viliam Simko[2], Rustem Dautov[3], Fotis Gonidis[3], Paraskevi Zerva[4], Ivan Paez Anaya[5], Ariana Polyviou[1]

[1]Department of European Projects, Singular Logic S.A. Athens, Greece
{khamadache, apolyviou}@singularlogic.eu
[2]Karlsruhe Institute of Technology Karlsruhe, Germany
viliam.simko@kit.edu
[3]South East European Research Centre, Thessaloniki, Greece
{rdautov, fgonidis}@seerc.org
[4]Department of Informatics King's College London, UK
paraskevi.zerva@kcl.ac.uk
[5]IRISA - University of Rennes 1 & INRIA, Rennes, France
ivan.paez_anaya@irisa.fr

Abstract. Cloud Computing provides simplicity to its consumers by saving them the efforts to deal with their own infrastructure, environments or software. This simplicity relies on the shifting of problems from the client to the provider, introducing new paradigms (virtualization, scalability, flexibility, pay-per-use, etc.). This simplicity comes with a price for the consumer that may accurately, or not, reflect the costs of the provider. In this paper we propose to identify the different points, in the Cloud Computing architecture, where the costs are generated, how their reduction/optimisation are considered, and finally we point-out which of these key points need to be further investigated, according to their foreseeable efficiency.

Keywords: Cost, Pricing Model, Cloud Computing, Cost Factors.

1 Introduction

Since its very emergence, cloud computing has been claiming to be the 5th utility [1] and to be adopting the pay-per-use principle, when the customer is only charged for the computational resources he actually consumes (e.g., CPU, memory, network, etc.). In practice, however, this principle typically results in a flat rate pricing scheme (e.g., time-based, volume-based, usage-based, location-based, etc. [2]). Such pricing scheme can be appealing for customers as they seem to offer good trade-off and often important reductions compared to the equivalent pay-as-you-go price. However, flat rates also assume that users' are able to assess accurately their needs and optimize their use to ensure the highest cost-effectiveness. In practice, this is almost never the case as predicting precisely and completely one's needs are almost impossible. The reality of flat prices is that they mainly benefit to providers as they allow them to anticipate and optimize their resource allocation as well as more efficiently prevent

SLA violations. However, such practice is dangerous for the cloud itself, hence, instead of offering the opportunity of highly customized services and SLAs, cloud providers tend to limit their offers to a basic set of bundles, available to everyone, with exactly the same constraints. Several researches have been conducted to propose dynamic pricing strategies for providers. For example, a recent research by Sewook Wee [3] analysed Amazon Spot Instances' run-time pricing mechanism (i.e., the price changes on an hourly basis based on the current workload of the data center) and clearly demonstrated that in reality with a flexible pricing scheme, the cloud provider can charge its customers up to 52% less money for exactly the same amount of computational resources. Intensive efforts have been made from the research perspective, to provide optimal dynamic pricing schemes, allowing one, in theory to save cost. Those schemes rely mainly on the adaptation to market's demand, users' profiles and needs, and desired profit margins [4], [5]. Despite interesting from the pricing and market consideration perspectives, these approaches are often misused by customers, either because of a gap between their expected and actual use of the cloud, or because of their inability to optimize their consumption but also in some case because of some misconception about these schemes. Unfortunately for consumers, the use of such pricing schemes often results in higher charges than if they had used flat rates and their reduction. In order for providers to be able to offer high-dynamic priced services, they also need to rationalize their costs and identify the necessary actions to be undertaken in order to efficiently and safely reduce them.

This paper places itself in this perspective and tries to consider the whole cloud ecosystem and evaluate, at the different levels of the system, what are the source of costs and how they can be tackled. This paper will follow a straightforward plan, after this introduction we will dive into the analysis of the Cloud ecosystem and identify its different source of cost. This analysis will provide us with the necessary material to discuss about the current state of the cloud and sketch some promising perspectives.

2 Cloud Ecosystem Cost Analysis

The Cloud Computing paradigm can be seen from multiple perspectives as the synthesis of years of research in different domains spanning from virtualization software to hardware optimisation, graph theory to help the management of grids, service computing, etc. Even if those techniques were not intended for the development of the Cloud, their recent evolution (especially the increase in computing performance of the last decade) has provided the base on which the Cloud Computing was able to emerge. In [7] the authors have identified the challenges for the cost of the cloud in data centers and they present a cost breakdown list including the server cost, infrastructure costs (power delivery and heat dissipation). This research doesn't provide an exhaustive list of the costs in the cloud and only emphasis on some specific aspects directly linked to data centres and then closer to hardware concerns. In this paper we try to consider the complete Cloud ecosystem, from data centers to service orchestration and consumption.

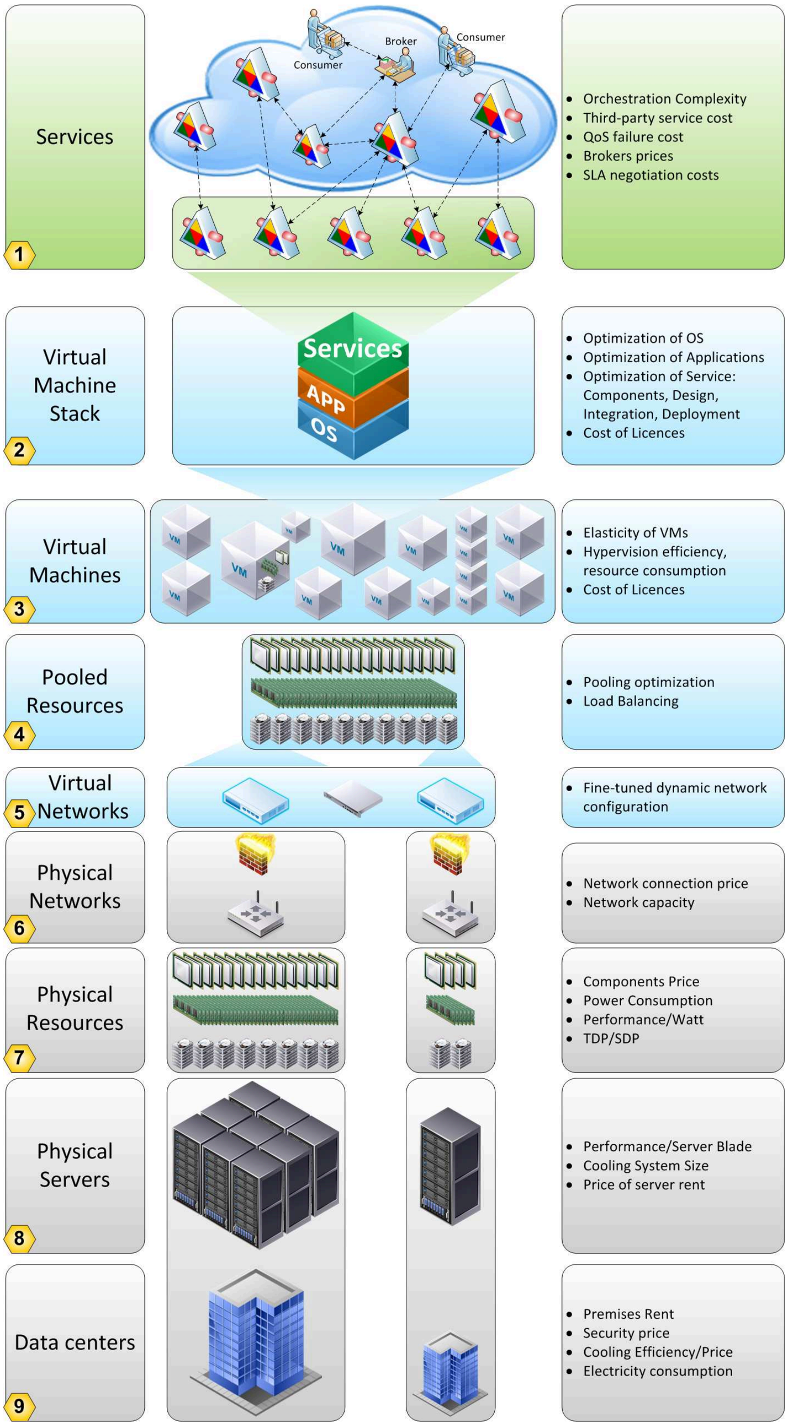

Figure 1 Costs in the Cloud

Figure 1depicts the usual organisation of a Cloud Provider and its interface with the rest of the Cloud. It will be used in the remainder of this section as the basis for our analysis.

2.1 Service

At the top of the schema we have the Cloud itself with its multitude of services hosted by different providers, as well as the consumers and brokers that are consuming and/or recommending them. This layer includes a large variety of automatic, semi-automatic and manual activities (service recommendation, selection, negotiation, composition, aggregation, orchestration, evaluation, etc.). The costs directly involved by this layer are the following:

- (Third party) Service price: this is the most basic cost in the cloud, it applies not only to end users/consumers but also, and maybe most importantly to providers that are relying on external services for a part of their offer (for instance, SaaS providers will be more than likely to start their activity by relying on a PaaS or IaaS provider, ensuring the flexibility and scalability that they may need at their very beginning.) Cost of such services are not limited to their actual price, as they may also involve an heightened security and performance risk, in some cases forcing the provider to overprovision or anticipate third-party service failures or disruption. In [20] the authors propose a very interesting approach to consider the costs involved by the composition and selection of services. In the end third-party service prices are an important and almost unavoidable cost for providers.
- Broker price; in order to help them navigating through the sea of service offers, service consumers tend to rely on the help of service brokers, entrusting them in finding the best offer, matching their requirements and possibly with the lowest possible price. However, this “Brokering service” itself has a cost, and it can influence the final price of a provider’s offer if he relies too heavily on brokers or if his brokering agent underperforms and cannot find cost-effective services.
- SLA/QoS Requirements convey a two-sided cost aspect: the first side is related to the constraints expressed by the requirements, forcing the provider to make the proper provisioning and configuration, involving costs for maintenance, monitoring and provisioning of resources. The other side is when a violation of SLA requirements occurs during the service consumption, resulting in a penalty for the service provider. This penalty can be partially balanced in the case where the violation is due to a sub-provider, with which the main provider has its own SLA. A consideration of this cost has to be taken while designing and deploying the service, evaluating the risk of failure and the cost of such failure.
- The composition and orchestration of service often implies a non-null cost, as it may require ad-hoc developments to deal with the particularities of some services or due to the use of a legacy system carrying a set of inconsistencies. The orchestration of services also implies a higher risk regarding the respect of SLA constraints, hence implying a potential penalty.
- On a more marginal note, the time and resources involved in the establishment of the SLA can, in some cases, be elevated, due the necessity of fine tuning of the

different parameters, especially when the provider has to reconsider some the SLA he agreed on with its own sub-providers, if it imply the orchestration of numerous services with different SLA, or simply if the negotiation between consumer and provider are tense.

2.2 Virtual Machine Stack

The second layer is the Virtual Machine Stack depicting the virtualisation and configuration of the operating system, applications and services installed on the underlying virtual machine. This layer comes with its own costs factors that are mainly linked to the optimisation of resource consumption and performance:

- OS optimization: this factor is highly critical for cost saving, as it imply limitations for the complete virtual machine, hence limiting the performance and optimization of installed applications and executed services.
- Applications optimisation: lying just on top of the OS, applications are required to be resource-efficient. This optimisation can be achieved through the development of the application, but almost always imply the fine tuning of its configuration (for instance preventing logs to fill the allocated disk space, endangering the complete VM behaviour).
- The optimisation of services is itself an important factor of cost reduction, for multiple aspects. Services have to be resource-efficient, even if they need large resource for their process, they have to consume as few as possible. Hence, even if cloud computing seems to offer unlimited resources, they come at a cost. This implies that optimisation have to take place on different levels:
    - Design: the design of the service has to take into consideration the consumption of resources, just as in the case of “classic computing”.
    - Integration: an inadequate integration of a service within an ecosystem can induce additional costs and the disruption of the provider.
    - Deployment: beyond the design and the integration, the deployment of the service can also influence the final cost of the service. Hence, an incorrect deployment can lead to the necessity of redeployment, potentially disrupting the consumption of the service and requiring manual intervention.
- Finally, the cost of licences is not to be forgotten. And even if the use of free software and components is to be encouraged to reduce the overall cost of the Cloud, licenced software are often hard to completely replace. In the end, it may be acceptable for providers to rely on licenced components, as long as they offer better performance, security and reliability.

2.3 Virtual Machines

Going one layer down we find the Virtual Machines with their own specificities and related costs. Virtual Machines are composed of a set of computing resources (cpu, ram, hdd/ssd) dedicated to the virtualisation of one device that will, in turn, run an OS, application and services. Such VM are managed through hypervisors that dynamically allocate resources to them according to their needs and the available pool of resources. This naturally implies a set of factors that impact the final cost of the service:

- The overall hypervision efficiency can be a source of additional costs, hence, different hypervisors can be more or less effective in allocating resources. They can also require different resources to ensure their
- The elasticity of VM (the speed with which they can adapt to the variations of consumers' requirements is a critical aspect of the cost in the Cloud. Hence, the more responsive to the changes, the more optimized can the allocation of VM be, saving resources for other VMs needing them, or quickly adapting to an increase in consumers' needs. Researches such as [16] can help improving the elasticity of VMs by allowing a fine consideration of underlying resource.
- The cost of licences is once again to take into consideration for hypervisors.

2.4 Pooled Resources

Closely linked to the allocation of VM, pooled resources are obtained via the aggregation of resources offered by multiple servers. They are in turn allocated/ dispatched among the different Virtual Machines managed by the hypervisor. This layer is usually heavily optimized along two aspects:

- Pooling: the pooling of resources is to be optimized as much as possible in order to reduce the cost that may be involved by: incomplete pooling (some resources are not added to the pool, inducing a direct loss of computing power), pooling resource consumption (the pooling itself has to be lightweight, in order not to consumer resources that could be used by services).
- Load balancing: related both to the allocation of virtual machine and ultimately to services executed on these machines, the load balancing of pooled resources should take care to prevent the overload of a single resource server by efficiently balancing the allocation of resources to VMs.

2.5 Virtual Networks

Somehow the deepest virtualisation layer, the virtual networks creates a substrate interconnecting the different server and resources to be pooled for the creation of VMs. The only source of cost at this level is the need for a fine-tuned and dynamic network configuration. In term of cost, the main concern of this layer can come from the risk of failing to respect some privacy-related SLA constraint caused by a variety of factors (inexpressive SLA, lack of responsiveness, failure to automate the creation of network, legal constraints, etc.).

2.6 Physical Networks

The physical network plays a critical role in the Cloud Computing architecture, as it as to ensure that the different services and resources can be connected. The cost factors of this layer are the following:

- Network connection: the price of the external connection of the datacentre to the Internet is rarely a limiting factor, but still has to be accounted among the costs of the cloud.
- Network capacity is not itself a cost, but a source of risk if it is undersized, potentially leading to an inefficient use of the datacentre resources, hence requiring the renting of resources or leading to SLA violations.

2.7 Physical Resources

The pooled resources used for VM and service are an aggregation of physical resources offered by the datacentre servers. These physical resources have a cost in themselves but also imply additional costs:

- Components price: the most basic source of cost, the price of components providing computing resources (it can be transparent when the servers are rented.). This includes CPU/GPU, RAM, mass storage, network adapters, etc.)
- Power consumption: an extremely important aspect of cost is the one of the electricity consumption. Indeed, if physical or virtual components can be bought and represent a long-term investment, the electricity is a permanently consumed resources that, in most cases, needs to be bought from power suppliers. In [8] the authors propose a controller framework for optimizing power consumption, performance benefits and a pricing function for the transient costs that come up by various adaptations with the aim to maximize utility.
- Performance/Watt of the resources represents, in the end, the energy-efficiency of the different components. The higher the performance/watt, the most energy and cost-effective (sometime a bit counter-balanced by the higher initial costs of components).
- The TDP/SDP (Thermal Design Power and Scenario Design Power) indicates the relative heat dissipation of the different components. This has a direct impact on the cooling solution required by the datacentre. High-performance components often imply high heat levels, requiring the installation of expensive cooling systems.

2.8 Physical Servers

Physical servers correspond to the actual blades and racks holding the different physical resources "together", they are related to costs largely induced by the components themselves:

- Performance/Server indicates the relative performance of a server blade/rack. It offers interesting information, as the multiplication of servers implies higher infrastructure costs (network, room space, cooling, wiring, etc.).
- The cooling system is also dramatically important for a server as it can generate important costs in term of maintenance, energy consumption and room space usage.
- In case the servers are rent, the price of the rent is the obvious cost associated with the server.

2.9 Data centers

At the very bottom of the diagram, the datacentres themselves have their own associated costs:

- Rent price: natural cost to be considered is the price for renting the datacentre premises.
- Potentially a marginal cost, but extremely important to end-users, the security of the premises have to be ensured by dedicated means (private security company or else) in order to avoid any data theft or service disruption/interruption.
- Cooling price: as mentioned above, the price for cooling the different servers and their components can be extremely elevated and represent unavoidable costs.
- Finally, the overall electricity consumption of the premises heavily impact the final price of the cloud services as it is the result of the combined resource consumption of services running within the datacentre.

3 Discussion and Future Works

This section is organized in two steps: first an analysis of the trends evoked so far in the previous section and toconclude this paper some light shedding on the most promising and efficient research trails to follow.

3.1 Current State

As illustrated, the cloud ecosystem relies on numerous layers, each of them is bringing its own complexity and inherent costs. The reduction of third-party service cost is largely (directly or indirectly) treated in the literature and with the emergence of new brokering platforms supporting the automatic discovery and orchestration of services, the rationalization of orchestration costs will come naturally. In a second time, SLAs have taken a specific place for the cloud computing world, as they represent the legal contract binding the providers to its obligations. As we have detailed, SLAs can be two-sided and a too constraining SLA can imply costs and reduce the final efficiency. Virtualized environments (OS, applications, services) have their own costs, and even if it is possible to adjust them, the final impact on the service cost can be limited (depending on the service usage). The provider of virtual machines are now relatively well established and the performance provided as well as the performance needed to manage the VMs is well controlled, however there are still progresses to be made in the anticipation of resources provisioning and liberating, especially in the case of sudden increase/decrease in requirements. On this matter, in [11] Armbrust et al. argue that elasticity – the ability to quickly scale up or down one's resource usage – is an important economic benefit of cloud computing as it transfers the responsibility of adjusting the available resources to the provider. Network bandwidth also plays an important role, as it can become the bottleneck of the datacentre, especially if the latter is installed in a remote area with low connection speed. From the physical resources perspective, a number of actions can be undertaken. The current evolution of components (processors, memory) is not anymore to blindly try to produce the fastest hardware. Indeed, a specific emphasis is currently taken on the power efficiency, the reduction of heat dissipation, allowing embedding components in smaller form factors (smartphone, tablets, wearable devices - fitness bracelet, watches, glasses, etc.). Another trend for the past years has been to change the way data centers or at least clusters are built. Indeed, especially in the domain of high-performance computing, GPU have made a remarked entrance on the market by proposing highly parallelized architecture, suited for accomplishing myriads of small tasks. In [17] the author proposes a cloud infrastructure based on GPUs. Without changing so radically their infrastructure, datacenters can still evolve to reduce their long term costs by investing into low-power processor. Indeed, even if CPU vendors are offering a large variety of processors, they noticeably propose processors with lower thermal and energy specifications, without compromising the performances. Such processors are more expensive, but prove to save costs in the end. In [14], Barroso and Hölzle illustrated the significance of different cost parameters such as energy and server procurement by looking at a number of case studies that represented different deployment options. Barroso and Hölzle predicted that in the future, the costs of the data center facility (including energy usage) will become significantly larger than the actual server procurement costs. According to [11], the combined power and infrastructure costs of a data centre, in 2008, have been estimated as 40% of its total costs. The report [5] estimated that by 2014 these cost would contribute 75% of the total cost, which is a significant shift from 20% in the early 90's. However, other reports, such as [7] and [6], estimated that by 2014 these cost would contribute 75% of the total cost of a data center. These statements imply that, on the long range, researches should rather focus on reducing the cost of building and exploiting data centres. With this perspective in mind, [9] shows that exploiting dynamic price change, external temperature and delayed scheduling can significantly reduce the price for running a geographically distributed data center specialized in running batch jobs. The authors proposed an algorithm for scheduling the incoming requests based on the time-varying price of electricity and outside temperature at a given location.

As an interesting move, some researchers have tried to take a step "back" and develop tools helping them evaluate the cost of their service. In this perspective, Li et al. [15] have designed a tool to model the Total Cost of Ownership (TCO) of setting up and maintaining a cloud by taking into account the costs of hardware, software, power, cooling, staff and real-estate. Calculating the TCO of a cloud starts by taking the required number of physical servers as an input. Lin et al. also provided a method for calculating a cloud's Utilization Cost. This allows cloud providers to calculate costs in a similar manner to TCO. However, rather than inputting the number of physical servers in-to the model, they can start by inputting the maximum number of virtual machines that they need in their cloud. Modelling the Utilization Cost can be useful for cloud providers as it allows them to see and analyse a detailed breakdown of the total costs of building a cloud. This type of research paves the way toward a better cost-awareness of providers, an element essential to the development of real "pay-what-you-use" commercial offers.

3.2 Research Trails

From our analysis, some research trails appear more promising, or at least more efficient than others to reduce costs. The dynamic pricing of the cloud services is both a goal to reach and an opportunity to lower the costs of services by permanently adapting the service provision to customers' requests instead of overprovisioning through the use of flat rates and bundles.

Trail 1: One of the fundamental factors to reduce the costs of cloud services is definitely the price of electricity, as illustrated through [9], [19] a specific interest should be given to the consideration of electricity price in different region of the globe. For instance, in many countries electricity costs less during night hours (e.g. from 10pm until 7am). Beyond this day/night variation, prices may also vary several times per day. For instance, in those parts of the U.S. with wholesale electricity markets, prices vary on an hourly basis and are often not well correlated at different locations. Moreover, these variations are substantial, as much as a factor of 10 from one hour to the next. This variation of prices creates opportunities for providers to lower their costs by running computations in data center during these low-cost windows. Amazon in 2009 was the only cloud provider to launch so-called spot instances and to offer cheaper rates based on the current cost of the electricity. Spot instances offer its customers an ability to postpone their (not time-critical) workload for cheaper (i.e. night) periods by specifying the highest possible price they agree to pay. One step further to the described case, would be the possibility to move the application workload to the region with cheapest electricity rate at a given moment – that is, migrating applications from, for example, Japan to the UK and then to the US as the day passes, thus achieving lowest possible rates for running applications in a data centre.

Trail 2: Going a bit further into the exploration of such "external" factor, the consideration of weather conditions can also, indirectly, impact the costs of the cloud. For instance, an important heat wave can increase the need for cooling within a data center, consuming more electricity, resulting in a higher cost.

Trail 3: Final research trail, but probably one of the first to explore is the need for advanced and efficient cost awareness and forecast mechanisms (such as the ones proposed in [11] and [18]). Indeed, a lower electricity price does not always imply a lower price in the end, the eventual cost and risks of migrating VMs to an external location have to be evaluated. Even without considering such "extreme" cases, a cloud provider has to be able to monitor and anticipate the costs that are generated by the services he offers in order to accurately adapt their deployment, evaluate the need for redesign, etc.

In this paper we have established some of the main research trails that should be followed in order to help reducing the costs in the cloud, with the perspective of reaching the real "pay-per-use " principle. As evoked previously,a real cost-awareness and forecasting is needed to help reduce the costs at the different layers of the ecosystems. Given this consideration, our future work will focus on the investigation and design of such tools.

Acknowledgements

The research presented in this paper is supported by the European Union within the FP7 Marie Curie Initial Training Network "RELATE".

References

1. Buyya, R., Yeo, C.S., Venugopal, S., Broberg, J., Brandic, I.: Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. Future Generation Computer Systems 25(6), 599-616 (2009)
2. Ruiz-Agundez, I., Penya, Y. K. and Bringas, P. G. "A Taxonomy of the Future Internet Accounting Process"'Proceedings of the Fourth International Conference on Advanced Engineering Computing and Applications in Sciences', Florence, Italy, 2010, pp. 111-117
3. Wee, S.: Debunking Real-Time Pricing in Cloud Computing. In: 2011 11th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing. pp. 585–590. IEEE (2011)
4. Anandasivam, A., Buschek, S., Buyya, R.: A heuristic approach for capacity control in clouds. In: Commerce and Enterprise Computing, 2009. CEC'09. IEEE Conference on. pp. 90–97. IEEE (2009)
5. Han, L.: Market acceptance of cloud computing: An analysis of market structure, price models and service requirements. Tech. rep., Bayreuther Arbeitspapiere zurWirtschaftsinformatik (2009)
6. D. Durkee, "Why cloud computing will never be free," Communications of the ACM, vol. 53, May. 2010, p. 62.
7. Greenberg, A., Hamilton, J., Maltz, D.A., Patel, P.: The cost of a cloud: research problems in data center networks. ACM SIGCOMM Computer Communication Review 39(1), 68–73 (Dec 2008)
8. Jung, G., Hiltunen, M.a., Joshi, K.R., Schlichting, R.D., Pu, C.: Mistral: Dynamically Managing Power, Performance, and Adaptation Cost in Cloud Infrastructures. 2010 IEEE 30th International Conference on Distributed Computing Systems pp. 62–73 (2010)
9. Guler, H., Cambazoglu, B.B., Ozkasap, O.: Cutting Down the Energy Cost of Geographically Distributed Cloud Data Centers. In: Pierson, J.M., Da Costa, G., Dittmann, L. (eds.) Energy Efficiency in Large Scale Distributed Systems, Lecture Notes in Computer Science, vol. 8046, pp. 279–286. Springer Berlin Heidelberg, Berlin, Heidelberg (2013)
10. Jung, G., Hiltunen, M.a., Joshi, K.R., Schlichting, R.D., Pu, C.: Mistral: Dynamically Managing Power, Performance, and Adaptation Cost in Cloud Infrastructures. 2010 IEEE 30th International Conference on Distributed Computing Systems pp. 62–73 (2010)
11. Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R.H., Konwinski, A., Lee, G., Patterson, D.A., Rabkin, A., Stoica, I., Zaharia, M.: Above the clouds: A berkeley view of cloud computing. Tech. Rep. UCB/EECS-2009-28, EECS Department, University of California, Berkeley (Feb 2009)
12. Goiri, I.n., Guitart, J., Torres, J.: Economic model of a Cloud provider operating in a federated Cloud. Information Systems Frontiers 14(4), 827–843 (Sep 2011)
13. Walker, E. 2009. The Real Cost of a CPU Hour. *Computer*, 42, 4, 35-41.
14. Barroso, L. A. and Hölzle, U. 2009. The Datacenter as a Computer: An Introduction to the Design of Warehouse- Scale Machines. Morgan & Claypool Publishers.
15. Li, X., Li, Y., Liu, T., Qiu, J. and Wang, F. 2009. The Method and Tool of Cost Analysis for Cloud Computing. In IEEE International Conference on Cloud Computing (CLOUD-II 2009), Bangalore, India, September 2009, 93-100.

16. Hindman, B., Konwinski, A., Zaharia, M., Ghodsi, A., Joseph, A. D., Katz, R., Shenker, S. and Stoica, I. (2011, March). Mesos: A platform for fine-grained resource sharing in the data center. In Proceedings of the 8th USENIX conference on Networked systems design and implementation (pp. 22-22).
17. Giunta, F., Montella, R., Laccetti, G., Isaila, F., & Blas, F. (2011). A GPU accelerated high performance cloud computing infrastructure for grid computing based virtual environmental laboratory. Advances in Grid Computing.
18. Kim, K. H., Beloglazov, A., & Buyya, R. (2009, November). Power-aware provisioning of cloud resources for real-time services. In Proceedings of the 7th International Workshop on Middleware for Grids, Clouds and e-Science (p. 1). ACM.
19. Le, K., Bianchini, R., Zhang, J., Jaluria, Y., Meng, J., & Nguyen, T. D. (2011, November). Reducing electricity cost through virtual machine placement in high performance computing clouds. In Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis (p. 22). ACM.
20. Martens, B., & Teuteberg, F. (2012). Decision-making in cloud computing environments: A cost and risk based approach. Information Systems Frontiers, 14(4), 871-893.