SERVICE LEVEL AGREEMENTS:
A Main Challenge For Next Generation Networks

Emmanuel Marilly[1], Olivier Martinot[1], Hélène Papini[1], Danny Goderis[2]

[1]ALCATEL
Route de Nozay, 91460 Marcoussis, France
E-mail: {Emmanuel.Marilly | Olivier.Martinot | Helene.Papini}@alcatel.fr

[2]ALCATEL
Francis Wellesplein 1, B-2018 Antwerp, Belgium.
E-mail: Danny.Goderis@alcatel.be

Abstract: Service Level Agreement (SLA) specification and management becomes a key differentiator in the service provider's offerings. SLA management will allow service providers to offer different levels of service guarantees and to differentiate himself from its competitors. It will improve its ability to satisfy customer expectations, as the customer will exactly know what to expect in terms of quality of service. SLAs are also a key requirement for the deployment of multimedia over the Internet, i.e. for Next Generation Networks. The Internet today is however a best-effort network, without the need for such strict contractual guarantees. Therefore numerous initiatives are currently undertaken to specify SLAs and to identify the new challenges for the provider's service and network management.
The aim of this paper is to present a high-level overview of the main SLA-related issues: the actors involved in SLA negotiations, the technical specification of SLAs and the mapping of SLAs to service classes. We give an overview of ongoing SLA-related initiatives in the industrial world, standardisation bodies and the research community.

Keywords: Service Level Agreement, Service Level Specification, Next Generation Networks, Service Management, IP, QoS.

1 Introduction

The telecom market is evolving towards services. New services will be proposed by the use of Internet capabilities, like multi media service, e-hotel, e-commerce, data transfer, unified messaging,... increasing the use of the network, dealing with convergence of data and voice networks.

Network traffic over the Internet in a multi-domain / multi-operator context is increasing as the number of users and applications (data, video, voice, ...). The results of this evolution is to increase the complexity of the network management.

Until now, IP has provided a "Best Effort" service in which network resources are shared equitably. But the weakness in this IP original design has been exposed. As a result, the Internet requires changes to accommodate the new applications requirements, and provide Quality of Service differentiation. Bandwidth is needed, but it is not enough. In the context of multi-domain / multi-operator it is necessary to define a Service Level Agreement.

The paper is organised in the following way: First of all, the SLA and the deriving elements will be defined. Then, a summary of the various views of service classifications will be presented. Finally, parameters to service assignment will be tackled. This paper concludes by an identification of major challenges that have to be faced to enforce SLA Management.

2 Service Level Agreement

2.1 Definition

A Service Level Agreement (SLA) is a contract between a network service provider and a customer that specifies, usually in measurable terms, what services the network service provider will furnish and what penalties will assess if the service provider cannot meet the established goals. Service providers differentiation will be driven by the reliability of the SLA Management and its monitoring during exploitation to contributing to the customers trust.

0-7803-7422-3/02/$17.00©2002 IEEE

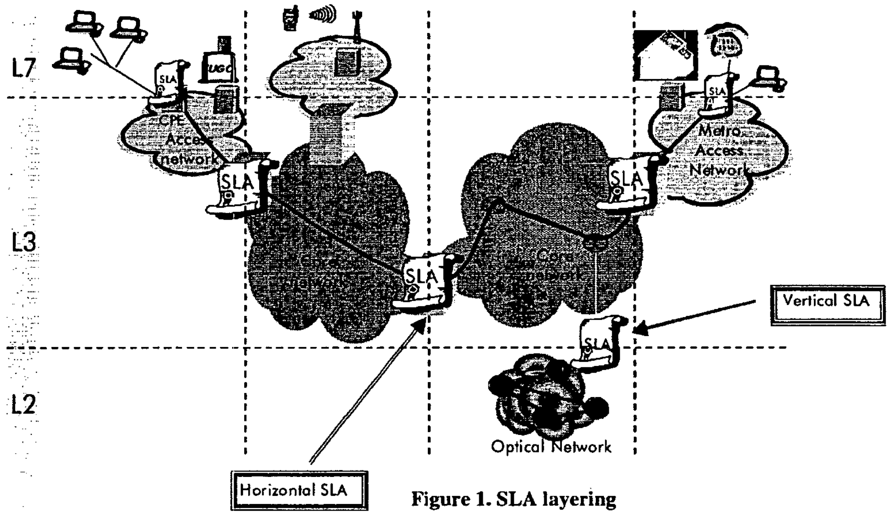

Figure 1. SLA layering

In order to guarantee customers the proper level of performance, service providers often offer services with SLAs, which provide customers with measurements of statistics like network availability, throughput, and latency.

2.2 SLA Market

An SLA by itself has no interesting value, if it is not managed efficiently.
When considering the expectations of customers regarding the service providers, most of the market studies highlight the following requirement :

- Reliable measurement of the quality of services.
- Provision of the expected quality of services.
- Optimisation of the resource usage.

SLA Managmenent will be a key player in the adoption of the Next Generation Services (like Video on Demand, Unified messaging,...) offered by the Next Generation Networks and GPRS/UMTS. SLA Management will enforce the confidence customer can have on their use, and support the transition from traditional usage of the services, to more elaborated usage of these new services.

2.3 SLA actors

The network service provider may be an operator, a carrier, an ISP or an ASP. The customer may be a carrier, an ISP, an enterprise or a subscriber (end user). SLA contractors are usually associated as follows:

CONSUMER		PROVIDER
carrier	$\leftrightarrow$	carrier
ISP/transport	$\leftrightarrow$	carrier
enterprise/operator	$\leftrightarrow$	ISP
enterprise	$\leftrightarrow$	ASP
subscriber	$\leftrightarrow$	ISP

Table 1. SLA actors

It has been difficult for service providers to offer service level management and, therefore, sophisticated SLAs. This difficulty is based on the fact that service providers must deal with several different network technologies and elements, often with parts of these network segments "owned" by different service providers.
The difficulty consists in the derivation of the SLA on the various network layers. Horizontal and vertical SLA can be distinguished.

- The horizontal SLA is a SLA between two Providers being at the same OSI layer (for instance two IP domains or two Optical Transport Network (OTN) domains).
- The vertical SLA is a SLA between two Providers at two different OSI layers (for instance between the core MPLS network and an optical network)

Figure 1 clearly illustrates the SLA management and layering issues.
A service provider who wishes to offer SLAs must provide an end-to-end service level management system that can accurately and granularly measure network performance.

3 Service Level Specifications

The Service Level Specifications (SLS) are the technical specifications deriving from the SLA. SLS can be a precise specification directly related to the SLA, but it can also be an interpretation of the SLA, an adaptation depending on the provider or on the service.
To propose an End-to-End solution to SLA/QoS management, it is necessary to define services, SLS parameters and a classification of these services depending on the SLS parameters. The focus on service level rather than on network level allows definition of service/SLA/QoS independently from the underlying network technology.

S1	voice
S2	videophone
S3	telephony services
S4	multimedia
S5	video on demand
S6	white board
S7	VPNx (IP VPN, optical VPN, tel. VPN= PMR)
S8	data real-time (telnet, ...)
S9	data interactive (web, transactions, e-commerce, email-server, ...)
S10	data streaming (FTP, bulk data transfer/retrieval, still image, ...)

Table 2. Service list

P1	loss
P2	delay
P3	jitter
P4	bandwidth
P5	consumer/provider
P6	monitoring
P7	admission control
P8	topology
P9	reliability/protection/security
P10	mono-direction/symmetric/asymmetric

Table 3. SLS parameters list

3.1 Services and SLS Parameters

Table 2 represents a non-exhaustive list of services giving a panel as large as possible (audio, video, data, tunnel). Table 3 represents a list of SLS parameters. These parameters are network technology independent.
The service parameters can be described with SLS parameters, but all the SLS parameters are not necessary. Only some parameters are used to define them. P1, P2 and P3 parameters are particularly end user oriented for QoS perception. Enterprise is more concerned with global parameters like P4 and P9. Other consumers (operators, ISP...) are concerned with P4, to P10 parameters, especially P4 and P9.

3.2 Mapping SLS parameters to Services

Services and SLS parameters have been identified. A mapping between these elements will give the most significant parameters defining the service. Four classes of performance have been chosen. All the relations between service and SLS parameters have been qualified by these classes.

++	very high performance	=	default performance
+	high performance	0	Indifferent

		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
		loss	delay	jitter	bandwidth	consumer/p rovider	monitoring	admission control	topology	reliability	direction
S1	Voice	=	++	++	+	++	=	++	=	+	2
S2	Videophone	=	++	+	++	++	=	++	=	+	2
S3	Telephony Services	=	=	0	0	++	=	++	=	+	1
S4	Multimedia	+	++	0	++	++	=	++	=	+	1.5
S5	video on demand	=	++	0	++	++	=	++	+	+	1
S6	White board	++	+	0	0	++	=	++	=	+	2
S7	VPN	depends on the propagated flow type						++	++	++	2
S8	Data real-time	++	++	0	+	++	=	++	=	+	2
S9	Data Interactive	++	+	0	0	++	=	++	=	+	1
S10	Data Streaming	++	=	0	0	++	=	++	=	+	1

Table 4. Mapping between Service and SLS parameters

Table 4 allows a regrouping of the various services in class of services. Many fora, standardisation bodies, and research projects have been working on this problem. The following chapter presents the most significant trends.

4. Services Classification

Services can be consolidated into several categories. There are many ways to consolidate distinct services. The classification depends on the SLS parameters. For an end-user the main parameters perceived are loss, delay, and jitter. For an enterprise, the global parameters corresponding to a level of billing are the bandwidth and the reliability (MDT, MTBF, MTTR). When the reliability or the provided bandwidth is not satisfactory, it can also be a determining reason to move to another provider. A perception-oriented classification is based on delay and loss parameters. In the following are enumerated proposed classifications (IETF, 3GPP, IST Tequila project, ETSI Tiphon project...).

4.1 Internet Engineering Task Force

The Internet needs bandwidth management and QoS. Adding QoS raises significant concerns, it enables technical characterisation of the QoS (e.g. SLA/SLS, reliability, delay, jitter, packet loss), mechanisms to provide QoS (e.g. RSVP, DiffServ), integration of QoS mechanisms at different levels and with different technologies (e.g. MPLS, ATM, Ethernet, Optic transport), management and assurance of QoS (traffic engineering).
IETF classification is mainly based on DiffServ [1][2] and DSCP [3] coding. There exists six kinds of per-hop-behavior (PHBs) : Expedited Forwarding (EF), Assured Forwarding 1 to 4 (AF1 to AF4), and Best Effort (BE).

EF PHBs is a high priority class of service (CoS) (in order to follow a minimum rate), with a strict maximum rate constraint above which traffic is dropped.

AF PHBs are defining four CoS, and for each of them, three drop precedences are associated. The semantic of the CoS is to be determined when installing a DiffServ network, the classification/metering/shaping depending on this semantic. Instead of dropping, AF PHBs actions may be a re-marking (from a drop precedence to another, from AF1 to AF4, from AF to BE...).
BE PHBs relies only on scheduling mechanisms of the routers, with no classification or shaping. It should only be verified that enough bandwidth may be used for BE traffic.

DiffServ is a very promising technology; however delivering real-time multimedia services on DiffServ-based IP networks still involves a lot of open research issues (ie : deployment, provisioning, authorization, fairness, congestion control, routing, security...). The main IETF groups working on this subject are:

- Internet Traffic Engineering (TEWG)
- Realtime Traffic Flow Measurement (RTFM)
- IP Performance Metric (IPPM)
- Remote Network Monitoring (RMONMIB)

4.2 3GPP

3GPP is the acronym of the project called "Third Generation Partnership Project". 3GPP provides Technical Specifications and Technical Reports for a 3rd Generation Mobile System based on evolved GSM core networks and the radio access technologies that they support (i.e., Universal Terrestrial Radio Access (UTRA) both Frequency Division Duplex (FDD) and Time Division Duplex (TDD) modes).
In 3GPP [4], four traffic classes are defined, based on delay requirement, each of them supporting error tolerant or error intolerant applications (table 5).

Traffic class	Conversational class	Streaming class	Interactive class	Background
	Conversational RT	streaming RT	Interactive best effort	Background best effort
	Delay < 150 ms.	Delay < 1 sec.	Delay < 1 sec.	Not guaranteed
Fundamental characteristics	Preserve time relation (variation) between information entities of the stream (stringent and low delay)	Preserve time relation (variation) between information entities of the stream	Request response pattern Preserve payload content	Destination is not expecting the data within a certain time Preserve payload content
Error tolerant applications	voice / video	streaming audio / video	voice messaging	Fax
Error intolerant applications	telnet, interactive games	FTP, still image, paging	Web browsing, e-commerce, e-mail server access	e-mail arrival, notification

Table 5. 3GPP Classification

4.3 Tequila

TEQUILA (Traffic Engineering for Quality of service in the Internet at Large) is a European research project with as primary goal the development of an integrated architecture and associated techniques for providing end-to-end QoS in a DiffServ-based IP network [5]. The Tequila Consortium consists of Alcatel, Algosystems S.A., FT-R&D, IMEC, NTUA, RACAL, UCL, TERENA and UniS. The project deals as well with the Service as with the Resource Management aspects. MPLS and IP-based techniques for traffic engineering are studied. TEQUILA concentrates on the definition and related QoS-provisioning of *IP connectivity services* like e.g. IP Virtual Leased Lines. An IP connectivity service is defined by (a set of) SLSs. End-user services such as Voice over IP or other multimedia applications (as e.g. defined in 3GPP) may then make use of the unambigious IP connectivity service API, which is specified at the IP transport level.
The TEQUILA project has taken the initiative in the IETF to propose a standard template for the IP-related parameters and semantics of an SLS [6]. Table 6 represents the tequila SLS parameter settings for various services. Some explanations ont the table 6 can be done.

The *topological scope* identifies the geographical region where the SLA contract is applicable by e.g. specifying ingress and egress interfaces. The *flow descriptor* uniquely identifies the packet stream of the SLA by e.g. specifying a packet filter (Differentiated Services Code Point, IP source address prefix, etc). The *traffic descriptor* describes the traffic envelop through e.g. a token bucket, allowing to identify in-and out-of-profile packets. *Excess treatment* then specifies the treatment of the out-of-profile packets at the network ingress edge including dropping, shaping and re-marking. The IP *performance parameters* specify the QoS network guarantees offered by the network to the customer (or the multimedia application) for in-profile packets such as delay, jitter, packet loss and throughput guarantees. *Service Schedule* finaly specifies when the contract is applicable by giving e.g. hours of the day, month, year.

Remark that the performance parameters might be either quantitative or qualitative. The latter allows for the definition of so-called Olympic services.

	Virtual Leased Line Service	**Bandwidth Pipe for Data Services**	**Minimum Rate Guaranteed Service**	**Qualitative Olympic Services**		**The Funnel Service**
Comments	Example of a uni-directional VLL, with quantitative guarantees	Service with only strict throughput guarantee. TC and ET are not defined but the operator might define one to use for protection.	It could be used for a bulk of ftp traffic, or adaptive video with min throughput requirements	They are meant to qualitatively differentiate between applications such as:		It is primarily a protection service; it restricts the amount of traffic entering a customer's network
				on-line web-browsing	e-mail traffic	
Topological Scope	(1|1)	(1|1)	(1|1)	(1|1) or (1|N)		(N|1) or (all|1)
Flow Descriptor	EF, S-D IP-A	S-D IP-A	AF1x	MBI		AF1x
Traffic Descriptor	(b, r) e.g. r=1	NA	(b, r)	(b, r), r indicates a minimum committed information rate		(b, r)
Excess Treatment	Dropping	NA	Remarking	Remarking		Dropping
Performance Parameters	D =20 (t=5, q=10e-3), L=0 (i.e. R = r)	R = 1	R = r	D=low L=low (gold/green)	D=med L=low (silver/green)	NA
Service Schedule	MBI, e.g. daily 9:00-17:00	MBI	MBI	MBI	MBI	MBI
Reliability	MBI, e.g. MDT = 2 days	MBI	MBI	MBI	MBI	MBI

(b, r): token bucket depth and rate (Mbps), p: peak rate, D: delay (ms), L: loss probability, R: throughput (Mbps), t: time interval (min), q: quantile, S-D: Source & Destination, IP-A: IP Address, MBI: May Be Indicated, NA: Not Applicable, MDT: Maximum Down Time (per year), ET: Excess Treatment, TC: Traffic Conformance

Table 6. Tequila Classification

4.4 Aquila

Aquila (Adaptive Resource Control for QoS Using an IP-based Layered Architecture) is a European research project. It defines, evaluates, and implements an enhanced architecture for QoS in the Internet.
There is a set of commonalties between the AQUILA and TEQUILA approaches [7]. The main difference is that the AQUILA consortium has introduced the concept of predefined SLS types that are based on the generic SLS definition. From the point of view of the applications, a predefined SLS type supports a range of applications that have similar communication behavior and therefore similar QoS requirements, such as for delay, packet loss, etc... Table 8 represents the predefined SLS types defined in the AQUILA project for various services:

Predefined SLS Type	**Premium CBR**	**Premium VBR**	**Premium Multimedia**	**Premium Mission Critical**
Service	Voice VLL -like	Video Teleconferencing	Streaming multimedia Premium FTP	Transaction oriented applications

Table 7. Aquila Classification

4.5 Tiphon

The ETSI Tiphon project's objective is to support the market for voice communication and related voiceband communication between users. It ensures that users connected to IP based networks can communicate with users in Switched Circuit Networks (SCN - such as PSTN/ISDN and GSM).

In ETSI Tiphon project [8] are defined four classes of service for speech:

Class	4 BEST	3 HIGH	2 MEDIUM	1 BEST EFFORT
End-to-end delay	<100ms	<100ms	<150ms	<400ms (target)

Table 8. Tiphon Classification

4.6 TeleManagement Forum

The TM Forum provides a performance reporting concept [9] and SLA management handbook [10]. The objective of the Handbook is to assist two parties in developing an SLA with a practical view of the fundamental issues.

The main TMF projects working on this subject are:

- QoS/SLA management Project
- SLA Management Project (Wholesale and Retail Services)

4.7 Partial Synthesis

The table 9 is a synthesis of all the collected classifications and of the Mapping SLS parameters / services.
This synthesis is obviously not complete but it identifies services offered in term of categories and provides a mapping between Class of Services and SLS parameters. It will be used to determine service performance measures relevant for efficient SLA monitoring.

Class of Applications	**Bandwidth needed**	**Loss sensibility**	**Delay sensibility**	**Jitter Sensibility**
Interactive applications (telnet, Database access, short web transactions, Xwindow, ...)	low	medium	high	Low
Interactive multimedia applications (voice or video over IP)	medium	medium	high	High
Non-interactive multimedia applications (Tele-learning, broadcast, ...)	high	medium	low	Low
Request oriented applications (client/server, NFS, ...)	medium	low	high	Low
Transfer applications (FTP, backup, tele-market, long web transactions)	high	low	low	Low
Network control (SNMP, RIP, OSPF, BGP, ...)	low	high	high	Low

Table 9. Synthesis Classification

5 Conclusion: SLA Management Main challenges

SLA Management must provide means/tools for reliably provision the service, monitor the SLA fulfilment during the service usage and anticipate any decrease of performance [11].

The Next Generation Networks can be seen as an assembly of different autonomous networks, each having their own role in the service provision (fixed access, mobile access, transport...), their own technology (multi-vendor equipment), and operated by separated authorities (multi-operators).

The first difficulty for service providers to offer service level management and, therefore, sophisticated SLAs comes from the fact that service providers must deal with several different network technologies and elements, often with parts of these network segments "owned" by different service providers. Then a generic SLA for various types of networks must be provided (i.e. standard).

The second difficulty is the derivation of the SLA on the various network layers. A service provider who wishes to offer SLAs must provide an end-to-end service level management system that can accurately and granularly measure network performance.

The main challenges of the SLA management can be developed in three parts. The SLA management itself, the fulfilment and the assurance in the SLA management.

5.1 Main Challenges in SLA Management

The open issues on SLA management focus on these three particular points:

- *The information model of SLS* (contents of the SLS). The SLS template or information model is not clearly defined. It is mandatory for the SLS to be defined by a template or an information model in order to allow cooperation/negotiation between entities. This model should also cope with various type of service (3GPP, Multimedia, IP-VPN ...).
- *The SLS negotiation protocol*. The negotiation allows cooperation/negotiation between OSS and can be divided into three aspects: The Functional aspect (i.e. the negotiation of SLSs, the modification of an implemented SLS, and the information about an implemented SLS), The Security aspect (i.e. authentication, access control, integrity and confidentiality), and The Inter-domain aspect (i.e. SLS negotiation between distinct administrative domains).
- *The end-to-end point of view*. This induces that agreements have to be established between providers, and between providers and customers. Therefore it is necessary to establish Out-Sourcing agreement (a SLA) with other networks providers to lease part of their network, that can be for instance defined as leased line, VPN, ...

Hence two scenarios can be developed to provide an end-to-end service to an End-user customer:

- o The End-User must manage different SLAs and is the only one who manages their interactions from end-to-end.
- o The End-User manages only one SLA with his service provider and all necessary information for management must be propagated into the network (Out-Sourcing) from end-to-end.

The automatic Management of the SLA/QoS (for instance automatic negotiation) does not exist yet. It requires the mapping of the SLA requirement into technical configuration of network equipment and the specification of tools to generate QoS parameters from SLA (in particular for the Optical-SLA). The SLA monitoring must be improved in order to determine service performance measurements relevant for efficient SLA monitoring, manage the network to maintain the SLA requirements (equipment reconfiguration) and optimise the network performance and the network usage.

5.2 Main Challenges in the Fulfilment part of the SLA management

In the fulfilment part, the main issues for industrial are:

- *The Resource Admission Control*. This function must take into account the whole process from subscription to activation for long and short-term provisioning. The long-term provisioning includes network optimising and uses SLS long-term subscription and the short-term provisioning needs different algorithms for optimising the network.
- *The Allocation Management*. It addresses the communication with the NEs and is necessary for any test and the issue concern the mediation towards vendor-specific NE.
- *The Resource Allocation Request Handling*. It must manage the reservation of resources in case of short-term services and the requests coming by signalling.

Other issues concern the link between the different elements (for instance architecture between fulfilment and assurance, SLS fulfilment...). The feedback from the Assurance to the fulfilment part, for short-term problem resolution, must be studied.

5.3 Main Challenges in the Assurance part of the SLA Management

The ultimate role of service quality management is to help match expected quality with perceived quality. This is accomplished by assuring that the achieved performance of service is in line with specifications and contracts. The main issues of assurance in SLA management are:

- *QoS Metrics* Computation and QoS Metrics Report Management. There are many standardization issues (for instance IETF, TMF...).
- *Performance Management*, data collection and network measurement to service level information.
- *Problem Management*, automatic handling, root cause analysis, trouble ticketing and traffic forecast. Automatic problem resolution via the *traffic engineering*.
- *Forecast function*. The recommendation for performance improvement (short/long term corrections) is still an open issue. The forecasting techniques exist but the use of these techniques in the telecom area and especially in service forecasting is not clearly defined (which techniques use and on which parameters).

5.4 Ultimate Conclusion

Nowadays, some providers begin to offer SLA contracts (e.g. with early discounting features). However, these offers are still basic and require development and research work that have been discussed in this paper. In conclusion we stress some key topics related to SLA management:

- SLA model suitable for a wide range of services,
- Automation: operators want to reduce the time for deploying services from several months to a few days,
- Scalability: networks are growing (1000 x larger in 5 years) so all solutions have to be scalable to meet operators requirements,
- Cross-domain: consequence of scalability, to be fulfilled, a service crosses several different domains.

References

[1] "An Architecture for Differentiated Services",- *IETF, RFC 2475*, December 1998.

[2] "Assured Forwarding PHB Group",- *IETF, RFC 2597*, June 1999.

[3] "Definition of the Differentiated Services Field (DS-Field) in the IPv4 and IPv6 headers",- *IETF, RFC 2474*, June 1998.

[4] "Study on PS domain services and capabilities",- *3GPP, TR-22.976, R2000-v2.0.0*, December 2000.

[5] "A Management and Control Architecture for Providing IP Differentiated Services in MPLS-based Networks", - *IEEE Communication Magazine* -, P. Trimintzios and all, May 2001.

[6] " Service Level Specification Semantics and Parameters ", *-IETF Draft-<draft-tequila-sls-00.txt>* -, D. Goderis and all, November 2000, http://www.ist-tequila.org

[7] "Definition and usage of SLSs in the AQUILA consortium", *-IETF Draft-<draft-salsano-aquila-sls-00.txt>* -, S. Salsano all, November 2000, http://www-st.inf.tu-dresden.de/aquila/

[8] "Definition of QoS classes", *-Tiphon project, ETSI TS 101 329-2 v1.1.1-*, July 2000.

[9] "Performance Reporting Concepts and Definitions", *-TMF 701-*,May 1999.

[10] "SLA Management Handbook", *-TMF GB971-*, June 2001.

[11] "SLA management: a key differentiator for service providers" Alcatel Telecom Review, G. Désoblin, H. Papini, 3rd quarter 2001.