INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

E.860

(06/2002)

SERIES E: OVERALL NETWORK OPERATION, TELEPHONE SERVICE, SERVICE OPERATION AND HUMAN FACTORS

Quality of telecommunication services: concepts, models, objectives and dependability planning – Use of quality of service objectives for planning of telecommunication networks

---

Framework of a service level agreement

ITU-T Recommendation E.860

ITU-T E-SERIES RECOMMENDATIONS

OVERALL NETWORK OPERATION, TELEPHONE SERVICE, SERVICE OPERATION AND HUMAN FACTORS

INTERNATIONAL OPERATION	
Definitions	E.100–E.103
General provisions concerning Administrations	E.104–E.119
General provisions concerning users	E.120–E.139
Operation of international telephone services	E.140–E.159
Numbering plan of the international telephone service	E.160–E.169
International routing plan	E.170–E.179
Tones in national signalling systems	E.180–E.189
Numbering plan of the international telephone service	E.190–E.199
Maritime mobile service and public land mobile service	E.200–E.229
OPERATIONAL PROVISIONS RELATING TO CHARGING AND ACCOUNTING IN THE INTERNATIONAL TELEPHONE SERVICE	
Charging in the international telephone service	E.230–E.249
Measuring and recording call durations for accounting purposes	E.260–E.269
UTILIZATION OF THE INTERNATIONAL TELEPHONE NETWORK FOR NON-TELEPHONY APPLICATIONS	
General	E.300–E.319
Phototelegraphy	E.320–E.329
ISDN PROVISIONS CONCERNING USERS	E.330–E.349
INTERNATIONAL ROUTING PLAN	E.350–E.399
NETWORK MANAGEMENT	
International service statistics	E.400–E.409
International network management	E.410–E.419
Checking the quality of the international telephone service	E.420–E.489
TRAFFIC ENGINEERING	
Measurement and recording of traffic	E.490–E.505
Forecasting of traffic	E.506–E.509
Determination of the number of circuits in manual operation	E.510–E.519
Determination of the number of circuits in automatic and semi-automatic operation	E.520–E.539
Grade of service	E.540–E.599
Definitions	E.600–E.649
Traffic engineering for IP-networks	E.650–E.699
ISDN traffic engineering	E.700–E.749
Mobile network traffic engineering	E.750–E.799
QUALITY OF TELECOMMUNICATION SERVICES: CONCEPTS, MODELS, OBJECTIVES AND DEPENDABILITY PLANNING	
Terms and definitions related to the quality of telecommunication services	E.800–E.809
Models for telecommunication services	E.810–E.844
Objectives for quality of service and related concepts of telecommunication services	E.845–E.859
Use of quality of service objectives for planning of telecommunication networks	**E.860–E.879**
Field data collection and evaluation on the performance of equipment, networks and services	E.880–E.899

For further details, please refer to the list of ITU-T Recommendations.

ITU-T Recommendation E.860

Framework of a service level agreement

Summary

The liberalization and deregulation process that started during the last decade in the telecommunication's environment is still running and it is cause of meaningful changes. Increasing competition, favoured also by customer performance requirements, produces big pressures upon service/network providers. The latter, after having faced especially cost reductions for several years, nowadays try to improve quality of service (QoS) in order to differentiate their products from those of their competitors.

In addition the situation is complicated by the increasing demand of global services which involve in their provisioning several service/network providers. Therefore, roles of all entities that take part to service provision and their relationships have to be described. The scope is to state responsibilities of each provider and to assure quality of service required from customer.

A useful tool in formalizing the mentioned inter-relationships between entities is *Service Level Agreement* (SLA), that is the result of a negotiation between two or more parties with the objective of reaching a common understanding about service delivered, its quality, responsibilities, priorities, etc.

This Recommendation describes a generic structure of SLA adopting an approach independent from the type of service and the technology used. That approach is useful especially in a multi-provider environment which today has become reality.

Source

ITU-T Recommendation E.860 was prepared by ITU-T Study Group 2 (2001-2004) and approved under the WTSA Resolution 1 procedure on 29 June 2002.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU [had/had not] received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

CONTENTS

ITU-T Recommendation E.860

Framework of a service level agreement

1 Introduction

The liberalization and deregulation process that started during the last decade in the telecommunication's environment is still running and it is cause of meaningful changes.

Increasing competition, favoured also by customer performance requirements, produces big pressures upon service/network providers. The latter, after having faced especially cost reductions for several years, nowadays try to improve quality of service (QoS) in order to differentiate their products from those of their competitors.

In addition the situation is complicated by the increasing demand of global services which involve in their provisioning several service/network providers.

Therefore, roles of all entities that take part to service provision and their relationships have to be described. The scope is to state responsibilities of each provider and to assure quality of service required from customer.

A useful tool in formalizing the mentioned inter-relationships between entities is *Service Level Agreement* (SLA), that is the result of a negotiation between two or more parties with the objective of reaching a common understanding about the service delivered, its quality, responsibilities, priorities, etc.

In the following a generic structure of SLA is described adopting an approach independent from the type of service and the technology used. That approach is useful especially in a multi-provider environment which today has become reality.

Therefore we begin with recalling QoS terms and definitions (clause 2) and the one stop responsibility concept (clause 4).

Afterwards the structure of a SLA is described in all its components (clause 5) and its application in a multi-provider environment is illustrated in clause 6.

2 Quality of Service definitions and terms

A standardization of terms and definitions in QoS is important for two main reasons:

- to avoid the confusion introduced by contrasting terms and definitions;
- to maintain the consistency between different groups involved in developing telecommunication standards.

In this clause terms and definitions for a QoS Framework are introduced, referring to [E.800] for the other QoS terms.

2.1 Entity

An *Entity* is a generic unit involved in using/delivering a service. It is characterized by its states and its transitions from a state to another (Figure 2-1). During a transition an entity can execute functions and interact with other entities through its outputs.

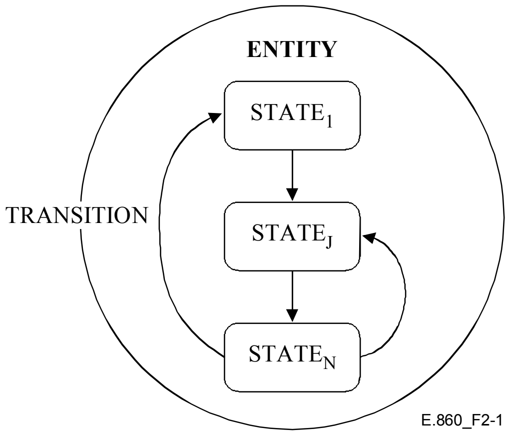


Figure 2-1/E.860 – States and transitions of an entity

An entity that delivers a service to another entity is called **provider** while the entity that receives the service is called **user**. The term user may indicate either an end user, a regulatory authority or a service provider. The latter receives a service from another service provider.

2.2 Interaction Points and Interfaces

An *Interaction Point* is a point where two entities can exchange information.

A group of interaction points at the logical boundary between two entities constitutes an *interface* (Figure 2-2).

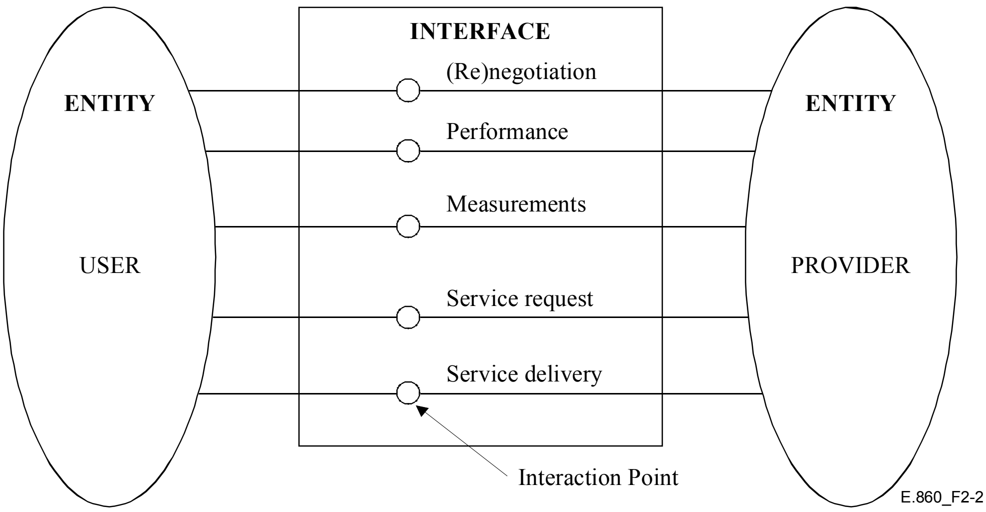


Figure 2-2/E.860 – Provider, user, service, interface, interaction points

Sometimes an interaction point between user and provider may not belong to their logical interface although this point remains under the control of the provider.

Interaction points can be located either on horizontal interface (between domains of the same functional level) or on vertical interface (between domains of different functional level) (Figure 2-3).

It is often useful to group more entities together in one; obviously the correspondent interfaces will be redefined in a convenient way.

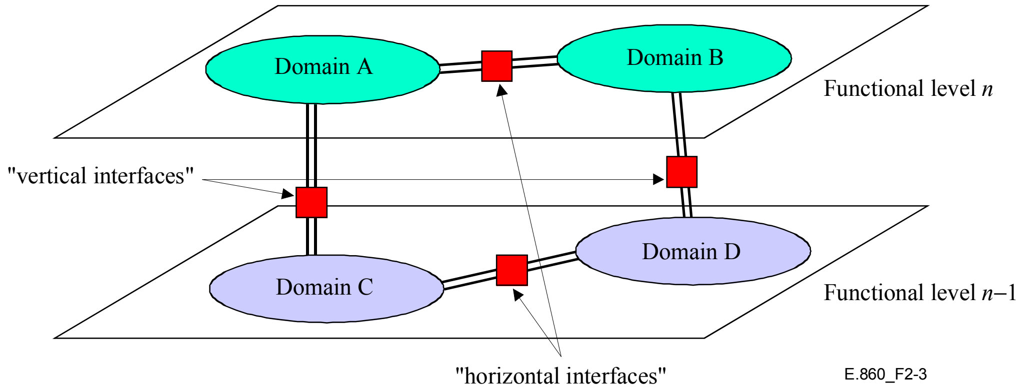
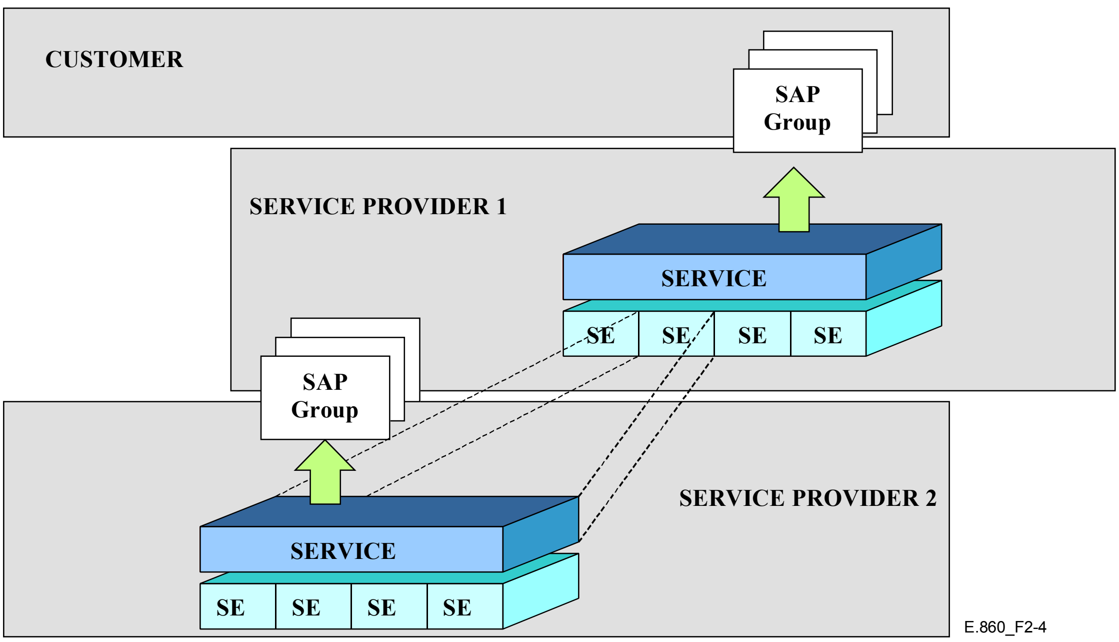


Figure 2-3/E.860 – "Vertical" and "horizontal" interfaces

2.3 Service, Service Element and Service Access Point

A *service* is a group of functions provided by an organization to a user through an interface [E.800].

The interaction points located in the interface between the SP's domain and the user's domain are called *Service Access Point* (SAP) and represent the points where service is delivered. SAP's definition is very important because all service components located between the SAPs in the SLA are under the SP's responsibility. In performance reporting, more SAPs may often be grouped in a *SAP Group*. The concept of layered architecture may be applied also in service definition and in performance estimation (e.g., Service Availability delivered from a provider to a user in correspondence of a SAP) (Figure 2-4).



SE Service Element
SAP Service Access Point

Figure 2-4/E.860 – Service composition

The example in Figure 2-4 shows a customer who buys a service from SP1 stipulating an agreement which contains responsibilities and priorities related to quality of service; in respect of this agreement performance parameters have to be provided in correspondence of SAP group by the provider to the customer.

To deliver the service, provider 1 combines more *service elements* (SE) which may be at its own disposal or bought from other providers.

Obviously, to reach the QoS level contracted with the customer, SP1 has to demand a suitable QoS from SP2 which delivers the SE used.

2.4 Quality of Service

The *Quality of Service* definition in [E.800] refers to *"the collective effect of service performance which determines the degree of satisfaction of a user of the service".*

However, considering the latest developments, in order to be compliant with E.800 and to have a useable definition in a contract, we will describe QoS as the *"degree of conformance of the service delivered to a user by a provider in accordance with an agreement between them",* a definition that can be considered as a measurable subset of the [E.800].

The latter definition, in fact, is more market-oriented even if QoS is measured from the customer's point of view in both definitions. This is, in fact, the factor that finally produces success or failure of the service.

2.5 Service Level Agreement (SLA)

A *Service Level Agreement* is a formal agreement between two or more entities that is reached after a negotiating activity with the scope to assess service characteristics, responsibilities and priorities of every part.

A SLA may include statements about performance, tariffing and billing, service delivery and compensations.

Every performance reporting may include only the QoS parameters agreed in the correspondent SLA.

2.6 Relationships between Quality of Service and Network Performance

The overall quality of a telecommunication service, as perceived from customer's point of view, is influenced by many factors which are correlated with network performance parameters.

Figure 2-5 shows such a relationship:

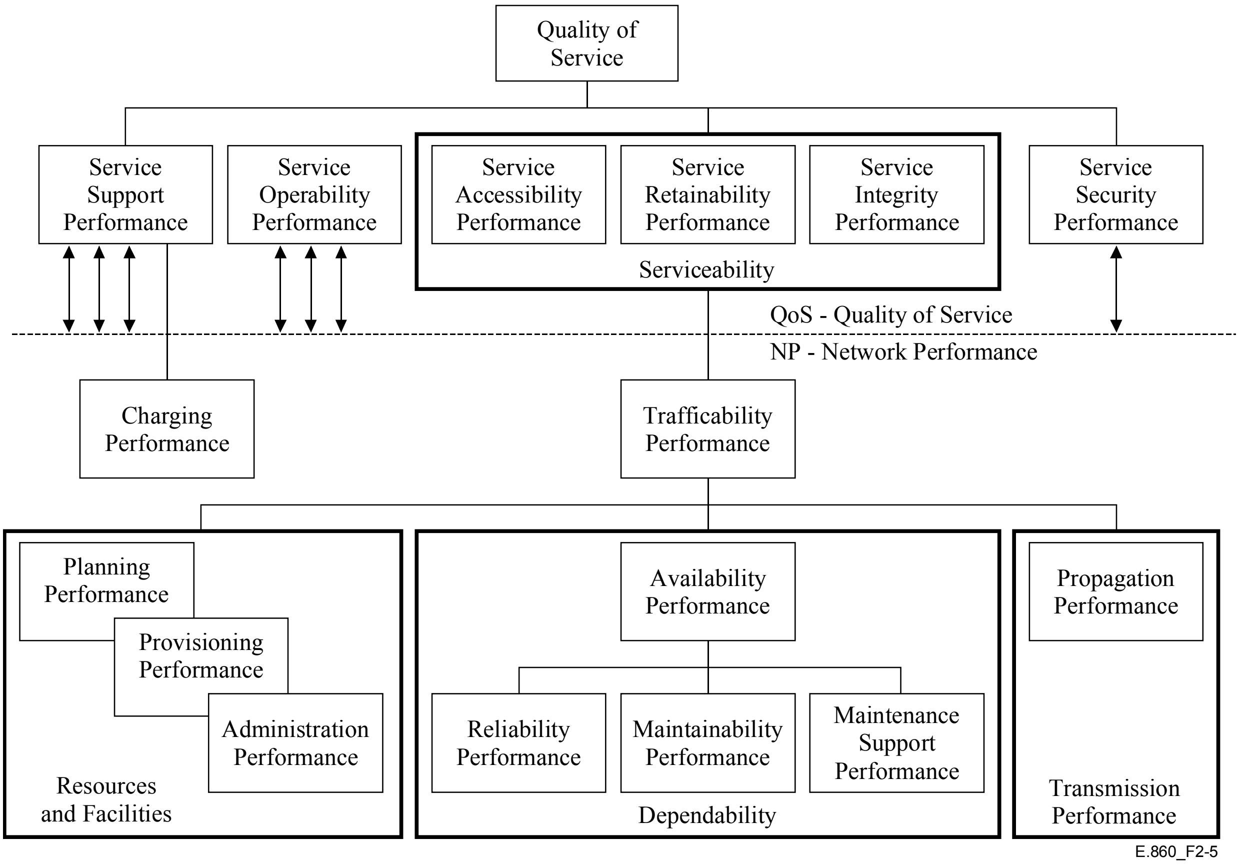


Figure 2-5/E.860 – QoS parameters organization, from [E.800]

The essential aspect, while estimating a service, is customer's opinion, and so its degree of satisfaction with the provider that is the only entity it directly interacts with. This level of satisfaction comes from the perception of the various service aspects (support, operability, serviceability, security), which are influenced by network characteristics.

For definitions of the concepts in Figure 2-5 see [E.800].

The scope of this Recommendation does not take into account the terminal performances, but it is referred exclusively to the end-to-end network quality as perceived from the final user.

3 Abbreviations

This Recommendation uses the following abbreviations:

BI	Business Interface
ITU-T	International Telecommunication Union – Telecommunication Standardization Sector
QoS	Quality of Service
SA	Service Availability
SAP	Service Access Point
SDF	Service Degradation Factor
SE	Service Element

SLA	Service Level Agreement
TI	Technical Interface
UA	UnAvailability

4 Approach to multi-provider environment

In a multi-provider environment, relationships that exist between SPs may be very complex.

Indeed a SP (primary), which wants to deliver a service to a customer, often uses service elements provided by other SPs, and consequently it becomes much more complex to assure QoS level stated in the SLA. Therefore it is necessary to define responsibilities of all the entities involved in service delivery and, above all, to co-ordinate all activities to reach the agreed QoS levels.

In order to simplify the resolution of the problem mentioned above, we can apply the concept of *"one stop responsibility"*[1].

The one stop responsibility concept is based on the SLA stipulated between two entities and, especially on QoS conditions. This latter part is called *QoS Agreement or Service Quality Agreement (SQA)*[2].

The possible content of a SLA is illustrated in Figure 4-1:

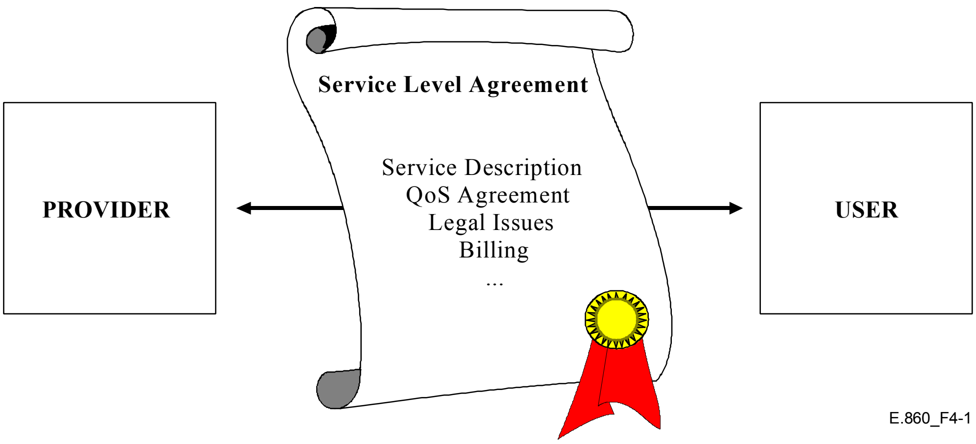


Figure 4-1/E.860 – One stop responsibility and Service Level Agreement

The one stop responsibility, agreed by a provider to a customer within a SLA, allows a user to retain a primary service provider, (with whom he agreed on the SLA) as the only responsible for the overall QoS received. In its turn, the primary provider, since occurring problems depend on services received by other SPs, can apply the same one stop responsibility to its sub-providers.

By applying the one stop responsibility in a recursive manner to all entities (provider and sub-providers) which take part in service provision (Figure 4-2), the service agreed with end user is guaranteed.

---

[1] From EURESCOM [P806-GI].

[2] This is the denomination used in [E.800].

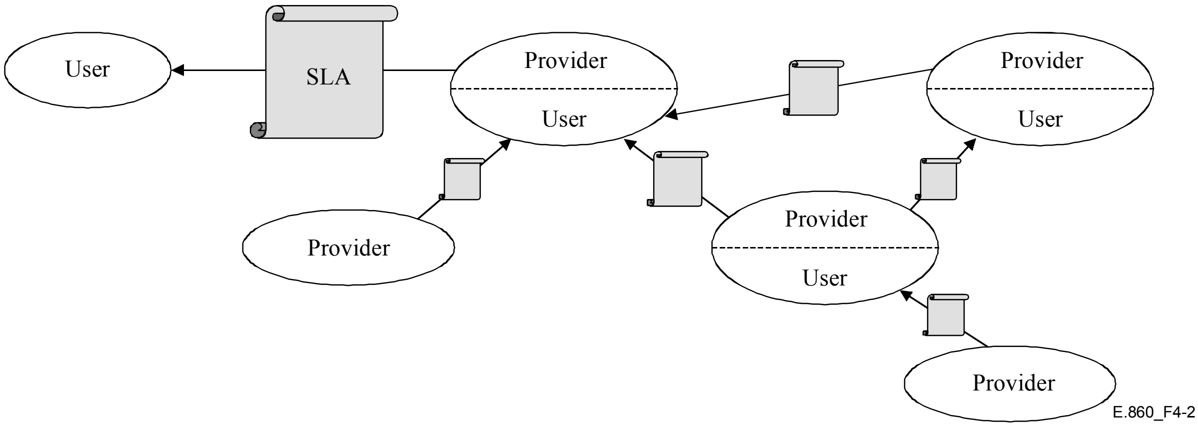


Figure 4-2/E.860 – Application of the one stop responsibility concept

So, thanks to the one stop responsibility and the recursive application of SLAs, the complex problem of service provisioning in a multi-provider environment is decomposed into elementary relationships between only two entities (user/provider pair).

However, this implies that primary SP has to have QoS flexibility with its customers because quality of service elements (from its sub-providers) may oscillate within the agreed ranges.

5 Service Level Agreement (SLA)

5.1 What is a SLA

A *Service Level Agreement* is a formal agreement between two or more entities that is reached after a negotiating activity with the scope to assess service characteristics, responsibilities and priorities of every part.

A SLA may include statements about performance, billing, service delivery but also legal and economic issues.

The part of SLA which refers to QoS is called *QoS Agreement* and includes formal program agreed between two entities to monitoring, measuring and deciding QoS parameters. The goal is to reach the QoS agreed with end user and then obtain its satisfaction.

In the following clauses a structure of SLA is proposed and all its components are described, in particular those of QoS Agreement.

5.2 Structure of a SLA

The generic structure of a Service Level Agreement is illustrated in Figure 5-1:

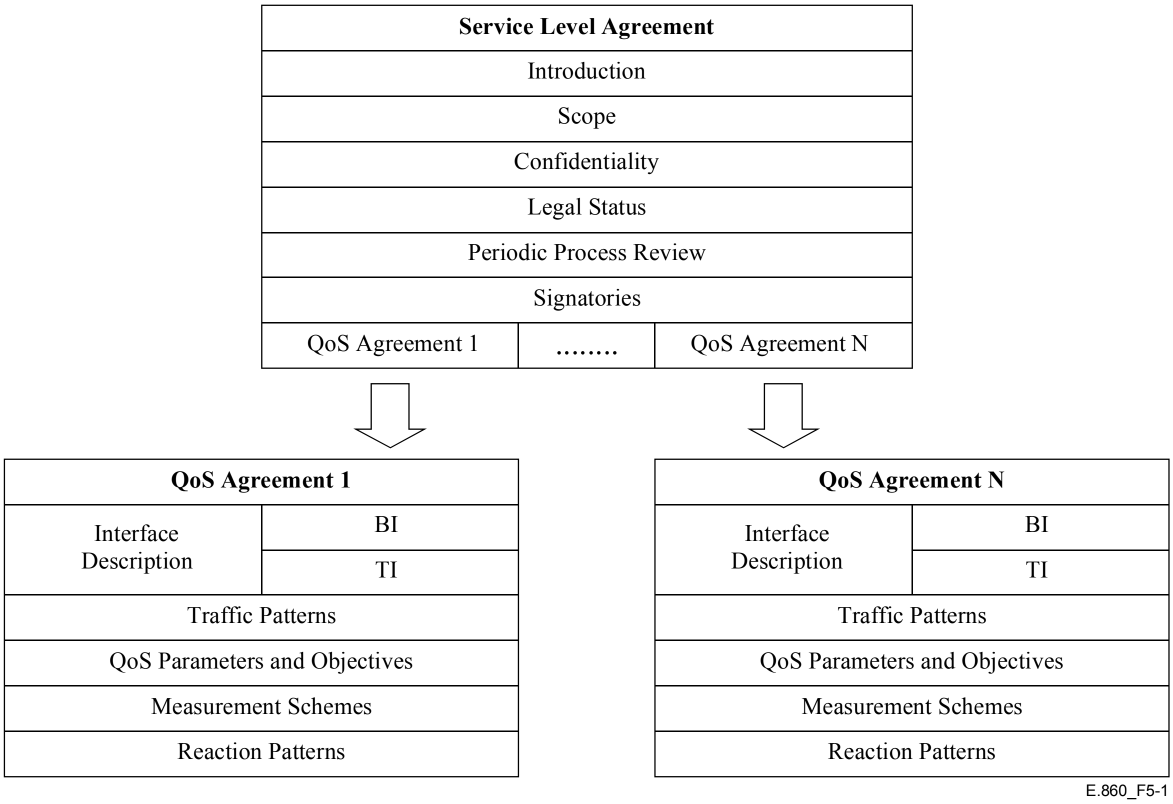


Figure 5-1/E.860 – Generic structure of a Service Level Agreement

As represented in Figure 5-1, a SLA refers to all services exchanged between two entities (multi-services SLA) and it is made up of one common part and of other service specific parts.

This approach avoids repetitions and simplifies the addition of new services within the SLA.

5.2.1 Introduction

It describes the purpose of the SLA that may be:

– to define service levels that all entities have to guarantee for customer's satisfaction;

– to assist two entities (User, Service Provider, Network Provider) in exchanging information with suitable QoS and Network Performance;

– to provide base notions of measurements and parameters for realization of the agreement.

5.2.2 Scope

It describes, in a general manner, services which the SLA deals with and their target performance.

5.2.3 Confidentiality

It specifies the treatment of the agreement and the sharing of information between the involved parties. It is in the interest of all parties that confidential information not be disclosed to entities which are not part of the agreement (e.g., a SP which is a competitor in the same market). This does not apply to public SLA signed with Regulatory authority.

5.2.4 Review process

It defines the frequency (daily, monthly, semi annual, etc.) and format (paper, electronic) with which QoS information has to be exchanged. It may specify also the frequency of review of the QoS Agreement, so it can be always up to date with actual technology and customer's expectations. This part may be optional.

5.2.5 Compensations

A SLA may include the compensations for an unreached level of quality as an economic issue of the contract.

5.2.6 Signatories

Authorized representatives of all parties should sign the agreement to ensure all obligations undertaken.

5.3 QoS agreement

5.3.1 Interface description

An interface is the logical boundary between two entities and it is composed of a group of interaction points. These points are always near the user's domain and enable it to exchange information with the service provider who, at least virtually, controls all interaction points.

Regarding the type of information exchanged, interface description is categorized as:

– Business Interface (BI); or

– Technical Interface (TI).

5.3.1.1 Business Interface (BI)

It is composed of interaction points always located between user and SP. These are used for specific QoS Agreement functions as well as (re)negotiation, performance reporting and reaction patterns which are triggered when the agreed QoS level is not provided (see 5.3.5).

5.3.1.2 Technical Interface (TI)

Its interaction points exchange service specific information and allow measurements from which QoS parameters are derived. Sometimes points may not regard primary SP if (a part of) service was delegated to a sub-SP.

Figures 5-2 and 5-3 help us to better understand the difference between technical (TI) and business (BI) points of view:

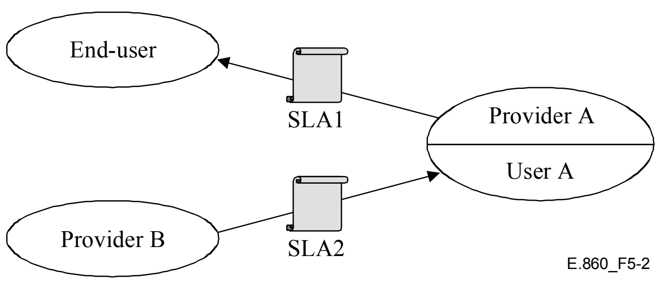


Figure 5-2/E.860 – Business relationships between entities

Figure 5-2 illustrates business relationships.

In this case end user negotiates an agreement (SLA1) with SP A for a service provision; on its side SP A buys one or more service elements from SP B and agrees SLA2 with the latter.

Naturally every SLA implies that a BI between the correspondent entities (user-provider pair) must exist.

Let us see now the situation from a technical point of view.

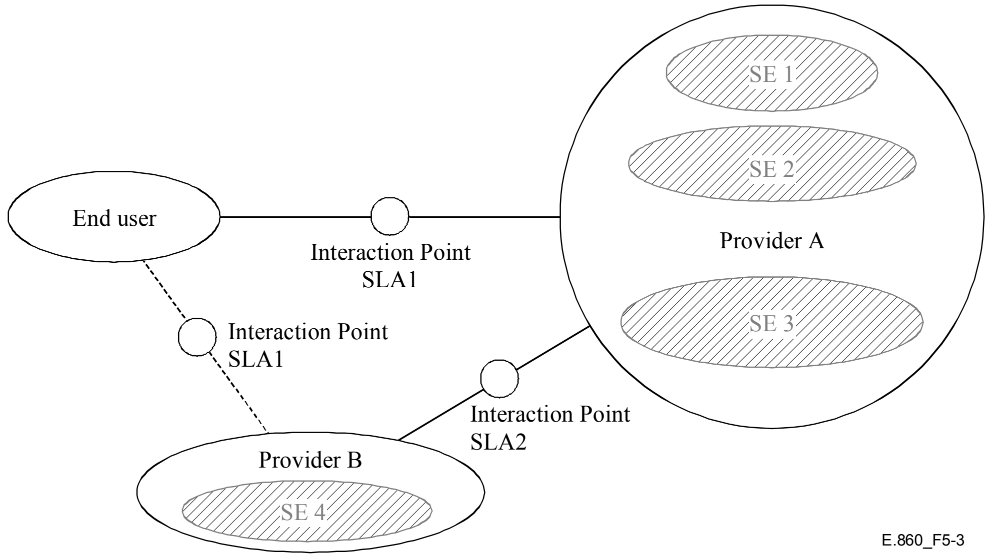


Figure 5-3/E.860 – Technical interaction points of SLA1 and SLA2

In Figure 5-3 interaction points of TIs relating to SLA1 and SLA2 are illustrated.

As regards SLA1, one can notice the presence of two interaction points: one between user and SP A and another between user and SP B.

The interaction between user and SP B refers to an indirect relationship since there is no business agreement between the two entities.

This means that the customer has to address SP A for all complaints dealing with service, even though he received part of it from SP B.

For this reason, indirect relationships with sub-providers are often hidden to customers.

5.3.2 Traffic patterns

In order to manage its own resources properly, every entity must know the characteristics of traffic that it receives from other entities (traffic at the ingress points). If we also consider that outgoing traffic for one entity is incoming traffic for another entity, then the statement "QoS Agreement must include description of all traffic exchanged" is easily justified. This is true both for application and management flows.

Also the conditions (thresholds) that enable activation of reaction patterns from the receiving entity have to be specified. In such a manner, when incoming traffic not conforms the agreed one, the receiving entity may react with mechanism as well as traffic shaping.

Finally description of traffic patterns should be well understood from entities at the both sides of the interface. In fact, only in this way every eventual reaction is simply justified to the penalized entity.

5.3.3 QoS parameters and objectives

Definition of QoS parameters is an essential moment in developing a SLA and, in particular, its correspondent QoS Agreement. Indeed behaviour of all entities which take part to QoS Agreement is influenced by those parameters, so every point of view must be considered in their definition.

For the same reason, once defined, the QoS parameters have to be expressed in a clear and convenient way, with a simple language for end user and a more technical one for providers.

A way to individuate QoS parameters in an interconnection between two entities may be the one proposed in [E.801] and reported in Figure 5-4, where both customer and network parameters are taken into consideration.

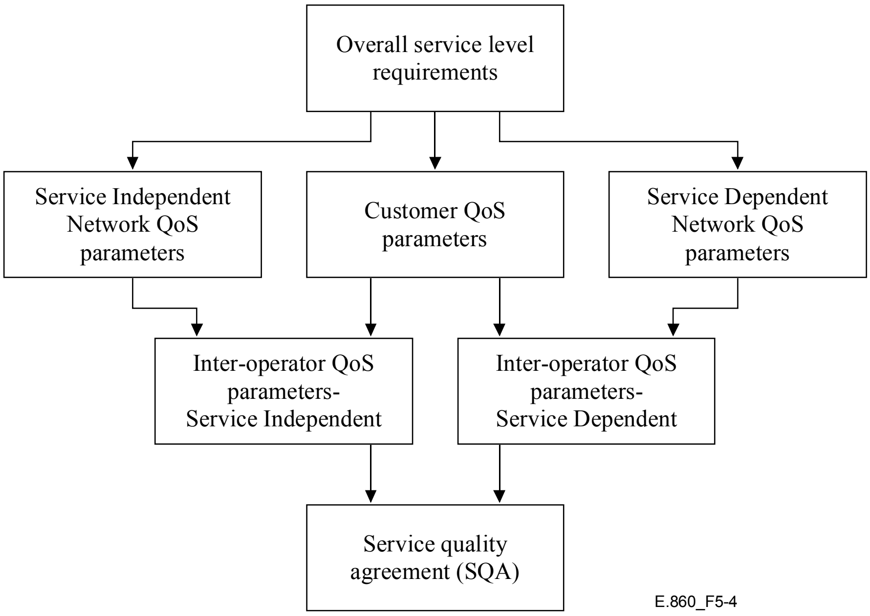


Figure 5-4/E.860 – Determination of QoS parameters, from [E.801]

The categorization of parameters into service dependent and service independent parameters allows to determine the last ones in a common manner for all services, which simplifies their utilization.

5.3.3.1 QoS parameter classification

QoS is assessed by assigning proper values to QoS parameters.

We have already seen in 2.5 that quality of service is influenced, in a more or less direct manner, by network performance. This observation leads us to classify QoS parameters as *direct* and *indirect*, where:

- A *direct* parameter refers to a specific service element and is determined collecting direct observations of event in correspondence of its interaction points.
- A *indirect* parameter is defined as a function of other direct parameters.

Events considered in definitions and measurements may refer to universal standards or be agreed by the parties in the SLA.

Once stated, QoS parameters may be algorithmically combined into a single index relating to the overall QoS in order to provide how close the service offered is to the contracted service.

A possible definition of such a quality index may be a weighted average of QoS parameters, where the single contribute would be represented by a weighting value which is agreed with the customer and specified in the SLA.

5.3.3.2 Timeline Model – Parameters

A telecommunication service is defined as a group of functions whose realisations are observed through direct or indirect analysis of the corresponding events.

If we consider all possible types of functions, a useful categorization of primary QoS parameters is the Timeline Model approach defined in ETSI and illustrated in Figure 5-5.

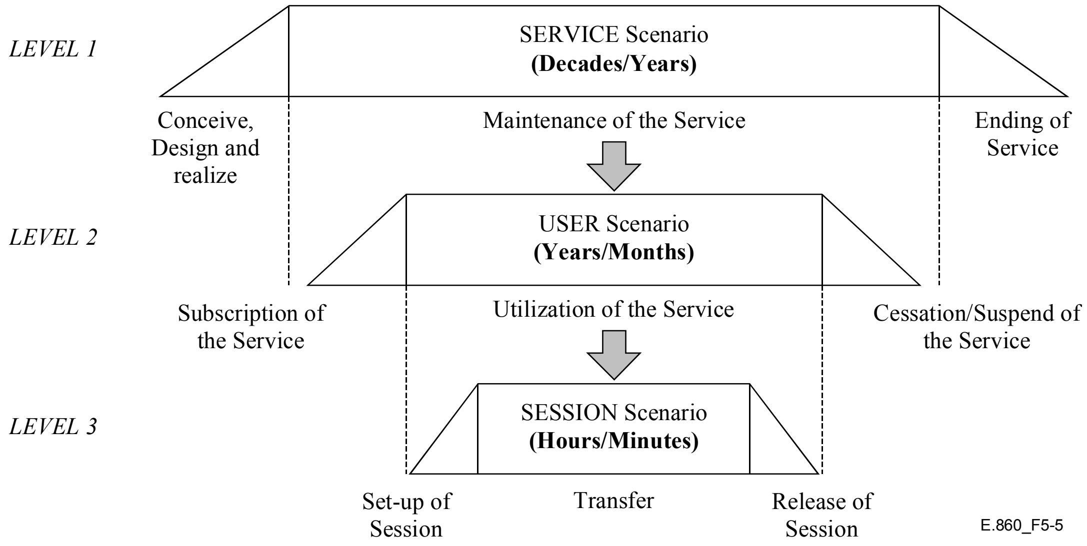


Figure 5-5/E.860 – QoS parameter levels in Timeline Model

The Timeline Model identifies three possible scenarios (levels) on the basis of the temporal scale and each scenario is then divided in three phases. So all the service delivery activities are included at the first level. In the second level, the focus is moved from the totality of the users to the service utilization of each single user, while in the third level the attention is focused on the single call.

Quality of Service depends on accuracy of the service functions done. As defined in [I.350], the quality of the function is valued on 3 criteria:

- Speed.
- Accuracy.
- Reliability.

All the parameters used to characterize the quality of the various phases above are then classified based on these criteria.

- **Speed**: It characterizes the aspects of temporal efficiency associated with a function. It is defined on measurements made on sets of time intervals.
- **Accuracy**: It characterizes the degree of correctness a particular function is realized with. This type of parameter is based on either the ratio of incorrect realizations on total attempts or the rate of incorrect realizations during an observation period.
- **Reliability**: It expresses the degree of certainty with which a function is performed. This type of parameter is based on either the ratio of failures on total attempts or the rate of failures during an observation period.

The results are some matrices PHASE/CRITERION in order to classify QoS parameters as shown in Table 5-1:

Table 5-1/E.860 – Classification of QoS parameters in Timeline Model

SERVICE Scenario:

CRITERION / PHASE	Reliability	Speed	Accuracy
Design, realization			
Maintenance			
Ending			

USER Scenario:

CRITERION / PHASE	Reliability	Speed	Accuracy
Subscription			
Utilization			
Cessation			

SESSION Scenario:

CRITERION / PHASE	Reliability	Speed	Accuracy
Set-up			
Transfer of information			
Release			

See Appendix I for examples of application.

Indirect Parameters are defined as functions either of values of primary QoS parameters or of decisions taken on the basis of the latter. A remarkable example is *Service Availability,* which is described in the following paragraphs.

Service Availability is a key parameter and of major interest for end user. So it has to be defined in a clear and convenient way in order to avoid every misunderstanding between the customer and the SP.

Service Availability refers to percentage of time (SA%) during which the contracted service is operational at the respective Service Access Points. The term "operational" means that the customer has the ability to use the service at the quality level specified in the SLA.

Service Availability is often derived from measurements of Service Unavailability (UA%) applying the formula:

$$SA\% = 100\% - UA\%$$

The expression that allows to calculate UA% is the following:

$$UA\% = \frac{\sum \text{outage interval}}{\text{Active time}} \times 100\%$$

In order to take into consideration complete service outage (service fully unavailable) and partial service outage (service degraded available), a Service Degradation Factor (SDF) is usually introduced for every outage interval, so the above formula becomes:

$$\text{UA\%} = \frac{\sum(\text{outage interval} \times \text{SDF})}{\text{Active time}} \times 100\%$$

where: $0 \leq \text{SDF} \leq 1$

A list of SDF values, with the corresponding types of events, may be added to the SLA.

5.3.3.3 QoS objectives

QoS objectives may be expressed by target values, thresholds and ranges set to QoS parameters.

Whenever this is the case, entities involved should also specify in the SLA if those values represent only indications or if, as they are not respected, reaction procedures will be activated.

Finally, since QoS objectives are closely related to measurements and reaction patterns, both measurement and reaction procedures should fit the granularity set to QoS objectives.

5.3.4 Measurements

Once QoS parameters and related target values have been agreed, entities should agree upon measurement definitions and schemes. This is especially so if there are agreed escalation methods.

Measurement descriptions should include a description of what, when, where and who should perform measurement procedures and test processes, while how measurements should be done need not be specified because they are technology dependent and outside the scope of the customer's interest.

The methodology to evaluate measurement results is also important and may be included in this part of SLA.

In a multi-provisioning contest, exchange of information between remote entities will be expected in order to collect all measures needed for QoS evaluation.

More generally, measurements for a primary QoS parameter are taken at specific *measurement points*. Such points are simply interaction points, where reference events or their outcomes can be observed, and may be located or not in the technical interface (e.g., when measurements are obtained from a sub-provider).

Once primary QoS parameters are obtained, parameters of second level can be derived as functions of these values.

Similar observations may be applied in measurements of every entity's incoming/outcoming traffic.

Typical network measurements may refer to:

– service provision;
– service restoration;
– fault occurrence rate (customer reported and/or network detected);
– availability of interconnection(s);
– customer trouble reports (complaints and/or faults);
– end-to-end testing (either non-intrusive or test calls);
– traffic performance;
– facsimile performance.

Furthermore, in order to consider QoS from customer's point of view, marketing surveys addressing customers are very useful.

Finally, a scheme for definition of measurement processes may be composed by the following points[3]:

- identification of every relevant measurement point;
- description of the measurement environment: service, relevant QoS parameters, traffic conditions;
- definition of the methodologies for obtaining the measured values;
- specification of the methodology to be used for taking decisions concerning the conformance of the measurements with the values agreed in the SLA.

5.3.5 Reaction patterns

A *reaction* is a process that is activated in a more or less automated way whenever commitments on traffic patterns and on QoS parameters are not fulfilled.

Typical examples are:

- provider's reaction to an incoming traffic that differs from the description in the SLA;
- user's behaviour when service provider does not provide QoS agreed in the SLA.

More generally, a useful description of a reaction may be realized depicting it as a process which is characterized by inputs, outputs and constraints (Figure 5-6).

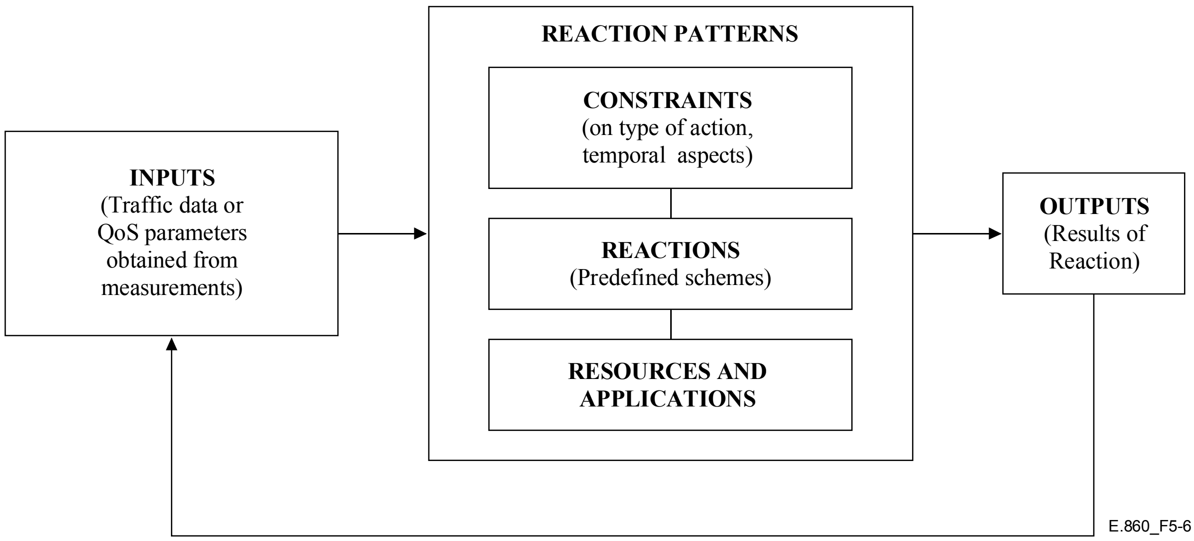


Figure 5-6/E.860 – Reaction patterns

On the input side, measurements are taken on traffic and on QoS in order to detect situations of outage, faults, incoming traffic which do not comply with the agreed traffic pattern, or if QoS is inadequate.

Once obtained, measures are compiled and compared with target values in the SLA (constraints).

Depending on comparison results and available resources, the output of the reaction is finally individuated.

[3] From EURESCOM [P806-GI].

Possible reaction outputs are the following:

- no action;
- monitoring the achieved QoS;
- traffic flow policing through traffic shaping and/or admission control;
- reallocating resources;
- warning signals to customer/SP when thresholds are being crossed;
- suspending or aborting the service.

The utility of the proposed schematization subsists in its capacity of describing accurately the causes and effects. These aspects, in fact, are useful in justifying an entity that starts a reaction.

In addition, the description of temporal aspects through diagrams, which report all reference events with the corresponding time durations, may efficiently contribute to the same scope.

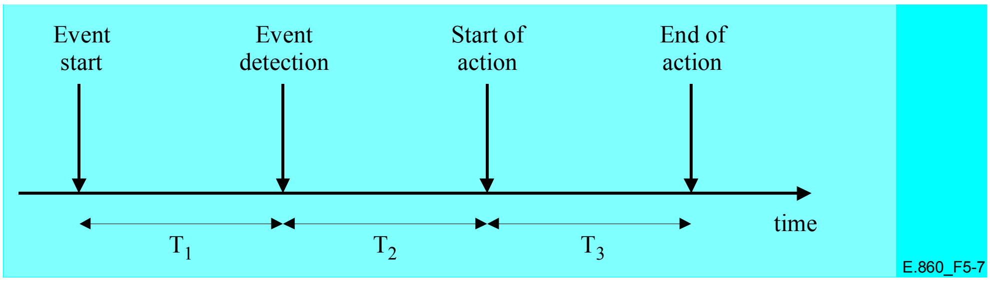


Figure 5-7/E.860 – Representation of events and actions with relating time duration

For example, main events of a reaction process and its time intervals may be illustrated as in Figure 5-7. In particular, T_1 is the index of the SP system efficiency and depends especially on the technology adopted; T_2 quantifies the quickness of the reaction against the detected event and depends especially on the efficiency of SP's human resources; T_3 is the index of the adopted reaction pattern efficacy and it is due to both the SP and the customer, since they both contributed to the definition of the reaction in the SLA.

Therefore we can conclude that, in this contest, the most important factor for a SLA is T_2 because it is the only parameter that discriminates the efficiency of a SP from the others. (T_1 does not suit the case because different SPs often adopt the same technology).

6 Applying Service Level Agreement in a multi-provider environment

As already seen, several service providers are often involved in a service provision and collaborate together in realizing the various service elements.

In this clause we deal with the methodology to use whatever service is provided in a multi-provider environment.

6.1 End-to-End QoS

Consider the case when a SLA between an end user and a provider, for a connection passing through several SP domains, is agreed. Thanks to the one stop responsibility, the end user will require the agreed QoS exclusively from the service provider with whom he agreed upon the SLA, while the latter will have to guarantee that QoS by signing, in its turn, suitable SLAs with its sub-providers.

A traditional approach to multi-provider environment consists in making an association of entities (SPs) (Figure 6.1) which all agreed upon a common document dealing with parameters, objectives and techniques of measurement for QoS.

Statements included into the standard are then utilised for formulating SLAs between each pair of entities and between entities and their customers. Such an approach assures that End-to-End QoS of connections, which pass through several SPs, will fulfil QoS agreed with end user in the SLA.

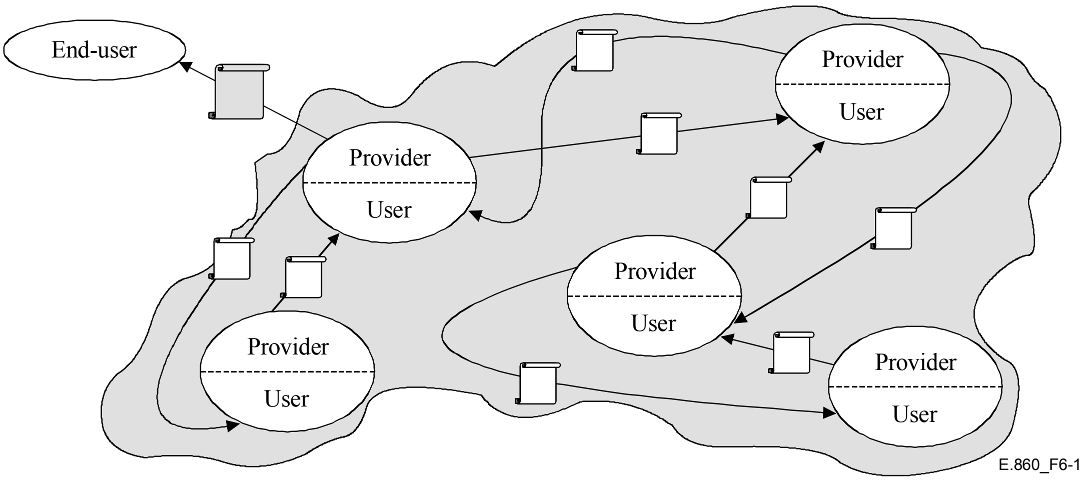


Figure 6-1/E.860 – Association of entities which are involved in realization of the End-to-End QoS

More generally another method, which contains as a particular case the preceding one, consists in making a chain of SLAs between each user-provider pair involved in the provisioning of the same service. The chain is constructed starting from the SLA between end user and primary provider which states the End-to-End QoS; in its turn the primary provider, taking in mind the promised performance, will contract SLAs with its sub-providers and so on in a recursive manner (Figure 6-2).

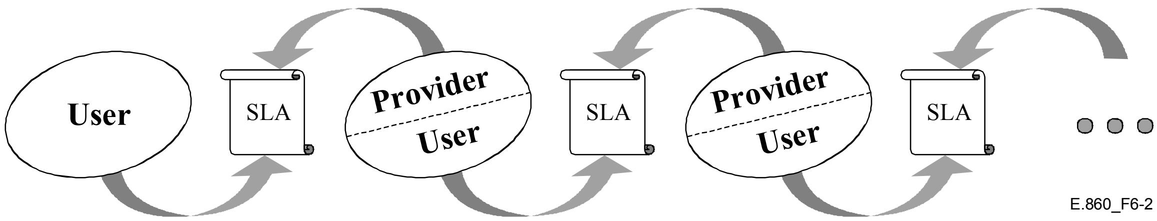


Figure 6-2/E.860 – Chain of Service Level Agreements

6.2 End-to-End SLA

Sometimes in a end-to-end connection, in addition to SLAs between each pair of entities involved, it is useful to have also an **End-to-End SLA**. Such a document is a SLA between all entities with the scope to reach a common understanding about QoS issues as well as business matters.

Some of the issues an End-to-End SLA deals with may be:

- type of service;
- definition of common processes for the Business Interface (e.g., NMF Business Process Model);
- technical constraints;
- definition of QoS/Performance parameters for end-to-end relationships;
- notification and action in case of problems;
- common management policies;

- security;
- various interfaces (accounting, trouble administration, etc.)
- ...

6.3 A general procedure

In general, a procedure which enables to apply in an efficient manner the SLA structure proposed in clause 5 may be that illustrated in Figure 6-3.

The input includes the service description, the entities involved, the description of their roles and their relationships (the so-called Business Model). The procedure individuates the service elements provided to primary provider by its sub-providers and all the interfaces (both BI and TI) that will be used. The outputs are the service delivery configuration with the description of a SLA for each pair user-provider identified during the procedure.

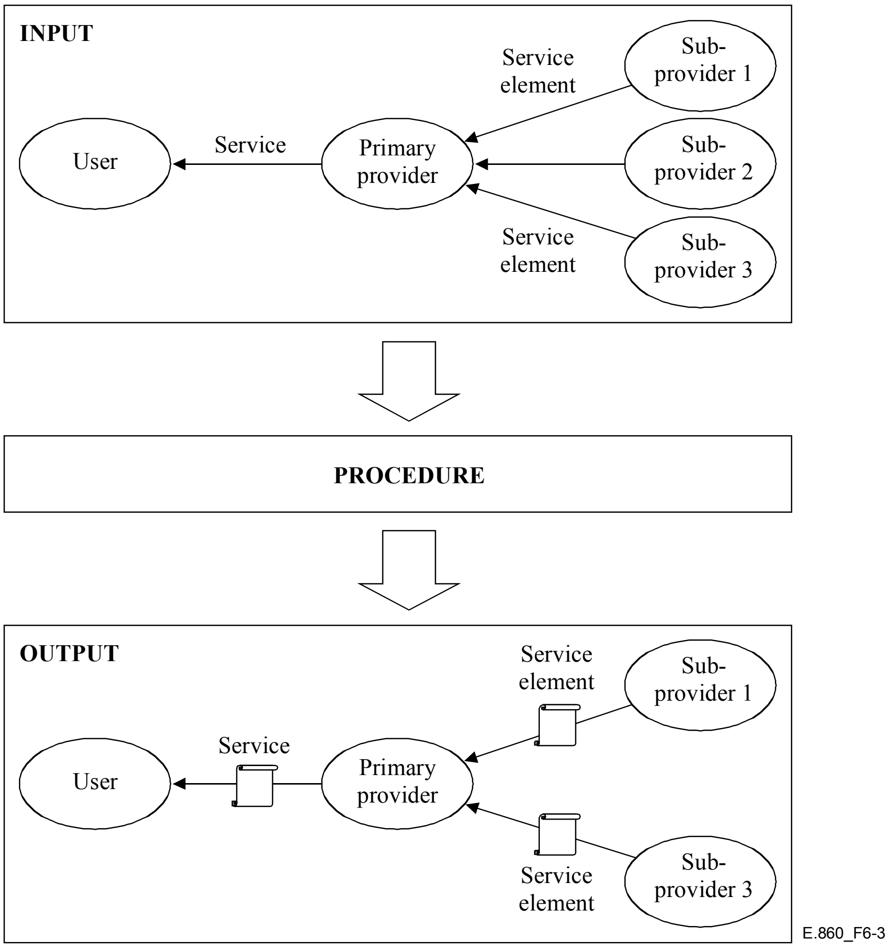


Figure 6-3/E.860 – Methodology for "SLA process"

A way to implement a similar procedure may be the one in Figures 6-4 and 6-5 where we can observe two main steps. The first step aims to identify the SLAs that will be agreed upon, while the second step settles the conditions in each SLA.

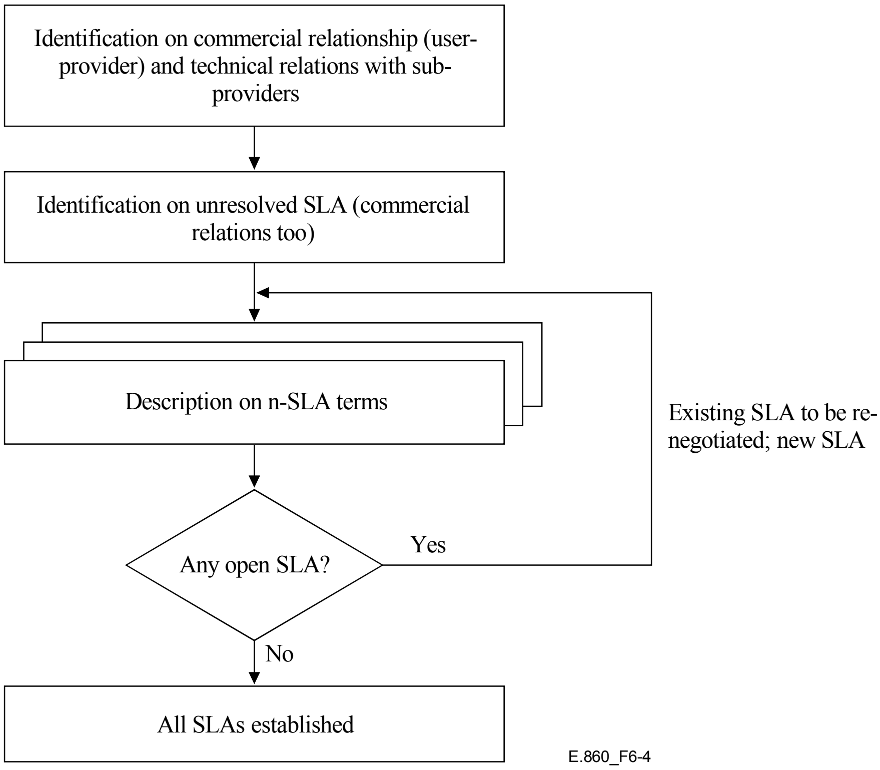


Figure 6-4/E.860 – Step 1: Individuation of relevant SLAs

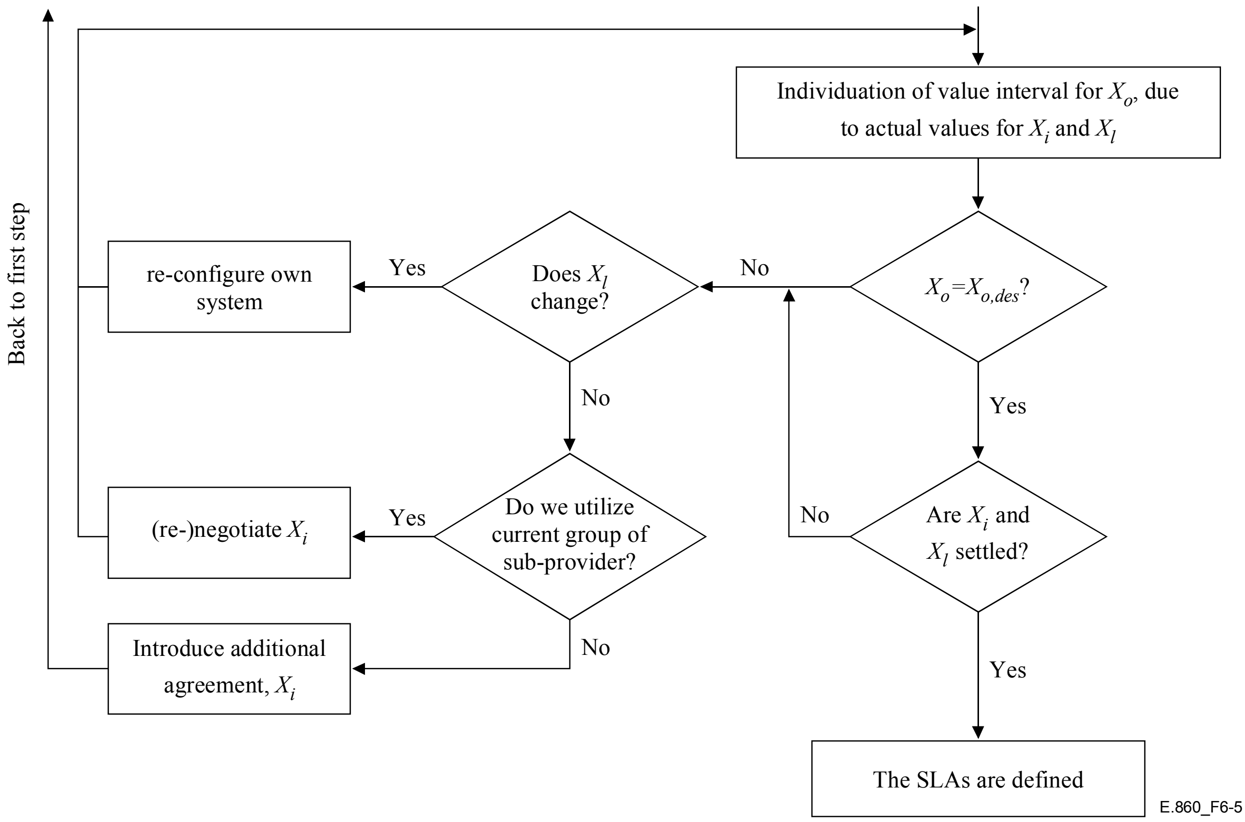


Figure 6-5/E.860 – Step 2: Individuation of the SLA content

In particular in Figure 6-5 $X_{o,des}$ is the level of QoS agreed between primary provider and end user, X_i and X_l represent the levels of QoS that primary provider obtains respectively from its sub-providers and by the resources in its domain, X_o is the QoS actually delivered to end user.

Note that the point of view is always the primary provider's point of view.

Appendix I

Examples of classification parameters

Table I.1/E.860 – Example of classification of PSTN parameters in Timeline Model

SERVICE Scenario:

CRITERION / PHASE	Reliability	Speed	Accuracy
Design, realization	–	–	–
Maintenance	MTBF	MTTR	
Ending	–	–	–

USER Scenario:

CRITERION / PHASE	Reliability	Speed	Accuracy
Subscription	–	Mean Activation Time	–
Utilization	Availability	–	–
Cessation		Mean Disactivation Time	–

SESSION Scenario:

CRITERION / PHASE	Reliability	Speed	Accuracy
Set-up	ASR, ABR, NER, others	PDD, PGAD, others	–
Transfer of information	Dropped calls	–	Quality index, CCI
Release	–	–	–

Table I.2/E.860 – Example of classification of packed switched parameters in Timeline Model

SERVICE Scenario:

CRITERION / PHASE	Reliability	Speed	Accuracy
Design, realization	–	–	–
Maintenance	Availability (network site), MTBF	MTTR	
Ending	–	–	–

USER Scenario:

CRITERION / PHASE	Reliability	Speed	Accuracy
Subscription	–	Mean Activation Time	–
Utilization	Availability (user site)	–	–
Cessation		Mean Disactivation Time	–

SESSION Scenario:

CRITERION / PHASE	Reliability	Speed	Accuracy
Set-up	ASR, ABR, Server and network congestion	–	–
Transfer of information	Dropped session, Delay variation (jitter, etc.)	One way delay, network congestion	Information loss (Packet loss, etc.)
Release	–	–	–

For reference of the parameters in this appendix, see ITU-T Handbook *on Quality of Service and Network Performance* (1993).

Bibliography

[E.800] ITU-T Recommendation E.800 (1994), *Terms and definitions related to quality of service and network performance including dependability*.

[E.801] ITU-T Recommendation E.801 (1996), *Framework for Service Quality Agreement*.

[I.350] ITU-T Recommendation I.350 (1993), *General aspects of quality of service and network performance in digital networks, including ISDNs*.

[NMF701] NMF 701: *Performance Reporting Definitions Document, Issue 1.0,* 04/97.

[P806-GI] EURESCOM P806-GI Deliverable 1: *The EQoS Framework,* January 1999, and EURESCOM P806-GI Deliverable 4: *How to apply EqoS? Case study VoIP. QoS handbook outline,* June 2000.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
Series D	General tariff principles
Series E	**Overall network operation, telephone service, service operation and human factors**
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks and open system communications
Series Y	Global information infrastructure and Internet protocol aspects
Series Z	Languages and general software aspects for telecommunication systems

22803

Printed in Switzerland
Geneva, 2003