Self-establishing a Service Level Agreement within Autonomic Cloud Networking Environment

Mohamad Hamze, Nader Mbarek, Olivier Togni
Le2i Laboratory UMR 6306 CNRS, University of Burgundy
Dijon, France
{Mohamad.Hamze, Nader.Mbarek, Olivier.Togni}@u-bourgogne.fr

***Abstract*—Today, cloud networking which is the ability to connect the user with his cloud services and to interconnect these services within an inter-cloud approach, is one of the recent research areas in the cloud computing research communities. The main drawback of cloud networking consists in the lack of Quality of Service (QoS) guarantee and management in conformance with a corresponding Service Level Agreement (SLA). Several research works have been proposed for the SLA establishing in cloud computing, but not in cloud networking. In this paper, we propose an architecture for self-establishing an end-to-end service level agreement between a Cloud Service User (CSU) and a Cloud Service Provider (CSP) in a cloud networking environment. We focus on QoS parameters for NaaS and IaaS services. The architecture ensures a self-establishing of the proposed SLA using autonomic cloud managers.**

I. Introduction

Cloud computing is a promising technology for the realization of large, scalable on-demand computing infrastructures. Many enterprises are adopting this technology to achieve high performance and scalability for their applications while maintaining low cost. Furthermore, an important differentiation element between cloud environments is the Quality of Service and Service Level Agreement. In general, a Cloud Service User (CSU) requires for his services an end-to-end Quality of Service (QoS) assurance, a high level reliability and a continued availability. However, there are limitations when users and/or applications rely on guaranteed QoS between deployed cloud infrastructures that are spread across geographically distributed sites or domains.

Therefore, inter-cloud systems are unavoidable as it is very difficult for a single Cloud Service Provider (CSP) to satisfy its customer requirements. In addition, if a cloud is out of resources, it could not meet the new demands of its users. Thus, geographically distributed data centers offer better end-to-end performance between CSU and CSP and improve reliability when failure occurs. On the other hand, the inter-cloud should be designed as a multi-vendor environment with the ability to migrate services from one provider to another and to locate the best resources not only in terms of computing capacity and storage, but also connectivity, bandwidth and delay. Therefore, the networking aspect of cloud computing is a critical factor for adopting this approach.

However, to achieve the cloud networking model, there are a number of open issues, such as the formality of a language for SLA between CSUs and CSPs, the interoperability of data formats, and communication using standard interfaces. In addition, there are major challenges on how to establish an SLA in cloud networking and how to choose suitable CSPs during the execution based on SLA and QoS requirements. Furthermore, the development of an autonomic cloud control aims to simplify the complexity and maximize efficiency while minimizing cost and user interactions.

978-1-4799-0913-1/14/$31.00 © 2014 IEEE

In this paper, we propose an architecture for self-establishing an end-to-end service level agreement for cloud networking. We propose at first a cloud networking architecture. We describe the splitting of SLA parameters and the interaction between different entities to establish the SLA. Then, we present the SLA self-establishing architecture using autonomic cloud managers.

The remainder of this paper is organized as follows: in Section II, we present a brief overview on the state of the art for cloud computing and networking. Section III highlights the most relevant research works and trends in this area. In Section IV, we describe the proposed architecture for establishing an end-to-end SLA in cloud networking. In Section V, we describe our self-establishing architecture. Lastly, Section VI concludes the paper and points out future work.

II. State of The Art

A. Cloud Computing

In 2011, the NIST [1] defined cloud computing as a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. It defined three service models: Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS).

B. Cloud Networking

The term cloud networking is introduced in a multi-administrative domain scenario, where network and data center domains must interact through defined interfaces to provide a service to CSU. Cloud networking allows specifying a service through high-level objectives or goals, rapid deployment and management of cloud services across data centers and operator networks, and autonomous optimization and management of underlying resources to provide and maintain requested QoS [2]. Cloud networking extends network virtualization beyond the data center to bring two new aspects to cloud computing: (a) connecting the user to services in the cloud, and

(b) interconnecting services that are geographically distributed across cloud infrastructures. Cloud networking users would be able to specify their needed virtual infrastructure and the desired networking properties to access these resources [3]. There are three kinds of networks in cloud environment [4]:

1) Intra-cloud network: this network connects local cloud infrastructures. The data centers require different traffic models to guarantee latency, packet loss and bandwidth requirements.

2) Core transport network (WAN/MAN): this network can be used by customers to access and consume cloud services deployed within the cloud provider's data center.

3) Inter-cloud network: it can be used to interconnect cloud infrastructures together. These cloud infrastructures may be owned by the same cloud provider or by different ones.

In addition, cloud networking services that provide network connectivity in the three networks above are named Network as a Service (NaaS). NaaS allows tenants to customize the services they receive from the network and the provider to decide how resources are allocated and shared between tenants. NaaS requirements include QoS, security and availability.

Indeed, the performance of data transfer in core transport and inter-cloud networks is difficult to guarantee. The most convenient solution is to use enough bandwidth to carry information in these networks. This bandwidth is generally provided by any entity that has the ability to provide network connectivity and reservation, e.g. Network Service Provider (NSP). Furthermore, the NaaS service offered by these NSPs is the Bandwidth on Demand (BoD). BoD is the ability to perform on-demand changes (increase, decrease) and instant provisioning of bandwidth on the particular links via standardized interfaces. The BoD supports the establishment of an end-to-end SLA between CSU and CSP [4].

C. Inter-Cloud

Inter-cloud computing allows on-demand assignment of cloud resources, including computing, storage and network, and the transfer of workload through interworking of cloud systems [4]. The communication between different clouds is established through the cloud networking that can provide a NaaS service (BoD). CSPs can interwork in Inter-cloud Service Broker (ISB) manner. It provides indirect interconnection between two or more CSPs and brokering services to CSUs or CSPs. The ISB provides and executes services of three categories [4]: (a) Service intermediation: the ISB enhances a given service by improving some specific capability and providing value-added services to CSUs. (b) Service aggregation: the ISB combines and integrates multiple services into one or more new services. (c) Service arbitrage: it is similar to service aggregation except that the services being aggregated are not fixed. In this case, the ISB has the flexibility to choose services from multiple sources.

III. Related Work

The research project SAIL (Scalable & Adaptive Internet Solutions) [5] describes a cloud networking architecture and focuses on security, but it does not consider the QoS and SLA. The research project Foundation of Self-governing ICT Infrastructures (FoSII) [6] is proposing solutions for autonomic management of SLAs in the cloud. It is investigating self-governing cloud computing infrastructures necessary for the attainment of established Service Level Agreements (SLAs). In addition, the research project Contrail [7] aims to vertically integrate an open-source distributed operating system for autonomous resource management in IaaS environments and PaaS systems. We consider our work to be very much in alignment with the objectives of these projects. But, our research work is innovative by considering the self-establishing of SLA in cloud networking environment, and by focusing on QoS guarantee for IaaS and NaaS services.

From standardization perspective, IEEE Cloud Computing formed the Inter-Cloud Working Group (ICWG). It announced the launch of two new standards development projects in April 2011: P2301 [8] and P2302 [9]. Open Grid Forum (OGF) is active in the definition of the Open Cloud Computing Interface (OCCI) [10] for the interoperability between clouds. GICTF (Global Inter-Cloud Technology Forum) [11] studies the standard Inter-Cloud interfaces to improve the reliability of the Clouds, and presents SLA metrics, including security attributes for Inter-Cloud environment. IBM recently presented CloudNaaS in [12], a cloud networking platform for enterprise applications. In our research work we propose to develop a cloud networking architecture and we aim to enable communication between, not only CSPs, but also broker, NSPs and CSPs. In addition, this communication is self-established using autonomic cloud managers and through agreed-upon interfaces. For that purpose, we use web services standard technologies.

Finally, there are many related research works on QoS, e.g. [13][14], but QoS mentioned in these works is for SaaS, PaaS or IaaS. In addition, several research works present SLA for cloud computing only: the project Mycloud [15] proposes Cloud Service Level Agreement (CSLA) and the authors in [16] propose to use Web Service Level agreement (WSLA) [17] in a cloud computing context. Therefore, our paper considers, in addition, NaaS services, QoS attributes and an SLA for autonomic cloud networking environment.

IV. Proposed Cloud Networking Architecture

In order to deliver the service level expected by CSU, the proper design, construction and management of a cloud networking architecture is very critical. Therefore, we propose to design an architecture for inter-cloud broker (Fig. 1). We assume an environment with multiple CSPs connected through several NSPs. The CSP can offer resources from one or several datacenters (DC). Each entity (broker, CSP, NSP) has a repository that contains information about resources and service levels. Within this cloud networking architecture, we must ensure consistency between the QoS requirements requested by the CSUs, and the SLA proposed by CSPs to allow multiple CSPs working together to meet the CSU requirements.

A. Cloud Networking Broker Architecture

In the proposed cloud networking broker architecture (Fig. 1), the CSU is connected to the cloud broker. The cloud broker is emerged as an intermediate entity between a CSU and CSPs to help the establishing of a service level that meets the CSU requirements and to simplify secure integration and lower

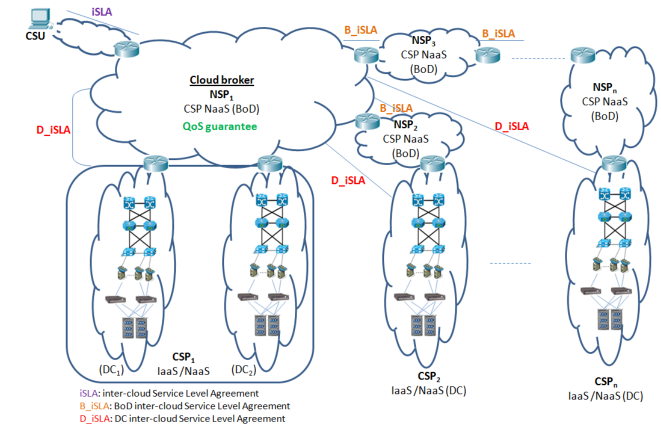

Fig. 1. Cloud Networking Broker Architecture

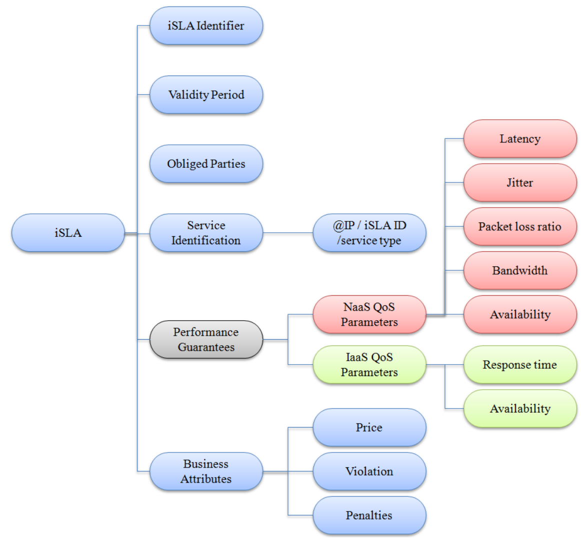

Fig. 2. Inter-cloud Service Level Agreement (iSLA) Structure

prices for CSU. It can help to move data between different CSPs. Thus, we propose the NSP as part of our architecture, providing BoD network service and playing the role of a NaaS CSP. It can be a cloud broker or a network operator. In addition, the CSP provides IaaS and NaaS services. The IaaS service concerns VM and storage resources and the NaaS service concerns network datacenter resources that connect local IaaS resources.

B. SLA description

In cloud environment, a Service Level Agreement (SLA) is a contract between CSU and CSP. It contains details of shared information and service level guarantees that are offered by CSP and required by CSU. An SLA contains both technical and non-technical parameters. The technical parameters constitute the negotiable part of the SLA, and are grouped together in a specification called Service Level Specification (SLS). The SLS is a set of parameters such as QoS parameters and their corresponding values which allow the definition of the service level offered to CSU. The SLA can be structured by: SLA Identifier, Validity Period, Obliged Parties, Service Identification, Guarantees offered on the Services, and Business Attributes (price, violation and penalties).

Several works have been done on the SLA establishing in cloud computing, but not in cloud networking. In addition, most of the current CSPs are limited only to the availability of resources and do not take into account many other important parameters such as latency and bandwidth [18]. In our cloud networking broker architecture we propose three types of SLA:

1) inter-cloud Service Level Agreement (iSLA): It is a contract between a CSU and a cloud broker with QoS parameters for NaaS and IaaS services. These QoS parameters could be quantitative or qualitative. The most important quantitative QoS parameters of an iSLA are shown in Fig. 2.

2) BoD inter-cloud Service Level Agreement (B_iSLA): It is a contract between a cloud broker and directly connected NSP or between several NSPs that enable a cloud broker to reach CSU and CSPs. It is for NaaS (BoD) services. QoS parameters are the same as iSLA NaaS QoS parameters.

3) Datacenter inter-cloud Service Level Agreement (D_iSLA): It is a contract between a cloud broker and a CSP for NaaS and IaaS services. QoS parameters are the same as specified within the iSLA Performance Guarantees attribute.

C. Interaction between entities

The interaction between CSU, Cloud Broker, CSPs and NSPs to achieve an end-to-end SLA is described in Fig. 3. At first, CSPs and directly connected NSPd describe IaaS and NaaS services for their available resources with different service levels and send this information to the cloud broker. In addition, if any changes occur within CSPs and NSPd, they send these changes to the cloud broker which updates its repository. Next, the CSU sends service requirements to the cloud broker to construct an iSLA specifying a service level that contains QoS parameters for NaaS and IaaS services. The cloud broker consults its repository and selects the appropriate CSPs and NSPd that meet the CSU QoS requirement. Afterward, the cloud broker sends a request to establish a D_iSLA (IaaS, NaaS) and a B_iSLA (NaaS BoD) with respectively selected CSPs and NSPd. Then, directly connected NSPd establish a B_iSLA with other NSPs that enable the cloud broker to reach CSU and CSPs. After this, the concerned CSPs reserve and configure VM and network resources to deliver IaaS and NaaS services, and the concerned NSPs reserve and configure network resources to deliver NaaS (BoD) service. Then, they update their repository with these changes and send confirmation to cloud broker. Finally, the cloud broker reserves and configures local network resources to deliver NaaS (BoD) service with QoS guarantee. Then, it updates its repository with these changes and establishes the iSLA with CSU.

V. AUTONOMIC CLOUD NETWORKING ARCHITECTURE

The availability of a reliable system for self-establishing the SLA reduces complexity and minimizes the user interaction. In this section, we propose to establish the iSLA, B_iSLA and D_iSLA corresponding to our cloud networking architecture autonomously using Autonomic cloud Managers (AMs).

In general, a domain refers to resource collections involving hardware and software managed by a single entity, e.g. cloud

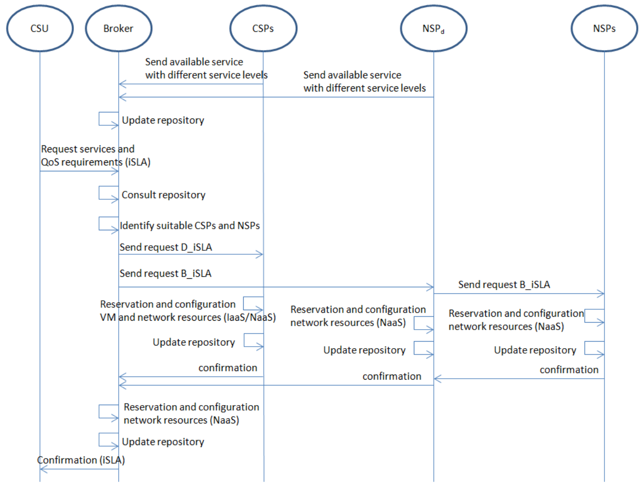

Fig. 3. MSC for Inter-cloud Broker Scenario

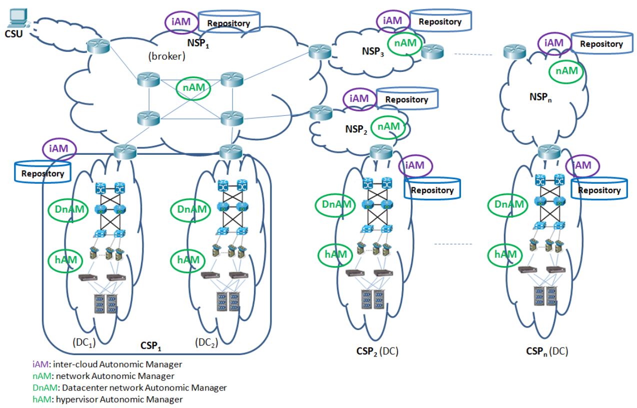

Fig. 4. Autonomic Cloud Networking Architecture

data center, or communication network. If we manage this domain in an autonomic manner, we call it Autonomic Domain (AD). In this context, we present the proposed cloud networking architecture with several Autonomic cloud Domains (Fig. 4). Each Autonomic cloud Domain (cloud broker, CSPs and NSPs), is under the authority of an inter-cloud Autonomic Manager (iAM). iAM communicates with other iAMs to achieve an agreement on a service level. In addition, it controls one or more AMs to configure resources in conformance with an agreed on service level.

These Autonomic cloud Managers are playing different roles within our autonomic cloud networking architecture:

1) network Autonomic Manager (nAM): it is responsible for creating and managing NSP virtual networks connecting CSU to CSP or connecting CSPs in conformance with the agreed B_iSLA, and monitoring workload and performance.

2) Datacenter network Autonomic Manager (DnAM): it is responsible for creating and managing virtual networks for the CSU within the cloud data center in conformance with the NaaS part of the D_iSLA, and regularly monitoring workload and performance.

3) hypervisor Autonomic Manager (hAM): it is responsible for creating, monitoring and controlling virtual machines (VMs) and storage capacity. Therefore, the CSP can consequently decide the allocation or deallocation of VMs to maintain the performance as close as possible to performance objective in conformance with the IaaS part of the D_iSLA.

We provide these autonomic cloud managers with the capability to achieve an agreement between the Autonomic cloud Domains.

VI. Conclusion

In this paper, we presented an architecture for self-establishing a service level agreement within cloud networking environment. At first, we have proposed an architecture of cloud networking. We have focused on QoS parameters for IaaS and NaaS services. In addition, we have presented the splitting of these QoS parameters and the establishing of an inter-cloud service level agreement (iSLA). Finally, we have proposed the use of autonomic cloud managers for the iSLA self-establishment.

As a future work, we intend to propose an algorithm for CSP and NSP selection, and present a concrete validation platform for our broker architecture. Then, we want to include security parameters in our proposed iSLA and study their impact on QoS parameters.

References

[1] NIST, ”The NIST Definition of Cloud Computing,” Special Publication 800-145, USA, 2011.

[2] P. Murray, A. Sefidcon, R. Steinert, V. Fusenig and J. Carapinha, ”Cloud Networking: An Infrastructure Service Architecture for the Wide Area,” Future Network & Mobile Summit 2012: 21st Annual Conference Proceedings, 2012.

[3] P. Schoo, V. Fusenig, V. Souza, M. Melo, P. Murray, H. Debar, H. Medhioub and D. Zeghlache, ”Challenges for Cloud Networking Security,” 2nd International ICST Conference on Mobile Networks and Management, pp.298-313, 2010.

[4] ITU-T, ”Focus Group on Cloud Computing, TR part 1–3,” 2012.

[5] ”SAIL Project,” http://www.sail-project.eu/.

[6] ”FoSII - Foundations of Self-Governing ICT Infrastructures,” http://www.infosys.tuwien.ac.at/linksites/FOSII/index.html.

[7] ”Contrail-project,” http://contrail-project.eu/.

[8] ”CPIP IEEE,” http://standards.ieee.org/develop/project/2301.html

[9] ”SIIF IEEE,” http://standards.ieee.org/develop/project/2302.html

[10] ”Open Cloud Computing Interface (OCCI),” http://occi-wg.org/.

[11] ”GICTF,” http://www.gictf.jp/index_e.html.

[12] T. Benson, A. Akella, A. Shaikh and S. Sahu, ”CloudNaaS: a cloud networking platform for enterprise applications,” In Proc of 2nd ACM Symposium on Cloud Computing, 2011.

[13] I. Goiri, F. Julia, J. Fito, M. Macas and J. Guitart, ”Resource-Level QoS Metric for CPU-Based Guarantees in Cloud Providers,” GECON 2010, LNCS 6296, pp. 34-47, 2010.

[14] G. Cicotti, L. Coppolino, R. Cristaldi, S. Antonio and L. Romano, ”QoS Monitoring in a Cloud Services Environment: the SRT-15 Approach,” Epsilon srl, Italy, pp. 15-24, 2011.

[15] Y. Kouki and T. Ledoux, ”CSLA: a Language for improving Cloud SLA Management,” In International Conference on Cloud Computing and Services Science, Portugal, April 2012.

[16] P. Patel, A. Ranabahu and A. Sheth, ”Service Level Agreement in Cloud Computing,” OOPSLA, 2009.

[17] H. Ludwig, A. Keller, A. Dan, R. King and R. Franck, ”Web service level agreement (WSLA),” IBM Corporation, 2003.

[18] I. Brandic and E. Feller, ”Energy and QoS-aware workload management in clouds,” Research report of COST group, 2011.