Chapter 11

SLA Driven Network Management

11.1. Introduction

Service Driven Network Management (SDNM) and Service Level Agreements (SLAs) have aroused great interest in the world network and telecom market in the past few years. We consider in this chapter SLAs regarding network services or those supported by a communication network such as the Internet. This will therefore exclude service level agreements relevant in other fields.

This chapter begins by describing the strong link between SLA-based network management on one hand and the policy management model and the QoS mechanisms studied in the preceding chapters on the other hand. It then proceeds to define the different components of an SLA. It is followed by a review of available SLA management systems and SLA modeling. The chapter ends with a description of a group of European projects which considered the provision of QoS based using SLAs.

11.2. Requirements for service driven management

The increasing impact of IT on the global performance of companies obliges IT managers to behave as true service providers. Also, telecommunication service

---

Chapter written by Issam AIB and Belkacem DAHEB.

providers and network operators competing in the current market are constrained to offer value-added services with tangible assurances to their customers. The differentiation of client requirements based on SLAs is then essential. An SLA is designed to establish a mutual agreement between the client (CL) and the service provider (SP), specifying the rights and responsibilities in which the CL and the SP engage regarding QoS in terms of negotiated price and duration. The CL-SP relationship founded in this contract requires that the provider has an exact knowledge of the level of service to be provided and a fine understanding of the users' subjective perception. It also requires transparent and regular communication to the client of the actual provided versus contracted service level.

11.2.1. *SLA and QoS*

QoS is a generic concept that covers several performance and statistical aspects. It is defined in [TMF 01] as "the collective effect of service performance that determines the degree of satisfaction of the user of the service" (for more details see Chapter 6). The final goal of an SLA is to specify the necessary parameters to ensure the client's desired QoS level.

11.2.2. *SLA and policy-based management (PBM)*

The set-up of an SLA between the CL and the SP implies the possibility of providing a QoS in a network that supports only best effort service. It therefore requires the installation of advanced mechanisms that make it possible to provide a more elaborated service than that provided by the IP architecture.

In this regard, work is being done by the IETF (Internet Engineering Task Force) in order to propose a model that ensures QoS enabled network control. This model, named PBM, consists of considering the network as a state machine on which actions take place. These actions emanate from policy rules resulting from the manner in which an SP wants to manage its network and therefore from the different SLAs with their clients. A design has been proposed for this model (Figure 11.1), more details of which can be found in Chapter 6.

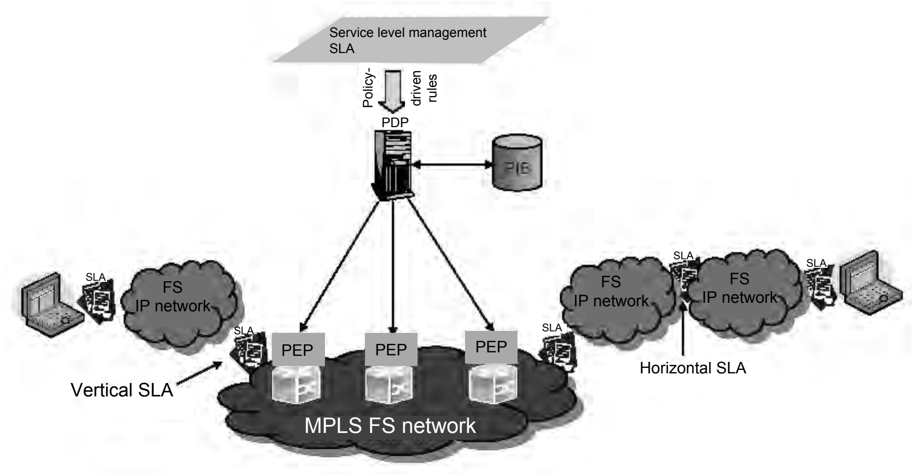


Figure 11.1. *Service driven management*

An SLA describes the high level business policy and is converted into Service Level Objectives (SLOs). An SLO consists of QoS parameters and their values. SLO values are decided after negotiations between the SP and the customer.

One way of providing the service agreed upon in the SLO is to derive a set of policy rules and manage the network according to these rules, [FAW 04]. The policy rules do not necessarily consist only of rules derived from the SLOs but they can have other sources such as the network administrator, etc.

11.3. The SLA

An SLA, to differentiate it from other types of contracts, is a formal measure between an SP and its CL which defines all the aspects to be provided by the service. Typically, an SLA covers availability, performance and invoicing, as well as specific measures in case of failure or malfunction of the service.

Consequently, an SLA constitutes a legal foundation, the basis on which the two participants (the customer and the SP) rely to plan their respective future affairs and relations. Eventually, the SLA serves as a means of judgment in case of dispute. This supposes the existence of a legal entity that certifies the contracts established between the clients and their SP.

An SLA can be composed of very basic elements that can be counted on the fingers of one hand, or can be composed of hundreds of lines of literary specifications surrounded by many ambiguities and complexities. A common error that many SLA providers and designers make is that they tend to think that the longer the SLA, the better it is.

Typically, an SLA contains the following components:

– a description of the contracting parties (the CL and the SP) as well as the legal party. This is called the WHO part of the SLA;

– a description of the service to be provided as well as the different components (the WHAT);

– the service scope: the duration of the validation of the contract and the geographical limits (the WHEN and the WHERE);

– the respective responsibilities of the contracting parties;

– the global QoS parameters of the expected service, as well as the QoS parameters of its components;

– the QoS parameters of an e-service concerning availability, performance, flow and rate of error;

– the procedure and method(s) of service billing;

– the procedures and mechanisms to use for the supervision of provided QoS and the manner in which the client receives a return to the QoS level;

– a detailed procedure concerning the maintenance and treatment of failures (rate of reactivity to the treatment of anomalies);

– a description of the exclusions in which the SLA is no longer valid;

– the consequences of disrespect of the contract:

- by the CL,

- by the SP.

In the following, we will explain in greater detail the components of an SLA. It is, however, true to say that there is not a global consensus among all service providers (network operators, Internet service providers, service provider applications, etc.) on what should be included in an SLA, nor on the semantic significance of an SLA and its components. This will be clarified in section 11.8.

11.3.1. *Designation of SLA contracting parties*

The designation of the CL and of the SP is used to identify the two parties as well as serving as a means of contact. Other attributes can be included, such as physical contact addresses, telephone and fax numbers, electronic addresses, etc.

The current tendency is to develop databases for user profiles. Profile databases make it possible to collect intelligent information on the client's use of services. Thus, the SP can propose a service adaptation according to its client's profile. In this case, an ID of the CL will be found in the identification field of the CL in the user profile database.

11.3.1.1. *Duration of the SLA*

This parameter tracks the global duration in which the SLA is considered valid by both parties. If the provided service is comprised of several sub-services, their respective periods of validity are included but are not necessarily equal to the global duration of the SLA.

11.3.1.2. *Service scope*

Considering that the SLA represents an assurance of QoS level, the SP must specify the physical extent of these guarantees. In fact, if an ISP can guarantee a level of QoS for their clients in the domain of its network operator, there is no way to guarantee the same QoS if the user flows travel through other domains, which is often the case. To remedy this difficulty, SLA chains with end-to-end guarantees have been introduced.

Nevertheless, there is a consensus on the fact that an SLA is a high level contract that describes the non-technical service parties. More details are given in the following section.

11.4. Specification of level of service (SLS)

The SLA is considered as the non-technical view of the service contract. The SLS represents the meticulous description of the technical parameters that constitute the service. The complete SLS is generally not accessible to the conventional client. This is principally due to the fact that it represents a network administrator view of the SLA. However, the client can receive a client SLS that contains technical parameters relevant to the user and make it possible for the client to confirm that the SP has adhered to the terms of the contract.

11.4.1. *Service availability guarantee*

Availability concerns more than just access to the service. It concerns a level of service on which the client can capitalize. For example, a website is considered accessible when the ping command sends an accessibility result. However, an IP telephony application considers that the communication service is inaccessible as soon as the available flow is clearly inferior to a predefined threshold.

For a composite service, an increase in its availability is, in general, a function of the increase in the availability of its components. However, it is important to know that the increase perceived by the client cannot correspond to that perceived by the network operator. Therefore, the semantics of service availability, as well as the SLS metrics used to verify the service, must be specified.

For example, considering the availability described in the IP service SLA offered to a company as proposed in [MAR 02]:

> The SLA specifies that the operator's network will be available 99.8% of the time evaluated over a period of a month and averaged with regard to the client company's n routers. The network is considered unavailable after 5 successive minutes of operation of failed pings; the frequency of pings is 20 sec with a 2 sec time-out.

It is clear that this offer of availability is very SP-oriented. In fact, in order to violate the availability metric, the network would have to be unavailable for approximately 14 hours/month. In fact, the network could well be "pingable" without really being available to the client and therefore be seen as unavailable by the client. Thus, an availability metric must consider the client's point of view rather than the SP's point of view to remain objective with regard to the SLA.

11.4.2. *Service performance guarantee*

The performance part of an SLS is principally the reason for its existence. The groups of precise parameters that determine the service's performance vary according to the service considered. We will see, in the section that addresses different SLA research projects, several typical examples of performance parameters.

For an IP network connectivity service, for example, a connection's performance is calculated with regard to four primary parameters: bandwidth, delay, jitter and rate of errors. Also, the manner in which these metrics are calculated must be specified as the supervision technique of the performance parameters considerably influences the final values.

[TMF 01] makes a distinction between performance parameters and performance events. An event is an instantaneous phenomenon that takes place within the network that is providing the service or in its environment and which affects the QoS provided to the CL. Examples of events are: errored seconds (ES), severe ES and severe errored periods.

Performance parameters are derived from performance event processing carried out during periods of defined measurements. The processing can consist of a calculation of an average, or of a comparison with a single threshold (for example: the maximum tolerated number of severe ES).

The exact determination of performance events and performance parameter calculation functions give the SP a practical means of determining and ensuring performance parameters proposed to its clients.

It can be noted that the supervision of performance parameters must be done with tools that the client trusts. If the client has no objective means of ensuring to himself of the reality and quality of provided services, he will not be able to ask for compensation in the case of a non-respected SLA. In certain cases, the SP and the CL call upon a neutral third party specialized in the measurements and which takes care of the supervision of performance parameter so that each party is more assured with regard to the validity of the performance reports.

11.4.3. *Billing methods*

There are multiple ways of billing a service. For example, an ISP can bill its customers at a fixed rate, by volume, by duration of use, by prepaid cards, or according to QoS requirements, etc. The CL must be well-informed on the billing method used in order to plan a budget and know how to select and manage the use of the service offered.

A billing system can be static or it can take into consideration the client's evolution and loyalty. However, the implementation of a complex billing structure is often costly and not justifiable unless the proposed services generate sufficient added value to ensure the perennial nature of the SP's system.

If we take, for example, the ISP market, the following tendency is shaped according to the importance of the ISP:

– For an ISP with few clients, notably during its implementation and as long as its client base does not exceed a few thousand, fixed billing is applied in most cases.

This does not require any particular billing software and does not incur excessive fees for the company.

– Beyond several thousand clients (say, between 2,000 and 10,000 clients), a billing software, preferably low-priced, becomes possible. These would generally be programs developed on the basis of Excel or File Maker Pro files, packed with sufficient functionalities to start proposing interesting service offers.

– For several tens of thousands of clients, a software package such as INFRANET from Portal becomes unavoidable. The cost of the licenses is justified in this case as it is compensated by the number of subscribers. These complete solutions have the necessary infrastructure to address the numerous evolving offers and the necessary developments for integration in the global ISP system. This would be a made-to-measure package.

11.4.4. *Addressing anomalies*

The SLS must contain clear and precise clauses regarding how service function anomalies are taken into consideration. Anomalies can originate with the SP (malfunction in the offered QoS) or with the CL (abuse of service use, non-conforming use of the service). The SLS also contains clauses that specify conditions that put the contract out of the field of the SP's QoS engagements.

In a policy management environment, anomalies are handled with the help of behavior-based rules that define the actions to take, often on the part of the SP in the case of malfunction, degradation of the level of service or poor service used by the CL. For example, in the case of excess bandwidth usage by an IP network CL, the SP can implement many policies to regulate the CL's usage. It can ignore excess packets, or mark them as un-guaranteed packets that can possibly be isolated in case of congestion. It can also alert the client of the excess, or simply bill the client for the excess packets. However, the rules for the management of anomalies of use and of service functions must be properly specified in the SLS so that neither party is surprised by the behavior of the other.

11.5. Service contract chains

The federation of services is an unavoidable consequence in a competitive environment, in which the best solution for the SPs' survival is to specialize and innovate in a precise (sub-)domain. Thus, in order to achieve a Final Client (FC) service, a chain of several CL-SP can take place, in which SPs become CLs for other SPs. For example, an ISP is an SP for the FCs, whereas it is a CL for its network operator. Similarly, a VoD (Video on Demand) server is a CL for its ISP. A CL can be a CL for its ISP for Internet access and during a period of time can also be a CL for a VoD service. This is how the problem of SLA chains and end-to-end SLA management arises.

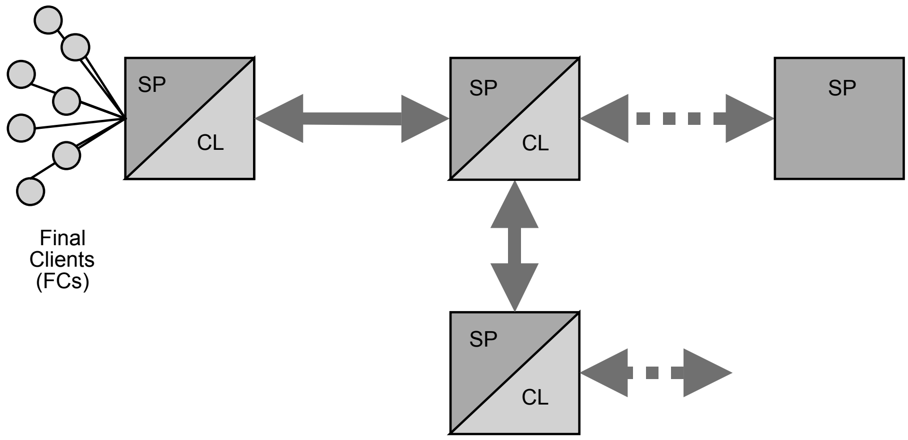


Figure 11.2. *SLA chains*

11.6. SLA types

The classification of SLAs can be done in various possible ways. We have already seen the differentiation of CL-SP and SP-SP SLAs. In other works, it was agreed that two main SLA types should be differentiated. The differentiation criteria concerns the OSI layers that are in question in the SLA.

11.6.1. *Horizontal SLA*

A horizontal SLA is an SLA between two SPs within the same OSI layer. In this category, for example, is an SLA between two IP domains or two request content providers.

11.6.2. *Vertical SLA*

A vertical SLA is an SLA between two SPs on two different OSI layers. In this category, for example, is an SLA between an MPLS network domain and an optic transport network (OTN), or an SLA between a VoD server and its ISP.

11.7. SLA management (SLM)

An SLA on its own is not sufficient if it has not been respected and properly managed. SLM represents the most difficult part of the Service-Oriented Architecture (SOA) which we proposed to study in this chapter. In fact, the challenge presented by an SP when it decides to adopt this architecture is to implement a management system that offers the contracted level of service to each client, as well as all the mechanisms and the gauges that make it possible to supervise the different service parameters, in order to be able to report on the performance and the reality of provided services. These mechanisms also make it possible to detect service failures. In short, an SLM system must manage [MAR 02] the individual life cycle of each client SLA, ensure the coherence of the SLA database as well as the coordination of end-to-end SLAs, as well as the observation of its business' global objectives.

We detail these three points in the following sections.

11.7.1. *An SLA life cycle*

An SP's SLM tool must support an SLA life cycle management. As with all modeling techniques that undergo several cycles, an SLA also goes through a group of development phases before its implementation. [TMF 01] identifies five principal phases: service development, negotiation and sale, implementation, execution and evaluation. Other divisions are possible according to the considered SLA's vision. [DAN 03] proposes four development phases: creation, deployment and provisioning, implementation and supervision, cancellation. Moreover, according to the economical scenario, there could be a varying number of sub-phases in each phase of the SLAs life cycle. Principally, this SLA life cycle contains the following phases [TMF 01].

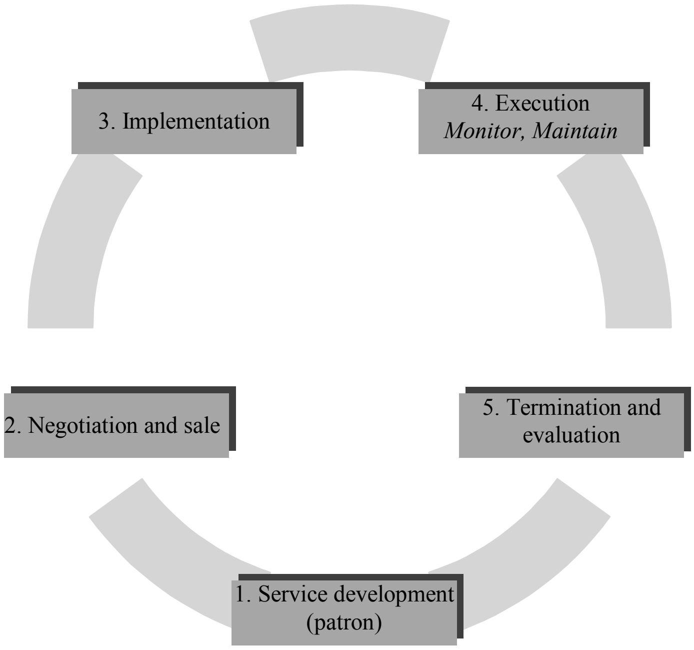


Figure 11.3. *SLA life cycle [TMF 01]*

11.7.1.1. *Service development phase*

The development of a new service offer begins with several external and internal sources.

The need for the development of new applications is especially caused by the aggressive competition on the current network service provider market (ISP, network operators, content providers, etc.). Also, the emergence of new communication technology implies the development of new services and new applications to support them. A last factor, which is just as important, is the group of client requirements for the development of new services, which are sometimes aimed at the specific needs of clients who are prepared to pay the corresponding costs.

In addition, the results of the evaluations of services and existing SLAs, of customer service and of the internal measures and reviews carried out on the revenues and existing service management efficiency contribute to the conception of new solutions to improve the image and quality of the services offered to the customers.

The development of the service itself is thus based on four principal components:

– the identification of the clients' needs;

– the identification of appropriate service characteristics: parameters, values, level of service;

– the identification of resource capabilities available to support the desired service;

– and finally, the development of a basic model for the service and the corresponding SLA.

The final result of this phase is therefore the description of the new service with its corresponding SLA template.

11.7.1.2. *Negotiation and sale*

This is the phase during which the client subscribes to a certain offer of service with a varying number of the concerned parameters specified. The client can subscribe to a global service offer of long duration (a year, for example) and he solicits from time to time an addition or modification to a new sub-service. The final result of this phase is a signed SLA between the SP and their CL.

11.7.1.3. *Deployment and implementation*

This is the phase during which the service is truly deployed (for example, the deployment of specific policies for the CL traffic), tested and activated.

11.7.1.4. *Execution and supervision*

This phase encompasses the agreed-upon duration of the validity of the SLA. It consists principally of tracking the proper functioning of the service, the real-time follow up of activity reports and service levels, as well as the detection and handling of eventual anomalies.

11.7.1.5. *Termination and evaluation*

This phase is composed of two principal parts. The first concerns evaluating the CL's level of satisfaction and identifying the evolution of the CL's needs. The latter represents the SP's internal evaluation to measure the overall satisfaction level of its clients, to group the performance results of the resources that were put into play to assure the SLAs, to compare the obtained performance results with regard to the SP's general goals and finally to propose improvements that will serve as a link to return to the first phase of the SLA life cycle.

11.7.2. *End-to-end SLM*

An SLA is considered an interface between a CL and an SP. In practice, there can be one or more domains and therefore several SPs between the final CL and the actual SP. Each SP domain is supposed to be independently and proprietarily managed.

To reduce the service deployment time, it is necessary to establish a standard for SLA representation and negotiation. Chapter 7 will handle more precisely the problems encountered in the establishment of these SLAs.

11.7.3. *Observing the SP's global objectives*

The observance of the SP's global objectives begins by respecting each SLA concluded with its clients. The terminology used for this is FAB (Fulfillment, Assurance, Billing), which constitutes the key components to ensure a successful management.

An SP must be able to provide, following the different concluded SLAs, a reliable provisioning of resources (routers, servers, processors, etc.). The existence of tools that make it possible to deduct the necessary provisioning operations from a group of SLAs is primordial and constitutes a major challenge to the SOA architecture.

The development of techniques that enable the automatic translation of SLAs into policy rules for the provisioning and management of resources but also the maintenance of QoS supervision and billing is the current solution in network and connectivity service management.

11.8. SLA modeling and representation

After the study we have just carried out on SLAs, we now understand that the starting point of an SOA solution that lends itself well to portability, life cycle management automation and SLA interoperability is that of a standard SLA model.

The standardization task is difficult enough, due to the complexity of the subject, and there currently is no unanimity for a precise SLA model.

We describe here the group of TMF recommendations [TMF 01] for the conception of an SLA. TMF considers that it is very difficult to define a generic model that encompasses all SLA types because it is not always easy to translate client QoS needs in terms that the SP uses in the SLA.

From the view of the TMF, an SP proposes a group of products and services to their clients. This group consists of a number of commercial offers. A commercial offer represents a pack of services offered by the SP to the client and can be a single service (for example, an ATM PVC) or a group of services (for example, an xDSL access with email and Web access).

This commercial offer is composed of a group of service elements. Each service element can be associated to a service class (for example, Gold or Silver). A service element models the capabilities specific to a particular technology (for example, IP connectivity, xDSL connectivity) or operational capabilities (for example, customer service).

Service Resources are the basic blocks that compose basic service elements and are often invisible to the user. Figure 11.4 shows such a service package [TMF 01].

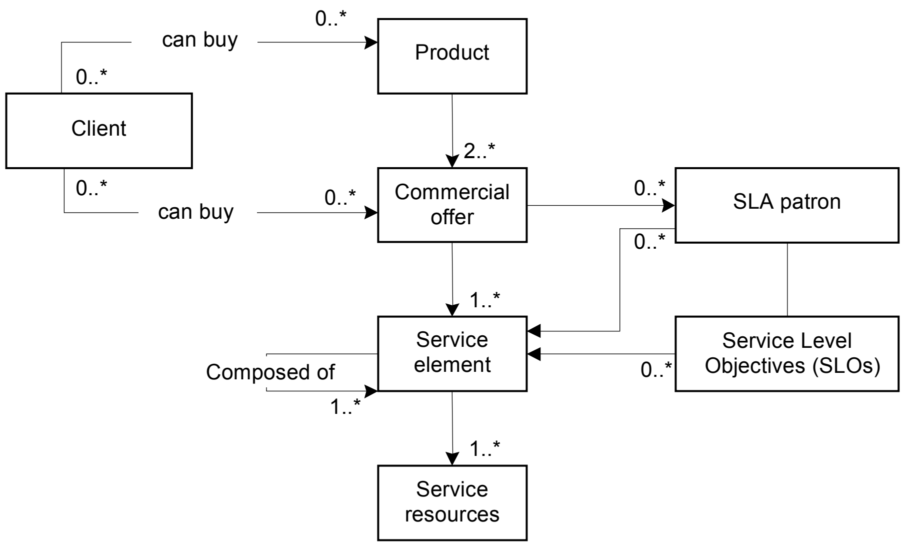


Figure 11.4. *Service composition [TMF 01]*

For example, a residential Internet connectivity service offer can comprise two SLA template instances: basic email and residential access. The basic email template uses the Service Email service element and the residential access template uses the

Access IP service element. Finally, these service elements rely on basic resources for their implementation, such as an email server, an authentication server, a DHCP server, access routers or an access modem.

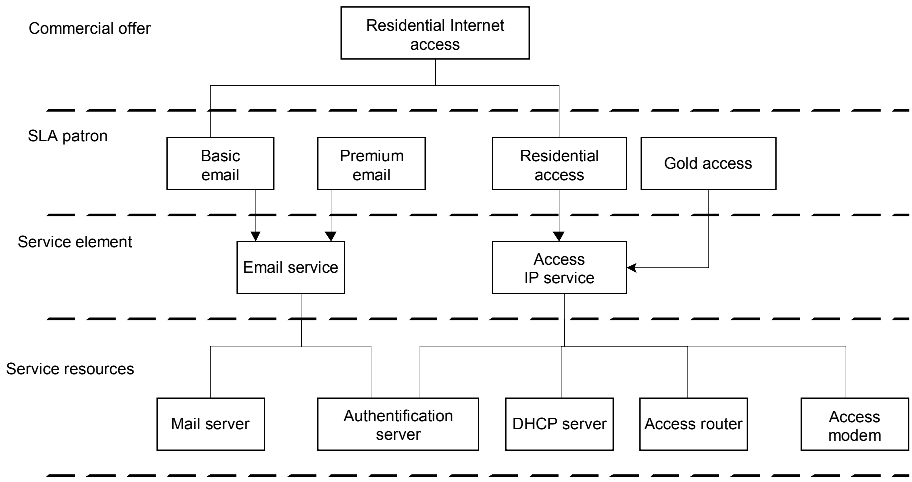


Figure 11.5. *Example of a service package*

An intrinsic part of the development of a service is the production of an SLA template. The role of an SLA template is to capture a group of SLOs for a certain service (see Figure 11.4). The objective of service level is the representation of that guaranteed level of service offered in the SLA.

11.9. Research projects and activities

This section gives a general overview of the principal projects that have addressed QoS and the field of defining SLAs.

There are many projects whose objective is to provide recommendations for constructing the future generation of data communication networks.

11.9.1. *The TEQUILA project*

The objective of the TEQUILA (Traffic Engineering for Quality of Service in the Internet, at Large Scale) project is to study, specify, implement and validate a group of service definitions and traffic engineering tools to obtain end-to-end QoS as required in the Internet [GOD 01].

The TEQUILA group differentiates between a qualitative and a quantitative SLS:

– a qualitative SLS is used to specify a service with relative QoS indicators, such as the packet loss rate;

– a quantitative SLS is used to represent a service for which the QoS indicators are measured exactly, for example, time of transit.

The group does not stop there. It has also defined different SLS attributes. These are described as follows:

– the span: uniquely identifies the geographical or topological region in which the QoS applies;

– flux identifier: identifies a flow of data with specific characteristics;

– conformity and traffic envelope: describes IP traffic flow characteristics;

– excess handling: describes how the service provider will handle excess traffic, in other words, traffic outside the profile;

– performance guarantee: the performance parameters describe guarantees of the services offered by the network to the user for profile traffic only;

– calendar of service: indicates the start and end times of service, for example, when the service is available;

– availability: indicates the average service breakdown allowed per year and the maximum time allowed for service re-establishment after breakdown.

This is, by and large, the SLS format defined by the TEQUILA consortium.

11.9.2. *The AQUILA project*

The AQUILA (Adaptive Resource Control for Quality of Service Using an IP-based Layered Architecture) consortium's objective is to define and implement a layered architecture for QoS support in IP networks [SAL 00]. To accomplish this, the AQUILA consortium has worked on defining the SLS based on the works of the TEQUILA project.

The consortium is in agreement with the need to have a formalized standard SLS representation between the client and the network. This SLS representation should be very general and able to express all the possible service offers based on the

DiffServ model [BLA 98]. The AQUILA consortium has signaled the need for a mechanism to simplify the generic SLS description. This has led to the definition of predefined SLS types [SAL 00].

From the application point of view, a predefined SLS can simplify the interaction between the client and the network. It makes it possible to support a range of applications with the same communication behavior and therefore similar requirements to QoS, such as delay, packet loss, jitter, etc. From the operator's point of view, this simplifies the network management and enables an efficient flow aggregation [SAL 00].

The SLS structure defined by AQUILA is composed of the following attributes:

– SLS type;

– span;

– flow identification;

– traffic description and conformity test;

– performance guarantees;

– service calendar.

As we can see, AQUILA's SLS omits some of the TEQUILA's SLS entries (excess handling, availability) while including a new attribute called the SLS type that distinguishes between a client SLS and a predefined SLS. The predefined SLS values proposed by AQUILA are:

– PCBR – Premium CBR;

– PVBR – Premium VBR;

– PMM – Premium Multimedia;

– PMC – Premium Mission Critical.

11.9.3. *The CADENUS project*

The goal of the CADENUS (Creation and Deployment of End-User Services in Premium IP Networks) project is to provide dynamic creation and configuration across the components linked to the user. This includes the authorization and registration to service components linked to the network along with the proper QoS and compatibility control [ROM 00].

The CADENUS project introduces the notion of dynamic service creation (DSC). This is obtained from communication with many IP service functions and components. These components can be firewalls, proxy servers, SMTP relays, etc. They define CADENUS' premium IP layer, the service creation layer, which manages entities and resources, negotiations and dynamic SLA creation [CAM 01].

CADENUS uses SLA and service templates during the negotiation phase and considers at least two different types of dynamic behaviors:

– user requirements that vary with time;

– network conditions that vary with time (the user is updated with regard to these changes via feedbacks).

Thus, four scenarios can be created by combining each of the types described above [ROM 00]:

– no user requirement variations/no network condition variations;

– user requirement variations/network condition variations;

– no user requirement variations/network condition variations;

– user requirement variations/no network condition variations.

Each of these scenarios needs a different SLA type. The first case needs a static SLA and the other cases need renegotiable SLAs.

11.9.4. *The SEQUIN project*

The objective of the SEQUIN (Service Quality across Independently Managed Networks) project is to define and implement an end-to-end QoS approach that will operate across multiple management fields and will exploit a combination of the two technologies, IP and ATM [CAM 01].

SEQUIN's work focuses on the development of a parameter definition of QoS. Four parameters are adopted:

– delay;

– IP packet delay variation;

– capacity;

– packet loss.

Different QoS classes have been defined based on these parameters.

With the goal of developing the QoS parameter definition cited above, SEQUIN used the works of the TEQUILA, CADENUS and AQUILA projects previously presented and has contributed to the analysis that follows in relation to SLAs.

Multi-domain SLA support requires a group of standardized semantics for SLSs negotiated in different locations [CAM 01]:

– between the client and the service provider;

– within an administrative domain (intra-domain SLS negotiation);

– between administrative domains (inter-domain SLS negotiation).

CADENUS approach attempts to contribute to the support of multi-domain SLAs by defining, initially, detailed and global SLA terms. The detailed SLA (DSLA) refers to the contract between a final user and a service provider. The global SLA (GSLA) refers to the inter-network contract of multi-domain scenarios created by service providers, with the purpose of supporting their final client SLAs. An GSLA takes into account traffic aggregates passing from one domain to another. In general, there is no direct relationship between DSLA and GSLA [CAM 01].

TEQUILA approach handles the support of multi-domain SLAs by supposing that the service negotiation process is composed of a service subscription and a service invocation phase. The following scenarios are possible in the negotiation of inter-domain SLS [CAM 01]:

– hop-by-hop SLS negotiation (where a jump is an autonomous system – AS);

– end-to-end SLS negotiation (with the pre-establishment of interAS conduits);

– local SLS negotiation.

Consequently, the SEQUIN project defines a simple SLA and an SLS architecture that conforms to its own end-to-end QoS development needs across multiple management domains.

The SLS can contain the following:

– two types of predefined SLS as described in the AQUILA approach:

- premium IP,

- IP+,

– a group of attributes as defined in TEQUILA and AQUILA;

– a group of predefined values for each of these attributes (by using the IETF's IP performance measures as metrics the values of which will be presented to the users as an assurance for the network's conformity with the current SLA);

– use of implementation directed by the data (flow) from a recursive service level negotiation between the transportation sub-domains, for the support of inter-domain SLAs, as described by CADENUS.

Finally, the SEQUIN project mentions that by comparing TEQUILA, AQUILA and CADENUS approaches to SLAs, we get the impression that they evolve from AQUILA's predefined static SLS to the entirely specified SLS creation architecture in TEQUILA and the most dynamic CADENUS aspect, thus making DSC possible [CAM 01].

11.9.5. *The EURESCOM project*

Founded in 1991 by 20 European network operators, EURESCOM (European Institute for Research and Strategic Studies in Telecommunications) is an institution leader in collaborative telecommunications R&D [EUR 01]. Some of the projects developed by EURESCOM have great importance in defining SLAs. Some of these projects are mentioned below.

11.9.5.1. *The QUASIMODO project*

Indicated as QoS MethODOlogies and solutions within the service framework: measuring managing and charging QoS.

The QUASIMODO project started in 1998 with a first objective of proposing a QoS model that considers user level QoS classes and network-level network performance parameters. Its second objective was to execute tests with a few significant services and applications with the purpose of finding the correlation between QoS classes and performance parameters [KRA 01].

Among the most pertinent results of this project, we can consider the development of a QoS model. The principal characteristics of this mode are summarized as follows [KRA 01]:

– a group of application categories is considered (this makes the model conscious of the applications);

– the users can choose a quality class. These classes are created such that a user will always see the difference between one class and another;

– a group of network performance parameters is to be managed in order to provide the required quality class to the user for a certain application category.

Laboratory tests (with experts) and acceptability tests (with real users) were carried out on certain significant applications belonging to different application categories (principally VoIP, Streaming and e-commerce on the Web). These tests were carried out to find the correlation between quality classes and network performance levels. The results of these tests show that, for each application category, it is possible to create two quality classes such that they present different network performance values (in terms of delay, jitter and loss) and are perceived differently by the user [KRA 01].

The project's results do not directly contribute to the definition of SLA, but it is interesting and useful to consider the conclusions concerning different quality classes being effectively perceived by the users.

11.9.5.2. *The "interoperator interfaces for ensuring end-to-end IP QoS" project*

This project started in 2000 with the intention of supporting the interests of European operations in the managed interaction of networks and IP services, especially in the framework of QoS. In the goal of perceiving QoS, end-to-end QoS must be supported. If the service is offered through many operator domains, these operators must cooperate to ensure that the client requirements are achieved. The processes of inter-domain management, interfaces and models are necessary for the support of this cooperation [BRU 01].

The main objectives of this project are [BRU 01]:

– to understand the new management requirements presented by the end-to-end IP services and to link these requirements to existing works and standards;

– to produce specifications for management processes, models and interfaces required to support the assurance of end-to-end IP QoS;

– to detect network performance monitoring requirements and parameter measuring of services required for the support of these processes, interfaces and models.

The results of this project are published in a group of documents, which are very precious for the project's objectives. A document called *Measurement of Performance Metrics and Service Events* is certainly interesting for defining and implementing SLAs. The goal of this document is to identify key-service performance and event metrics that need to be controlled for supporting IP QoS inter-domain management and to identify the different abstraction views and levels required by the users and receivers of this information [BRU 01].

The document skips through the traffic profile and the service performance metrics that the vendor and the buyer of quality IP services can use to verify the accomplishment of the SLA. Many primary metrics have been described by using the IETF's and ITU-T's contemporary definition. The metrics are derived from transit and end-to-end services, including the definitions of QoS availability.

11.9.6. *The Internet2 – Qbone project*

Internet2 – Qbone is a consortium that receives a network of QoS test in the Internet. This group defines the concept of Bandwidth Broker. From the point of view of Internet2 – Qbone, policy control, policy admission control, AAA functions, network management functions, intra- and inter-domain routing are affected by or affect the bandwidth broker.

The project's goal is not to define a general SLA or a standard detailed SLS. Its first goal is to attempt to provide Premium Service based on the EF PHB of the DiffServ architecture.

Among the parameters agreed upon by the provider and its customer for a Premium Service, we include:

– the start and end time of the service;

– the source and the destination;

– the size of the MTU;

– the peak flow.

The Premium Service guarantees are:

– no loss caused by congestion;

– the delay is not guaranteed;

– the IP Delay Variation is guaranteed on the condition that the IP route is not changed.

Premium Service is unidirectional and the excess traffic is rejected.

In this project, we suppose that the SLA/SLS between two adjacent domains (peer domains) is statically established and configured. Then, the bandwidth broker receives requests that require a Premium service, called resource allocation requests. Based on these requests, the definition of Premium service, and the SLA/SLS between two domains, the bandwidth broker will deduct the necessary PHB and the traffic conditioning mechanism, and will configure the boarding routers.

More specifically, the SLS declares that the traffic of a certain class must follow certain traffic polishing conditions, enter the domain by a specific link, be handled with certain PHB(s) and whether the traffic's destination is in the domain that receives the traffic.

However, the TCS (Traffic Conditioning Specification) is defined as an entity that specifies the classifier rules, the traffic profiles, the marking, reject and modeling rules that are applicable to a traffic selected by a classifier. Here are some of the parameters that can be included in a TCS:

– performance: flow, reject probability, delay;

– span: domain entrance and exit points;

– traffic profile: token bucket parameters;

– how to handle excess traffic;

– marking;

– modeling;

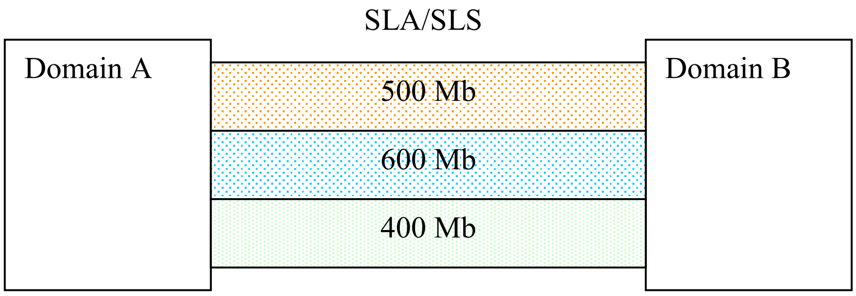
– mapping to a well-known DSCP service.

Internet2 – QBON distinguishes between SLA/SLS and resource reservation. SLA/SLS is a means of making it possible for a company to have a number of resources and resource reservation is the means for company members to obtain resources within the limit of available company resources. This idea is illustrated in Figure 11.6.

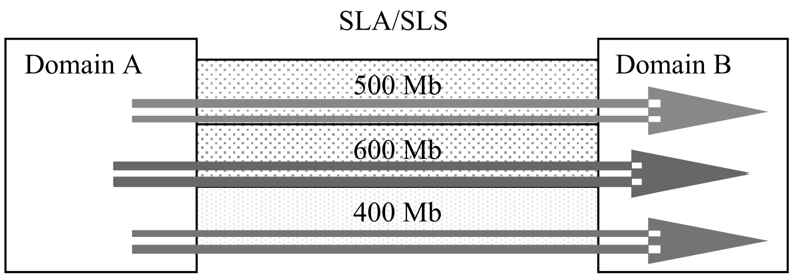
An SIBBS (Simple Inter Bandwidth Broker Signaling) protocol enables the discussion between two bandwidth brokers to fulfill dynamic resource reservation. SLS and resource reservation are both based on well-known services in the DiffServ environment.

11.9.7. *The ARCADE project (ARchitecture for Adaptive Control of IP Environment)*

The ARCADE project has reunited academics and R&D centers (LIP6, INRIA) and industries (France Télécom R&D, Thomson-CSF Communications, QoSMIC).



(a) SLA/SLS makes it possible for domain A to have:
500 Mb of type Resource 1
600 Mb of type Resource 2
400 Mb of type Resource 3
and domain A pays domain B for these resources



(b) Resource reservation makes it possible for domain A to dynamically assign its resources to its traffic

Figure 11.6. *SLA/SLS and resource reservation*

The goal of the ARCADE project [ARC 01] is to draw up a general model that makes it possible to control IP networks. This control is based on the determination of a profile for each user and client, with the aim of allowing it to communicate adapted resources. These resources can by dynamically allocated. The control is carried out on the security, mobility and QoS. The part of the architecture that was conceived and developed in this project concerns the policy servers and the definition of an intelligent interface (an extension of the COPS protocol) between the policy server and the IP network nodes. This protocol is called COPS-SLS.

COPS-SLS is a protocol defined by LIP6, ENST and Alcatel [NGU 02] for the dynamic negotiation of a level of service. The idea of COPS-SLS is to define a means of communication between the client and the network and also between networks to establish an end-to-end QoS through a policy management system. The result of the SLS negotiation will be implemented by PDPs (Policy Decision Points) that take care of resource allocation in the concerned domains.

COPS-SLS can be used to dynamically provide an intra- or inter-domain SLA while dynamically establishing an end-to-end QoS.

The SLS parameters used in COPS-SLS are based on the parameters defined in TEQUILA, apart from the generally static parameters such as reliability. In addition, COPS-SLS makes it possible to negotiate predefined services (for example, well-known services) as well as unpredefined services. The rest of this part presents defined and potentially defined parameters in COPS-SLS.

The QoS level in a data flow is expressed in terms of:

– service ID: this parameter can be used for a well-known service or a defined service in the predefined SLS mode. When the Service ID is used, it is not necessary to specify duration, jitter and packet loss parameters in the message;

– delay: the duration required in the negotiation phase;

– jitter: the jitter required in the negotiation phase;

– packet loss: the packet loss rate required in the negotiation phase.

These parameters can be gathered from the applications that are QoS conscious. For example, the size of the playback memory in streaming applications can define duration and jitter needs.

Traffic is characterized by different types of rates and parameters according to the conformance algorithm and the QoS level required by the data flow:

– conformance algorithm: simple token bucket, tri-color token bucket, etc.;

– peak flow;

– average flow;

– maximum flow;

– maximum burst size;

– excess handling.

The extent of negotiations identifies the starting point and the end of the QoS guarantee. It makes it possible for a domain to know if it is final or not.

– Starting point: [IP address(es)] the point(s) from which traffic needs a service level guarantee.

– End point: [IP address(es)] the point(s) where the service level guarantee ends.

The identification of flows defines parameters for recognizing the flow at a domain's entrance. These parameters depend on the type of data flow for which the client wants to negotiate a service: IP address, source sub-network's mask length, port numbers, protocol number, DSCP, MPLS label, IPv6 flow label.

The service calendar specifies the start and end time of the SLS: hour, date, etc.

Each COPS message can transport more than one SLS. Each SLS is identified by an SLS ID. The client-handle value is used to identify different requests sent by the SLS client.

11.10. Conclusion

Significant steps in SLA design have been completed by TEQUILA, AQUILA, CADENUS, SEQUIN and ARCADE. Their results represent a good basis for defining any type of network SLA.

It is currently possible to define flexible and more dynamic SLAs addressing certain changes in one or more parts, notably the SLS. The SLA no longer concerns a single domain but several domains. This is how the inter-domain SLA was born, which created an end-to-end reflection involving a chain of SLAs. Each can be dynamically negotiated via specific negotiation protocols.

11.11. Abbreviations and acronyms

AAA	Authentication, Authorization, Accounting
CBR	Constant Bit Rate
CL	Client
COPS	Common Open Policy Service
COPS-SLS	Common Open Policy Server – Service Level Specification
DiffServ	Differentiated Services
DSC	Dynamic Service Creation
DSCP	Differentiated Service Code Point

DSLA	Detailed SLA
ENST	Ecole National Supérieure des Télécommunications
e-Service	Electronic Service
FC	Final Client (end user)
GSLA	Global SLA
ID	Identifier
IETF	Internet Engineering Task Force
IP	Internet Protocol
ISP	Internet Service Provider
ITU-T	International Telecommunication Union – Telecommunication Standardization Sector
LIP6	Informatics Laboratory, University Paris 6
MPLS	Multi-Protocol Label Switching
MTU	Maximum Transport Unit
PBM	Policy-Based Management
PHB	Per Hop Behavior
QoS	Quality of Service
SDNM	Service Driven Network Management
SLA	Service Level Agreement
SLO	Service Level Objective
SLS	Service Level Specifications
SP	Service Provider
TCS	Traffic Conditioning Specification
VBR	Variable Bit Rate
VoD	Video on Demand
VoIP	Voice over IP

11.12. Bibliography

[ARC 01] http://www-rp.lip6.fr/arcade.

[BLA 98] BLAKE S., BLACK D., CARLSON M., DAVIES E., WANG Z., WEISS W., "An Architecture for Differentiated Services", *Request for Comments: 2475*, December 1998.

[BRU 01] BRÜGGEMANN H., "Inter-operator Interfaces for Ensuring End to End IP QoS", EURESCOM P1008 project, http://www.eurescom.de/public/projects/P1000-series/p1008/default.asp.

[CAM 01] CAMPANELLA M., CHIVALIER P., SEVASTI A., SIMAR N., SEQUIN "Deliverable 2.1-Quality of Service Definition", http://www.dante.net/sequin/QoS-def-Apr01.pdf.

[CIM 99] Common Information Model (CIM) Specification, version 2.2, DMTF, June 1999.

[DAN 03] DAN A., LUDWIG H., PACIFICI G., "Web Services Differentiation with Service Level Agreements", *IBM Software Group Web Services Web site*, 2003.

[EUR 01] EURESCOM home page, http://www.eurescom.de.

[FAW 04] FAWAZ W., DAHEB B., AUDOUIN O., BERDE B., VIGOUREUX M., PUJOLLE G., "Service Level Agreement and Provisioning in Hybrid Photonic Networks", *IEEE Communications Magazine Special issue on "Management Of Optical Networks"*, January 2004.

[GOD 01] GODERIS D., T'JOENS Y., JACQUENET C., MEMENIOS G., PAVLOU G., EGAN R., GRIFFIN D., GEORGATOS P., GEORGIADIS L., VAN HEUVEN P., "Service Level Specification Semantics and Parameters", *Internet Draft*, *draft-tequila-sls-01.txt,* work in progress, June 2001.

[KRA 01] KRAMPELL M., "QUASIMODO – QUAlity of ServIce MethODOlogies and Solutions Within the Service Framework: Measuring, Managing and Charging QoS", EURESCOM P906 project, http://www.eurescom.de/public/projects/P900-series/p906/default.asp.

[MAR 02] MARILLY E. *et al.*, "Requirements for Service Level Agreement Management", *IEEE*, 2002.

[MAR 02] MARTIN J., NILSSON A., "On Service Level Agreements for IP Networks", *IEEE INFOCOM 2002*, 2002.

[MEL 01] MELIN J. L., "*Qualité de Service sur IP*", Editions Eyrolles, 2001.

[NGU 02] NGUYEN T. M. T. *et al.*, "COPS Usage for SLS Negotiation", *Internet Draft*, <draft-nguyen-rap-cops-sls-03.txt>, July 2002.

[RAI 03] RÄISÄNEN V., *Implementing Service Quality in IP Networks*, John Wiley & Sons, 2003.

[ROM 00] ROMANO S. P., ESPOSITO M., VENTRE G., CORTESE G., "Service Level Agreements for Premium IP Networks", *Internet Draft*, draft-cadenus-sla-00.txt, November 2000.

[SAL 00] SALSANO S., RICCIATO F., WINTER M., EICHLER G., THOMAS A., FUENFSTUECK F., ZIEGLER T., BRANDAUER C., "Definition and Usage of SLSs in the AQUILA Consortium", *Internet Draft*, draft-salsano-aquila-sls-00.txt, November 2000.

[TMF 01] TMF, "SLA Management Handbook", *Public Evaluation 1.5*, June 2001.