

Initial Thoughts on Business-driven IT Management Challenges in Cloud Computing Providers

J. Oriol Fitó and Jordi Guitart

Barcelona Supercomputing Center and Technical University of Catalonia

Barcelona, Spain

{josep.oriol, jordi.guitart}@bsc.es

Abstract—Nowadays Cloud computing is recognized as the most emerging computing paradigm. Because of its promising benefits, every day more and more enterprises are relying on Cloud systems. Furthermore, new Cloud business models are appearing, most of them within the SaaS marketplace, which fully depend on PaaS and IaaS providers. In any case, the expectation from businesses that IT (Cloud) services and infrastructures should bring them closer to the achievement of their Business-Level Objectives (BLOs) is spreading. Due to this fact, the presence in Cloud providers of a self-management of Cloud services and infrastructures driven by business-level aspects is mandatory. In this direction, the Business-Driven IT Management (BDIM) discipline has been evolving as the most promising way in the sense of aligning IT (low-level) management decisions with business-level objectives coming from providers themselves, as well as from their users. In this paper, we expose several BDIM challenges on the Cloud computing paradigm. Consequently, we outline key issues for the inclusion of BDIM-related features into the core operation of Cloud providers.

I. INTRODUCTION

Up to now, Cloud computing is emerging as the most popular computing paradigm of nowadays [1]. The use of Cloud systems leads to promising business models and, actually, benefits for both stakeholders, i.e. providers and end-users, are very clear [2]. Furthermore, Cloud providers are led to participate in Cloud federations, where resources from disparate providers (offered as public Clouds) are combined in novel ways. In these scenarios, each participating provider can considerably extend the bounded computing capacity of its own resources (private Cloud) through outsourcing deals with third-party providers, as well as improve the profitability of its private Cloud by insourcing resources from others.

Besides, businesses in general are need of new ways to efficiently achieve their high-level goals, known as Business-Level Objectives (BLOs). Rather than optimizations of managerial appearance, the adoption of other groundbreaking ways to improve business interests, strategies and goals must be explored. In actual fact, the reliance of businesses on Cloud systems is also growing significantly. In this regard a recent Gartner report [3] points Cloud computing as the technology that will be the most strategic for companies of today. Thus, the contribution that Cloud entities could have in improving the fulfillment of BLOs seems extremely promising. To that end, an efficient self-management of these Cloud entities aware of business interests is a research challenge that needs to be totally approached. The leading motivations for

developing such management are: (1) Cloud infrastructures are evolving into very complex systems in terms of dimension and management; (2) there are several aspects to be tackled in order to efficiently manage Cloud services during their whole life cycle; (3) adopting a BDIM model is a significant requirement for an effective IT governance in businesses; and, as mentioned before, (4) the increasing expectation from businesses that IT systems should assist in achieving their BLOs. In this direction, the Business-Driven IT Management (BDIM) discipline is clearly very appropriate to be widely used in Cloud providers. It aims at rethinking IT management from a business point of view.

In this paper, we outline some key research challenges in BDIM for Cloud computing. In particular, we focus on how Cloud providers driven by business-level objectives should (self-)manage their Cloud entities in order to aggregate significant values to the business on top of it, as well as to their users (which actually may be other providers). In brief, what is sought is to carry out IT-level management decisions aligned with business objectives. As stated before, business-driven self-management processes should be supported by proper BDIM models. The latter are responsible of producing *Key Performance Indicators (KPI)* of the business, also known as BDIM metrics, by evaluating, in a synergistic way, disparate business-level parameters (BLPs). These can be financial, such as revenue, costs, loss, etc.; or non-financial like risk, trust and energy efficiency. Beyond this, risks (and uncertainties) involved in each possible management decision deserve an special attention. In reality they can be considered either as an opportunity or as a threat to the organization's success.

II. RELATED WORK

A main research agenda for business-driven IT management is presented in [4], where both key BDIM-related concepts and crucial research challenges are outlined. In particular, ideas on BDIM research in both autonomic computing and outsourcing deals somehow serve as the basis of BDIM challenges in the Cloud computing. Moreover, a vision for a business-driven adaptive IT infrastructure is described in [5]. Actually, almost all ideas presented in this paper also represent the foundation of BDIM issues specific to Cloud computing.

Up to now, autonomous IT management processes driven by business-level aspects have been proposed in several works. Aiber et al. [6] present a general architecture, and a set

of technologies and methodologies enabling autonomic self-optimization according to BLOs. However, they only consider the goal of maximizing the income of a single e-commerce site. The ability to deal with changes in the IT infrastructure is explained. Nevertheless, they avoid to present a validation test about how to recognize and react to these changes in order to constantly keep the IT infrastructure optimized with regard to BLOs. For this reason, approaches like this must be extended in order to be used by Cloud providers driven by several (disparate) BLOs. Furthermore, some research efforts have used BDIM methodologies to also increase the business value of e-commerce applications. For instance, [7] and [8] face the problem of designing IT infrastructures for such e-commerce contexts by taking into account their business impact, such as the revenue loss due to IT infrastructure failures. Notably, these approaches do not provide dynamism when allocating resources to services, which is extremely required to deal with typical changes in the environment (e.g. demand variations).

Regarding specification and management of SLAs, Sauvé et al. [9] present a method whereby the best values for Service Level Objectives (SLOs) of an SLA are chosen by considering the business perspective. However, the impact model proposed only includes business losses due to failure or performance degradation of IT components. In addition, they consider static resources allocations, thus dealing only with the IT services' provisioning (deployment) phase. Afterward, they pose the problem of planning and scheduling changes in business-driven IT management [10]. An architecture which follows the principles of BDIM is proposed. Unfortunately, the evaluation is based on a numerical illustration, and they only address the BLO of minimizing business loss. In a similar research direction, Abi et al. [11] introduce a business-driven framework for the management of utility computing environments. It is aimed to drive the management of IT entities from the business point of view. However, these last two works need human intervention when applying management actions.

The virtualization technology, which is the basis of any Cloud environment, allows to overcome such limitations. Cloud providers are able to dynamically allocate resources to services in order to meet changing conditions (i.e. time-varying workloads) that affect their BLOs. More to the point, the services' life cycle can be self-managed (automated) through corrective low-level actions. In any case, the BDIM discipline is becoming a promising way for an accurate and efficient IT management in recent and future years, and in several disparate Cloud scenarios. Nevertheless, there is a lot to be researched before BDIM methods can become widely applicable in the Cloud computing realm.

III. BDIM IN THE CLOUD COMPUTING

In essence, BDIM can be seen as IT Service Management (ITSM), which is based on concepts and practices of the Information Technology Infrastructure Library (ITIL) approach [12], together with business metrics. BDIM is defined in [13] as a new, evolutionary and comprehensive IT management approach aimed to improve IT infrastructure,

Quality of Service (QoS) and business results at the same time. Therefore, it involves a new culture, tools and decision-making processes that help businesses in achieving their BLOs. And in this sense the Cloud computing paradigm opens a wide range of opportunities in terms of extending traditional BDIM approaches, which mainly lack automation capability.

As pointed in the previous section, IT resources have been traditionally allocated to applications in an static and dedicated manner. This fact has meant several drawbacks for classic IT providers, since they have to allocate the amount of resources required to support peak demands (which usually cannot be predicted a priori). This archaic approach implies huge monetary investments for providers. Otherwise, the QoS offered is rather poor and, thus, both customers' satisfaction and provider's reputation are clearly diminished. In this regard, the Cloud computing and its inherent technologies, mainly the virtualization one, offer unprecedented possibilities in terms of completely self-manage Cloud systems and their entities: provides isolation to Cloud services while they share the same physical infrastructure, thus enabling their consolidation; allows dynamic resources allocation to services during their operation; offers the capability to migrate services (virtual machines) from one node to another, etc. All these low-level IT management capabilities give support to a wide range of new and promising BDIM scenarios for Cloud providers, which certainly will lead them to improve the fulfillment of their BLOs. In a nutshell, the use of virtualized resources as a whole represents a shift from traditional human-based to automated BDIM. Besides, there is a shift from offline to online IT management. The latter is now possible due to the aforementioned feature of reassigning resources to services on the fly. Among others, these shifts mean that any action needed to constantly adapt IT (Cloud) systems, according to changes in the business environment, can be done more efficiently.

More to the point, the federation of Clouds has to be taken into account. Any provider involved in this multi-Cloud scenario can insource and/or outsource workloads (resources) from/to other participants according to business and services requirements. The business impact of these operations is another BDIM challenge in the Cloud that needs to be addressed by the research community. Moreover, and as shown in next subsection IV-A, traditional BDIM models can be partially, or even totally, reused by Cloud providers. Such models together with new self-management optimization processes, relying on virtualization-level actuators, seem to be very promising for the success of Cloud providers guided by BLOs.

A. Adapting SLAs to New Cloud business Models

On one hand, the success of Cloud computing and relationships between Cloud providers and their customers (or between Cloud providers themselves) requires the definition of concrete conditions of service (e.g. QoS), which are commonly expressed into SLAs. Moreover, and in cases of a federation of Clouds, it is also essential to well-establish the QoS that providers offer between themselves. On the other hand, there are emerging new Cloud business models that fully rely

on external providers, such as online gaming services. This kind of knowledgeable customers, which are aware of Cloud advantages and dangers, will want to make decisions based on the highest profitability for them, as well as at the lowest risk and highest trust. In fact, they need quite special service's conditions since their whole business is based on external Cloud providers. In addition, several billing models are arising together with such paradigm. In this regard, Enomaly [14] is now assisting their customers on defining personalized pricing schemes. For instance, a lot of its customers choose to use quota based month subscription models. This represents a chance for Cloud providers because it generates revenue even if those customers never use their assigned Cloud capacity.

Due to these facts, an extension of current SLAs in order to determine business-level conditions (e.g. economic-related parameters and models, strict QoS, risk to be taken) into them is needed. These extended SLAs will facilitate the understanding that both the business on top of Cloud providers and their end-users have with respect to governing conditions of services. Furthermore, it will be very useful for outsourcing deals, where the outsourcer (an end-user or another provider of services) needs to agree high-level constraints. These can be determined by either SLAs' statements already agreed with its end-users or from its BLOs themselves.

IV. BUSINESS-DRIVEN CLOUD COMPUTING PROVIDERS

We aim to design and implement self-management of Cloud entities, i.e. services and infrastructures, aware of BLOs. Its main goal is to fulfill those high-level goals, while administrating and managing Cloud entities in an efficient way through the analysis of determinant business-level parameters (BLPs).

A. BDIM model

The core management operation in business-driven Cloud providers undoubtedly needs the support of suitable BDIM models for the scenario considered. In this sense our proposal (depicted in Figure 1) is envisioned to be enough generic and adaptable to be used by any Cloud provider driven by BLOs.

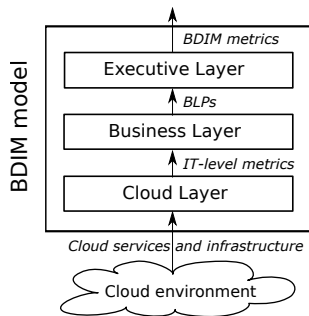


Fig. 1. BDIM model architecture.

The topmost layer, i.e. *Executive Layer*, captures variations in BDIM metrics, e.g. profit and clients' satisfaction, caused by fluctuations on BLPs relevant to each objective considered. This high-level impact is assessed by an upper linkage model $Link_{up}$, and resulting BDIM metrics are actually the result of

the whole model and must be understood by business people. The complexity of this layer increases when multiple BLOs are to be achieved, since different BDIM metrics are assessed and, therefore, taken into account in the decision making process. In these convoluted cases, several tradeoffs may arise and must be solved by suitable business-driven management policies. This multiple definition is a key property for contributing with a general and extensible BDIM model which may be used for Cloud providers with any kind and any number of BLOs.

The intermediate level, namely *Business Layer*, models how IT-related metrics (such as throughput and energy consumption) or events (new services deployment, SLA violations, resource failures, ...) affect changes in BLPs, like revenues or losses, energy efficiency, risk and trust. Consequently, it provides to the upper layer a determination of BLPs that are significant for BLOs determined by executives. The matching between those high- and low-level metrics (business- and IT-related, respectively) is assessed by the lower linkage model $Link_{low}$. Moreover, SLAs have an important role in this in-between layer, because this $Link_{low}$ is mainly supported by both statements and parameters specified in them.

Finally, the *Cloud Layer* extracts, by means of a virtualization platform-independent monitoring system, IT-level metrics concerning both services deployed on the provider and the underlying federated infrastructure used. This kind of monitoring tool is needed due to the use of resources from third-party providers, which may be offered with different virtualization technologies. Examples of these low-level metrics are service's response time (seconds), service's availability (percentage of annual uptime), service's completion deadline (date), and private Cloud's capacity, to name but a few.

Cloud providers adopting this model will be able to assess at any moment the achievement or not of BLOs through the quantification, expressed with final business metrics (i.e. BDIM metrics), of business impacts due to IT-level events.

B. Business-driven IT Self-management Optimization Loop

Beyond the competences and outcomes of the BDIM model, we consider its involvement in an optimization procedure that is essential to perform an efficient and accurate IT self-management driven by business aspects. Its main goal is to foresee the consequences on business results, in terms of BDIM metrics, of possible management solutions that seem apposite to ameliorate organization's BLOs when IT-related events occur. To that end, BLOs, business-driven policies and the BDIM model must be defined a priori by the business. Notice that this is the minimum and inevitable human intervention in the whole optimization process, which operates completely autonomously once those business aspects are determined. Actually, both BDIM model and policies can be modified during the process without affecting its proper operation.

The optimization loop is as follows: when either an IT-level event takes place or an end-user asks for a new service to be deployed, an adaptation of the Cloud environment is required. For this purpose, a policy management framework decides all possible management solutions to face this adaptation

with regard to both the current status of the Cloud environment and its own policies. Afterward, the BDIM model, which also considers the current infrastructure status, uses IT-level metrics in order to provide the foreseen impacts that these solutions will have on organization's BLOs if they are performed. The determination of these predictions is very important, since it simulates the business impact (i.e. instantaneous BLOs fulfillment) of each candidate solution. Based on these, business-driven policies determine and apply the most efficient low-level management action in terms of business results and, moreover, solve all possible tradeoffs among conflictive objectives. To put an example, the maximization of the ecological efficiency can be achieved by consolidating services and powering off underused nodes. However, these actions might lead to an undesired minimization of the QoS offered and, consequently, clients' satisfaction might be also diminished. The idea is that management policies decide to carry out the action that contributes better to the overall BLOs fulfillment. Anyway, once the Cloud environment is adapted, the model recalculates BDIM metrics and sends them to business executives, who may decide to modify BLOs, business-driven policies and the BDIM model.

C. Use Case: Business's Profit Maximization

In particular to this use case, the BDIM model is composed by two linkage models: the upper one assesses the profit as the sum of incomes minus the amount of costs and losses; while, therefore, the lower one provides a quantification of these BLPs by evaluating both IT-level status and metrics of Cloud systems. Moreover, an efficient business-driven management policy should consider these business-level metrics (i.e. profit, loss due to SLA violations, etc.) when applying management actions. This is very important since placing a new service in a given node can affect the performance of other ones already running on it, and thus incurs SLA penalties due to poor QoS. Anyway, when an IT-level event takes place, the BDIM model provides to the policy management framework the profit that will be earned in the case of applying each possible management action. Then, and based on these disparate business results, the policy determines and applies the most efficient low-level management action in terms of fulfilling the BLO.

Besides, imagine a case where the business decides to add a new objective: eco-efficiency maximization. So, now we have a situation with multiple BLOs and several tradeoffs may appear. For instance, accepting a lot of services would lead to profit maximization. However, this fact would imply to have powered on all in-house nodes, with the corresponding undesired minimization of the ecological efficiency. In any case, this kind of tradeoffs must be solved by the policy management framework and based on weights of importance that the business give to each business-level objective.

V. CONCLUSIONS

Due to both the promising emergence of the Cloud computing paradigm and the rising expectation from businesses that IT systems should contribute to the achievement of their

strategies and goals, the research community needs to address the inclusion of the BDIM discipline in the Cloud computing realm. In particular, the management of Cloud services and infrastructures carried out by Cloud providers have to be governed by business-level aspects. Additionally, innovative and suitable self-management optimization procedures driven by business objectives, like the one presented in this paper, are mandatory to ensure businesses success. Our future work will be focused on the research directions explained herein.

ACKNOWLEDGMENT

This work is supported by the Ministry of Science and Technology of Spain and the European Union (FEDER funds) under contract TIN2007-60625, by the Generalitat de Catalunya under contract 2009-SGR-980, and by the European Commission under FP7-ICT-2009-5 contract 257115 (OPTIMIS).

REFERENCES

- [1] R. Buyya, C. Yeo, and S. Venugopal, "Market-oriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities," in *10th IEEE International Conference on High Performance Computing and Communications, (HPCC'08)*, 2008, pp. 5–13.
- [2] M. Armbrust, A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica *et al.*, "Above the clouds: A Berkeley view of cloud computing," *EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2009-28*, 2009.
- [3] Gartner, Inc. analysts, "Gartner Identifies the Top 10 Strategic Technologies for 2010," *Gartner Report (press release)*, October, 2009.
- [4] A. Moura, J. Sauvé, and C. Bartolini, "Research challenges of Business-Driven IT management," in *2nd IEEE/IFIP International Workshop on Business-Driven IT Management, (BDIM'07)*, 2007, pp. 19–28.
- [5] V. Machiraju, C. Bartolini, and F. Casati, "Technologies for business-driven it management," *Extending Web Services Technologies: the Use of Multi-Agent Approaches*, Kluwer Academic Publishers, 2004.
- [6] S. Aiber, D. Gilat, A. Landau, N. Razinkov, A. Sela, and S. Wasserkrug, "Autonomic self-optimization according to business objectives," in *Proceedings of International Conference on Autonomic Computing, (ICAC'04)*, 2004, pp. 206–213.
- [7] J. Sauvé, F. Marques, A. Moura, M. Sampaio, J. Jornada, and E. Radziuk, "Optimal Design of E-Commerce Site Infrastructure from a Business Perspective," in *Proceedings of the 39th Annual Hawaii International Conference on System Sciences, (HICSS'06)*, vol. 8. IEEE, 2006, p. 178c.
- [8] F. Marques, J. Sauvé, and A. Moura, "Business-oriented capacity planning of IT infrastructure to handle load surges," in *10th IEEE/IFIP Network Operations and Management Symposium, (NOMS'06)*. IEEE, 2006, pp. 1–4.
- [9] J. Sauvé, F. Marques, A. Moura, M. Sampaio, J. Jornada, and E. Radziuk, "SLA Design from a Business Perspective," in *Proceedings of the 16th IFIP/IEEE International Workshop on Distributed Systems: Operations and Management, (DSOM'05)*. Citeseer, 2005, p. 72.
- [10] J. Sauvé, R. Rebouças, A. Moura, C. Bartolini, A. Boulmakoul, and D. Trastour, "Business-driven decision support for change management: planning and scheduling of changes," *Large Scale Management of Distributed Systems, 17th IFIP/IEEE International Workshop on Distributed Systems: Operations and Management, (DSOM'06)*, pp. 173–184, 2006.
- [11] I. Aib, M. Salle, C. Bartolini, and A. Boulmakoul, "A business driven management framework for utility computing environments," in *Proceedings of the Ninth IFIP/IEEE International Symposium on Integrated Network Management (IM'05)*. IEEE, 2005, pp. 16–19.
- [12] I. ITIL, "IT Infrastructure Library for Service Management."
- [13] J. Sauvé, A. Moura, M. Sampaio, J. Jornada, and E. Radziuk, "An introductory overview and survey of business-driven it management," in *The First IEEE/IFIP International Workshop on Business-Driven IT Management, (BDIM'06)*, 2006, pp. 1–10.
- [14] "Enomaly ECP Service Provider Edition," 2011, <http://www.enomaly.com/>.