A Vision for Reliable Network Services Architecture

Jean Benjamin Belinga Hanane Oumina Daniel Ranc

Institut National des Télécommunications
Département Logiciels et Réseaux
9 rue Charles Fourier, 91011 Evry, France
{jean-benjamin.belinga,daniel.ranc}@int-evry.fr
hanane@huawei.com

Abstract. The increasing complexity, heterogeneity and dynamism of networks, systems and services have made our informational infrastructure unmanageable and insecure. The last events in the European telecom operator and IT landscapes showed inherent limits of current network and systems management architectures with respect to availability, resiliency and QoS. This paper presents a new architectural vision, merging well working technologies to tackle the problem complexity of network services management. Particular attention is drawn on the multiple stakeholders environment, where realistically many parties collaborate to fulfill end-to-end service to end users.

1 Introduction

Networks and systems have become so large, complex, and fast, and have assumed so many important tasks that when things go wrong, overburdened IT staffs often can't implement fixes fast enough to avoid mission-critical problems. The complexity of systems and networks keeps growing even beyond human capabilities to handle management tasks for achieving the best benefits from such systems. The last major events, which have occurred in telecommunication world, show an example of some problems the operators face today. These difficulties may become even worse when considering future complex multiple mobile stakeholder environments associating access, content providing, multi-domain transport and metropolitan services to the usual network management context. The aim of this paper is to propose an architecture that attempts to simplify the complexity of this problem area.

2 Problem Definition

Analyzing causes of major network breakdowns in 2004 that have intervened within the telecom operator and IT environment in Europe in term of QoS, availability and denial of service, the first remark to be made is that all impacted systems show some degree of monopolistic architecture either by their centralized structure or by their IT composition.

The development of IT infrastructures have aggravated network, system and services management complexity due to:

- The emergence of new networking technologies, such as adhoc networking, which are combined with established technologies and that may interface to each other
- The rapidly increasing size of individual networks and the Internet as a whole
- The accelerated development of new technologies, which forces companies to restructure their IT systems more frequently
- The growing pressure on operators of time to market constraints that contradicts a careful and integral validation and testing of systems
- The lack of experience on the multiple stakeholder environment: each operator remains single-oriented; and the fact, that services run under different organizations and administration policies
- The heterogeneity and the independence of resources and components required by these services.

This indicates that IT infrastructures in large companies will grow even more complex in the future. How to adequately administer network services and systems with the inherent complexity, and insure end-to-end services with availability and required quality?

This question calls to a reflexion on the manner of managing the future networks and systems. In this paper, we particularly focus on the multi-operator, multi-stakeholder context: the place where complexity interacts with complexity, with highest constraints on confidentiality and availability.

This research aims at proposing a global architecture to insure the end-to-end service delivery in term of QoS, availability and resiliency; at achieving this aspired solution by building grid-like network with some autonomic features. For this purpose, the research is aimed at merging the following technologies:

- TMF-NGOSS that provides the reference architecture of our system (the eTOM Framework and the Service Framework);
- The features of Grid and Peer-to-Peer systems for achieving interoperability, redundancy and reliability;
- Some primitives of autonomic computing for achieving self-management;

3 Background

3.1 TMF-NGOSS

The TMF-NGOSS (New Generation OSS) [6][7] provides a comprehensive, integrated framework for developing, procuring and deploying operational and business support systems and software. NGOSS principles and tools apply to service providers, software suppliers and system integrators, providing de facto standards for business process mapping and automation.

eTOM. enhanced Telecom Operation Map (eTOM) - is the business view of the NGOSS [5][6]. The particularity of this model is that eTOM considers external actors especially the costumer in the model. It differentiates the long term and the short term business processes in the life cycle of products and services. Two operators communicate with each other in a Customer/Provider way.

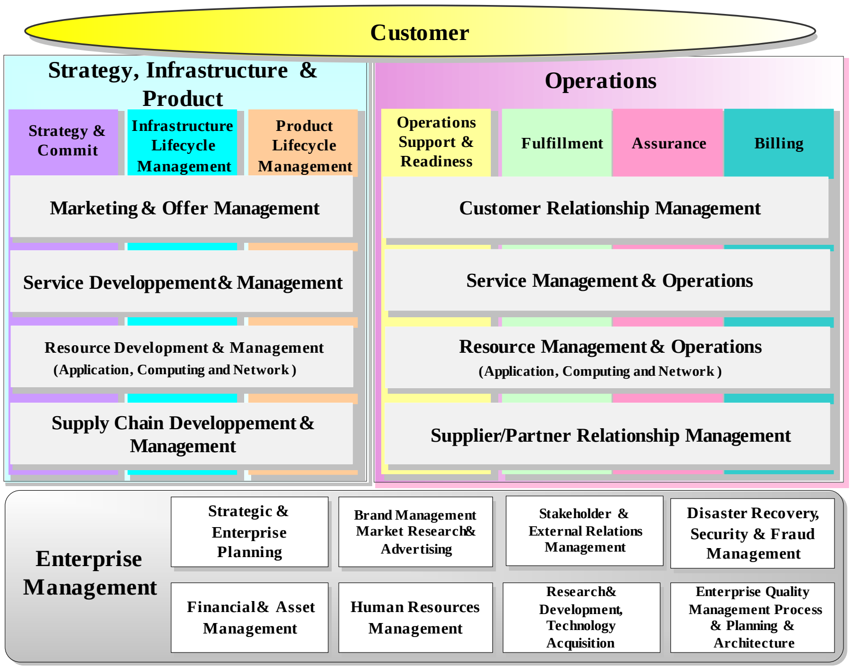

Fig. 1. eTOM Model

3.2 Service architecture

Multi-stakeholder environment. Many emerging real-time and low-jitter oriented services e.g. Video on Demand (VoD) involve a set of distributed stakeholders contributing to the end-to-end service, such as end-user, resource provider, content provider, network operator, information broker, service portal, information carrier. This plurality generates new requirements to the collaborative information exchanges between the telecom stakeholders in order to insure adequate service delivery, to guarantee the contract agreements and also to satisfy the end-user requirements. And what arises from the study of existent monopolistic systems is the lack of experience on the multiple stakeholder environment: each operator remains single-oriented. This "single-minded" point of view is not realistic anymore. Our belief is that in fact many parties are involved in an end-to-end service delivery.

Definition of service. A service is recursively defined as composed of sub-services [8], each sub-service is provided by a single service provider. The key word here is the collaboration of multiple stakeholders to provide a final service to the customer that has the freedom to formulate the properties of its service. In other words, there is no fixed pre-defined portfolio of services offered by the network. But the network has to be able to correlate the sub-services published by the broker to constitute and offer the service required by the customer. A customer may compose its service as he likes. In this context, a final service can have more than one composition according to the sub-services involved to constitute it. A subgoal is to constitute the service, minimizing subservices, with respect to the quality of service that the customer waits for.

3.3 Grid and Peer-to-Peer Systems

Grid systems. Although GRIDs [19] are still not publicly widespread as of today, their architecture seems to represent a promising advance with respect to distribution and scalability in terms of calculation power and storage. Grid technologies support the sharing and coordinated use of diverse resources in dynamic, distributed "virtual organizations" (VOs) – that is, the creation, from geographically and organizationally distributed components, of virtual computing systems that are sufficiently integrated to deliver the desired QoS [17].

Peer-to-Peer systems. Widespread peer to peer services insure flexible and efficient delivery of video and audio documents. A careful inspection of their architecture shows a particular solid resilience against a) disruption of servers (by automatic subcontracting to other servers) and b) disruption of another client (the peer is then exchanged with a working one).

Features of autonomic computing. Autonomic computing is the technology that aims at enabling computing systems to be self-managed with minimal external intervention [15]. IBM defined four main characteristics of autonomic computing systems [16]:

- Self-configuration: With the ability to dynamically configure itself on the fly, a system can adapt immediately with minimal external intervention to the deployment of new components or changes in the IT environment. The system configuration should be done automatically, as well as dynamic adjustments to that configuration to handle changing environments.
- Self-protection: Self-protecting system lets the right people access the right data at the right time and can take appropriate actions automatically to make itself less vulnerable to attacks on its runtime infrastructure and business data. Then system should be capable of detecting and protecting resources from both internal and external attacks, thus maintaining overall system integrity.
- Self-healing: Self-healing systems can detect problematic operations (either proactively through predictions or otherwise) and then initiate corrective action

without disrupting system applications. Then system should be able to recover from faults that might cause some parts of it to malfunction.

– Self-optimization: Self-optimization refers to the ability of the IT environment to efficiently maximize resource allocation and utilization to meet end users' needs with minimal human intervention. Then system should monitor its components and look for ways to optimize its working, like resource allocations, load balancing, and different network traffic optimizations.

3.4 Architectural remarks

From the definition of service, as mentioned earlier, we argue that the deployment issue has to be considered as a critical part of a management framework. This brings two problems for management in multiple stakeholder environment. Firstly, technological differences between domains make it difficult to find a common management mechanism to be used by all operators. Operators utilize different tools and deploy different policies to manage their local networks. Secondly, even if commonly agreed mechanisms and policies exist, the commercial interests (or financial constraints) of operators may prevent their wide deployment.

4 Approach overview

4.1 Computational point of view

The network becomes by itself a true system, and the evolutionary advent of new applications of real time internet gives place to open and dynamic systems based on cooperative distributed networks and other stakeholders like portals, information brokers, content providers... This shift resulted notably in the success of OSS, Business Process based approaches such as the eTOM.

4.2 Framework architecture

Considering the overall network as a federation of multiple collaborating stakeholder networks, each of them insuring a part of end-to-end service, we describe here our vision of the management of such a complex network.

We introduce a new concept based on Grid Computing technology: the Grid Network Management [13][14].

Grid Network Management is a way to simplify the problem of the complexity of the centralized network management and services in a multi-stakeholder environment, while placing a higher notch in architecture, and by applying the principles of virtualization to a system of services and networks management. The objective is to design and to implement an open architecture for the systems and services management in a multiple stakeholder environment, by developing principles generalized for the grids being able to be applied to specific fields. This architecture is directed towards delivering services, and insuring efficient creation of new services within time-to-market constraints.

The architecture reference of our management system is the eTOM architecture. The main idea is to develop a multi-eTOM-based platform making it possible to manage the heterogeneity of network services, to control the problem of interconnection between various stakeholders of the environment, to provide mechanisms making it possible to draw up SLAs and QoS contracts "on-the-fly" ensuring the interconnection between stakeholders and end-to-end service quality management strategy, to design and create functions and applications capable to circumvent various problems resulting from possible flaws in a given stakeholder. In this context, a summarizing term of self-management network may be adequate.

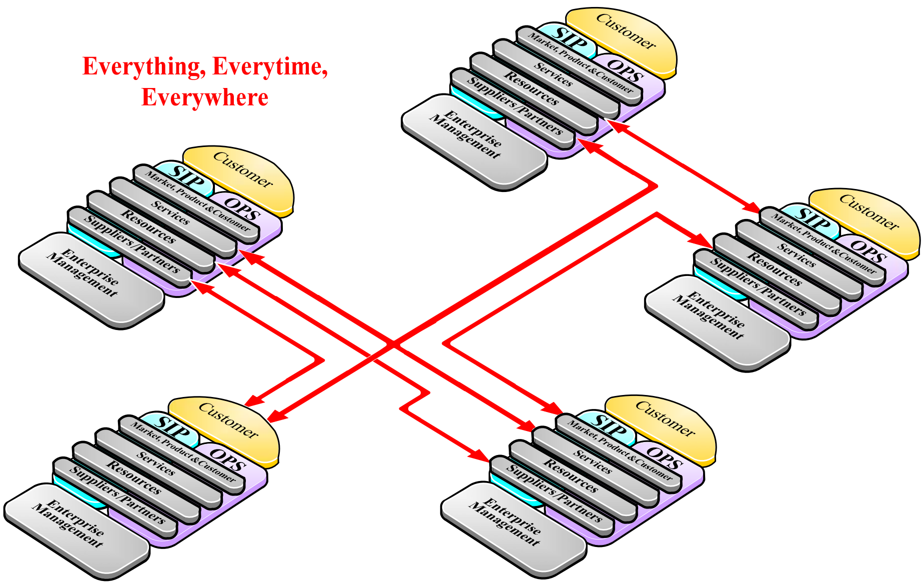

Fig. 2. Partial example of multi-eTOM environment

4.3 Information point of view

At the highest level of the information model is the SLA, which derives information from specifications and cost agreement. It has to be noted that classical SLAs are human-readable but are not powerful enough to provide an end-to-end QoS set-up. There is an emphasis on the SLA concept and its implementation in terms of Service Level Specifications (SLSs), which contain all the technical parameters needed to set-up the corresponding QoS inside a global network like multi-eTOM environment. The proposed approach attempts to cover this emphasis enabling to have instantly provisioned services based on dynamically negotiated SLAs, like in [9][10].

4.4 "Virtualized" service framework

Service management is concerned with end-to-end service delivery quality and management. The focus is on the monitoring and management of SLAs between the customer and internal and external suppliers involved in the service delivery value chain; this make it difficult to adopt an end-to-end strategy for Service Quality Management.We aim to improve the service management by applying the principles of virtualization to a system of services and networks management.

By virtualization [18], we understand the abstraction of our reference architecture – the eTOM framework. It is a set of transformation processes in the virtualization layer during which associations between virtualized entities and "underlying" ones are established and changed.

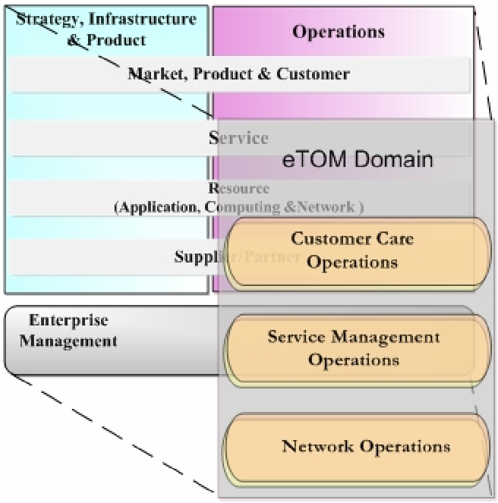

Fig. 3. Virtualization at eTOM layers

The picture above depicts the most noticeable operation layers: Customer Care Operations, Services Operations and Network Operations.

The use of this "virtualized" service framework allows service specification information to be shared between multiple service providers without the need to declare proprietary information. The relevance of a common service framework therefore is fundamental; it not only impacts the practical physical monitoring of performance metrics but also affects the interactions and contractual relationships between all parties involved in the service delivery chain. The following figure considers a scenario of a VoD delivery and illustrates the complexity of multi-party interactions between identified service providers:

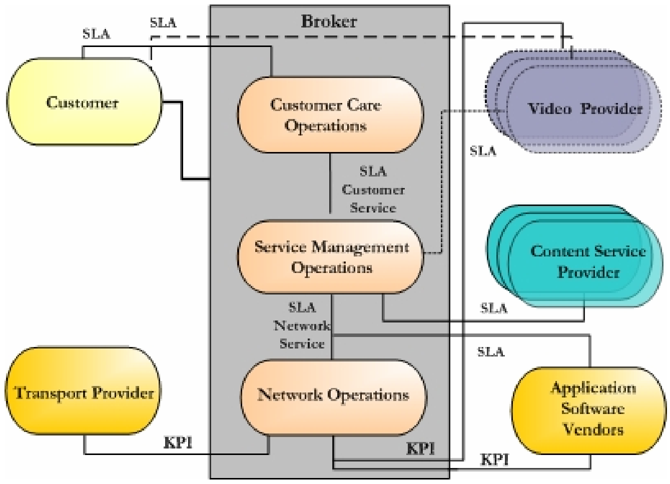

Fig. 4. A scenario of VoD delivery

4.4 Implementations considerations

Our platform is a global distributed architecture which represents a peer-to-peer relationship between two or more partners involved to achieve a common goal. It is characterized by a community that includes all partners, a contract defining the terms of the relationship, and a policy to rule the relationship's lifecycle. As we said above, SLA is one of the fundamental concepts to be taken into account in the architecture in order to offer interoperability.

The agent technology - the Manlets. To fulfill the requirements of SLAs , for the first implementations of the project, we will make use of the agent technology called Manlet (Management Applet), detailed in the project ImaGEnS [11] [12]. Manlets are components of the relevant information that circulates on the (global) network.
Manlets feature following main properties:

– Mobility: Manlet is the entity making a Business Process live,
– Polymorphism: their shape, behaviour and properties change according to their location within their lifecycle,
– Autonomy: Manlets follow their path in constant independence,
– Intelligence: Manlets have embedded intelligence that enables adaptative interaction with the environment, lifecycle management, re-routing, re-birth, self-destruction etc.

Manlets all inherit from a generic object implementing all the common and basic features listed above. As shown in the figure above, it can be cloned, can dispatch itself, display itself, can be disposed or can dispose itself, and finally can be archived or archive itself.

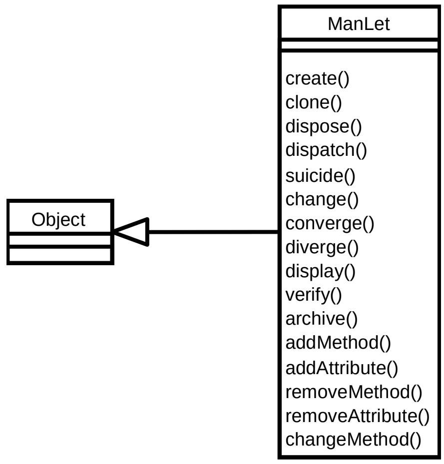

Fig. 5. Manlet generic model - partial UML diagram

Manlets (including all the constraints and requirements of the contract) will index all the partners entering the chain of provisioning the service, and will dynamically carry out the negotiation, the establishment of the contracts with all the requirements of quality of service – thus forming a field of composition of service. It is clear that the dynamicity of ManLet technology makes Service Management very different from traditional Enterprise Management. By the virtualization of domains, the real management in our environment can be formed on a per transaction basis, thus implying dynamics several orders of magnitude larger than found in today's network management systems, whose topology tends to be fairly static. In fact, in our environment the main management resources convert to services, SLAs and QoS requirements across different networking systems. Consequently, the goal of end-to-end management, which has been the target of several management efforts, must become achievable in a multi-eTOM environment.

In our exemple of VoD delivery, we use Manlet as one of the way to chain domains into end-to-end connection. Each SLA interaction in figure 4 depicts a ManLet activity at a particular status of its lifecycle. The chain of all steps and ManLet lifecycle states is to be compared to the notion of TMF Business Process, in the context of multistakeholder environment. It does allow the process to be more efficient in terms of automation, security, availability.

5 Conclusion and related work

The question now is how much various stakeholders can correlate their offers and services to satisfy customer needs and what is the cost of management and optimization of resources in such environment? At last, what is the future of stakeholder's competition and market targets? Can we think on the standardization of service composition that can regulate the offer and the demand of services or are we about to tend to create an on-fly automated on-demand telecom environment?

In this paper, we presented a novel vision to implement a management infrastructure based on a multiple stakeholder approach with availability as a constant focus, and combination of principles with can be applied to network management such as Grid Computer, and peer-to-peer systems. The first implementations using ManLet technology shown in this paper is only one of many ways to chain different eTOM-based domains into an end-to-end connection. We hope that the study present in this paper will lead to excellent results.We presented the arguments for our first implementation – the interaction beetween actors towards a specific agent technology: Manlet.

We are currently implementing the way the actors are connected through Manlet/SLA, and how this is related to the eTOM framework.

References

[1]. Baldonado, M., Chang, C.-C.K., Gravano, L., Paepcke, A.: The Stanford Digital Library Metadata Architecture. Int. J. Digit. Libr. 1 (1997) 108-121

[2]. Bruce, K.B., Cardelli, L., Pierce, B.C.: Comparing Object Encodings. In: Abadi, M., Ito, T. (eds.): Theoretical Aspects of Computer Software. Lecture Notes in Computer Science, Vol. 1281. Springer-Verlag, Berlin Heidelberg New York (1997) 415-438

[3]. Van Leeuwen, J. (ed.): Computer Science Today. Recent Trends and Developments. Lecture Notes in Computer Science, Vol. 1000. Springer-Verlag, Berlin Heidelberg New York (1995)

[4]. Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs. 3rd edn. Springer-Verlag, Berlin Heidelberg New York (1996)

[5]. enhanced Telecom Operations Map - The Business Process Framework, TeleManagement Forum 2004, GB921 v4.0, March 2004.

[6]. TeleManagement forum www.TMForum.org

[7]. New Generation Operational Support System (NGOSS), NGOSS Release 1.5. TMF GB920. TeleManagement Forum, November 2000.

[8]. Service Model Framework, Release 4.5. TMF GB924. TeleManagement Forum, November 2004.

[9]. SLA Management Handbook, TeleManagement Forum 2004, GB 917-1 v2.0, Member Evaluation version, July 2004.

[10]. Dugeon, O., Diaconescu, A. From SLA to SLS up to QoS Control: the CADENUS Frame-work, WTC'2002

[11]. Ranc, D.,, Vaculova, L., Elmejjad, S., Maknavicius L., Touré F., Daurensan V., Une organisation orienté cycle de vie de l'information pour la gestion des réseaux et services. GRES, Février 2003, Fortaleza-CE-Brésil.

[12]. Ranc D., Vaculova L., Elmejjad S. ImaGEnS: Deliverable v1, v2, v3 An ALCATEL/INT Strategic Collaboration, 2002-2003.

[13]. M. Libicki, Who Runs What in the Global Information Grid: Ways to Share Local and Global Responsibility. RAND Corporation, 2000.

[14]. A. Bodertsky, D. Folk, Knowledge Management for Wireless Grid Operation Centers. 35th Hawaii International Conference on System Sciences. 2002.

[15]. Herrmann, K. , Mhl G. , Geihs K. , Self Management: The Solution to Complexity or Just Another Problem? IEEE Distributed Systems online 1541-4922 2005 Published by the IEEE Computer Society Vol. 6, No. 1; January 2005.

[16]. Kephart,J. O. and Chess,D. M. . The Vision of Autonomic Computing. Computer, January 2003.

[17]. Foster, I., Kesselman, C., and Tuecke, S. The Anatomy of the Grid: Enabling Scalable Virtual Organisations. International Journal of High Performance Computing Applications, 15(3).200-222.2001. www.globus.org/research/papers/anatomy.pdf.

[18]. Graupner S., Knig R., and al.: Impact of Virtualization on Management Systems. HP Laboratories. 33 (2003)

[19]. www.globus.org/