Towards the Future Internet
G. Tselentis et al. (Eds.)
IOS Press, 2009

doi:10.3233/978-1-60750-007-0-193

Multi-level SLAs for Harmonized Management in the Future Internet*

Wolfgang Theilmann[1] and Luciano Baresi[2] **

[1] SAP Research – CEC Karlsruhe
Vincenz Priessnitz Strasse 1, 76131 Karlsruhe, Germany
wolfgang.theilmann@sap.com
[2] Dipartimento di Elettronica e Informazione Politecnico di Milano
P.zza L. Da Vinci, 32 - 20133 Milano, Italy
baresi@elet.polimi.it

Abstract. The Future Internet is about to fundamentally change social and economic interactions at a global scale. The integrated access to people, media, services, and things will enable new styles of interaction at unprecedented scale, flexibility, and quality. However, this also calls for a well-defined and sound approach for management and governance that allows for clear harmonization and translation of issues across domains and layers. This paper presents a proposal that aims to blend management and governance issues at business, software, infrastructure, and network level, and introduces a multi-level SLA management approach to bridge these issues across different layers. It also sketches some insights on management and governance practise and requirements in various industrial domains.

1 Introduction

The integrated access to people, media, services, and things, provided by the *Future Internet*, will enable new styles of societal and economic interactions at unprecedented scale, flexibility, and quality. The Future Internet, through the metaphor of *Internet of Things*, will provide location independent, interoperable, scalable, secure, and efficient access to a coordinated set of services [3], but such broad vision demands for a sound and well-defined approach for management and governance. This approach has to harmonize and bridge the various views

* The research leading to these results is partially supported by the European Community's Seventh Framework Programme ([FP7/2001-2013]) under grant agreement n. 216556.

** On behalf of the SLA@SOI consortium [7] which includes T. Ellahi, H. Li, W. Theilman (SAP), F. Torelli (Engineering), J. Kennedy (Intel), M. Alvarez, A. Castro, J. Lambea, S. Aleman (Telefonica Investigacin y Desarrollo), C. Kotsokalis (University of Dortmund), M. Trifu, C. Momm (Research Centre Karlsruhe), A. Marconi, M. Pistore (Fondazione Bruno Kessler), L. Baresi, E. Di Nitto, S. Guinea (Politechnico Milano), G. Spanoudakis (CITY University), R. Perrot, T. Harmer (Queens University Belfast), G. Pipan (XLAB), G. Armellin (GPI) and M. Evenson (eTel).

and layers of the Future Internet following the subsidiary principle: as many issues as possible should be dealt with locally, while as few issues as possible are to be managed in a more integrated way.

This multi-layered view requires that supplied services[3] be managed coherently. Service Level Agreements (SLAs) and policies are becoming common means for doing this. SLAs specify the conditions under which services are provisioned, but current management frameworks typically only focus on single service interfaces. They neither use SLAs for managing the implementation and delivery of services, nor they recognize/support the fact that many services may be composed of lower-level services, involve third-party providers, and rely on a possibly complex business/IT stack [2]. While SLAs are routed in the respective customer requirements, policies are provider-specific means to express constraints and rules for their internal operations. These rules may be independent of any particular customer.

The paper presents the proposal of the EC project SLA@SOI to integrate and blend management and governance at the different levels. The core idea is to use SLAs as primary means to express management and governance concerns and to share and translate these concerns across different views, perspectives, and layers of the Future Internet.

The main innovation of the framework will be the integration of the following features: (1) standardized models for SLA descriptions at the different layers, (2) an automated e-contracting framework, (3) systematic grounding of SLAs from the business level down to the infrastructure, (4) methods and tools for multi-layer SLA management, including planning, optimization, and provisioning, (5) methods and tools for monitoring and accounting services and SLAs through standardized interfaces, (6) exploitation of virtualization technologies at infrastructure level for SLA enforcement.

The remainder of this paper is organized as follows. Sections 2-5 introduce the most important management and governance issues at business, software, infrastructure, and network level, respectively. Section 6 provides details on how the multi-level SLA management can contribute to a harmonized and holistic approach across these layers. Section 7 provides insights on the practise and requirements of different industrial domains, and Section 8 concludes the paper.

2 Business Level

Nowadays, service providers (like Google, Amazon, and Facebook) must mediate between the business challenges enabled by network and IT convergence and users demanding for more and more new value-added services [8].

The ideal solution would be the emergence of a new service *marketplace*, formed by the convergence of Internet, media, and telecommunication industries. This new marketplace should support the exposition of multiple heterogeneous services, from different providers and industries, control and guarantee their quality, and maximize customers and revenues. New services could be

[3] In this paper, we consider services as means to deliver value to customers.

added by composing those supplied by different providers, and the marketplace should also help automatically establish and manage business agreements and relationships among the participants, and expand marketing possibilities beyond advertisement-based business models (e.g., subscriptions or pay-per-use models). The end-to-end manageability of the service lifecycle —from creation to monetization— becomes a key issue, and must work seamlessly also for those services that span different providers.

The use of SLAs is indispensable for this new marketplace to emerge. An SLA may specify for instance the levels of availability, serviceability, performance, operation, or other attributes (e.g., billing) and even the penalties when the agreement is violated [10]. At this level, it is worth distinguishing between: *supplier/partner SLAs*, which include the conditions under which a third-party service can be contracted and reused by other service providers, and *customer SLAs*, which include the conditions under which the service can be contracted by a customer.

Service conditions are formalized using SLA templates (e.g., based on WS-Agreement [4]) and the production of these templates must become an intrinsic part of service development. SLA templates contain default values specified during the service development phase. Once a service provider or customer enters into negotiations to contract a service, the SLA template is used as a blueprint for the SLA [9]. If a service depends on other services, supplied by different providers, customer SLA templates must take into account these dependencies and ensure the consistency of SLA parameters and values.

3 Software Level

Management and governance at this layer primarily consist of two aspects: a centralized or federated configuration management database (CMDB), populated with the live state of the service and software landscapes, and a set of processes carrying out the activities of the service lifecycle. These processes are executed in accordance with the information available in the CMDB. Additionally, these processes produce further information being pushed to the CMDB for satisfying the requirement of predictive autonomic management.

The CMDB holds information, SLAs, rules, policies, and various models to capture diverse aspects of the software and service landscapes. More in detail, it comprises:

- The **software landscape**, which contains software component models based on standards like SCA (Service Component Architecture). These artifacts are extended with non-functional properties of the service/software components.
- The **software configuration models**, which contain information about the deployment-time configurations of the various software components.
- The **service landscape**, which offers a comprehensive information model containing detailed descriptions of the various elements: services, SLAs associated with them, their software components, and the required execution

infrastructure (e.g., middleware, application servers, and databases). This landscape also considers the relationships between these elements and those of the business and infrastructure layers.
- The **operational rules** and **policies**, which are consulted by the processes during the execution and operation phase of the service lifecycle. These rules and policies ensure that the service landscape be in a consistent state and adheres to set business policies.

The processes are carried out in various phases of the lifecycle in accordance with the information fetched from the CMDB. These processes can operate only at this level or work together with the other levels to ensure the fulfillment of multi-layer SLAs. These processes include:

- **Deployment/Provisioning** activities to identify the required set of software components essential for delivering the services. Moreover, some capacity planning and sizing procedures help identify the required setting of logical resources (lower levels) to ensure the service levels as constrained by the SLAs. These activities rely on the SoA models defined in the software landscape.
- **Configuration** activities to focus on the identification and setting of "knobs and switches" used during the later phases of the lifecycle for SLA management. These activities are performed by utilizing the configuration models from the software landscape and information derived from higher-level SLAs.
- **Monitoring** activities to help configure knobs and switches. These are the most critical processes for management and governance. Monitored information is stored onto the CMDB for historical purposes and for the sake of autonomic predictive management.
- **Adaptation/Change Management** activities to analyze the information gathered from monitoring and take appropriate actions if the current configuration could lead to possible SLA violations. Effective governance requires proper change management procedures; the service landscape models can facilitate change impact analysis to understand the implications before performing the actual changes.

4 Infrastructure Level

This level allows the provider to advertise and manage the infrastructure, reserve and provision it as per negotiated SLAs, and stage the software images. Management at this level centers on the infrastructure landscape: what virtual machines (VMs) are provisioned where, and what commitments and reservations have been made into the future. Governance is concerned with higher-level business rules defined by the infrastructure provider. These policies can span areas such as security, energy-consumption, data privacy and isolation, business continuity, and geographical and legal constraints. Both management and governance are being enabled by a policy-driven architecture to allow the overall system to

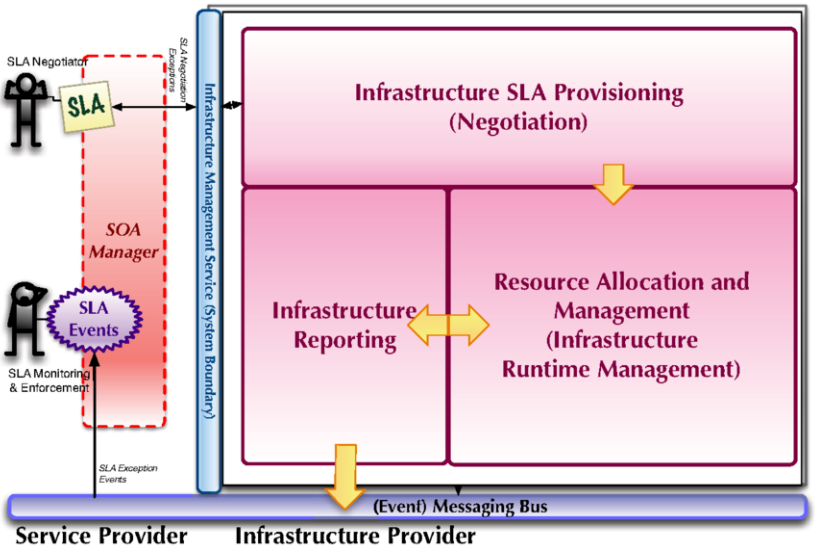

Fig. 1. Infrastructure Management Architecture.

be kept under control despite the potentially diverse needs of the individual SLAs.

Figure 1 provides a high-level view of this architecture. **Infrastructure SLA provisioning** includes the functionality that is required prior to actual resource allocation. It translates requests from potentially high-level infrastructure requirements originating from the software-level to technical requirements relevant to the infrastructure level. It also maps these to appropriate bundles of resources, and verifies that the infrastructure is available to accommodate the actual requests. Policies at this stage could mandate provisioning VMs on physical machines dedicated only to that customer, or mandate particular failover resources being reserved in parallel with the main request.

Resource allocation and management performs the actual provisioning. Hardware, either physical or virtual, is provisioned by either internal or external resource providers. This functional component includes an *autonomic management* capability to identify if and when re-provisioning should occur to properly accommodate defined policies, or SLA conditions that are in danger of being violated. Once the infrastructure is provisioned, appropriate event logging, correlation, and monitoring is instantiated in the **Infrastructure reporting** component to allow all relevant conditions or events to be identified and escalated, either internally for immediate disposition, or externally in case of, for example, the unavoidable violation of an SLA. Again, policies are used extensively to define what information needs to be escalated internally and to where, notwithstanding the additional external commitments agreed in the SLAs.

5 Network Level

Nowadays, the underlying network must allow users to stay connected permanently from different environments (e.g., when at home, or at work) with a wide range of devices. This scenario must be paired with the growth of the telecommunication sector, with the evolution of both mobile communications (e.g., UMTS and HSDPA) and fixed broadband access (e.g., ADSL and FTTH). Future network infrastructures will support the convergence and interoperability of both network technologies. The ultimate goal is to optimize the transmission of voice, data, and media to and among users, no matter of their locations or devices. In the more immediate future, this convergence means that a single service can be executed through and switched between different wired and wireless networks.

Historically, the services provided to users have been tightly coupled with specific technologies and networks, but this solution is not suitable to cover the needs of the above scenario. Services and networks must be decoupled, but we must prepare the network to cover this need. Network and service providers are studying how to evolve their infrastructures to new architectures like NGN (New Generation Network). The new concept of network management must move from vertically integrated services (silos-oriented service management) to horizontally integrated services, with many common capabilities available at the same time.

Once networks and services are decoupled, we can start organizing a service-based business environment. This new multi-domain/multi-provider environment is challenging to network management systems, which now should provide a complete supervision of the services through different domains and different networks. In addition, it is mandatory not to forget customers and their perception of offered services. In this scenario SLAs play a key role. The agreements between customers and service providers must specify the conditions under which the service is provided and consumed. These parameters must be built up from the parameters of the individual networks or domains. The agreement becomes the basis for monitoring the network through the various layers.

6 Multi-level SLAs for Harmonized Management

Currently the integration of the different management activities described so far is mostly based on human activities. For instance, experts are usually able to recognize if a degradation in performance is due to a wrong software configuration or to insufficient hardware resources. In the former case, they can solve the problem by using their expertise, while in the latter case, they raise the issue to the infrastructure manager to solve it. If the problem is not solved directly by the infrastructure manager, the issue is raised to the business level to decide, for example, the activities and systems with higher priority when assigning resources.

As the complexity of applications grows, this human-based process becomes more and more challenging. Difficulties increase when applications cross the boundaries of a single organization, and when independent human-based and

automated services are integrated over the Internet [1]. Approaches and tools for automating the integration of management activities become mandatory.

The proposal presented in this paper is based on a conceptual and a software framework for the harmonized management of all application levels based on the unifying concept of Service Level Agreement (SLA). Several SLA management frameworks have been presented in the past, but they typically focus on the specification of processes and models, giving no support to automation [6]. Moreover, they focus on the level of single service interfaces and do not recognize/support the fact that many applications are composed of a hierarchy of services at different levels [2]. Composition of services is one of the basic assumptions of the SoA approach [5], nevertheless current relevant methodologies only pay attention to the composition of software services without considering neither their business-level representations nor the infrastructure on which these services are provisioned.

Our proposal applies the concept of SLA in a uniform way, not only to software level services, but to all levels of the application stack. Usually SLAs are the result of a negotiation between a customer and the service provider. The proposal regards each application level as a service provider for the higher levels and as service consumer for the lower ones. Each provider tries to fulfill its policy rules to satisfy internal requirements, while negotiating with the potential consumers to satisfy external requirements.

SLAs and negotiation work for both human and automated services, for internal and external use, and at all levels to improve the management of IT applications. The possibility that applications be distributed among specialized providers —in a more reliable and collaborative way— will encourage the use of the SoA model at global level, allowing for a more efficient use of all kinds of available resources.

Figure 2 gives a simplified overview of the SLA management process. As today's business systems typically consist of complex layered systems, user-level SLAs cannot be directly mapped onto the physical infrastructure. Services might be composed of other more fundamental services that could be even provided by external parties. Consequently, a stepwise mapping of higher-level SLA requirements onto the lower levels, and the aggregation of lower-level capabilities into higher-level aggregations, is crucial for grounding user-level SLAs onto the infrastructure. This vertical information flow must carefully reflect service interdependencies as well as the originating business context. In addition to SLAs, the vertical information flow also covers monitoring, tracking, and accounting data and must support brokering and negotiation processes at each layer. As shown in the figure, the overall SLA management process may include different stakeholders, namely customers, service, and infrastructure providers, and also various business steps such as business assessment, contracting, and sale. The overview is intentionally simplified in the sense that no service chains are visualized. Such chains would represent the cases where service providers rely on additional external providers.

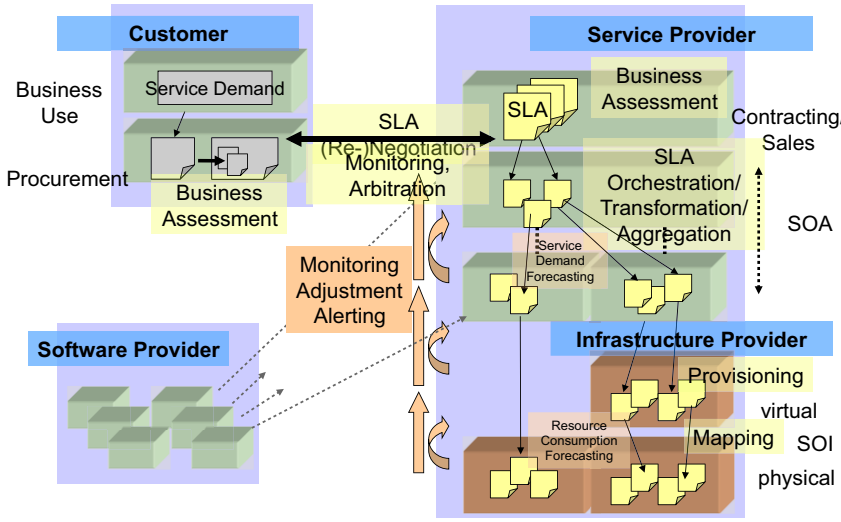

Fig. 2. Envisaged interactions of SLA stakeholders.

A service provider offers services with differentiated, dependable, and adjustable SLAs and can negotiate concrete SLAs with (individual or groups of) customers in an automated fashion. This business goal imposes requirements on software providers (to provide components with predictable non-functional behaviour), on infrastructure providers (to support SLA-aware management of resources), and also on the service provider (to translate and manage SLAs from business level along the IT stack down to the network). Of course, complete business value chains can be easily composed on top of this setup.

7 Industrial Practise and Requirements

This section briefly presents the industrial practice in the financial, e-Government, and ERP domains, which are used as case studies in the project.

A financial service provider must comply with management restrictions. The general financial service regulator typically stipulates conditions in which sharing of data may take place on an individual server or server site. For example, trading and customer accounts data cannot be held or processed on the same server. The regulator may specify that certain portfolio data may only be held and/or processed within county boundaries, geographic regions or ranges, or in fact any combination of these. For example, US government pension data may only be held within the US and within 50 miles of New York.

Individual financial service customers may have their own management restrictions as far as precise security infrastructures are concerned depending on the nature of their business and the service they are availing of. For example, a back test prediction service may not require security encryption on data trans-

mission of historical data, but may require full authentication/authorization access controls to profit prediction services and security encryption of transmitted results. Financial customers may also require that their data are held and processed on an exclusive server or even site to ensure that data sharing does not happen with other customers.

The design and management of e-Government services for citizens, such as the provision of medical and social assistance to elderly or disabled people, is a complex process that involves several actors with different roles and expectations. The citizen, who is the service beneficiary, expects efficient health and social services for addressing his/her own specific need. The governance, which is the service customer, has the necessity of monitoring and analysing the costs and quality of the health and social system as a whole. The different organizations that contribute as services providers (e.g., hospitals, nursing homes, private companies and charities) have their specific business goals to fulfill and need to negotiate with the governance the quantity and quality levels of provided services (e.g., number of medical treatments offered per day by an hospital).

The fundamental role of SLAs in e-Government service provisioning has been already recognized in the context of the so called G2G (government-to-government) services. Their adoption is, in contrast, still very limited in scenarios such as the medical and social care, which also require G2B (government-to-business) and G2C (government-to-citizen) service provisioning. The particular challenge in this context is that SLAs are not only based on market rules, but they are most often driven by "social" agreements between public bodies and citizens. As a consequence, the SLA negotiation (both between public bodies and citizens and between public bodies and private service providers) is different than in market-oriented domains. Another challenge of this domain is that it requires an integration of human-based services (e.g., home care, medical assistance at home, and transport services) with IT services; the underlying service oriented infrastructure is hence not only a technological infrastructure, but also a social and organizational one.

ERP services typically constitute the core of the IT architecture. They tend to be long-lasting and typically require lengthy implementation phases. However, also ERP services are more-and-more offered in an on-demand fashion, both to allow customers to have a dynamic consumption with low capital expenses but also to support the flexible construction of more complex business value chains.

Management and governance in these environments lead to significant challenges in terms of interoperability of solutions, but also in a proper negotiation and resolution of possibly conflicting policies. Furthermore, policies often do not just apply to a single service but to a complete system of services that somehow constitute a separate business context. Current practise in this area is still rather primitive. SLAs are partially formally expressed at the business level, however never translated or correlated with other SLAs on service/system layers. Policy support is mainly realized for local perspectives (e.g., for the application administrator of a set of software services) with poor interlinkage to other roles, views or organizations.

8 Conclusions and Future Work

This paper discusses the proposal of the EC project SLA@SOI [7] to integrate and blend management and governance at the different levels of the IT stack. The core idea of this paper is to use SLAs as primary means for expressing management and governance concerns and to share and translate these concerns across the different views, perspectives and layers of the Future Internet. The main goal of the project is to provide an SLA management framework that allows for consistent specification and management of SLAs in a multi level environment. The framework is designed for integration into different service-oriented infrastructures and will be evaluated within various complementary industrial case studies.

References

1. Chesbrough H. and Spohrer J.. A Research Manifesto for Services Science. Communications of the ACM, 2006. 49(7): p. 35-40.
2. CoreGRID. Using SLA for Resource Management and Scheduling - A Survey, TR 0096. August 2007, www.coregrid.net/mambo/images/stories/TechnicalReports/tr-0096.pdf.
3. Future Internet Assembly. Bled Declaration on Future Internet. April 2008, www.future-internet.eu/index.php?id=47.
4. Open Grid Forum. Web Services Agreement Specification, March 2007, www.ogf.org/documents/GFD.107.pdf.
5. Papazoglou M.P. and Van Den Heuvel W.J. Web Services Management: A Survey. IEEE Internet Computing 9 (2005) 58-64.
6. Schaaf T. Frameworks for Business-driven Service Level Management: A Criteria-based Comparison of ITIL and NGOSS. Proceedings of BDIM 2007, 65-74, Munich, Germany.
7. SLA@SOI consortium. SLA@SOI project (IST- 216556; Empowering the Service Economy with SLA-aware Infrastructures). www.sla-at-soi.eu
8. tmforum. Service Delivery Framework Overview (TR139 Version 2.0, July 2008)
9. tmforum. SLA Handbook Solution Suite v2.5 www.tmforum.org/page30754.aspx
10. Wikipedia en.wikipedia.org/wiki/Service_level_agreement