A service-level agreement (SLA) is a negotiated contract between two parties, the customer and the

service provider. The agreement can be legally binding or informal and specifies the services that

the customer receive rather than how the service provider delivers the services. The objectives of the

agreement are:

• Identify and define customers’ needs and constraints, including the level of resources, security,

timing, and quality of service.

• Provide a framework for understanding. A critical aspect of this framework is a clear definition of

classes of service and costs.

• Simplify complex issues; for example, clarify the boundaries between the responsibilities of the

clients and those of the provider of service in case of failures.

• Reduce areas of conflict.

• Encourage dialogue in the event of disputes.

• Eliminate unrealistic expectations.

An SLA records a common understanding in several areas: (i) services, (ii) priorities, (iii) responsibilities,

(iv) guarantees, and (v) warranties. An agreement usually covers: services to be delivered,

performance, tracking and reporting, problem management, legal compliance and resolution of disputes,

customer duties and responsibilities, security, handling of confidential information, and termination.

Each area of service in cloud computing should define a “target level of service” or a “minimum

level of service” and specify the levels of availability, serviceability, performance, operation, or other

attributes of the service, such as billing. Penaltiesmay also be specified in the case of noncompliance with

the SLA. It is expected that any service-oriented architecture (SOA) will eventually include middleware

supporting SLAmanagement. The Framework 7 project supported by the European Union is researching

this area (see http://sla-at-soi.eu/).

The commonmetrics specified by an SLA are service-specific. For example, themetrics used by a call

center usually are: (i) abandonment rate: percentage of calls abandoned while waiting to be answered;

(ii) average speed to answer: average time before the service desk answers a call; (iii) time service factor:

percentage of calls answered within a definite time frame; (iv) first-call resolution: percentage of incoming

calls that can be resolved without a callback; and (v) turnaround time: time to complete a certain task.

There are two well-differentiated phases in SLA management: the negotiation of the contract and

the monitoring of its fulfillment in real time. In turn, automated negotiation has three main components:

(i) the object of negotiation, which defines the attributes and constraints under negotiation; (ii) the

negotiation protocols, which describe the interaction between negotiating parties; and (iii) the decision

models responsible for processing proposals and generating counterproposals.

The concept of compliance in cloud computing is discussed in [55] in the context of the user’s

ability to select a provider of service. The selection process is subject to customizable compliance with

user requirements, such as security, deadlines, and costs. The authors propose an infrastructure called

Compliant Cloud Computing (C3) consisting of: (i) a language to express user requirements and the

compliance level agreements (CLAs) and (ii) the middleware for managing CLAs.

The Web Service Agreement Specification (WS-Agreement) [20] uses an XML-based language

to define a protocol for creating an agreement using a predefined template with some customizable

aspects. It only supports one-round negotiation without counterproposals. A policy-based framework

for automated SLA negotiation for a virtual computing environment is described in [379].