Cloud computing is widely regarded as one of the most disruptive technologies of the 21st century. As with any new technology that finds itself in widespread use, this, too, needs careful assessment and a pellucid understanding of what it offers to enterprises and individuals alike. This book has been authored to meet one of such requirements that encompass the service levels expected out of implementing this technology. The book provides an insight into how service level agreements are created and managed in cloud data centers and related environments.

The authors, Mr. Rajesh Bose, PhD scholar at the University of Kalyani; Mr. Satadru Sengupta, a professional content writer; and Mr. Sandip Ray, research scholar at the University of Kalyani have jointly and individually made significant contributions in the fields of Cloud Computing, IoT and related fields.

978-620-2-19960-5

SLA FOR CLOUD COMPUTING

Rajesh Bose
Satadru Sengupta
Sandip Roy

Interpreting SLA and related nomenclature in terms of Cloud Computing

A layered approach to understanding service level agreements in the context of cloud computing

LAP LAMBERT Academic Publishing

Bose, Sengupta, Roy

INTERPRETING SLA AND RELATED NOMENCLATURE IN TERMS OF CLOUD COMPUTING

Rajesh Bose
Satadru Sengupta, Sandip Roy

About the Authors

Rajesh Bose is an IT professional employed as Deputy Manager with Simplex Infrastructures Limited, Data Center, Kolkata. He graduated with a B.E. in Computer Science and Engineering from Biju Patnaik University of Technology (BPUT), Rourkela, Orissa, India in 2004. He went on to complete his degree in M.Tech. In mobile communication and networking from Maulana Abul Kalam Azad University of Technology, West Bengal (Formerly known as West Bengal University of Technology) - WBUT, India in 2007. He has also several global certifications under his belt. These are CCNA, CCNP-BCRAN, and CCA (Citrix Certified Administrator for Citrix Access Gateway 9 Enterprise Edition), CCA (Citrix Certified Administrator for Citrix Xen App 5 for Windows Server 2008). Mr. Bose has submitted his doctoral thesis in cloud computing from University of Kalyani. His research interests include cloud computing, IoT, wireless communication and networking. He has published more than 42 referred papers and 3 books in different journals and conferences. Mr. Bose also has experience in teaching at college level for a number of years prior to moving on to becoming a full time IT professional at a prominent construction company in India where he has been administering virtual platforms and systems at the company's data center.

Satadru Sengupta is somewhat of an outlier in the field of computing and technology in general, and particularly insofar as cloud computing, IoT, and wireless sensor technologies are concerned. His forte is his unique ability to grasp technicalities and to resolve issues given his ITIL certification and occupation as an IT Help Desk manager. However, it is for his methodical approach in coalescing empirical observations from research experiments into plain English that makes him an invaluable and an indispensable addition to almost any research team, and especially those related to the field of cloud computing. He has been a content writer for the most part of his career, and is a member of the Data Center team at Simplex Infrastructures Limited, Kolkata. His clarity in understanding programming languages, too, makes him an asset in research areas that require software coding. This book is an acknowledgement of his tireless and sincere dedication to present meaningful contribution in texts involving cloud computing, IoT, and Smart Technologies.

Sandip Roy is currently pursuing PhD from University of Kalyani, Kalyani, India. He is an Assistant Professor and Teacher-in-Charge of Department of Computer Science & Engineering of Brainware Group of Institutions-SDET, Kolkata, West Bengal, India. He received M.Tech. degree in Computer Science & Engineering in 2011, and a B.Tech. in Information Technology in 2008 from Maulana Abul Kalam Azad University of Technology, West Bengal (Formerly known as West Bengal University of Technology). He has authored over 20 papers in peer-reviewed journals, conferences, and is a recipient of the Best Paper Award from ICACEA in 2015. His main areas of research interests are Cloud Computing, Internet of Things, Cloud of Things and Smart Technologies.

ACKNOWLEDGEMENT

The purpose of a book of this nature is to disseminate knowledge. I firmly believe that without the good grace of the Almighty it would not have been possible to sow the seeds of the book's foundation in the first place. I thank the Lord for having granted me the gift to write. But above all, I thank Him for having given me the patience and endurance to do what it takes to give this book the content, shape and form in which it has been presented for the readers for whom the book was intended.

We had the good fortune of interacting with some of the most knowledgeable and experienced professionals in the IT industry during compiling this book. Arguably, most of them have provided interesting information and deep insights as to how cloud computing has evolved to the stage where it is now.

My co-authors, Mr. Satadru Sengupta and Mr. Sandip Roy, have been instrumental in giving shape and substance to the book. Without their invaluable contributions, much of the research that has gone into this book would not have been possible. Although experts in their respective fields, they have reposed much faith and confidence in my ability to play a leading role in putting this book together.

Without my parents, my aims and ambition would have remained unrealized. My mother, the Late Jharna Bose, whose dream it was that I should devote my energy to inspire those who depend on me for their own academic sustenance, would have found it adequate to see fruits of my labour take shape thus. After my mother passed away, my ailing father, the Late Sisir Kumar Bose, made it a mission of his life to ensure that I never for once deviated from the goal set out before me. With their blessings and kind words which I cherish, I gathered the strength to surge through successive waves of trials and tribulations. They were the ones to realize the value of technology years before. I value their simplicity and far-sightedness which has enabled me to achieve this modest goal.

My personal contributions, however, would not have materialized if it had not been for my family who supported me in my quest and dream to write my first book. I owe a great deal to my wife, Swati, who has steadfastly stood beside me in my attempts to put together the initial sketches and flow of the chapters. Without her unflinching support and dedication, it would not have been possible to expand my horizons and push forward my career. This book is dedicated to her.

To my son, Pablo, I offer my deepest love for having made me smile all the while. For having understood that his father needed time to write the book instead of playing games with him when he found me hammering away at a laptop keyboard. I hope one day he would realize why I spent countless weekends and holidays on this book.

Last but not the least, I take this opportunity of thanking my friends and colleagues who have given their enthusiastic support at every stage of writing this book.

SUMMARY

Cloud computing, the model for providing on-demand access to a pool of shared resources with minimum provider interference, is emerging as a substitute to common IT infrastructure. As increasing numbers of cloud consumers dispatch their workloads to cloud providers, Cloud Service Level Agreements (Cloud SLAs) form an important component of the contractual relationship between a cloud service customer and a cloud service provider of a cloud service. Given the global nature of the cloud, SLAs usually span many jurisdictions, with often varying applicable legal requirements, in particular with respect to the protection of the personal data hosted in the cloud service. Furthermore different cloud services and deployment models will require different approaches to SLAs, adding to the complexity of SLAs. SLA contains an explanation of the agreed service, parameters of the level of service, the guarantees regarding the Quality of Service, arrangements and cures for all cases of violations. To attract companies to outsource their services to clouds, providers need to offer Service Level Objectives specified in SLAs for their customers. The content of such Service Level Objectives is a key reason for the successful usage of cloud computing and consists of Key Performance Indicators. Due to the dynamic character and complex nature of the cloud environment, creating SLAs for the cloud can be very difficult. Delivering effective service level management (SLM) is a vital requirement for today's Cloud services providers (CSPs). This book looks at the importance of SLM and the key requirements for effectively and profitably delivering these services.

This book provides a practical reference to help enterprise information technology (IT) and business decision makers analyze cloud service agreements (CSAs) from different cloud service providers. The book informs decision makers of what to expect and what criteria to use as they evaluate CSAs from such potential suppliers CSAs are primarily written to set clear expectations for service between the cloud customer (buyer) and the cloud provider (seller), but should also exist between a customer and other cloud entities, such as the cloud carrier, the cloud broker and even the cloud auditor. This Guide focuses primarily on the CSA details between the cloud customer and cloud provider. The aim of this Book is to present how SLAs are created, managed and used in cloud computing environment.

WHO IS THIS BOOK FOR?

Around the globe, corporate policy-makers and those entrusted with the task of pushing the frontiers of their respective enterprises' information technology framework, have begun to adopt cloud technology in a big way. At a time when it was a nascent technology, service providers and end-users pushed across the spectrum of desired and achieved performance levels. Mostly it was a case of trial and error.

With rising competition, most if not all businesses find they performing a balancing act on a knife-edge. It has become essential, therefore, for managers and Data Center professionals to zero in on a cloud solution that is not only cheap to operate, but is also flexible. For the entrepreneur and student alike, the subject of service level management, service level objectives, and levels of assurances can be a confusing ill-digested mass of ideas. This book has been written to supplement and clarify many of these topics concerning cloud computing service level agreements.

Organizations and start-ups on a budget would find the section on cloud service billing of interest. For security professionals and Data Center experts, the sections on implementing security standards and disaster recovery plans, we hope, would form the core of this book.

We look forward to having views and opinions of our readers. The authors wrote this book after drawing inspiration from students, IT professionals, and those entrusted with the task of laying down the path towards migrating to cloud. By no means is this book complete. However, we hope that it shall help those interested in cloud computing to see silver linings no matter how complex and how dark the clouds of computing may be.

Table of Contents

1.CHAPTER I

Understanding the metrics of cloud computing services: a fundamental discussion on service level agreements.

1.1. Introduction

Cloud-based services are increasingly becoming common place. These services include infrastructure as-a-service (IaaS), platform-as-a-service (PaaS), and software-as-a-service (SaaS). Each service is typically accompanied by a service level agreement (SLA) which defines the minimal guarantees that a provider offers to its customers. The lack of standardization in cloud-based services implies a corresponding lack of clarity in the service level agreements offered by different providers.

For cloud computing, the quality and reliability of the services become an important aspect, as customers have no direct influence on the services. Therefore Service Level Agreements are fundamental to effective cloud utilization and especially business customers need them to ensure risks and service qualities are prevented respectively provided in the way they want. For this purpose, the expected service qualities are documented legally binding in contracts between provider and customer. Due to significant variation in consumer needs, SLAs have to be created individually by a negotiation process. The confirmed SLAs serve as a basis for compliance and monitoring of the QoS. Due to the dynamic cloud character, the QoS attributes must be monitored and managed consistently.

As the aforementioned cloud service model matures and becomes ubiquitous, it raises the possibility of improving the way services are provisioned and managed, thus allowing providers to address the (diverse) needs of consumers. In this context, Service Level Agreements (SLAs) emerge as a key aspect, since they serve as the foundation for the expected quality level of the service between the consumer and the provider. Nevertheless, the diversity of the proposed SLAs by providers (with marginal overlaps), has led to multiple different definitions of cloud SLAs. Furthermore, misconceptions exist on what is (if there is) the difference between SLAs and contract, what is the borderline, what are the terms included in each one of these documents and if and how are these linked. We provide the following definitions according to ITIL:

A **Service Level Agreement** (SLA) is a formal, negotiated document that defines (or attempts to define) in quantitative (and perhaps qualitative) terms the service being offered to a Customer. Any metrics included in a SLA should be capable of being measured on a regular basis and the SLA should record by whom.

A **Contract** is a legally binding agreement between two or more parties. Contracts are subject to specific legal interpretations.

A Service Level Agreement (SLA) is a formal negotiated agreement between two parties. It is a contract that exists between the Service Provider (SP) and the Customer. It is designed to create a common understanding about Quality of Service (QoS), priorities, responsibilities, etc. SLAs can cover many aspects of the relationship between the Customer and the SP, such as performance of services, customer care, billing, service provisioning, etc. However, although a SLA can cover such aspects, agreement on the level of service is the primary purpose of a SLA .

Service Level Agreements play a central role in the service lifecycle, since by capturing service expectations and entities responsibilities they drive both engineering decisions at conception level (during for example service design) and operational decisions (during for example service usage and delivery). SLAs enable participating entities to agree on what services will be offered, how will the services be delivered and who will be responsible for execution, completion, potential failures and privacy aspects.

1.2. SLA in Cloud Computing

A service-level agreement is an agreement between two or more parties, where one is the customer and the others are service providers. This can be a legally binding formal or an informal "contract" (for example, internal department relationships). Contracts between the service provider and other third parties are often (incorrectly) called SLAs – because the level of service has been set by the (principal) customer, there can be no "agreement" between third parties; these agreements are simply "contracts." Operational-level agreements or OLAs, however, may be used by internal groups to support SLAs.

SLAs commonly include segments to address: a definition of services, performance measurement, problem management, customer duties, warranties, disaster recovery, and termination of agreement. In order to ensure that SLAs are consistently met, these agreements are often designed with specific lines of demarcation and the parties involved are required to meet regularly to create an open forum for communication. Contract enforcement (rewards and penalties) should be rigidly enforced, but most SLAs also leave room for annual visitation so that it is possible to make changes based on new information.

SLAs have been used since late 1980s by fixed line telecom operators as part of their contracts with their corporate customers. This practice has spread such that now it is common for a customer to engage a service provider by including a service level agreement in a wide range of service contracts in practically all industries and markets. Internal departments (such as IT, HR, and real estate) in larger organizations have adopted the idea of using service-level agreements with their "internal" customers — users in other departments within the same organization. One benefit of this can be to enable the quality of service to be benchmarked with that agreed to across multiple locations or between different business units. This internal benchmarking can also be used to market test and provide a value comparison between an in-house department and an external service provider. Service level agreements are, by their nature, "output" based – the result of the service as received by the customer is the subject of the "agreement." The (expert) service provider can demonstrate their value by organizing themselves with ingenuity, capability, and knowledge to deliver the service required, perhaps in an innovative way. Organizations can also specify the way the service is to be delivered, through a specification (a service level specification) and using subordinate "objectives" other than those related to the level of service. This type of agreement is known as an "input" SLA. This latter type of requirement is becoming obsolete as organizations become more demanding and shift the delivery methodology risk on to the service provider. Service level agreements are also defined at different levels:

a) Customer-based SLA: An agreement with an individual customer group, covering all the services they use. For example, an SLA between a supplier (IT service provider) and the finance department of a large organization for the services such as finance system, payroll system, billing system, procurement/purchase system, etc.

b) **Service-based SLA**: An agreement for all customers using the services being delivered by the service provider. For example:

A car service station offers a routine service to all the customers and offers certain maintenance as a part of offer with the universal charging. A mobile service provider offers a routine service to all the customers and offers certain maintenance as a part of offer with the universal charging An email system for the entire organization. There are chances of difficulties arising in this type of SLA as level of the services being offered may vary for different customers (for example, head office staff may use high-speed LAN connections while local offices may have to use a lower speed leased line).

c) **Multilevel SLA:** The SLA is split into the different levels, each addressing different set of customers for the same services, in the same SLA.

d) **Corporate-level SLA:** Covering all the generic service level management (often abbreviated as SLM) issues appropriate to every customer throughout the organization. These issues are likely to be less volatile and so updates (SLA reviews) are less frequently required. Customer-level SLA: covering all SLM issues relevant to the particular customer group, regardless of the services being used. Service-level SLA: covering all SLM issue relevant to the specific services, in relation to this specific customer group. The underlying benefit of cloud computing is shared resources, which is supported by the underlying nature of a shared infrastructure environment. Thus, service level agreements span across the cloud and are offered by service providers as a service based agreement rather than a customer based agreement. Measuring, monitoring and reporting on cloud performance is based upon an end user experience or the end users ability to consume resources. The downside of cloud computing, relative to SLAs, is the difficultly in determining root cause for service interruptions due to the complex nature of the environment. As applications are moved from dedicated hardware into the cloud these applications need to achieve the same or even more demanding levels of service as classical installations. SLAs for cloud services focus on characteristics of the data center and more recently include characteristics of the network to support end-to-end SLAs. Any SLA management strategy considers two well-differentiated phases: the negotiation of the contract and the monitoring of its fulfilment in real-time. Thus, SLA Management encompasses the SLA contract definition: basic schema with the QoS (quality of service) parameters; SLA negotiation; SLA monitoring; and SLA enforcement—according to defined policies.

The creation of Service Level Agreements provides certain requirements to customers and providers. Customers need to be able to meet certain requirements in order to successfully define SLAs, which are listed briefly here. A customer must:

Understand the roles and responsibilities that are regulated by the SLA.

Be able to describe precisely and specific the service to be controlled by the SLA.

Know the requirements of the controlled services, and define the matching key figures.

Specify service levels based on the critical performance characteristics of the service.

Understand the process and procedures of regulated service.

These requirements are necessary so that the customer is able to put in the correct SLAs values, and to understand implications of his decisions. Furthermore, a SLA should fulfill the following tasks:

Describe the services accurately.

Specify the service quality to be provided in detail.

Describe detailed the key performance indicators, metrics and service levels.

Breakdown transparently all the costs.

• **Service guarantee time period** describes the duration over which a service guarantee should be met. The time period can be a billing month or time elapsed since the last claim was filed. The time period can also be small, e.g., one hour. The smaller the time period, the more stringent is the service guarantee.

• **Service guarantee granularity** describes the resource scale on which a provider specifies a service guarantee. For example, the granularity can be on a per service, per data center, per instance, or per transaction basis. Similar to time period, the service guarantee can be stringent if the granularity of service guarantee is fine-grained. Service guarantee granularity can also be calculated as an aggregate of the considered resources, such as instances or transactions. For example, aggregate uptime of all running instances must be greater than 99.95%. However, such a guarantee implies that some instances in the aggregate SLA computation can potentially have a lower percentage uptime than 99.95% while still meeting the aggregate SLA. As a consequence, aggregate SLA computation leaves provider the wiggle room to better manage its offered services.

• **Service guarantee** exclusions are the instances that are excluded from service guarantee metric calculations. These exclusions typically include abuse of the system by a customer, or any downtime associated with the scheduled maintenance.

• **Service credit** is the amount credited to the customer or applied towards future payments if the service guarantee is not met. The amount can be a complete or partial credit of the customer payment for the affected service.

• **Service violation** measurement and reporting describes how and who measures and reports the violation of service guarantee, respectively.

1.3. Cloud Computing

Based on the observation of the essence of what Clouds are promising to be, Buyya et. al. (2009) propose the following definition: "*A Cloud is a type of parallel and distributed system consisting of a collection of inter-connected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resource(s) based on service-level agreements established through negotiation between the service provider and consumer*". Hence, Clouds fit well into the definition of utility computing.

Figure 1 shows the layered design of Cloud computing architecture. Physical Cloud resources along with core middleware capabilities form the bottom layer needed for delivering IaaS. The user-level middleware aims at providing PaaS capabilities. The top layer focuses on application services (SaaS) by making use of services provided by the lower layer services. PaaS/SaaS services are often provided by 3rd party service providers, who are different from IaaS providers.

User-Level Applications: this layer includes the software applications, such as social computing applications and enterprise applications, which will be deployed by PaaS providers renting re-source from IaaS providers.

User-Level Middle wire: Cloud programming environments and tools are included in this layer facilitate creation of applications and their mapping to resources using Core Middleware Layer services.

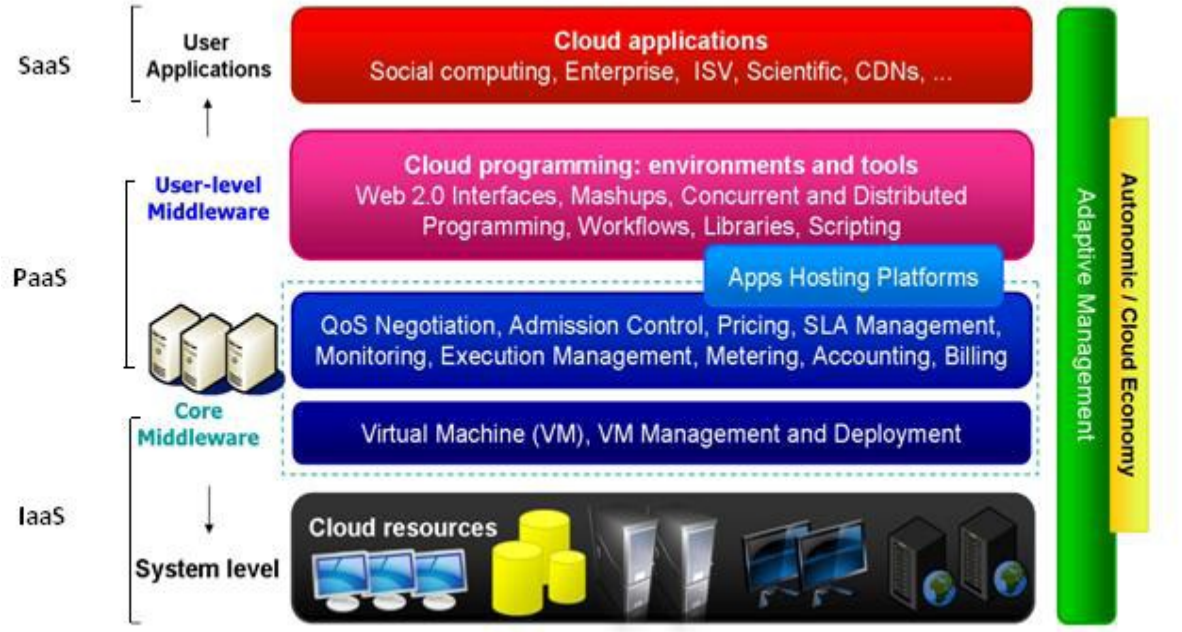

Figure 1: Layered Cloud computing architecture.

Core Middleware: this layer provides runtime environment enabling capabilities to application services built using User-Level Middleware. Dynamic SLA management, Accounting, Monitoring and Billing are examples of core services in this layer. The commercial examples for this layer are Google App Engine and Aneka.

System Level: physical resources including physical machines and virtual machines sit in this layer. These resources are transparently managed by higher level virtualization services and toolkits that allow sharing of their capacity among virtual instances of servers.

Cloud users require SLAs to identify the technical performance demands satisfied by a cloud supplier. SLAs can accomplish conditions about the superiority of service, security, and cures for the facing malfunctions. A cloud supplier could also state in the SLAs a group of guarantees that are not prepared to users clearly, i.e. restrictions, and duties that cloud users have to approve on. A cloud user can select a cloud supplier with preferable pricing and more complimentary conditions. Normally, a cloud supplier's pricing strategy and SLAs are non-discussable, except if the user looks forward to intensive employment and can be able to discuss for superior convention.

Relying on the services demanded the actions and employment situations can be diverse over cloud users.

The clients of SaaS might be corporations that offer their participants with entrance to software applications, end clients who immediately exploit software applications, or software application directors who constitute applications for the clients. SaaS expenses can be paid according to the number of the end clients, the usage time, the network bandwidth spent, the quantity of information kept or the period of keeping information.

Cloud clients of PaaS can exploit the instruments and the resources supplied by cloud suppliers to progress, examine, install and administer the applications presented in a cloud medium. PaaS clients can be application designers who develop and accomplish application software. Also, they can be application examiner who execute and examine applications in cloud-based locations. They can be application publishers who distribute applications through the cloud, or can be application managers who constitute and control applications. PaaS expenses can be paid based on, operation, database space, network resources used by the PaaS application, or the period of the platform convention.

Clients of IaaS have an entrance to virtual computing machines, network storing space, network groundwork elements, and other essential resources on which they can install and operate random software. The clients of IaaS can be system designers or system managers who are concerned in making, running, organizing and controlling services for IT groundwork processes. IaaS users are provided with the abilities to enter these resources, and are paid depending on the quantity or time period of the resources used like; CPU hours consumed by virtual computing machines, capacity, network bandwidth used, and quantity of IP addresses utilized for particular periods. Cloud users want an SLA before delivering their groundwork of cloud information stations to have confidence about the resources supplied and to have the facility to get the preferred level of efficiency.

The cloud supplier is an individual or an institute that is accountable for providing accessible service to concerned actors. A Cloud Supplier develops and administers the computing groundwork needed for supplying the services, operates the cloud software that supplies the services, and makes procedure to transport the services to the Cloud Clients through network entrance.

For SaaS, the cloud supplier installs, constitutes, preserves and improves the process of the software applications on the cloud groundwork so that the services are provided at the estimated service levels to cloud clients. The supplier of SaaS considers many of the tasks in handling and monitoring the applications and the groundwork, while the cloud users have partial managerial monitoring of the applications.

For PaaS, the Cloud supplier organizes the computing groundwork for the platform and operates the software that supplies the elements of the platform like; software implementation stack, databases, and other elements. Moreover, the PaaS Cloud supplier usually provisions the improvement, organization and administration procedure of the PaaS cloud user by supplying instruments like; integrated development environments (IDEs), improvement form of cloud software, software development kits (SDKs), distribution and organization instruments. The PaaS Cloud user has monitoring on the applications and probably some of the introducing locations settings, but has no or restricted entrance to the groundwork likes the network, servers, operating systems (OS), or storage.

For IaaS, the Cloud supplier obtains the tangible resources of the service, such as the servers, networks, storage and hosting groundwork. The Cloud supplier operates the software needed to have computing resources available to the IaaS cloud user through a group of service interfaces and computing resource ideas such as virtual computers and virtual network interfaces. The IaaS Cloud

users exploit these resources, such as a virtual machine for their essential computing requirements suitable for SaaS and PaaS Cloud users. An IaaS Cloud user has entrance to more essential shapes of computing resources and has more monitoring on the software elements inside the application, OS and network. The IaaS Cloud supplier has control on the tangible hardware and software that makes the providing of these groundwork services probable.

Cloud computing that supplies less cost and gives price-as-you-use resources is quickly earning activity as a substitution to the conventional IT Groundwork. As users come up to exploit cloud computing, the superiority and consistency of the services come to be significant parts. But the requirements of the service clients differ meaningfully, so equilibrium has to be done through the negotiation procedure. At the conclusion of the negotiation procedure, supplier and user reach to a covenant (SLA). This SLA assists as the basis for the predictable level of service between the user and the supplier. The QoS characteristics that are commonly portion of an SLA alters frequently, so the parameters have to be carefully controlled and observed

1.4. The architecture model of CSLA (Cloud Service Level Agreement)

To ensure the quality of service, One SLA (Service Level Agreement) usually is signed between users and service providers. The SLA definition from TMF (Tele Management Forum): SLA is a formal agreement negotiated by two entities, a legally binding contract, the respective responsibilities and other aspects of the consensus and agreements between the service provider and the customer. In short, CSLA is the guarantee terms and mishandling for QoCS (Quality of Cloud Service) signed between the cloud service providers and cloud service consumers. CSLA includes five parts: cloud service, cloud technology, cloud quality report, and cloud security, cloud business, as illustrated in Figure 2.

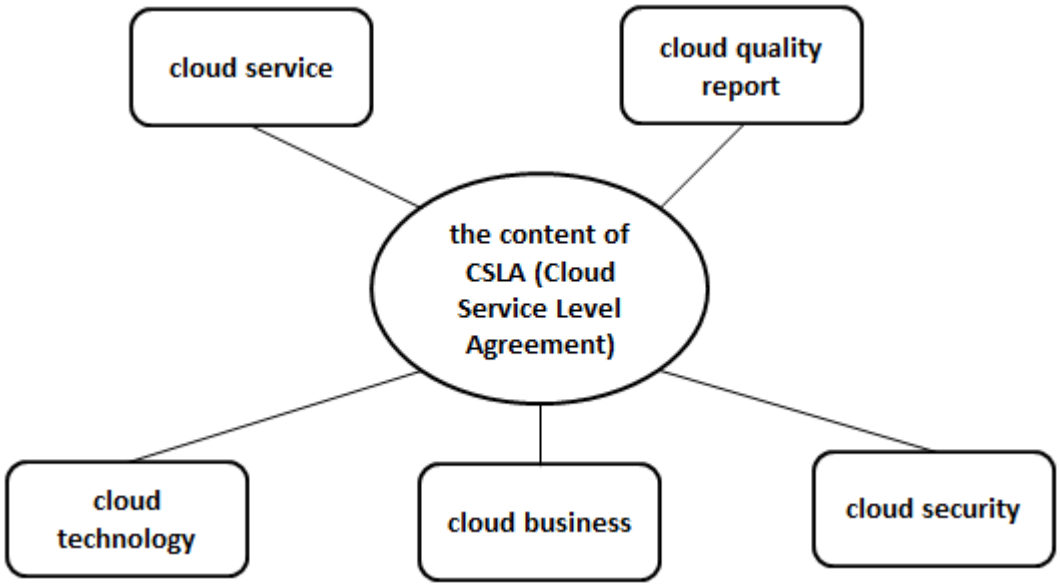

Figure 2: CSLA structure

1.5. Principles for the development of Service Level Agreement Standards for Cloud Computing

The Internet and other advances in computing have spawned a global digital economy and the continuing evolution of cloud computing has added a new and rapidly growing dynamic. While cloud computing is increasing in maturity, it is still in its nascent stages and the related technologies, business models and polices will undoubtedly evolve over a number of years.

There are a number of efforts underway to facilitate adoption of cloud computing by adding clarity to the agreements between cloud service customers and cloud service providers, thus making them more comparable and comprehensible. These efforts are valuable but at the same time it is important to not constrain the technical and business innovation of cloud computing.

The following is a set of principles that can assist organizations, through the development of standards and guidelines for cloud SLAs and other governing documents. These principles are not intended to be limiting or to even set model terms.

1.5.1. Technology Neutral

Essential hallmarks of cloud computing are flexibility and extensibility for which technology neutrality is a necessary foundation. Cloud services can be built using any number of technologies and a particular technology stack should not be assumed.

For example, many cloud services expose REST interfaces or APIs but they can also use technologies such as Web Services to receive data and interoperate with other services.

In another example being technology neutral is important because cloud services commonly run on virtualized hardware platforms but virtualization should not be assumed.

Continuous improvement to deliver increasing value is critical to the future of cloud computing and the freedom to innovate technically is key to that. Cloud services are built on open source software and proprietary software alike. There can also be a variety of hardware platforms underlying cloud services.

1.5.2. Business Model Neutral

A particular business model for cloud services should not be assumed. Cloud services may be funded by any number of methods such as pay per use, long term contracts, advertising, public funds and others. Remedies for failure to achieve cloud service level objectives (SLOs) stated in the SLA can also take different forms such as refunds on charges, free services or other forms of compensation.

1.5.3. World-wide applicability

The Internet is a global communications channel and it is built on standards that are respected worldwide. Likewise, cloud services have a global audience of governments, small businesses, enterprises, NGOs and individuals. Agreements that govern cloud services must account for regional, national and local laws, regulations and policies but everyone benefits from globally common concepts, vocabulary and globally accessible technology.

1.5.4. Unambiguous definitions

Keeping the definition of service level objectives well-defined and unambiguous is important to ensure the effective standardization of cloud SLAs and to enable clear communication between cloud service providers and cloud service customers. As technology develops and new terminology is developed it will also be important to ensure definitions are up-to-date and consistent with an evolving cloud services landscape.

1.5.5. Comparable Service Level Objectives

Service Level objectives ('SLO') are often quantitative and have related measurements. For cloud service customers to make informed decisions when choosing cloud services, it is best if the service level objectives offered by each cloud service provider for similar services can be easily compared. Measurements should also be comparable since reduced comparability impedes adoption. However, from case to case reviewing less-quantitative or qualitative SLOs and comparing different services may provide extra insights for making such informed decision.

To be comparable, service level objectives need not be determined by identical means but sufficient information about the SLO needs to be provided by cloud service providers. Standardized terminology, metrics and templates can be helpful in documenting how a particular SLO is determined.

Service level objectives are often associated with metrics. A metric is a defined measurement method and measurement scale, which is used in relation to a quantitative service level objective.

Metrics are used to set the boundaries and margins of errors which apply to the behaviour of the cloud service and any limitations. Metrics may be used at runtime for service monitoring, balancing, or remediation. Using a standard set of metrics or metric templates in the cloud SLA makes it easier and faster to define a cloud SLA and service level objectives, and simplifies the task of comparing one cloud SLA to another.

It is often true that a given SLO may have multiple different metrics which can be used. It is important that an SLA makes it clear which metric(s) are being used for each quantitative SLO.

1.5.6. Conformance through disclosure

Since standards and guidelines for cloud SLAs should be technology and business model neutral, they should not mandate a specific approach for any concept. For example, service availability can be measured in different ways4, some of which depend on the specific cloud service. A compute service is different than a cloud email service and service availability for each will be computed differently.

Cloud service providers should document their method of achieving SLOs for each concept in their cloud SLA based on standard concepts and vocabulary.

1.5.7. Standards and Guidelines which span customer types

Cloud services are valuable to both enterprises with thousands of users as well as small businesses with just a few users. In many cases, the cloud service is a highly standardised offering that relies heavily on uniformity to enable economies of scale and offer customers benefits, such as low prices. In some cases, the cloud SLA and other governing documents may be negotiated between the cloud service customer and the cloud service

provider but such a negotiation cannot be assumed by default. In many cases, cloud service customers are offered a fixed standard agreement by the cloud service provider, which they can either choose to accept, or they can choose a different cloud service provider that offers different terms and conditions.

Standards and guidelines for cloud SLAs must be able to span from the smallest cloud service customer to the largest. Useful standards and guidelines exist, produced by organisations such as ENISA, NIST or ISO/IEC. For example, in the field of security, relevant work is using the approach to analyze and refine an individual control into one of more security SLOs, which are then associated with metrics and measurements that can be either quantitative or qualitative.

However, it is not possible to list exhaustively relevant standards, guidelines or certifications and many other useful specification initiatives exist.

1.5.8. Cloud Essential Characteristics

While cloud computing is a form of distributed computing, there are differences between traditional on premise and outsourced computing and cloud computing. These differences are best described in ISO/IEC 17788 'Cloud Computing Overview and Vocabulary' as:

- Broad Network Access
- Measured Service
- Multi-tenancy
- On-demand self-service
- Rapid elasticity and scalability
- Resource pooling

While many of the concepts from traditional distributed computing SLAs apply to cloud SLAs, the specific needs of cloud computing must be recognized and accounted for.

1.5.9. Proof Points

Any effort to develop standards and guidelines for cloud SLAs should take into account the state-of-the-art and to some degree represent the capabilities of the cloud services industry.

The state-of-the-art should not necessarily limit the introduction of new ideas or the re-use of long standing concepts but they should be considered relative to industry's capabilities including the cloud essential characteristics. Before introducing a particular concept into a standard or guideline for cloud SLAs the organization[9] should look for proof points to ensure the concept is viable from both technical and business perspectives.

1.5.10. Information Rather Than Structure

Standards and guidelines for cloud SLAs should not specify the structure of the SLA, instead they should illustrate and specify the concepts that should be addressed.

What is valuable is information that helps business and technical stakeholders understand the non-legal concepts and vocabulary used in cloud SLAs. Some of the concepts mentioned in this book may not be part of the standard offering for all cloud computing services, given the important differences between cloud services models (IaaS, PaaS, SaaS, xaaS), as well as the many different cloud services provided within such group of cloud services models.

The fact that an SLO is not implemented does not necessarily imply that the service is of lower quality or performing worse. There may also be cases where similar information could be derived from other SLOs.

A cloud SLA can be a part of an overall Master Service Agreement (MSA). The SLA describes and sets service level objectives for the cloud service. However, the organization and the names used for the MSA and its associated documents can vary considerably and the location of a particular service level objective within the document set can also vary. These documents may include, but are not limited to:

- ✓ Master Service Agreement (MSA)
- ✓ Service Level Agreement (SLA)
- ✓ Service Agreement
- ✓ Acceptable Use Policy
- ✓ Privacy Policy
- ✓ Security Policy
- ✓ Business Continuity Policy
- ✓ Service Description

1.5.11. Leave the Legal Agreement to Attorneys

Standards and guidelines for SLAs should specify the concepts and definitions necessary for the cloud service provider to describe the cloud service and its attributes. The agreement between the cloud service provider and cloud service customer can refer to the clearly defined information in the SLA, but the agreement itself must meet local legal requirements and those must be left to the discretion of qualified attorneys.

Furthermore, the purpose of this guideline is to inform both cloud service customers and cloud service providers about some considerations when understanding or comparing SLAs in the context of their particular situation.

1.6. Anatomy of a Typical Cloud SLA

A typical SLA of a cloud provider has the following components.

• **Service guarantee** specifies the metrics which a provider strives to meet over a service guarantee time period. Failure to achieve those metrics will result in a service credit to the customer. Availability (e.g., 99.9%), response time (e.g., less than 50ms), disaster recovery, and fault

• **Service guarantee time period** describes the duration over which a service guarantee should be met. The time period can be a billing month or time elapsed since the last claim was filed. The time period can also be small, e.g., one hour. The smaller the time period, the more stringent is the service guarantee.

Purpose – mentions why SLA is formed.

Parties – mentions the parties included in the SLA and their jobs.

Scope – describes the services mentioned in the SLA; SLA structure should illustrate the service so that the consumer can simply recognize the services procedure.

Restrictions – states the essential steps to be done in order to supply the required service levels.

Service-level objectives – The service levels that are approved by the customer and the providers. It contains a group of service level indicators such as; availability, performance, and reliability. Each part of the service level, like availability will have

a target level to complete. Service Level objectives have day-time restrictions related to them to describe their validity.

Service-level indicators – those indicators are used to measure these levels of service.

Penalties – describes what is to be done when the provider cannot achieve the goals in the SLA. If the SLA is taken with an external provider, there should be a choice of concluding the contract.

Optional services – services that are not ordinarily needed by the customer, but might be needed as exclusion.

Exclusions – states what is not included in the SLA.

Administration – defines the procedures formed in the SLA to achieve and measure its goals.

SLAs have been utilized for a long period in IT fields to determine the demands of the clients of IT services. An SLA specified the anticipations of the service client and provider. It is frequent for providers to transport services at variable levels of quality depending on the cost of the service. An SLA is precious for assisting all actors to recognize the trade-offs between cost, plan, and quality. Same as any kind of contract, an SLA cannot assure that all commitments will be maintained, but it describes what will take place if those commitments are not met. Guaranteeing the superiority of services supplied over the internet is a large challenge, because the internet is dynamic. Some of these challenges are:

- Low performance of typical protocols.
- Security cases.
- Infrastructure malfunctions.

"An SLA cannot guarantee that you will get the service it describes, any more than a warranty can guarantee that your car will never break down. In particular, an SLA cannot make a good service out of a bad one. At the same time, an SLA can avoid the risk of choosing a bad service". A "good quality" service is one that mitigates the requirements of the client that include goodness and appropriateness. The way to design SLAs is to supply sufficient data or metrics for a client to preselect services depending on the preferred stage of superiority. Usually, SLAs are stated in basic content, using forms or toolkits. Providers design their systems in a way that measurements are gathered and then matched to the metrics determined in the SLA.

There are three major SLA categories:

1. **Basic** – an SLA with well-organized metrics that are calculated and/or confirmed. The gathering of these metrics is usually completed physically.

2. **Medium** – a multi-stage superiority depending on the cost of the service. The goal is to equalize the stages of superiority and cost.

3. **Advanced** – dynamic distribution of resources to achieve requirements.

1.7. Users of the Sla Life Cycle

Domain expert: It is an entity to represent the domain knowledge in the cloud SLA. There are various people involved such as business management, finance authorities, clients of the consumer,

software architects, development and maintenance team, legal experts, security and privacy experts and many more. As cloud SLA spans with many domain concepts, forming ontology is necessary to gather the knowledge. Three SLA ontologies indicating the service domains such SaaS, PaaS and IaaS are developed using protégé editor.

Cloud Consumer: In the pre-negotiation phase, consumer attempt to learn the Cloud Standards Customer Council-CSCC ten step processes by answering the sequence of questions, thus matching the provider SLA in the registry. Monitoring and validation activities require the consumer participation.

Cloud provider: It starts with the creation of SLA templates and registering into the repository. Provider is the core actor involved in almost all the activities in the life cycle.

1.8. SLA Life Cycle

SLA has six main stages to be completed. These stages are as follows; development of both service and SLA templates, discovery and negotiation of an SLA, service provisioning and deployment, execution of the service, assessment and corrective actions during execution, and both termination and decommission of the Service. The SLA lifecycle was described by the Tele Management Forum as shown in Figure 3.

Figure 3: SLA Life Cycle

1.8.1. Development of Service and SLA Templates

This stage includes the identification of customer requirements and needs, the network capabilities, the identification of the suitable service features and parameters, service's levels, service executional environment, and the implementation of the standard of SLA templates.

1.8.2. Discovery and Negotiation of an SLA

Discovery stage consists of; the negotiation of an SLA with the consumer to select the values of SLA parameters related to specific services, the costs gained from the service customer after signing the SLA, the costs incurred by the service provider when the SLA is violated, the definition and at last periodicity of the reports associated with service to be delivered to the service customer.

1.8.3. Service Provisioning and Deployment

This stage include the service's resource provisioning, where the service is enabled and prepared for the service shopper consumption, configuration of the network which might be to achieve specific requirements in the service, or to support the service network overall,

and service activation. Service provisioning and deployment stage may need the reconfiguration of the service resources to support the executional stage which will lead to a successful achievement of the SLA parameters.

1.8.4. Execution of the Service

This stage is the actual test of the service. It consists of three main phases, The first is service execution and monitoring, Then the real time of reporting and at last the validation of QOS which refers to the quality of service. The final phase of this stage is SLA violation processing.

1.8.5. Assessment and Corrective Actions during Execution

SLA assessment stage consists of two parts, the assessment with the individual customer, and the overall service assessment. The SLA assessment of the customer includes reviewing the Quality of Customer Service (QoS), customer gratification, achieving the possible enhancements, and altering requirements are examined for each SLA. Overall service assessment for major activities are readjusting of service goals, service operations modifying, defining the support problems of the service, and finally establishing different service levels.

1.8.6. Termination and Decommission Of the Service

Termination and Decommission of the Service stage in charge with the termination of the service. This termination may be a result of different reasons; it might be an issue in the contract, expiration, or violation. The decommissioning of discontinued services can cause termination to the SLA.

1.9. SLA Content

The structure of service level agreements are generally very scenario specific and can-not be easily generalized. However, there are some basic elements that should be present in every SLA. The following remarks are not intended to be used to create an universal pattern for SLAs, but rather give a guideline for most current contents of SLAs.

The contents of a SLA can be divided into the following four categories: agreement-related elements, service related elements, document-related elements and management related elements.

The agreement-related elements contain the basic rules of the agreement and include, among others, the subject of SLAs, objectives, partners, as well as the scope, entry into force, duration and termination of SLAs. Often these elements are shown in practice in the form of a preamble or introduction. The subject of SLAs introduction here describes the content and context as well as a description and demarcation of the services being controlled by the SLA. The objectives of the SLAs reflect the specific objectives of both parties and serve, among other things, as a basis for future success control.

The service-related elements represent those elements which describe the regulation of a service. These must be specified individually for each service. The content is basically to describe who, when, where, and what services are provided. The description of the service should be generally understandable. The description of the quality of a service is the central role of the SLA. The negotiated quality of service is defined by Key Performance Indicators (KPIs), which is the basis for the Service Level Objectives" (SLOs). These indicators include a label next to the calculation or metric, and a reference area and measurement point. Similarly here, the cost of services to be provided are defined.

Document-related elements include administrative and editorial elements, which play a minor role inside a SLA and are mainly there to improve the handling, understanding and readability. These

elements are, e.g., version, the date of last modification, revision history, table of contents, the index or glossary. These elements increase the readability by underpinning the context and explain the background.

The management-related elements include the aspects that have to do with the administration and control of SLAs. These represent a very important section of the contents of a SLA, since both the customer notification and the procedure in case of problems or failures to meet the service levels are regulated.

Furthermore, penalties and compensation in case of damage which may occur due to deviations from service levels are regulated.

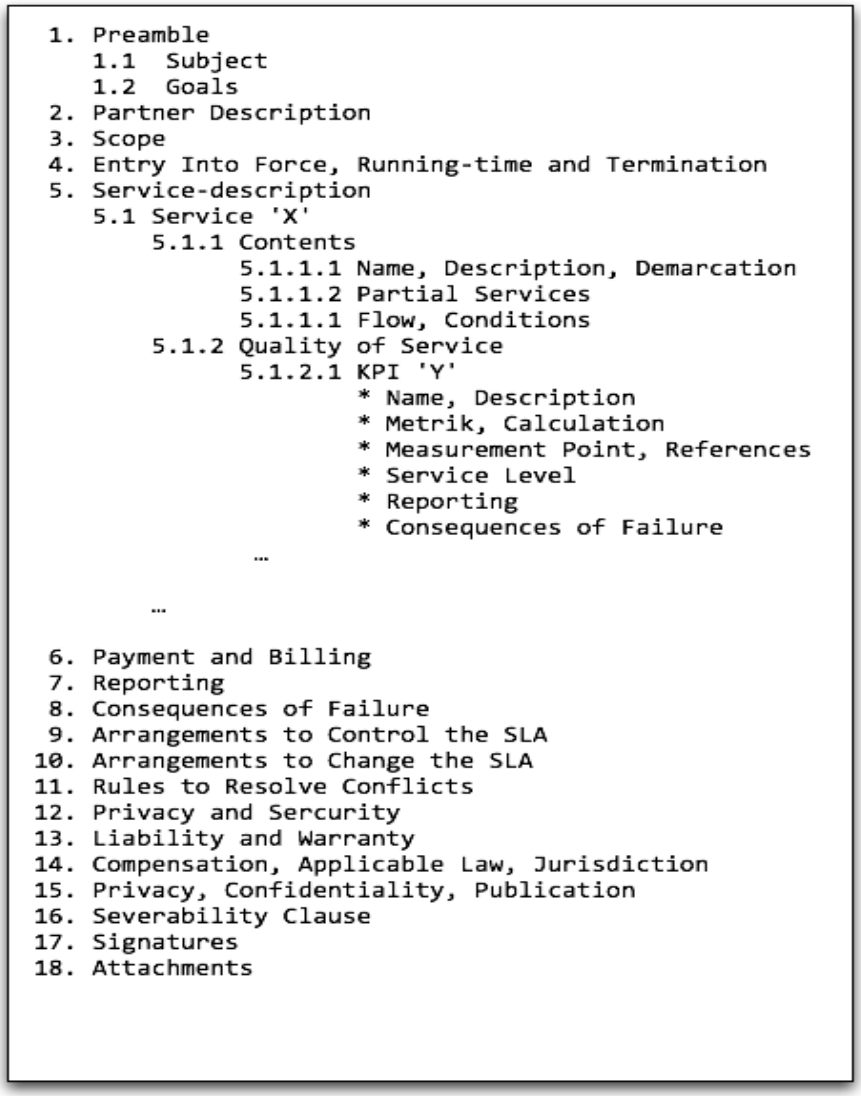

Figure 4: SLA Structure

Based on the presented elements, an exemplary structure of an SLA can be created. This can be seen in Figure 4 above. Here, it is clear that the service descriptions, or service level objectives are the central aspect of each SLA. These and their contents are described in more detail in the following

sections. Likewise, it comes clear that even small SLAs mean large administrative overhead and the creation is a lot of work.

1.10. Cloud SLA Metrics

SLA parameters are determined by metrics. These metrics state how service parameters can be calculated. Also determines estimations of quantifiable parameters. The planned SLA metrics for cloud computing examine the four kinds of cloud services which are (SaaS, PaaS, IaaS, and Storage as a Service). For every branch of the SLA they state the mainly significant parameters that users can utilize to make a consistent form of compromise with the service supplier.

1.10.1. SLA Metrics for IaaS

Firms such as amazon.com supply infrastructure as a service. Many clients don not know clearly which significant parameter must be declared in the hardware side of the SLA. The study mentioned the mainly significant parameters for clients who are concerned in utilizing cloud as an infrastructure service as shown in Table 1.

Parameter	Description
CPU capacity	CPU speed for VM (Virtual Machine)
Memory size	Cash memory size for VM
Boot time	Time for MV to be ready for use
Storage	Storage size of data for short or long term of contract
Scale up	Maximum of VMs for one user
Scale down	Minimum number of VMs for one user
Scale up time	Time to increase a specific number of VMs
Scale down time	Time to decrease a specific number of VMs
Auto scaling	Boolean value for auto scaling feature
Max number can be configured on physical server	Maximum number of VMs that can be run on individual server
Availability	Uptime of service in specific time
Response time	Time to complete and receive the process

Table 1: SLA Metrics for IAAS

1.10.2. SLA Metrics for PaaS

In platform as a service case, developers who exploit PaaS do not need to install instruments or organize hardware to do the developing jobs. For SLA metrics associated to PaaS, the study illustrates the key parameters that can be utilized as an essential principle when developers wish for compromising with PaaS suppliers as shown in Table 2.

Parameter	Description
Integration	Integration with e-services and other platforms.
Scalability	Degree of use with a large number of online users
Pay as you go billing	Charging based on resources or time of service
Environments of deployment	Supporting offline and cloud systems
Browsers	Firefox, Explorer, etc.
Number of developers	How many developers can access to the platform

Table 2: SLA Metrics for PaaS

1.10.3. SLA Metrics for SaaS

Superior examples of SaaS are mail, calendar and social web sites supplied by Google, Yahoo and Microsoft. The study shows the familiar metrics and parameters for SaaS as an illustration of metrics for this kind of cloud service as shown in Table 3.

Parameter	Description
Reliability	Ability to keep operating in most cases
Usability	Easy built-in user interfaces
Scalability	Used with individual or large organizations
Availability	Uptime of software for users in specific time
Customizability	Flexible to use with different types of users

Table 3: SLA Metrics for SaaS

1.10.4. SLA Metrics for Storage as a Service

Online clients enter their information from diverse places. Some time ago, online storage suppliers were not is able to preserve a huge amount of information because there was no enough area in storage disks, network, and information supervision systems. Now, information storage service suppliers like S3 by amazon.com build up big numbers of storage hardware. Also they can handle and provide millions of clients powerfully with their technique of information delivering and guaranteeing that information are suitable for diverse kinds of applications. The parameters for information storage service metrics are fundamental necessities to compromise with storage suppliers as shown in Table 4.

Parameter	Description
Geographic location	Available zones in which data are stored
Scalability	Ability to increase or decrease storage space
Storage space	Quantity of units of data storage
Storage billing	How the cost of storage is calculated
Security	Cryptography for storage, transferring data, authentication, and authorization
Privacy	How the data will be stored and transferred
Backup	How and where images of data are stored
Recovery	Ability to recover data in disasters or failures
System throughput	Amount of data that can be retrieved from system in a specific unit of time
Transferring bandwidth	The capacity of communication channels
Data life cycle management	Managing data in data centers, and using network infrastructure

Table 4. SLA Metrics for Storage as a Service

1.11. SLA Exclusions, and Availability

For a typical customer cloud SLA is non-negotiable and it includes CSPs' promised service availability, service availability determination, and exclusions to service availability. The following table 5 summarizes cloud SLA of Amazon and Microsoft: cloud compute and storage services.

	Amazon		Microsoft	
Cloud product and service	Elastic Compute (EC2)	Simple Storage Service (S3)	Windows Azure Compute	Windows Azure Storage
Availability computation	Annually	Monthly	Monthly	Monthly
Guaranteed availability	At least 99.95%	At least 99.9%	At least 99.95%	At least 99.9%
Availability period interval consideration	5 minutes	5 minutes	none	1 hour
Exclusions to service availability	Factors outside of Amazon's reasonable control(demarcation point)	Factors outside of Amazon's reasonable control(demarcation point)	Factors outside of Microsoft's reasonable control(demarcation point)	Factors outside of Microsoft's reasonable control(demarcation point)

Table 5: Availability related components of cloud SLA of Amazon and Microsoft: cloud

Determination of EC2 Annual Uptime Percentage and S3 Monthly Uptime Percentage exclude downtime resulting directly or indirectly from any Amazon EC2 or S3 cloud SLA exclusions. Amazon's cloud SLA exclusions include among others; unavailability caused by factors outside of

Amazon's reasonable control, including any force majeure event or Internet access or related problems beyond the demarcation point of Amazon EC2.

The following equations show service uptime percentage determinations:

Amazon EC2 Annual Uptime Percentage = 100% - (% of 5 minute periods during the Service Year in which Amazon EC2 was in the state of "Region Unavailable")

Amazon S3 Monthly Uptime Percentage = 100% - (the average of the Error Rates from each five minute period in the monthly billing cycle)

1.12. SLA oriented resource allocation in Cloud computing

There are basically four main entities involved in SLA oriented resource allocation in Cloud computing.

- ✓ **Users/Brokers**
- ✓ **SLA Resource Allocator**
- ✓ **Virtual Machines (VMs)**
- ✓ **Physical Machines**

Users/Brokers: In general, the user interact with the Cloud management systems through an automatic systems such as brokers or schedulers who act on users behalf to submit service requests from anywhere in the world to the Clouds to be processed.

SLA Resource Allocator: The SLA Resource Allocator acts as the interface between the Cloud computing infrastructure and external users/brokers. It requires the interaction of the following mechanisms to support SLA-oriented resource management:

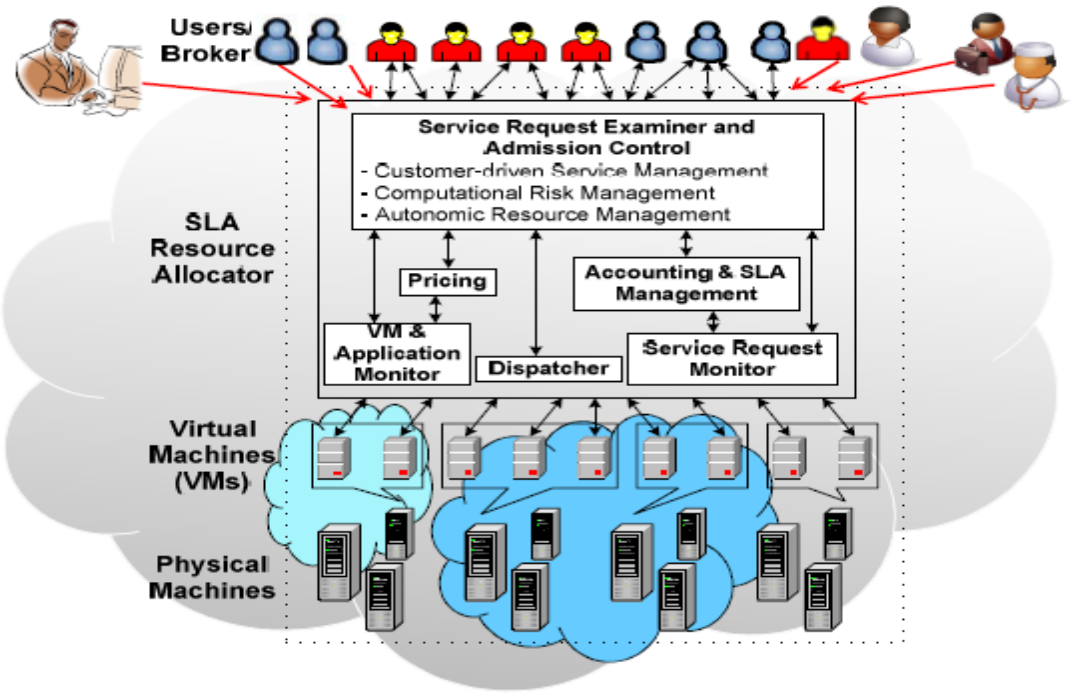

Figure 5: High-level system architectural framework.

o ***Service Request Examiner and Admission Control:*** *The user* service request is first interpreted by the Service Request Examiner and Admission Control mechanism that understands the QoS requirements before determining whether to accept or reject the request. It ensures no SLA violation by reducing the chances of resource overloading whereby many service requests cannot be fulfilled successfully due to limited resources available. Therefore, it also needs the latest status information regarding resource availability (from VM Monitor mechanism) and workload processing (from Service Request Monitor mechanism) in order to make resource allocation decisions effectively. Then, it assigns requests to VMs and determines resource entitlements for allocated VMs.

Autonomic Resource Management*:* This is the key mechanism that ensures that Cloud providers can serve large amount of requests without violating SLA terms. It dynamically manages the resources by using VM migration and consolidation. For instance, when an application requires low amount of resources, its VM is migrated to a host with lower capability, so that new requests can be served.

o ***Pricing*****:** The Pricing mechanism is a way to manage the service demand on the Cloud resources and maximize the profit of the Cloud provider. There are several ways in which service requests can be charged. For instance, requests can be charged based on submission time (peak/off-peak), pricing rates (fixed/changing) or availability of resources (supply/demand). Pricing also serves as a basis for managing computing resources within the data center and facilitates in prioritizing resource allocations effectively. Therefore, Cloud providers offer sometimes same/similar services at different pricing models and QoS levels. The two of the most prominent ones which are practically employed by Cloud providers: posted pricing and spot market.

o ***Accounting and SLA Management*****:** SLA Management is the component that keeps track of SLAs of customers with Cloud providers and their fulfilment history. Based on SLA terms, the Accounting mechanism maintains the actual usage of resources by requests so that the final cost can be computed and charged from the users. In addition, the maintained historical usage information can be utilized by the Service Request Examiner and Admission Control mechanism to improve resource allocation decisions.

o ***VM and Application Monitor*****:** Depending on the services provided, the resource management system has to keep the track of performance and status of resources at different levels. If service provided is compute resources, the VM Monitor mechanism keeps track of the availability of VMs and their resource entitlements. While in the case of application software services, the performance is continuously monitored to identify any breach in SLA and send a notification trigger to SLA Resource Allocator for taking appropriate action.

o ***Dispatcher*****:** The Dispatcher deploys the application on appropriate virtual resource. It also takes the responsibility of creating Virtual machine image and their initiation on selected physical hosts.

o ***Service Request Monitor:*** The Service Request Monitor mechanism keeps track of the execution progress of service requests.

• **Virtual Machines (VMs):** Multiple VMs can be started and stopped dynamically to meet accepted service requests, hence providing maximum flexibility to configure various partitions of resources on the same physical machine to different specific requirements of service requests. In addition, multiple VMs can concurrently run applications based on different operating system environments on a single physical machine since every VM is completely isolated from one another on the same physical machine.

• **Physical Machines:** The data center comprises multiple computing servers that provide resources to meet service demands.

1.13. Future of Cloud SLAs

In this section, we consider how a cloud provider may define SLAs for cloud services in the future.

1.13.1. Service guarantee:

The considered cloud providers only provide uptime guarantees for IaaS services. The cloud providers may also want to offer other guarantees such as performance, security,and ticket resolution time. Providing a performance guarantee becomes necessary if cloud providers oversubscribe the resources of physical servers to decrease the number of physical servers used and increase their utilization. The over-subscription of the physical servers implies that performance of virtual machines running on physical servers may become a concern. Further, co-location of a virtual machine with other workloads may also impact the CPU, disk, network, and memory performance of a VM. Moreover, enterprises purchasing cloud based services may demand a minimal level of performance guarantee. Therefore, it may be necessary for a cloud provider to offer performance based SLAs for its IaaS compute services with a tiered pricing model, and charge a premium for guaranteed performance.

1.13.2. Service guarantee time period and granularity:

The service guarantee time period and granularity determine how stringent is the underlying service guarantee. A service guarantee is stringent if the metric is performance based for a fine-grained resource over a small time period, e.g. 99.9% of memory transactions in a five minute interval must complete within one micro second. Such a stringent guarantee can be loosened by aggregating the service guarantee over a group of resources (e.g., aggregate uptime percentage of all instances must be greater than 99.5%). Providers can use a combination of service guarantee granularity and service guarantee time period to price their services appropriately. For enterprise and mission critical workloads, a cloud provider may have no choice but to provide finer service guarantees.

1.13.3. Service violation detection and credit:

None of the considered providers automatically detect SLA violation and leave the burden of providing the violation proof on the customer. This aspect may not be acceptable to customers with mission critical or enterprise workloads. A cloud provider can differentiate the pricing of its offering if it automatically detects and credits the customer for SLA violation. However, the tooling cost to automatically measure, record, and audit SLA metrics can be a concern.

1.13.4. Outcome based SLAs:

The cloud providers considered in this paper offer IaaS and PaaS services. Using these services, a customer can deploy her own applications in the cloud. However, in the future, cloud providers may offer outcome based services on top of cloud, where a provider delivers a complete solution for a customer using cloud. For outcome based services, a cloud provider needs to define SLAs for the promised outcomes and how those SLAs map to the underlying IaaS and PaaS infrastructure it provides.

1.13.5. Standardization of SLAs:

The lack of standardization in cloud SLAs makes it difficult for a customer to effectively compare them. As cloud services mature, and as the vision of utility computing is realized, the standardization of SLA is likely to take center stage. Structured representation of SLAs (e.g., in XML) may be necessary for standardized SLAs.

2.CHAPTER II

Reconciling metrics between service provider and client to maintain consistency in service: a perspective into service level management (SLM)

2.1. Introduction to Service Level Management (SLM) of Cloud Computing

Service Level Management (SLM) defines, negotiates, controls, reports and monitors agreed-upon service levels within predefined standard service parameters. Usually, effective IT service delivery is considered adequate when system issues are swiftly redressed to the satisfaction of users. An entity's ability to sustain appropriate IT service is heavily dependent on building service commitments and managing service levels.

SLM deployments can flounder because IT management skews service focus towards technology centric measurements specific to categorized domains. Correctively, the IT service department should provide circumspective insight into service levels that management understands. Furthermore, objective achievement should reflect building and measuring service-based contractual arrangements. Not only do service-based negotiations encourage directed dialog between IT and business units, but also promote IT practices unification across configuration items supporting computer applications and business processes.

SLA management is the integrated process of managing various SLAs from start to assessment. SLA management can be categorized into three groups; business level management, service level management, and network level management. The SLA service level management consists of several functions starting with SLA creation, negotiation, provisioning, monitoring, maintenance, reporting and assessment.

As monitor ability is one important SLA requirement. This specific function needs to be further looked into.

Monitor ability denotes that the service provider and the client can observe and manage the behaviour of the service related to the SLA, or employ a trusted third party to do so. Without this requirement, it would be impossible for a party to state that there is an SLA violation. Therefore its terms may be overlooked by the service provider. The problem faced when monitoring compliance with unanimous performance metrics is a big challenge for SLA engineers. The SLA must be designed to guarantee high monitor ability, and decrease the probability of low compliance.

The network level management consists mainly of network monitoring. Network monitoring is the process of the value of the network performance metrics (NPM) by different network monitoring tools and techniques. There are three known methods to monitor a network:

- **The first method** is the active monitoring which is traditionally used to measure loss, connectivity and delay. Active monitoring sends extra traffic between machines after setting up those test machines where measurements is to be taken to obtain the current status of the network. Active monitoring uses simple and easy tools, such as ping and trace-route. The system load is very low with active monitoring because the quantity of generated traffic is small compared with that of other methods. However, the generated test packets may be lost due to their low priority which makes it difficult to obtain the exact network status sometimes.

- **The second method** is the passive monitoring which relies on capturing the packets to obtain the current network status. This is why passive network is ideal to measure NPMs (network performance metrics) like utilization and throughput.

- **The third and the last method** to measure the status of the network is by using SNMP agents. Although this method is practical and simple, it is limited to measure the throughput and the functionality of NPMs.

As easy as it might seem, obtaining different NPMs using the described network monitoring methods above, it rather challenges applying the values obtained directly to QoS parameters. SLAs are constructed in terms of QoS while actual measurements are NPMs. This is why a mapping mechanism is needed and NPMs must be defined before deciding on the QoS parameters in an SLA. The mapping between QoS parameters and the measured NPMs depends mainly on the type of the provided service. It can be complicated and its outcome presentation form should be clear and understandable to the user (in QoS terms not NPMs).

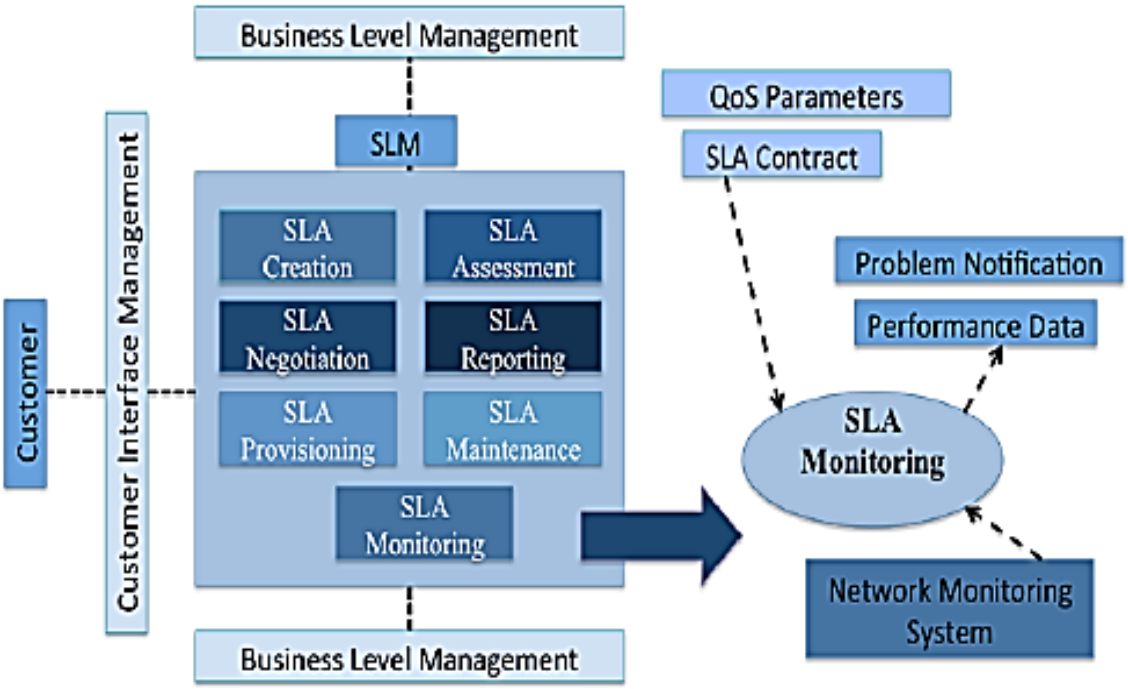

Figure 6: SLM and SLA Monitoring

SLM can be considered QoS **monitoring** and **management** based on key performance indicators (KPIs). QoS KPIs can range from generic availability and usage statistics to entity-centric per-interaction indicators. Adequate SLM requires potential problems identification -- such as gradual performance degradation -- and alerts creation enabling downtime risk minimization. Consequently, SLM practices should include comparing actual performance to pre-defined expectations, determining appropriate actions and generating expressive reports to permit service improvement.

Service monitoring is performed as part of the service level management. Through service monitoring, data related to performance is reclaimed from service resources for each one of the promised services. These reclaimed instances are then collated and integrated to form KQI for both service resource and product KQI as shown in Figure 7.

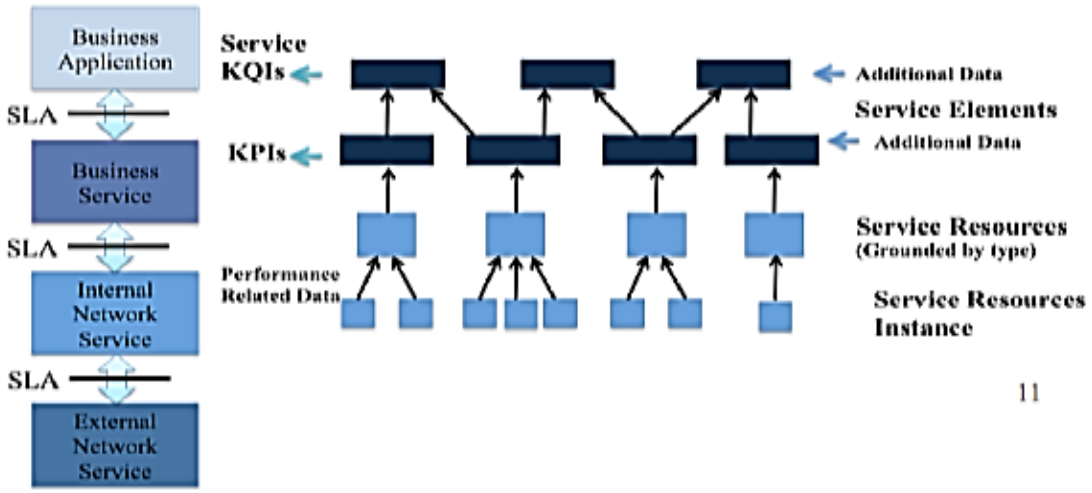

Figure 7: Relationship between Service Resources, KQI, and KPI

Another aspect of the service management process is SLA reporting. It is obvious now that the generated management reports are to be seen by more than one functional group that may include a senior management, SLA engineers, a financial group for handling both charging and billing, and at last end users. This is why the output format of these reports should be appropriate and understandable by all audience. New reporting tools are used and it is applicable to allow users to develop reports on their own.

2.1.1. Applying IT Service Management to Cloud Computing

IT assets are complex to manage and continually change due to the nature of technology and changing business requirements. Effective life cycle management of hardware, software licenses and service agreements; as well as permanent and contracted human resources are critical success factors (CSFs) not only for optimizing the IT cost-base, but also for managing changes, minimizing service incidents and assuring a reliable quality of service (QoS).

As suggested by International Business Machines (IBM), cloud computing enables entities to provision reliable, on-demand services in a flexible and affordable manner; thus, offering the benefits of open standards, scalable systems and service oriented architecture. However, there are potential challenges associated with managing a cloud environment, including:

- Rapid growth of virtualized resources across multiple domains
- Linkage of dynamic resources to underlying IT infrastructure
- Operational monitoring and problem determination across the physical and virtualized infrastructure

Usually the rapid growth of **virtualized resources across multiple domains** begets heightened IT service delivery expectations. To reconcile this perspective, management normally insists on increased quality, functionality and ease of use; decreased deployment time; and continuously improving service levels -- with multilateral cost containment or abatement.

For the entity's IT service delivery personnel, business expectations generally translate into providing appropriate SLM of cloud computing. Typically, SLM is considered the primary IT managerial area that ensures promised services are delivered when and where expected at agreed-upon cost. As with most managerial endeavors, there should be a well formulated plan. Consequently, assisting in actualizing expectations for SLM processes is the Service Quality Plan (SQP) addressing specific managerial objectives.

2.1.2. Defining IT Service Delivery and Support Levels

To enable SLM, customers as well as internal and external suppliers should be identified and managed. For most service providers, cloud computing infrastructure consists of services delivered through central sites utilizing configured servers. Whereby, IT services often appear as single access points to clients.

Descriptively, establishing sound SLM necessitates clear service specifications and interfaces defined with customers (Service Level Requirements (SLRs)). Furthermore, internal Operational Level Agreements (OLAs) and contracts with external suppliers will facilitate adherence to negotiated SLAs.

2.2. The requirements for delivering SLM

For companies already engaged in managed services, or for those who are in some related practice of delivering IT solutions, the decision to offer SLM can be a natural one. As CSPs(Cloud service provider's) become more proactive, the natural tendency should be to offer higher value, higher margin services that will resonate clearly with the client. SLM is one such service. Following are a few of the most vital capabilities for delivering SLM:

• **Tools**. The ability to deliver effective SLM demands that sophisticated tools be used. While basic IT device monitoring and management tools are prerequisites for delivering managed services, SLM requires tools that examine and measure atypical objects within the IT stack. Areas like custom applications, Web servers, portals, extranets and anything else that users interact with are monitoring subjects for SLM. In order to achieve this result, the appropriate tools and processes must be acquired.

• **Processes**. While tools are important, without the right processes in place, SLM can never take place. Once the data is collected, the MSP or IT provider must have a process that will take that information and turn it into action. The process of taking monitoring data and transforming it into a plan that can actually detect and prevent problems from occurring is the essence of managed services and SLM. Only by developing a functional process for extracting and analyzing data can an MSP take preventative action on behalf of their clients. For example, data indicating poor site performance must lead to tests of the implicated objects. Tests must yield more information about how those objects are working (or failing). Proactive steps must then be taken in order to remediate any issues before they affect users. In order to prevent user or customer dissatisfaction, proactive IT management must be employed. In order to accomplish this, every MSP must have a process to follow.

• **Business consulting**. Especially among smaller MSPs, the concept of providing business consulting (or at the very least relevant data) is integral to delivering SLM. The practice of reactive IT management is antithetical to SLM. The delivery of SLM cannot be accomplished without providing a minimal amount of proactive management in order to maintain end-user satisfaction. This notion is lost on many MSPs who view IT management as more of a technical undertaking. It goes without saying that technology is an important component, but without the business data and advice the client is at the mercy of an IT network that may or may not be functioning optimally.

• **SLAs**. SLAs play an important role in the relationship between MSP and client. If an MSP is tasked with providing SLM on a particular application or appliance, the SLA should dictate the level of expectation and performance that can be expected. Quite a few consumers are left on their own when it comes to monitoring SLA performance. Without an MSP or SLM, business executives may never know if their infrastructure is performing as it should.

2.3. SLM ISSUES IN CLOUD COMPUTING

First of all, let us illustrate the relationship of Service Level Agreement (SLA) and SLM. SLA is a statement of service promise to customers, which are measured by service metrics or Service Level Objectives (SLO), and enforced by payments in front of filled promises of and penalties in front of unfilled ones.

SLM is the process through which a SLA is negotiated and service levels are controlled. Specifically, IT Service Management Processes, Operational Level Agreements, and Underpinning Contracts are managed to support the agreed Service Level Targets. SLM monitors and reports Service Levels, through regular Customer reviews. Of course, a robust SLM needs dedicated resources. Specific organization roles and management processes, various data repositories and systems to probe, manage and report performances.

SLM and SLA are to be used jointly (and actually are); SLM provides the way to create SLA, to provision system resources and to manage system performance.

However, a traditional super-structured SLM is rather inconsistent with the dynamics in the cloud, where customers have to select, even on demand, the right service provided among CSPs. So, a mutual trust should be established between customers and CSPs and new SLM issues emerge as we list here below:

2.3.1. Variable price and performance

CSPs deliver services in various forms, e.g. a service with similar functions may have different billing schemas in different CSPs. The service can also be provided on various performances to maximize revenues. Moreover, the same services provided by different CSPs may have different performances and normally overcommit their services in terms of capability and performance. In these cases, customers may be confused in selecting a service complying with their needs.

2.3.2. Untrusted collaboration

Cloud Computing provides a way to establish an IT facility without physical investments. In some cases, CSPs may provide a service of a quality lower than in SLAs. As

a consequence, users may not only lose the control of their IT resources but also stick in a situation where the service paid is not what they should receive.

2.3.3. SLA deviation

Generally, cloud service performance is measured by SLM. Ideally, results should not change regardless the side. However, the calculation may affect the QoS. For instance, the annual performance is normally higher than the monthly one. So, users need a fair-and-square performance measurement.

2.3.4. Negotiation

SLAs is widely used to define QoS. However, if neglecting SLA in SaaS, IaaS and PaaS, users will select cloud services by cost. As a result, the negotiation process of each trade will be either slow or unfair (controlled by CSPs).

2.3.5. Comments

The service portfolio is controlled by CSPs. Therefore, users cannot know the real capability and performance of services. Also, though users can comment service, comments are not published. So in current market, CSPs can sell immature services with unrepeatable or low performance.

3.CHAPTER III

Assessing performance of cloud service provider through service level objectives (SLO)

3.1. Introduction to SERVICE LEVEL OBJECTIVES (SLO)

Service Level Objectives (SLOs) are a central element of every service level agreements (SLA), which include the negotiated service qualities (service level) and the corresponding Key Performance Indicators. SLOs contain the specific and measurable properties of the service, such as availability, throughput or response time and often consist of combined or composed attributes. SLOs should thereby have the following characteristics:

- Achievable / attainable
- Repeatable
- Measurable
- Understandable
- Significant
- Controllable
- Affordable
- Mutually acceptable
- Influential

A SLOs should always contain a target value or service level, a metric and corresponding measurement period, as well as the type and location of the measurement. For this purpose, KPIs with associated service level values are stated. The KPIs contain information about the measurement process, place and unit as well. A valid SLO specification might, for instance, look like this: The IT system should achieve an availability of 98% over the measurement period of one month. The availability represents thereby the ratio of the time in which the service works with a response time of less than 100ms plus the planned downtime to the total service time, measured at the server itself. From such a description, the actual performance values can be compared with the reference values of the SLOs and the achievement is calculated. Based on this, further measures can carried out to for correction if necessary.

To choose the correct KPIs for a service a wide knowledge of the service and its usage is required. To give an insight into possible cloud-specific KPIs, the most common ones are listed briefly below without going into much detail. The following KPIs provide specifically for cloud computing selected guarantees but also may overlap in part with traditional KPIs, as the essential services requirements do not differ from other general services .

A. General Service KPI (key performance indicators)

Service Level Agreements must always be tailored to the service to be controlled. Nevertheless, there are some KPIs, which rules can be used in various SLA. These KPIs represent the basic needs of each service to run efficiently. These include, for example the availability, security aspects, service times and helpdesk, as well as monitoring and reporting. These are basic requirements for every purchased service.

1. Basic Services: The basic services include the availability which is defined at the time the service is usable the maintenance time relative to total time. Deemed usable here is if the system can handle request within a specified response time. Also included are the KPIs Mean Time Between Failure and Mean Time To Repair, which specify the time intervals at which to expect failures and how long it takes to repair them.

2. Security: Security KPIs regulate for example which software version levels shall be used, how long it should take until an update is implemented, as well as the scope and frequency of security

audits. Other important KPIs control the encryption of data, the use and timeliness of anti virus software and the isolation and logging.

3. Service and Helpdesk: Service and Helpdesk KPI control including the times at which assistance is provided, which support methods are applied or how many calls are received per week. Similarly, the qualification of the support personnel and the duration is given to problem solving.

4. Monitoring: Monitoring KPIs to define in which values are determined intervals to monitor and how to handle the resulting reports. The arrangements of these KPIs can be reused in the other categories.

B. Network Service KPIs

Particularly for cloud computing, the network has a strong meaning, as all provided resources and services are available through a network. Here, the network has to be considered both as pure transmission medium for other services as well as independent service itself. For the KPIs described here, the entry point of the provider network is usually chosen as measured point, as the guarantees of the provider refer only to this area.

Round Trip Time: Time of a network packet to travel from sender to receiver and back. Specifies how long the transmission of one packet needs within the network limits. Usually measured in milliseconds.

Response Time: Time taken by a request until the arrival of the response at the requesting interface. Here the time for the processing of the request is included as opposed to the pure orbital period of the round trip time. The type of the request and the behaviour of the processing have to be concretely defined for this.

Packet Loss: Percentage of lost packets in the total of transmissions. Formula:

$$\frac{Number\ \ of\ \ lost\ \ packets}{Number\ \ of\ \ total\ \ packets\ \ *\ \ 100} \qquad (1)$$

The value of this indicator should kept as low as possible since for example and a loss rate of 5% to 10% significantly affects the quality of VoIP applications.

Bandwidth: Gross capacity of the connection. Amount of data which cloud be transmitted within a time unit. Here, not the actual capacity is specified but the rated maximum capacity.

Throughput: Number of transmitted data per time unit.Only the pure transmitted data is taken into account, thus the capacity available to the user is specified. Measured in Mbit/s or / Gbit/s.

Network Utilization: Proportion of the throughput to the bandwidth. Here, it can be seen how busy the connection is.

Formula:

$$\frac{Throughput}{Bandwidth\ \ *\ \ 100} \qquad (2)$$

Latency: Time interval between submitting a packet and arrival at its destination. Is usually considered together with Jitter: The difference in the latency of a packet and the average / minimum / maximum run time. The run time variations are problematic especially in real-time applications, since packages may arrive too late or too early.

C. *Cloud Storage KPIs*

The term storage can be distinguished within cloud computing in two basic types. First, Storage as a service itself, that's obtained as a memory for pre-existing infrastructures. On the other hand storage can be used as part of another service such as a backup or data storage for cloud services.

Response Time: Time interval between sending a request to the storage and the arrival of the response at the output interface. Usually measured in milliseconds.

Throughput: Number of transmitted data per time unit. Here, a specified amount of data is transferred to the storage and measured the needed time from a given point. The size of the data set and package sizes are important factors for the validity of this measure. Furthermore, the network and its utilization must be considered.

Average Read Speed: In contrast to the throughput, the average reading speed usually refers to an individual hard drive. This value indicates how fast data can be read from the hardware. In RAID systems or virtual storage solutions, this figure is expected to interconnected hard drives.

Average Write Speed: Just like the reading speed it refers to the write speed to the hard drive. This value thus indicates how quickly data can be written from a source to the hardware.

Random Input / Outputs per second (IOPS): Number of possible random input / output operations per second for different block sizes. The higher the IOPS value, the faster the disk. This value is also important to measure how many concurrent accesses can be handled by the system.

Sequential Input / Outputs per second (IOPS): Number of possible sequential input / output operations per second for different block sizes. Free Disk Space Usable free capacity in % of the total capacity or remaining free space in MB, GB, or TB. This indicator can be very useful since thus it can be defined how much memory must always be at minimum available on the system.

Provisioning Type: Type of provisioning where at "thin provisioning" the client gets the storage not permanently assigned but it is dynamically allocated at runtime. In contrast, the thick provisioned storage is allocated to the customer immediately.

Average Provisioning Time: Time, the provider needs to provide a defined amount of data volume growth.

D. *Backup and Restore KPIs*

Backup and Restore KPIs refer to the storage, i.e., the stored data, as well as services, for example, VMs or SaaS services. Below, important KPIs are presented.

Backup Interval: The time interval in which a backup is performed. Here, an exact specification is given to the provider along with the backup type and a description of the scoop.

Backup Type: Definition of the backup type, e.g., full backup or incremental backup. Backup types ca relate to individual systems or whole service alliances.

E. Infrastructure as a Service KPIs

Infrastructure as a Service refers not only to the service itself but also to the virtual machines used. For this, additional VM KPIs are specified in this section.

VM CPUs Number and type of CPUs used by the virtual machine. Additionally information about the overbooking of the provided CPU resources shall be given. Here the shared resources are allocated with more capacity than is physically available. Thus, no real physical allocation of resources takes place. Actual performance is dependent on the overall consumption of the system.

CPU Utilization Proportion of CPU resources in use to the total number of resources provided per time unit. Also the CPU queue, which indicates the number of open requests to the CPU should be considered.

VM Memory Amount and type of the provided memory. This may relate to physical memory or virtual memory. Information about the overbooking of allocated memory resources should be stated.

Memory Utilization Proportion of the memory resources used to the total amount of memory made available to the VM.

Minimum Number of VMs Guaranteed number of the provided VMs with the specified specs stated in the previous points.

Migration Time :Time that is needed to move a VM from two predefined resources.

Migration Interruption Time Maximum time in which a customer has no access to migration to the resource.

Logging Retention of log data. Specifies how long log data to be stored by the provider and specification of what level to be logged. (e.g., INFO, DEBUG, etc.)

3.2. Performance Service Level Objectives(SLO) Overview

This section covers the common service level objectives that relate to the performance of the cloud service and the performance of related aspects of the interface between the cloud service customer and the cloud service provider. The set of service level objectives is not exhaustive, but not all the service level objectives are applicable to all cloud services.

3.2.1. Availability

Description of the context or of the requirement

Availability is the property of being accessible and usable upon demand by an authorized entity.

Description of the need for SLOs, in addition to information available through certification

Availability is usually covered by certification at a general level. Availability is a key service level objective, since it describes whether the cloud service can actually be used, and it is

typically necessary to specify numeric values for availability to make meaningful statements that are useful for cloud service customers.

The question of what "usable" means is a complex matter, which depends on the cloud service concerned. A service can be up and available, but perform so poorly that it is effectively unusable. Similarly, the service can be up, but respond with errors for valid requests. It can be valuable for the SLA to provide clear information on these aspects of service availability.

Description of relevant SLOs

Level of uptime (Often termed "availability")	describes the time in a defined period the service was available, over the total possible available time, expressed as a percentage.[15] Some cloud services specify that the service will be unavailable for specified periods for maintenance. It is common for the stated level of uptime to exclude these maintenance periods. In this case Uptime = Total Possible Available Time – (Total Downtime – Maintenance Downtime).
Percentage of successful requests	describes the number of requests processed by the service without an error over the total number of submitted requests, expressed as a percentage.
Percentage of timely service provisioning requests	describes the number of service provisioning requests completed within a defined time period over the total number of service provisioning requests, expressed as a percentage. Provisioning of cloud services may vary greatly depending on the type of service being considered – from storage provisioning to user account provisioning. It is thus expected that this SLO will need to be tailored to the particular service being considered.

3.2.2. Response Time

Description of the context or of the requirement

Response time is the time interval between a cloud service customer initiated event (stimulus) and a cloud service provider initiated event in response to that stimulus. The response time SLOs can vary depending on the point at which the customer stimulus is measured. For example, the measurement may start from when the customer initiates the stimulus on their device, or it may start from the point where when the request from the customer arrives at the cloud service provider's endpoint – the difference being the network transit time, which may be outside the control of the cloud service provider. Similarly, the point at which the response is measured can vary.

Description of the need for SLOs, in addition to information available through certification

Response time can be a highly significant aspect of the user experience of a cloud service – for some requests; response times that are greater than some threshold are regarded as unacceptable and can make the cloud service effectively unusable. Rarely are response times dealt with directly by certifications and furthermore, response times can vary depending on the nature of the request concerned or the type of the service being considered.

A factor that needs to be considered is that many cloud services support multiple different operations and that it is likely that the response time will differ for the different operations. As a result, response time SLOs need to clearly state which operation(s) are concerned.

Description of relevant SLOs

Average response time	refers to the statistical mean over a set of cloud service response time observations for a particular form of request.
Maximum response time	refers to the maximum response time target for a given particular form of request.

3.2.3. Capacity

Description of the context or of the requirement

Capacity is the maximum amount of some property of a cloud service. It is often an important value for cloud service customers to know when using a cloud service. The relevant properties vary depending on the capabilities offered by the cloud service and it is often the case that multiple different capacities are relevant for a given cloud service.

Description of the need for SLOs, in addition to information available through certification

Capacities are rarely the subject of certification and must be stated clearly in the SLA for a cloud service. Note that capacity SLOs refer to the capacities as seen by an individual cloud service customer and do not reflect the overall capacities supported by the cloud service provider – indeed it is commonly the case that the customer can change the capacity limits for their cloud service(s) by requesting a change in their subscription.

Description of relevant SLOs

There are a number of SLOs, which relate to the capacity of a cloud service

Number of simultaneous connections	refers to the maximum number of separate connections to the cloud service at one time.
Number of simultaneous cloud service users	refers to a target for the maximum number of separate cloud service customer users that can be using the cloud service at one time.
Maximum resource capacity	refers to the maximum amount of a given resource available to an instance of the cloud service for a particular cloud service customer. Example resources include data storage, memory, number of CPU cores.
Service Throughput	refers to the minimum number of specified requests that can be processed by the cloud service in a stated time period. (e.g. Requests per minute).

3.2.4. Capability Indicators

Description of the context or of the requirement

Capability indicators are service level objectives which promise specific functionality relating to the cloud service.

Description of the need for SLOs, in addition to information available through certification

Capabilities can be essential to the use of the cloud service from the perspective of the cloud service customer.

Description of relevant SLOs

External connectivity	specifies capabilities of the cloud service to connect to systems and services which are external to the cloud service. The systems and services involved may be other cloud services or they may be outside cloud computing (e.g. in-house customer systems).

3.2.5. Support

Description of the context or of the requirement

Support is an interface made available by the cloud service provider to handle issues and queries raised by the cloud service customer.

Description of the need for SLOs, in addition to information available through certification

Support capabilities may be required by certification, but the details are typically not covered by certification and must instead be described by SLOs.

Description of relevant SLOs

Support hours	specifies the hours during which the cloud service provider provides a cloud service customer support interface that accepts general inquiries and requests from the cloud service customer.
Support responsiveness	specifies the maximum time the cloud service provider will take to acknowledge a cloud service customer inquiry or request. It is typical for responsiveness to vary depending on a severity level which is attached to the customer request, with a shorter response time associated with higher severity levels.[16]
Resolution time	refers to the target resolution time for customer requests – in other words, the time taken to complete any necessary actions as a result of the request. This target time can vary depending on the severity level of the customer request, with shorter times attached to requests of higher severity.

3.2.6. Reversibility and the Termination Process

Description of the context or of the requirement

The termination process takes place when a cloud service customer or a cloud service provider elect to terminate the agreement. The termination process includes a series of steps which enable the customer to retrieve their cloud service customer data within a stated period

of time before the cloud service provider deletes the cloud service customer data from the provider's systems (including backup copies, which may be done possibly on a different schedule). The cloud service provider can potentially delete or aggregate any cloud service derived data (this is limited to derived data related to operations) that relates to the customer and their use of the cloud service, although such deletion may be limited in scope.

Description of the need for SLOs, in addition to information available through certification

Certification may require a well-defined termination process but does not typically define aspects such as the time periods involved.

Description of relevant SLOs

Data retrieval period	specifies the length of time in which the customer can retrieve a copy of their cloud service customer data from the cloud service.
Data retention period	refers to the length of time which the cloud service provider will retain backup copies of the cloud service customer data during the termination process (in case of problems with the retrieval process or for legal purposes). This period may be subject to legal or regulatory requirements, which can place lower or upper bounds on the length of time that the provider can retain copies of cloud service customer data.
Residual data retention	refers to a description of any data relating to the cloud service customer which is retained after the end of the termination process – typically this will be cloud service derived data, which could be subject to regulatory controls.

3.3. Security Service Level Objectives Overview

Specifying measurable security level objectives in SLAs is useful to improve both assurance and transparency. At the same time, it allows for establishing a common semantics in order to manage cloud security from two perspectives, namely (i) the security level being offered by a cloud service provider and, (ii) the security level requested by a cloud service customer.

The approach used in this section consists of analysing security controls from well-known frameworks into one or more security SLOs, when appropriate. These SLOs can be either quantitative or qualitative. This section focuses on the definition of possible security SLOs. Eight categories are provided, each with one or more SLOs.

The categories are representative of some important security requirements. However not all security requirement categories are reflected below, as relevant SLOs may not exist for each of them. For example resilience, business continuity and disaster recovery are important aspects of security, specific controls and measures are usually put in place by CSPs, but no SLO has been derived for these security aspects.

For each category, the SLOs are meant to provide more quantitative and qualitative information relevant to a specific control, in addition to what is usually assessed in the context of an audit for a certification .

It should be noted that the list of SLOs is not meant to be considered as exhaustive and that the SLOs proposed are not meant to be considered as applicable in all individual cases. The applicability of a particular SLO can depend on the type of service offered (in terms of both of service functionally and service model) and pricing of it (free service, paid, premium). It is important to understand that some of the SLOs relevant to security also have relevance in the areas of Data Management, Performance and Data Privacy .

3.3.1. Service Reliability

Description of the context or of the requirement

Service reliability is the property of a cloud service to perform its function correctly and without failure, typically over some period of time. This category is usually related to the security controls implementing business continuity management and disaster recovery in frameworks like ISO/IEC 27002. Allowable downtime, which accounts for scheduled maintenance and any other element carved out in the agreement, should be taken into account for this SLO.

Note that reliability also covers the capability of the cloud service to deal with failures and to avoid loss of service or loss of data in the face of such failures.

Description of the need for SLOs, in addition to information available through certification

Reliability is sometimes covered by certification, but the target for reliability needs to be stated so that the cloud service customer can assess whether the particular cloud service meets their business requirements.

Description of relevant SLOs

Level of redundancy	describes the level of redundancy of the cloud service supply chain, possibly taking into account the percentage of the components or service that have fail over mechanism. Redundancy varies also on the type of cloud service provided (IaaS versus SaaS for example)).
Service reliability	describes the ability of the cloud service to perform its function correctly and without failure over some specified period.

3.3.2. Authentication & Authorization

Description of the context or of the requirement

Authentication is the verification of the claimed identity of an entity (typically for cloud computing the entity is a cloud service user). Authorization is the process of verifying

that an entity has permission to access and use a particular resource based on predefined user privileges. Authentication and authorization are key elements of information security which apply to the use of cloud services.

Description of the need for SLOs, in addition to information available through certification

Certification generally validates that authentication and authorization mechanisms are in place for a system, but do not in general provide details of how those mechanisms are provided, which can be essential information for the cloud service customer.

Description of relevant SLOs

User authentication and identity assurance level	measures the Level of Assurance (LoA) of the mechanism used to authenticate a user accessing a resource. The LoA can be based on relevant standards like NIST SP 800-63 (Electronic Authentication Guidelines), ISO/IEC 29115 (Entity Authentication Assurance Framework) or the Kantara Initiative's Identity Assurance Framework (IAF).
Authentication	specifies the available authentication mechanisms supported by the CSP on its offered cloud services. In some cases the customer might need to analyse along with the CSP, those mechanisms allowing interoperability among their authentication schemes (e.g., cross-certification in the case of digital certificate-based authentication).
Mean time required to revoke user access	is the arithmetic average of the times required to revoke users' access to the cloud service on request over a specified period of time.
User access storage protection	describes the mechanisms used to protect cloud service user access credentials
Third party authentication support	specifies whether third party authentication is supported by the cloud service and defines which technologies can be used for third party authentication[18]. This SLO complements the previously defined "Authentication", and is the basis for interoperable authentication/identity management solutions between customer and providers.

3.3.3. Cryptography

Description of the context or of the requirement

Cryptography is a discipline which embodies principles; means and methods for the transformation of data in order to hide its information content prevent its undetected modification and/or prevent its unauthorized use. Also known by the term encryption.

Description of the need for SLOs, in addition to information available through certification

While many certification approaches require the use of data encryption in a variety of circumstances, there are many encryption methods in use and these methods vary in their strength and also vary in their cost - either in terms of performance or of the necessary processing power to use them. It is necessary for the SLA to describe specifics relating to encryption methods in order for the cloud service customer to evaluate a cloud service fully, since few certifications require the use of specific encryption methods.

Description of relevant SLOs

Cryptographic brute force resistance	expresses the strength of a cryptographic protection applied to a resource based on its key length, for example using the ECRYPT II security level recommendations [19] or the FIPS security levels[20] for encryption. Instead of using key lengths alone, which are not always directly comparable from one algorithm to another, this normalizing scale allows comparison of the strengths of different types of cryptographic algorithms.
Key access control policy	describes how strongly a cryptographic key is protected from access, when it is used to provide security to the cloud service (or assets within the cloud service).
Cryptographic hardware module	describes the level of protection that is afforded to cryptographic operations in the cloud service through the use of cryptographic hardware

3.3.4. Security Incident management and reporting

An information security incident is a single or a series of unwanted or unexpected information security events that have a significant probability of compromising business operations and threatening information security. Information security incident management are the processes for detecting, reporting, assessing, responding to, dealing with, and learning from information security incidents.

Description of the need for SLOs, in addition to information available through certification

How information security incidents are handled by a cloud service provider is of great concern to cloud service customers, since an information security incident relating to the cloud service is also an information security incident for the cloud service customer.

Description of relevant SLOs

Percentage of timely incident reports	describes the defined incidents to the cloud service which are reported to the customer in a timely fashion. This is represented as a percentage by the number of defined incidents reported within a predefined time limit after discovery, over the total number of defined incidents to the cloud service which are reported within a predefined period (i.e. month, week, year, etc).
Percentage of timely incident responses	describes the defined incidents that are assessed and acknowledged by the cloud service provider in a timely fashion. This is represented as a percentage by the number of defined incidents assessed and acknowledged by the cloud service provider within a predefined time limit after discovery, over the total number of defined incidents to the cloud service within a predefined period. (iei.e. month, week, year, etc)).
Percentage of timely incident resolutions	describes the percentage of defined incidents against the cloud service that are resolved within a predefined time limit after discovery.

3.3.5. Logging and Monitoring

Description of the context or of the requirement

Logging is the recording of data related to the operation and use of a cloud service. Monitoring means determining the status of one or more parameters of a cloud service. Logging and monitoring are ordinarily the responsibility of the cloud service provider.

Description of the need for SLOs, in addition to information available through certification

Log file entries are important to cloud service customers when analysing incidents such as security breaches and service failures as well as in monitoring the customer's day-to-day use of the service. It is necessary for there to be service level objective relating to logging and monitoring in order to fully describe the cloud service and its related capabilities.

Description of relevant SLOs

Logging parameters	describes the parameters that are captured in the cloud service log files .
Log access availability	describes what log file entries the cloud service customer has access to.
Logs retention period	describes the period of time during which logs are available for analysis (e.g. the period of time that log files are available for use by the cloud service customer)).

3.3.6. Auditing and security verification

Description of the context or of the requirement

Auditing is the systematic, independent and documented process for obtaining audit evidence about a cloud service and evaluating it objectively to determine the extent to which the audit criteria are fulfilled. The audit evidence required and the audit criteria are usually

determined by the audit scheme or certification scheme which is used to perform the audit. Certification is one of many ways to address audits.

Description of the need for SLOs, in addition to information available through certification

Audits are a means by which the cloud service provider can offer independent evidence that a cloud service meets particular criteria of interest to the cloud service customer – aiming to increase trust in the cloud service.

Description of relevant SLOs

Certifications applicable	refers to a list of certifications held by the cloud service provider for a cloud service, including the certifying body, the expiration date of each certification and the renewal period[21].

3.3.7. Vulnerability Management

Description of the context or of the requirement

Vulnerability is a weakness in an information system, system security procedures, internal controls, or implementation that could be exploited or triggered by a threat.

Management of vulnerabilities means that information about technical vulnerabilities of information systems being used should be obtained in a timely fashion, the organization's exposure to such vulnerabilities evaluated and appropriate measures taken to address the associated risk.

Description of the need for SLOs, in addition to information available through certification

Many of the information systems associated with a cloud service belong to the cloud service provider with the result that the cloud service customer is dependent on the provider for appropriate and timely management of vulnerabilities of those systems. SLOs for vulnerability management provide transparency for the customer.

Description of relevant SLOs

Percentage of timely vulnerability corrections	describes the number of vulnerability corrections performed by the cloud service provider, and is represented as a percentage by the number of vulnerability corrections performed within a predefined time limit, over the total number of vulnerability corrections to the cloud service which are reported within a predefined period (i.e. month, week, year, etc).
Percentage of timely vulnerability reports	describes the number of vulnerability reports by the cloud service provider to the cloud service customer, and is represented as a percentage by the number of vulnerability reports within a predefined time limit, over the total number of vulnerability reports to the cloud service which are reported within a predefined period (i.e. month, week, year, etc).
Reports of vulnerability corrections	is a description of the mechanism by which the cloud service provider informs the customer of vulnerability corrections applied to the provider's systems, including the frequency of the reports.

3.3.8. Governance

Governance is system by which cloud service is directed and controlled. The main area of concern is the way in which changes and updates to a cloud service are managed, whether the change request originates with the cloud service customer or originates with the cloud service provider.

3.3.8.1. Service changes

Description of the context or of the requirement

Cloud services may change from time to time. Examples of service changes include changes to functionality, changes to the service's interfaces and the application of software updates. Change to a particular service can be reflected in the SLA or in another contractual document.

Description of the need for SLOs, in addition to information available through certification

Cloud service customers need a reasonable notification period before changes to a cloud service take effect so that they can plan appropriately.

Description of relevant SLOs

Cloud service change	describes the type of change (such as SLA change or functional change),
reporting notifications	mechanism and period for the cloud service provider to notify cloud service customers of planned changes to the cloud service.
Percentage of timely cloud service change notifications	The number of change notifications made within a specified period of time over the total number of change notifications, expressed as a percentage.

3.4. Data Management Service Level Objectives Overview

As companies transition to cloud computing, the traditional methods of securing and managing data are challenged by cloud-based architectures. Elasticity, multi-tenancy, new physical and logical architectures, and abstracted controls require new data security strategies. Managing data and information in the era of cloud computing can affect all organizations. It begins with managing internal data and cloud migrations and extends to securing information in diffuse, cross-organization applications and services.

The data management SLOs presented in this section cope with important quantitative and qualitative indicators related with data life cycle management, and can be considered as complementary to existing and applicable security and data protection certifications offered by the cloud service provider.

Presented data management SLOs are subdivided in four (4) different top-level categories covering all aspects of the identified data life-cycle. Each category is subdivided in one of more SLOs that are applicable to that specific category. Not all SLOs may be relevant for each cloud service, in particular depending on the type of cloud service such as IaaS, PaaS or SaaS.

3.4.1. Data classification

Description of the context or of the requirement

Data classification is a description of the classes of data which are associated with the cloud service:

- cloud service customer data
- cloud service provider data
- cloud service derived data

Cloud service customer data is a class of data objects under the control of the cloud service customer. Cloud service customer data includes data input into the cloud service by the cloud service customer and the results of the cloud service customer's use of the cloud service, unless the master service agreement specifically defines a different scope.

Description of the need for SLOs, in addition to information available through certification

The following SLOs contain a specific list of data uses (provider and derived), that can be applied to compare different CSP offers in a concrete manner. This information is usually difficult to deduce in such a specific and concrete way from relevant security/data protection certifications. Customers should use this information to make informed decisions about their choice of CSP – e.g. are the CSP's listed "customer data uses" compliant with my requirements?

Description of relevant SLOs

Cloud service customer data use by the provider	describes stated policy for any intended use of cloud service customer data
Cloud service derived data use	describes what derived data is created by the cloud service provider from cloud service customer data, the intended uses for the derived data and what rights the cloud service customer has to inspect the derived data

3.4.2. Cloud Service Customer Data Mirroring, Backup & Restore

Description of the context or of the requirement

This SLO category deals with the actual mechanisms used to guarantee that the customers' data is available (online or offline) in case of failures forbidding access to it. The mechanisms falling under the scope of this SLO are divided in two widely-used categories (i) data mirroring, (ii) backup/restore.

Description of the need for SLOs, in addition to information available through certification

Widely used security certification[22] contains specific security controls that are implemented to avoid data loss. However, in many cases the information that can be extracted from those certifications rarely contains the basic measurements that can be used by the cloud service customer to assess/monitor if the implemented data security controls actually fulfil her requirements. In particular with refer to SLOs in the following areas:

- The timeliness of the mirroring mechanisms, which might be directly related with the geographical location of the cloud service provider's data centres,
- Concrete details related with to the frequency and method used by the cloud service provider's backup and recovery mechanism(s).

Proposed SLOs allow customers e.g., to fine-tune their risk assessment and business continuity procedures.

The SLOs can assist the cloud service customer in putting in place Recovery Point Objective and Recovery Time Objective when using the cloud service.

Recovery Point Objective is the maximum allowable time between recovery points. RPO does not specify the amount of acceptable data loss, only the acceptable time window. In particular, RPO affects data redundancy and backup. A small RPO suggests mirrored storage of both transient and persistent data while a larger window allows for a periodic backup approach. As with RTO, cloud service customers should determine their acceptable RPO for each cloud service they use and ensure that the cloud service provider's and their own disaster recovery plans meet their objectives.

Recovery Time Objective is the maximum amount of time a business process may be disrupted, after a disaster, without suffering unacceptable business consequences. Cloud services can be critical components of business processes. Cloud service customers must determine the RTO for each of their cloud service dependent business processes and likewise determine whether the cloud service provider's and the cloud service customer's disaster recovery plans are sufficient

Description of relevant SLOs

Data Mirroring	refers to the difference between the time data is placed on primary

Latency	storage and the time the same data is placed on mirrored storage.
Data Backup Method	refers to a list of method(s) used to backup cloud service customer data.
Data Backup Frequency	refers to the period of time between complete backups of cloud service customer data.
Backup Retention Time	refers to the period of time a given backup is available for use in data restoration .
Backup Generations	refers to the number of backup generations available for use in data restoration.
Maximum Data Restoration time	refers to the committed time taken to restore cloud service customer data from a backup.
Percentage of Successful Data Restorations	refers to the committed success rate for data restorations, expressed as the number of data restorations performed for the customer without errors over the total number of data restorations, expressed as a percentage.

3.4.3. Data Lifecycle

Description of the context or of the requirement

The following list of SLOs is related with the efficiency and effectiveness of the provider's data-life cycle practices, with a particular focus on the practices and mechanisms for data handling and deletion.

Description of the need for SLOs, in addition to information available through certification

Despite widely-used security certifications schemes usually deal with the topic of secure disposal24 usually the CSP-specific information related with the deletion and storage controls is not easy to extract. On one hand, the following list of SLOs provides information related with the assurance and timeliness associated with the deletion mechanism. On the other hand, are also presented quantitative SLOs associated with the reliability of the storage service (data retrievability and stored data's durability). Furthermore it may be of interest for the cloud service customer to be able to retrieve data after a deletion request has been posted and to have SLOs associated with that. Cloud service customers are expected to use the following list of SLOs to e.g., decide on the choice of available cloud storage mechanisms offered by the CSP.

Description of relevant SLOs

Data deletion type	describes the quality of data deletion, ranging from 'weak' deletion where only the reference to the data is removed, to 'strong' sanitization techniques to ensure that deleted data cannot be easily recovered[25].
Percentage of timely effective deletions	refers to the number of cloud service customer data deletion requests completed within a predefined time limit over the total number of deletion requests, expressed as a percentage.
Percentage of tested storage retrievability	refers to the amount of cloud service customer data that has been verified to be retrievable during the measurement period, after the data has been deleted.

3.4.4. Data Portability

Description of the context or of the requirement

The following list of SLOs is related with the CSP capabilities to export data, so it can still be used by the customer e.g., in the event of terminating the contract.

Description of the need for SLOs, in addition to information available through certification

In related security controls frameworks and certifications the implementation of data portability controls usually focuses on the specification of applicable CSP policies, which makes it difficult (and sometimes impossible) for cloud service customers to extract the specific indicators related with available formats, interfaces and transfer rates. The following list of SLOs focuses on these three basic aspects of the CSP data portability features, which can be used by the customer e.g., to negotiate the technical features associated with the provider's termination process.

Description of relevant SLOs

Data portability format	specifies the electronic format(s) in which cloud service customer data can be transferred to/accessed from the cloud service.
Data portability interface	specifies the mechanisms which can be used to transfer cloud service customer data to and from the cloud service. This specification potentially includes the specification of transport protocols and the specification of APIs or of any other mechanism that is supported.
Data transfer rate	refers to the minimum rate at which cloud service customer data can be transferred to/from the cloud service using the mechanism(s) stated in the data interface.

3.5. Personal Data Protection Service Level Objectives Overview

This paragraph focuses on the definition of appropriate SLOs with reference to the cases where the cloud service provider acts as a data processor, on behalf of its customer (data controller), which typically applies to B2B services26. Providers that act as data controllers or joint controllers (notably by processing personal data for their own purposes, outside of an explicit mandate from the customer) may still make reference to this document, but they and their customers need to ensure compliance with legal obligations that may derive from their controller role.

Besides, this paragraph concentrates on data protection measures that are suitable for being translated into SLOs, i.e. into objectives that must be achieved by the provider. Other data protection measures and obligations can be better managed through other instruments, such as adherence to a code of conduct, certification against an approved standard and the relevant contract and/or service agreement and applicable law.

In this context, it should be mentioned that there is on-going initiative of the C-SIG Code of Conduct Subgroup on the Data Protection Code of Conduct for cloud service providers27. In order to align both initiatives, this paragraph of the SLA Standardization Guidelines will be revised and updated after receiving the approval of the final version of the Code from the Article 29 Working Party.

3.5.1. Codes of conduct, standards and certification mechanisms

Description of the context of the requirement

The cloud service customer, as data controller, must accept responsibility for abiding by the applicable data protection legislation. Notably, the cloud service customer has an obligation

to assess the lawfulness of the processing of personal data in the cloud and to select a cloud service provider that facilitates compliance with the applicable legislation.

In this regard, the cloud service provider should make available all the necessary information, also in adherence to the principle of transparency, as described hereinafter. Such information includes information that may assist in the assessment of the service, such as the data protection codes of conduct, standards or certification schemes that the service complies with.

Description of the need for SLOs, in addition to information available through certification, adherence to codes of conduct, etc.

In the context of the above mentioned obligations, the information indicated hereinafter is useful in order to let the customer assess the cloud service's level of compliance with the applicable regulatory framework.

3.5.2. Purpose specification

Description of the context of the requirement

The principle of purpose specification and limitation requires that personal data must be collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes. Therefore, the purposes of the processing must be determined, prior to the collection of personal data, by the data controller, who must also inform the data subject thereof.

When the data controller decides to process the data in the cloud, it must be ensured that personal data are not (illegally) processed for further purposes by the cloud service provider, or one of his subcontractors.

Description of the need for SLOs, in addition to information available through certification, adherence to codes of conduct, etc.

In general, the cloud service provider may not process personal data, pursuant to the service agreement with its customer, for its own purposes, without the express permission of the customer. Otherwise, a cloud service provider that process the customers' personal data for its own purposes outside of an explicit mandate from its customers (e.g. in order to do market analysis or scientific analysis, to profile data subjects, or to improve direct marketing, all for its own account), will qualify as a data controller in its own right and must fulfil all the relevant obligations.

It is therefore important that the list of processing purposes (if any), which are beyond those requested by the customer, is defined.

3.5.3. Data minimization

Description of the context of the requirement

The cloud service customer is responsible for ensuring that personal data are erased (by the provider and any subcontractors) from wherever they are stored as soon as they are no longer necessary for the specific purposes.

Furthermore temporary data can be created during the operation of the cloud service, and may not be immediately deleted once they become unused for technical reasons. Periodic checks should ensure that such temporary data is effectively deleted after a predefined period.

Description of the need for SLOs, in addition to information available through certification, adherence to codes of conduct, etc.

The contract between the cloud service customer and the provider must include clear provisions for the erasure of personal data. Furthermore, since personal data may be kept redundantly on different servers at different locations, it must be ensured that each instance of them is erased irretrievably (i.e., previous versions, temporary files, etc.).

The following SLOs complement these indications, by translating them in a measurable objective that applies the data minimization principle in the course of the service.

Temporary data retention period	The maximum period of time that temporary data is retained after identification that the temporary data is unused.
Cloud service customer data retention period	The maximum period of time that cloud service customer data is retained before destruction by the cloud service provider and after acknowledgment of a request to delete the data or termination of the contract.

Description of relevant SLOs

3.5.4. Use, retention and disclosure limitation

Description of the context of the requirement

The cloud service provider, in its capacity as data processor, should inform the customer, in the most expedient time possible under the circumstances, of any legally binding request for which the provider is compelled to disclose the personal data by a law enforcement or governmental authority, unless otherwise prohibited, such as a legal prohibition to preserve the confidentiality of an investigation.

Description of the need for SLOs, in addition to information available through certification, adherence to codes of conduct, etc.

Besides the above mentioned obligation to inform the customer, the following SLOs aims to quantify the disclosures to law enforcement authorities over a period of time; this may also permit the customer to compare multiple offerings by different providers.

Description of relevant SLOs

Number of customer data law enforcement disclosures	refers to the number of personal data disclosures to law enforcement authorities over a predefined period of time (applicable only if the communication of such disclosures is permitted by law).
Number of personal data disclosure notifications	refers to the number of personal data disclosures to law enforcement authorities actually notified to the customer over a predefined period of time (applicable only if the communication of such disclosures is permitted by law).

3.5.5. Openness, transparency and notice

Description of the context of the requirement

Only if the provider informs the customer about all relevant issues, the cloud service customer is capable of fulfilling its obligation as data controller to assess the lawfulness of the processing of personal data in the cloud. Moreover, the cloud service provider shall make available the information that enable the customer to provide the data subjects with an adequate notice about the processing of their personal data, as required by law.

Notably, transparency in the cloud means it is necessary for the cloud service customer to be made aware of cloud service providers' subcontractors contributing to the provision of the respective cloud service.

Description of the need for SLOs, in addition to information available through certification, adherence to codes of conduct, etc.

Regarding the transfer of customer's personal data to the provider's subcontractors, the WP Opinion highlights the necessity that contracts between the cloud service provider and its subcontractors reflect, in terms of data protection provisions, the stipulations of the contract between cloud service customer and provider.

Furthermore, the cloud service customer consent (which can take the form of a general prior consent) is necessary for subcontracting and the customer may object to changes in the list of the subcontractors. In order to implement these provisions, the list of subcontractors must be made available to the customer.

The processing of certain special categories of data may require compliance with specific regulatory provisions, which may not be covered by standards or certifications schemes of general application. Therefore, it should be specified within the service agreement the possible special categories of data that the service is suitable for.

Description of relevant SLOs

List of tier 1 subcontractors	refers to the cloud service provider's subcontractors involved in the processing of the cloud service customer's personal data.
Special categories of data	refers to the list of the specific categories of personal data (if any), e.g. health-related or financial data or otherwise sensitive data, that the cloud service is suitable for processing, according to applicable standards or regulations.

3.5.6. Accountability

Description of the context of the requirement

In the field of data protection, accountability often takes a broad meaning and describes the ability of parties to demonstrate that they took appropriate steps to ensure that data protection principles have been implemented.

In this context, IT accountability is particularly important in order to investigate personal data breaches; to this end, the cloud platform should provide reliable monitoring and logging mechanisms, as described in the relevant sections of these Guidelines.

Moreover, cloud service providers should provide documentary evidence of appropriate and effective measures that are designed to deliver the outcomes of the data protection principles (e.g. procedures designed to ensure the identification of all data processing operations, to respond to access requests, designation of data protection officers, etc.). In addition, cloud service customers, as data controllers, should ensure that they are prepared to demonstrate the setting up of the necessary measures to the competent supervisory authority, upon request.

Description of the need for SLOs, in addition to information available through certification, adherence to codes of conduct, etc.

The cloud service provider must notify the cloud service customer in the event of a data breach that affects the customer data. To this end, the cloud service provider shall implement a data breach management policy which will specify the procedures for establishing and communicating data breaches. In this context, the first of the following SLOs implements these principles and allows the customer to evaluate the suitability of the provider's policy.

The second SLO relates to the need to be prepared to demonstrate the setting up of the necessary measures to the competent supervisory authorities, upon request.

Description of relevant SLOs

Personal data breach policy	describes the policy of the cloud service provider regarding data breach .
Documentation	refers to the list of the documents that the provider makes available, in order to demonstrate compliance to data protection requirements and obligations (e.g. procedures to respond to access request, designation of data protection officers, certifications, etc.).

3.5.7. Geographical location of cloud service customer data

Description of the context of the requirement

Personal data processed in the cloud may be transferred, also by subcontracting, to third countries, whose legislation do not guarantee an adequate level of data protection. This also implies that personal data may be disclosed to foreign law enforcement agency, without a valid EU legal basis.

To minimize these risks, the cloud service customer should verify that the provider guarantees lawfulness of cross-border data transfers, e.g. by framing such transfers with safe harbour arrangements, EC model clauses or binding corporate rules, as appropriate.

To this end, the cloud service customer shall be made aware of the location of data processed in the cloud, as required also by the above-mentioned principles of openness and transparency.

Description of the need for SLOs, in addition to information available through certification, adherence to codes of conduct, etc.

In this context, the following SLOs represent the instruments based on which the cloud service customer is allowed to control the location of its data.

Description of relevant SLOs

Data geolocation list	specifies the geographical location(s) where the cloud service customer data may be stored and processed by the cloud service provider .
Data geolocation selection	specifies whether cloud service customer can choose a given geographical location for the storage of the cloud service customer data.

3.5.8. Intervenability

Description of the context of the requirement

Directive 95/46/EC gives the data subject the rights of access, rectification, erasure, blocking and objection. Therefore, the cloud service customer must verify that the cloud service provider does not impose technical and organisational obstacles to these requirements, including in cases when data is further processed by subcontractors.

Description of the need for SLOs, in addition to information available through certification, adherence to codes of conduct, etc.

The contract between the cloud service customer and the cloud service provider should stipulate that the provider is obliged to support the customer in facilitating the exercise of data subject rights in a timely and efficient manner. The following SLO aims to define an objective term of reference for these activities.

Description of relevant SLOs

Access request response time	refers to the time period within which the provider shall communicate the information necessary to allow the customer to respond to access requests by the data subjects

4.CHAPTER IV

Engaging services of cloud service provider: defining servicing metrics, terms and conditions of cloud service agreement (CSA)

4.1. Introduction to Cloud service agreement (CSA)

Terminology changes have been made; specifically, the term service level agreement (SLA) has been replaced by cloud service agreement (CSA) to reference the broad agreement that is established between cloud customers and providers. The term SLA is now used to reference that part of the broader CSA that deals specifically with service levels.

The Current CSA Landscape

CSAs are a set of documents or agreements that contain the terms governing the relationship between the cloud customer and the cloud service provider. Because the cloud computing market is still developing, cloud customers should be aware that there may be a mismatch between their expectations and the cloud providers' actual service terms. For example, a CSA may not specify the geographic location where customer data will be stored. This could be a showstopper for customers subject to export restrictions of certain types of data from the U.S., or the export of "personal data" from the European Economic Area (EEA).

It is common for disputes to arise over the structure of the agreements, thus cloud customers must pay close attention to the language and clauses of the CSA. Large cloud providers can be inflexible with their CSAs, while small cloud providers may seem more flexible, but tend to over promise in order to obtain clients.

In general, the CSA is comprised of three major artifacts:

- *Customer Agreement*
- *Acceptable Use Policy (AUP)*
- *Service Level Agreement (SLA)*

This classification is not complete, nor is it uniformly adopted by the cloud industry: no standard nomenclature is used across the various cloud providers to specify their CSAs. Furthermore, cloud providers can modify their contract structure and terms at any time.

The *Customer Agreement* section of the CSA describes the overall relationship between the customer and provider. Since service management includes the processes and procedures used by the cloud provider, explicit definitions of the roles, responsibilities and execution of processes need to be formally agreed upon. The "Customer Agreement" fulfills this need. Various synonyms such as "Master Agreement," "Terms of Service," or simply "Agreement" may be used by certain providers.

An *Acceptable Use Policy (AUP)* is commonplace within a CSA. The AUP prohibits activities that providers consider to be an improper or outright illegal use of their service. This is one area of a CSA where there is considerable consistency across cloud providers. Although specific details of acceptable use will vary among IaaS, SaaS and PaaS providers, the scope and effect of these policies is the same, and these provisions typically generate the least concerns or resistance.

A typical *Service Level Agreement (SLA)* within the CSA describes levels of service using various attributes such as availability, serviceability or performance. The SLA specifies thresholds and financial penalties associated with violations of these thresholds. Well-designed SLAs can significantly contribute to avoiding conflict and can facilitate the resolution of an issue before it escalates into a dispute.

To guarantee an agreed service level, service providers must measure and monitor relevant metrics. There is often a mismatch between the metrics collected and monitored by the service provider and the higher-level functional (or "end-to-end") metric relevant to customers. This issue is common across service models, but is more acute for SaaS since customers want service levels to be met at the application level where they can be impacted by many factors. This is one reason why CSAs for SaaS usually lack stringent service level guarantees.

Service level guarantees for IaaS are better defined than for SaaS or PaaS, but that does not mean that they meet the customer's expectations. Most public cloud infrastructure services are available only through non-negotiable standard contracts which strictly limit the provider's liability. As a result, the remedies offered in case of non-compliance do not match the cost to the customer of the potential service disruptions. Furthermore, most IaaS providers put the burden of SLA violation notification and credit request on their customers.

In many cases, cloud SLAs do not offer refunds of charges but rather service credits against future use. Whether the relief is in the form of a credit or a refund, it is usually subject to a cap such as one month's standard billing. Credits against future billing will be of little or no benefit to customers that decide to switch providers following unsatisfactory service – and they clearly are meant to encourage the customer to stay with the current provider.

This rather biased situation is starting to evolve. As customers become more knowledgeable and competition increases, cloud providers are beginning to offer different service options that better shield customers from such risks.

For cloud customers, size also matters. In general, the larger the customer deployment, which translates to higher setup and monthly fees, the more power the customer can exert in negotiating more favorable CSAs, even with SaaS providers. No such improvements may be offered to small and medium businesses, but over time we expect the changes imposed by larger customers to trickle down to all other customers. Better CSAs will inevitably become a competitive factor. Eventually, customers of all sizes will be able to choose from a range of service terms that are more favorable and more flexible.

4.2. Guide for Evaluating Cloud Service Agreements

Before getting to the point of evaluating any CSA, customers must first perform a number of strategic steps (develop a comprehensive business case and strategy, select cloud service and deployment models, etc.).

With this strategic analysis as a prerequisite, this section provides a prescriptive series of steps that should be taken by cloud customers to evaluate CSAs in order to compare multiple cloud providers or to negotiate terms with a selected provider. The following steps are discussed in detail:

1) Understand roles and responsibilities
2) Evaluate business level policies
3) Understand service and deployment model differences
4) Identify critical performance objectives
5) Evaluate security and privacy requirements
6) Identify service management requirements

7) Prepare for service failure management
8) Understand the disaster recovery plan
9) Develop an effective governance process
10) Understand the exit process

Requirements and best practices are highlighted for each step. In addition, each step takes into account the realities of today's cloud computing landscape and postulates how this space is likely to evolve in the future, including the important role that standards will play to improve interoperability and comparability across providers.

Step 1: Understand Roles & Responsibilities

From the cloud service customer perspective, one of the significant areas of risk involved with cloud computing is associated with the division of activities and responsibilities between the cloud service customer and the cloud service provider. It is necessary to have a full understanding of who is responsible for which activities to ensure that there are no gaps which could lead to problems when using cloud services.

The ISO/IEC 17789 cloud computing reference architecture standard1 has 3 main roles for cloud computing:

- Cloud service customer
- Cloud service provider
- Cloud service partner

The cloud service provider and the cloud service customer are the most significant roles in the provision and use of cloud services while the cloud service partner is a party engaged in support of the activities of the cloud service customer and/or the cloud service provider.

There are a number of sub roles of each of the major roles – the sub roles are shown in Figure 8:

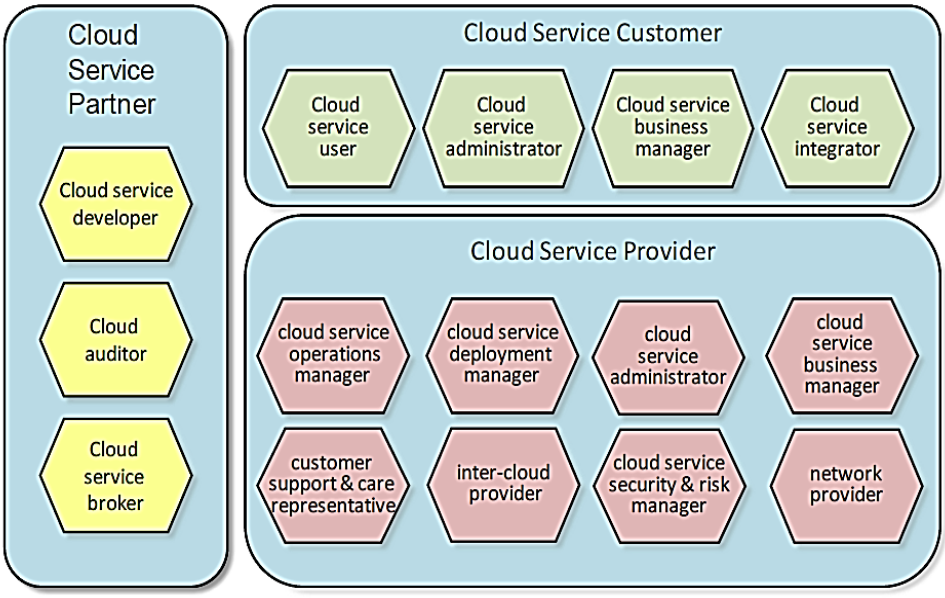

Figure 8: Cloud Computing Roles and Subroles

Each of the subroles in Figure 1 has a set of activities and responsibilities which are described in high-level terms in ISO/IEC 17789. There are also relationships between the subroles - for example, the cloud service administrator of the customer may interact with the customer support and care representative of the provider in cases where customer personnel experience problems using the cloud service.

Some of the subroles may appear in a CSA, or they may have a direct or indirect relationship to some aspects of the CSA. The subroles of the cloud service customer and the cloud service provider, in particular, are involved in the split of responsibilities that is typical for cloud services - the CSA should make clear statements about those responsibilities. Cloud service customers need to understand the activities and responsibilities of the various subroles and ensure that the CSA and its associated SLA contains appropriate commitments and service level targets to address those activities and responsibilities for the cloud service(s) covered by the CSA.

One important area for customers to consider is who is responsible for detecting and then reporting incidents where the cloud service fails to meet some aspect of the CSA or SLA. This can include outages where the cloud service is unavailable, or may include cases where performance fails to meet stated service levels (for example, response times are too long). How such incidents are detected must be established - it may be the responsibility of the customer and the customer may need to put in place appropriate monitoring technology. It is also necessary to be clear about how incidents are then reported and tracked until resolved.

One partner role that is particularly relevant to the CSA and to SLAs is the cloud auditor. It is unlikely that the cloud service customer has direct insight into the operations of the cloud service provider, particularly regarding aspects such as security and the protection of sensitive data such as personally identifiable information (PII). It is typical for cloud service providers to offer assurances about these aspects of their cloud services through certifications or attestations which are provided by third party cloud auditors who inspect the cloud service provider’s operations and issue reports typically based on one or more standards or certification schemes.

Each CSA may be unique based upon the customers’ requirements and the cloud services under consideration. CSAs can contain various elements and are not limited to quantitative measures, but can include other qualitative aspects such as alignment with standards and data protection. It is strongly recommended that cloud service customers gain a solid understanding of the spectrum of CSAs that currently exist for cloud service providers in order to compare cloud services offered by different providers and assess trade offs between cost and service levels.

It is important to recognize that the content of a CSA and associated SLA is likely to vary depending on the category of the cloud service. Considerations for an IaaS service offering compute and storage infrastructure are likely to be very different from those for a SaaS service that offers complete application functionality for some business functions. At the very least, the split of responsibilities between the provider and the customer are going to be different for these different cases, and this is necessarily reflected in differences in the CSA and SLA.

The following sections, which cover the cloud CSA evaluation steps in detail, each elaborate on the expected responsibilities of the customer and the provider for both business level and service level objectives. In order to make sound business decisions, it is important that customers understand what to expect from their cloud service provider. This, in turn, will help them clarify their own responsibilities and help them assess the true cost of moving to cloud computing.

Step 2: Evaluate Business Level Policies

Customers must consider the policy and compliance requirements relevant to them when reviewing a CSA since there are interdependencies between the policies expressed in the CSA and the business strategy and policies developed across the lines of business. The *data policies* of the cloud provider, as expressed in the CSA, are perhaps the most critical business level policies and should be carefully evaluated.

The obligations a cloud provider has to its customers and their data is governed by a potentially complex combination of:

- customer requirements,
- the data protection legislation applicable to the customer as well as to its individual users (which may not be under the same jurisdiction in a multinational company)
- The laws and regulations applicable where the data resides or is made available.

Customers should carefully consider these legal requirements and how the CSA deals with issues such as movement of data when redundancy across multiple sites means subjecting the data to different jurisdictions at different times. The issue of jurisdiction takes on additional complexity when global compliance is taken into consideration and more than one cloud provider is used. In these instances the customer may have to coordinate negotiations between providers to ensure the necessary data management.

Table 6 highlights the critical data policies that need to be considered and included in the cloud CSA.

Table 6: CSA Data Policies

Data Policy	Description / Guidance
Data Preservation and Redundancy	• Timely and efficient capturing and preservation of data is critical to maintaining the organizational memory of a business or the general user. Customers should therefore ensure they have an appropriate data preservation strategy that addresses redundancy within the system. • Cloud customers should ensure the CSA supports their data preservation strategy that includes sources, scheduling, backup, restore, integrity checks, etc. They should be concerned as to the protections offered or omitted by the service provider. • It must be possible to test the CSA to demonstrate the required level of service availability.

Data Policy	Description / Guidance
Data Location	• CSAs that cover locations under different jurisdictions are challenging. • Customers should consider how the CSA specifies where their data resides, where it is processed, and how this meets the various applicable regulations. Customers should also understand where the data is viewed or delivered, and whether this results in a transborder data flow with regulatory or tax implications. • For example, can the provider truly deliver a sound technical solution when sensitive data spans several jurisdictions with conflicting laws? Does the provider commit, in the CSA, to the specific location(s) where the customer's data will be stored? • If the provider reserves the right to add new locations or change data movement policies, will they give the customer advice notice? Preferably, will they obtain the customer's permission to relocated its data? • Is there a means to verify the current location of a data set?
Data Seizure	• Legal powers enable law enforcement and other government agencies to seize data under certain circumstances. Customers should ensure the CSA provides for sufficient notification of such events. • Customers should also ensure there are arrangements in place to make their data available in the event that their provider goes out of business. • in the event that the provider locks access to its systems because of a billing dispute or a security issue, the customer's date should not be "held hostage" while the issue is being resolved.
Data Privacy	• The provider's data privacy policy should be included in the CSA, and should ensure that the provider will conduct business in compliance with applicable laws on data privacy protection. • This includes identifying the data sets gathered, data retention policies, how the data is communicated, how personal data is stored and used, etc. • Data privacy in a cloud context is not just about the protection of the information about the customer's agents in its dealing with the provider (this is the narrow meaning in many existing Service Level Agreements), it also includes the privacy of the information that may be stored about the customer's own customers. • Refer to the Privacy section within Step 5 for more information.
Data Availability	• Assess whether the provider's maintenance schedules might interfere with business processes subject to external constraints, such as financial reporting or the business's hours of operation in certain regions.
Change Management and Notification	• The change management and change notification obligations of the provider should be carefully reviewed, especially the amount of time allowed to prepare for a change. The provider may also ask the customer to provide certain change notifications, which is a good opportunity to strengthen the customer's own change management policies.

In addition to data policies, there are a number of other business level policies expressed in the CSA that require careful evaluation. Uptime and availability are another area where customer requirements and policies may not match up with the language of the vendor, and where location and jurisdiction may come into play. For example, if the uptime guarantee is for "regular business hours," then organizations with multiple locations in different time zones need to clarify whether the guarantee covers only the headquarters location or all

regions. Similarly, “week-ends” or “holidays” have different meanings in different countries. For some multinational customers with offices in all continents, the sun literally never sets on their empire, and the provider may not be ready to commit to supporting them 24x365.

All of these policies will impact and influence the customer’s cloud strategy and business case. In many cases, these policies, as defined in the CSA, are non-negotiable and are similar across different cloud providers. However, there will be instances where some of these policies can be negotiated and/or some of these policies differ sufficiently across different cloud providers to warrant careful consideration from customers.

Table 7 below highlights the critical business level policies that need to be considered and addressed in the CSA.

Table 7: CSA Business Level Policies

Policy	Description / Guidance
Guarantees	• CSA guarantees should be defined, objective and measurable with an appropriately scaled penalty matrix that matches the impact of non-performance by the provider.[3] The CSA should clarify: o What constitutes excused or excluded performance o Escalation procedures o How service-level bonuses and penalties are administered o Remedy circumstances and mechanisms • Guarantees should be expressed as a measurable number, for example a percentage such as 99.999% for service availabliity, denoting the amount of time the service is guaranteed to be working. Other guarantees will refer to matrics in other units, such as time-to-repair in minutes, etc. • Availability measures need to include the measurement window.

Policy	Description / Guidance
Acceptable Use Policy	• The acceptable use policy will clearly describe how the customer may use the service and the agreement generally will describe what actions the provider may take in the event of a breach. • In today's cloud environment, this policy is typically non-negotiable and the terms generally favor the cloud provider. • Customers need to understand the impact of such policies if they use the cloud solution to in turn provide a service to end users over whom they have limited control,
List of Services Not Covered	• The CSA will state under what conditions and with which described services the customer is supported. The CSA may also state what is excluded and what constitutes illegal use. • Customers should look for explicitly stated exceptions and understand why the provider has excluded them.
Excess Usage	• Providers operate business models to drive revenue. While elasticity is a fundamental benefit of using the cloud, customers may find that usage above their contracted thresholds will incur high incremental rates which can be punitive and disrupt their budgets. • Customers should correctly size their usage requirements, reduce the opportunity for usage creep and consider and understand the "what-ifs" of exceeding their usage thresholds.
Activation	• The time at which the service becomes active must be defined precisely, in order to provide a "reference starting point" for the measurement of performance. This is important to measure certain metrics that are associated with a specific time window (e.g., number of outages per 30-day period). It impacts whether an "event" triggers a penalty clause. • From a CSA compliance perspective, it is important for customers to understand the trigger points under the CSA so they can independently measure event timing.
Payment and penalty models	• The CSA should clarify when/how payment is to be made. Provider payment models vary. Monthly recurring or "pay as you use" models are typical. • There may be credit terms that require advanced payment or payment every 30 days. "Just in time" service providers are sensitive to poor credit control and are likely to be more diligent in suspending service. • Equally, the customer needs to be diligent in obtaining service credit payments for outages.
Governance / Versioning	• Provider services evolve. New features may be added, others will go out of warranty, and some may persist indefinitely. Where the assumptions or conditions under which the CSA was initially accepted are changed, the customer should review the impact on their specific situation. • A good provider will maintain a proactive policy of advising customers of changes to their CSA and practice version control. • Customers should ensure that there is a mechanism to inform them of changes and, if not, amend their contract to put the onus on the provider to provide reasonable advance notice of updates.

<table>
<tr><th>Policy</th><th>Description / Guidance</th></tr>
<tr><td>Renewals</td><td>

Renewals are an opportunity to bargain for better rates or services levels, or relocate to another provider if necessary.
Providers may write in their contracts an automatic renewal clause that kick in in the absence of a 90-day cancellation notice before the contract's anniversary date. It is common for customers to overlook this deadline and be obligated to renew the contract without having had a chance to negotiate changes or even cancel the service.
Customers should read the terms and conditions of the renewal arrangements, and consider the conditions under which a provider may vary the service terms (or revise prices) upon renewal.

</td></tr>
<tr><td>Transferability</td><td>

Customers should consider the potential need to transfer an agreement in the event their business is sold.
Conversely, if the provider's business is acquired, the customer may not wish to so business with the new owner, and should have the option to migrate to a new service without penalties.
Customers may operate several accounts with a provider and want to offset account credits between accounts. Is this supported in the provider's contract terms?

</td></tr>
<tr><td>Support</td><td>

Customers must follow the provider's rules to report problems, in order to ensure that the support terms specified in the CSA are activated, and that the "clock starts ticking" for appropriate escalation and penalties.
An example of a support and escalation matrix related to service availability is provided below. All four target times in the table are associated with the commencement "time stamp" of the service or the notification of a service affecting event.

<table>
<tr><th>Priority</th><th>Description</th><th>Target Response Time</th><th>Target Update Time</th><th>Target Fix Time</th></tr>
<tr><td>P1</td><td>Production software unusable/Production cloud servers inaccessible</td><td>1 hour, Provider's executive notified of issue</td><td>1hr</td><td>Immediate - work commences and continues until issue resolved or workaround deployed</td></tr>
<tr><td>P2</td><td>Partial software functionality unusable/Partial service unavailable</td><td>4 hours</td><td>1day</td><td>2 days, subject to available maintenance slot</td></tr>
<tr><td>P3</td><td>Cosmetic issue</td><td>1 working day</td><td>1 working day</td><td>Next software release/service update</td></tr>
<tr><td>P4</td><td>Information request</td><td>2 working days</td><td>2 working days</td><td>n/a</td></tr>
</table>
</td></tr>
</table>

Policy	Description / Guidance
Planned Maintenance	• All systems require maintenance. Complex systems may be designed to include sufficient redundancy so that maintenances can be carried out without affecting the service. • The CSA may, however, describe "uptime" as an availability percentage (e.g. 99.90%). This is the equivalent of 8.5 hours downtime per year. CSAs may state that this does not include "planned maintenance." Thus, the provider may have a service outage for 8.5 hours. plus maintenance time, and the customer is not entitled to compensation under the CSA. This highlights the importance of defining the measurement window. If the availability percentage is measured each month, this allows 12 outages nut each of them cannot last more than 42 minutes without triggering a penalty.
Subcontracted Services	• Providers sometimes include in their CSA a clause that the CSA of an upstream (subcontracted) provider will govern the services provided by the sibcontracot, and that the only available penalties are those from the upstream provider even though its CSA may be less rigorous. The customer's expectation, based on reviewing the CSA of their immediate provider, may thus be violated. • Therefore, the customer should ensure that the immediate provider CSA states unambiguously that its CSA applies to the complete service, regardles whether parts of the service come from third parties.
Licensed Software	• Cloud services may include third party licensed software which is sold on a monthly licensed basis under a service provider license agreement. Such software is updated regularly by its manufacturer. • Providers may opt to pass the responsibility for updating the licensed software over to the customer once they have started to use the service. This absolves the provider of the risk of disrupting the customer's operation through an unforeseen software conflict or bug. • Alternately, the provider may "push" the update, in which case the CSA should require that the customer be given advance notice of the update. The customer should have the ability to opt out, or at least to defer the update. However, the supplier may be unwilling to continue to support older versions indefinitely, and there should be a legitimate exception for updates that correct serious security vulnerabilities.
Industry Specific Standards	• Regulated industries, like government, financial services, and healthcare, are subject to specific and often quite onerous standards which must be addressed in the CSA and implementation. • Customers who operate in these regulated industries should ensure that the their legal team is fully involved on the negotiation of the CSA.
Additional Terms for Different Geographic Region or Countries	• Customers should consider the provider's origins and primary market. Detailed refinements to the home market CSA may be required to properly cover customers who are located in remote markets. • Data protection legislation is one aspect of this, but customers should not limit their examination of the agreement to this sole aspect.

Step 3: Understand Service and Deployment Model Differences

Services offered by cloud providers typically fall into one of the three major groups of service models: Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). For each category, there are significant differences in the levels of cloud resource abstraction, service level objectives, and key performance indicators that will potentially be included in a CSA. In addition, the level of clarity varies significantly for each service model. To increase effectiveness, specific components of the CSA should be stated in measurable terms and should include:

- The service to be performed and outcome expectations
- Key Performance Indicators (KPIs) and the level of service that is acceptable for each
- The manner by which service is to be measured
- The parties involved and their responsibilities
- The reporting guidelines and requirements
- Incentives for the service provider to meet the agreed upon target levels of quality

The CSA is often the best indicator of how, and how often, the provider expects their service to fail. Therefore, customers must remember that downtime, poor performance, security breaches and data loses are their risks to bear. It's important that customers select a cloud provider who will help them with the fine details in supporting their workloads as they transition to cloud computing.

Table 8 highlights the different CSA considerations for each of the cloud service models.

Table 8: CSA Considerations for Service Models

Service Model	CSA Considerations
IaaS	• Cloud IaaS CSAs are similar to SLAs for network services, hosting, and data center outsourcing. The main issues concern the mapping of high level application requirements on infrastructure services levels. • Metrics are well understood across the IaaS abstractions (compute, network, and storage). Customers should expect to find a subset of the following metrics in their cloud SLA. o Compute metrics: *availability, outage length, server reboot time* o Network metrics: *availability, packet loss, bandwidth, latency, mean/max jitter* o Storage metrics: *availability, input/output per second, max restore time, processing time, latency with internal compute resource* • Compute metrics usually exclude service levels for compute performance. Customers are simply guaranteed availability of the compute resources for which they paid. • Customers must distinguish between IaaS development environments and IaaS production environments when reviewing their cloud IaaS service agreements. IaaS production environments will typically require more stringent service level objectives than IaaS development environments.

	• Network metrics in a cloud SLA generally cover the cloud provider's data center connectivity to the Internet as a whole, not to any specific provider or customer. • There are several standardization efforts within the IaaS space which help describe and manage the services offered at this level.[4] Whenever possible, customers should ensure the CSA includes provisions requiring their cloud providers to support open standard interfaces, formats and protocols to increase interoperability and portability.
PaaS	• Two main approaches exist for building PaaS solutions: *integrated solutions* and *deploy-based solutions*. When reviewing the PaaS service agreement, customers should consider tradeoffs in flexibility, control, and ease of use to determine which approach best meets their business needs. o Integrated solutions are web accessible development environments which enable developers to build an application using the infrastructure and middleware services supported by the cloud provider. Management of the application and its execution is primarily controlled by the cloud provider. Typically, service developers only have access to a provider-defined set of APIs which offer limited control on the coordination of code execution. o Deploy-based solutions enable deployment of middleware on top of resources acquired from an IaaS cloud provider, offering deployment services to the customers which automate the process of installation and configuration of the middleware.[5] These PaaS solutions offer a rich set of management capability including the ability to automatically change the number of machines assigned to an application, and self-scaling according to the application's usage. • At a minimum IaaS SLA's should roll into PaaS SLA's. • Customers must distinguish between PaaS development environments and PaaS production environments when reviewing their cloud PaaS service agreements. PaaS production environments will typically require more stringent service level objectives than PaaS development environments. • Standards are emerging to help identify PaaS services offered by cloud providers and standard interfaces for communicating with PaaS providers to provision or manage PaaS

	environments. Standards, like OASIS Topology and Orchestration Specification for Cloud Applications (TOSCA)[6] have come about to address portability and interoperability across providers. In addition, PaaS open source offerings such as Cloud Foundry and OpenShift are starting to build momentum in the market. • Customers should ensure their CSA includes support for open standards, as they become available, to reduce vendor lock in.
SaaS	• Customers should insist on flexible CSAs that are measurable against their objectives, not the cloud providers' reporting needs. • Given the wide variation of services provided at the SaaS level, it is difficult to provide a comprehensive and representative list of SaaS service level objectives for customers to look out for in their CSAs. • Customers should expect general SaaS service level objectives like *monthly cumulative application downtime*, *application response time*, *persistence of customer information*, and *automatic scalability* to be included in their CSA. • Customers should ensure that data maintained on the provider's cloud resources be stored using standard formats to ensure data portability in the event that a move to a different provider is required.

In addition to service models, service deployment terms should be included in a CSA. These terms should clarify to both parties signing the CSA the information required to verify the correctness of deployment actions. Specifically, these terms should identify:

- Deployment model
- Deployment technologies adopted

The deployment model included in the CSA should clearly specify one of the following options: *Private*, *Community*, *Public*, or *Hybrid*. Customers must be well educated on the characteristics and differences in each of these deployment models since potential value and risk varies significantly.

Table 9 highlights the different CSA considerations across the deployment models.

Table 9: CSA Considerations for Deployment Models

Deployment Model	CSA Considerations
Private (On-site)	• CSA considerations for Private (On-site) are similar to those of a traditional enterprise IT SLA. However, given that data center resources may be shared by a larger number of internal users, customers must ensure that critical service objectives like availability and response time are met via ongoing measurement and tracking.
Private (Outsourced)	• CSA considerations for Private (Outsourced) are similar to Private (On-site) except cloud services are now being provided by an external cloud provider. The fact that IT resources from the provider are dedicated to a single customer mitigates potential security and availability risks. • Customers should ensure the CSA specifies security techniques for protecting the provider's perimeter and the communications link with the provider. • Customers should consider the criticalness of the service being deployed to justify the added expense of this model over the Public model.
Public	• CSA considerations for the Public model are greater than the Private (Outsourced) model since the provider's IT resources are now shared across multiple customers. • As a result, customers should carefully review the CSA to understand how the provider addresses the added security, availability, reliability and performance risks introduced by multi-tenancy. • The ability to measure and track specific service level objectives becomes more important in the Public deployment model. Customers should also ensure the CSA provides adequate methods and processes for ongoing measurement.
Hybrid	• CSA considerations for the Hybrid model are similar to the Public model with the increased likelihood for unique integration requirements between cloud and enterprise services. • Customers should ensure the CSA adequately covers their service and data integration requirements. It is recommended to use a specific and standard document that describes the nature of the interface (along with quality level metrics and performance characteristics associated with the interface) and any security requirements. For example, if the interface is a web service, there may be authentication and authorization requirements with implications on the LDAP mechanism.

In addition to specifying the deployment model, the CSA should clarify how a service is made available to service users on a given cloud provider, for example:

• A web application is deployed on an application server as a Web application ARchive (WAR) file.

• A grid application is deployed on a grid container as a Grid ARchive (GAR) file.

• A virtual machine is deployed on an IaaS provider as a virtual machine disk image that may be represented in one of many different formats. Adoption and support for standards like the Distributed Management Task Force (DMTF) Open Virtualization Format (OVF) is recommended.

When CSAs are signed, a clear description of the technologies involved in the deployment of services should be specified. Note that there is a close relationship between deployment technologies and the kind of services being offered.

Step 4: Identify Critical Performance Objectives

Performance goals within the context of cloud computing are directly related to the efficiency and accuracy of service delivery by the cloud provider. Typical performance considerations include availability, response time and processing speed, but they can include many other performance and system quality perspectives. Cloud customers must decide which measures are most critical to their specific cloud environments and ensure these measures are included in their SLA.

Performance statements that are important to the cloud customer should be measureable and auditable, like all metrics, and documented in the SLA in order to provide for rational discussions between the parties. The relevant performance factors depend on the service model (IaaS, PaaS or SaaS) and the type of services provided within that model (for example, network, storage and computing services for IaaS). In order to assess performance objectively and establish trust between the parties, clear and consistent measurements are required. It must be clear how each metric will be used and what decisions will be made from the measurements to align service performance to specific business and technical goals and objectives.

This section will focus on two performance metrics: *availability* and *response time*. The intention is to provide a basic framework to identify and define meaningful and consistent cloud metrics. This framework can then be applied to other potential metrics not covered in this document. While many of the metrics may already be supported by your cloud provider; they may interpret the definition differently than you do. An agreed definition in the context of a specific cloud solution is critical. Some calibration may be required if a measurement captured by a provider does not exactly match the definition included as part of the SLA.

Industry standards should be used when possible to improve consistency. For instance, IEEE has good measurement definitions and categorizations for activities such as maintenance.

Here are the generally accepted definitions for the two metrics of interest:

• *Availability*. Percentage of uptime for a service in a given observation period.

• *Response time*. Elapsed time from when a service is invoked to when it is completed (typically measured in milliseconds).

Table 10 describes three different example scenarios (network availability, storage availability, and service response time) and the specific performance information required for each.

Table 10: Availability and Response Time Examples

	Network Availability (example)	Storage Availability (example)	Service Response Time (example)
Metric Name in SLA	Network Percentage Available Critical Business Hours	Storage Percentage Available	Service XXX Response Time in a Given Hour; Service YYY Response Time in a Given Hour.
Constraints	Critical time is defined as 12AM GMT to 12PM GMT Monday through Friday	None	Response times will only be evaluated for services XXX and YYY, which are PaaS reuseable services that will be invoked by our applications.
Collection Method	Machine	Machine	Machine
Collection Description	Using the DMTF, OGF[12], or other standard to consistently collect the measures.	Using the DMTF, OGF, or other standard to consistently collect the measures.	Using the DMTF, OGF, or other standard to consistently collect the measures.
Frequency of Collection	The network is "pinged" every one minute.	Specific storage services (read and update) are randomly "pinged" every one minute.	For each XXX and YYY service invoked, the response time is collected every five minutes.

Other Information	60 seconds of uptime will be recorded for each successful "ping."	60 seconds of uptime will be recorded for each successful "ping."	Each service will be reported separately. Hourly averages will be calculated.
Clarification	No reference to quality or availability of specific service. This is exclusively a measure of network availability.	No reference to quality or availability of specific service. This is exclusively a measure of storage availability.	No individual service reporting is needed (for example, listing of all services that exceeded SLA agreed response time).
Usage 1 in SLA	Network availability shall be 99.5% or higher between 12AM GMT to 12PM GMT Monday thru Friday.	Storage availability shall be 99.9% or higher	Response time for XXX service shall be less than 500 ms, YYY service less than 200 ms.
Usage 2 in SLA	For any day when network availability is less than 99.5%, a 20% discount will be applied for the entire day's network charges.	For any day when storage availability is less than 99.9%, a 50% discount will be applied for the entire day's storage charges.	If in any given hour the response times as stated are not met, all services of that type during that hour will be processed at no charge.

Both hardware and facilities should be considered when assessing critical performance levels in an IaaS context. Hardware includes: computers (CPU and memory), networks (routers, firewalls, switches, network links and interfaces), storage components (hard disks), and any other physical computing infrastructure elements. Facilities include: heating, ventilation and air conditioning (HVAC), power consumption and dissipation, communications, backup, and other aspects of the physical plant. In the case of PaaS or SaaS solutions, it can be presumed that the unavailability or sub-par performance of any of these components will affect the overall services, therefore it is not necessary to specify them – the measurements should be "end to end" expressed in terms of the user experience.

Moreover, particularly in the IaaS case, higher level business objectives may dictate what critical resources fall within the scope of the metrics. For example, the power consumption or the heat dissipation may or may not be included, depending whether the customer has established a corporate carbon footprint objective.

In summary, when considering performance metrics in a cloud SLA, it is recommended that consumers:

o Understand the business level performance objectives (for example, reduce cost and time to market per unit of software functionality).

o Identify the metrics that are critical to achieving and managing the business level performance objectives.

o Ensure these metrics are defined at the right level of granularity that can be monitored on a continuous basis (in a cost-effective manner).

o Identify standards that provide consistency in metric definitions and methods of collection.

o Analyze and leverage the metrics on an ongoing basis as a tool for influencing business decisions.

Step 5: Evaluate Security and Privacy Requirements

Security controls in cloud computing are, for the most part, no different than security controls in any IT environment. However, because of the cloud service models employed, the operational models, and the technologies used to enable cloud services, cloud computing may present different risks to an organization from traditional IT solutions.

There are two asset categories that require security and privacy consideration for cloud computing:

- Information (which belongs to the customer but has been moved into the provider's cloud)
- Applications, functions or processes (being executed In the cloud to provide the required service to the customer)

A required foundation for security, regardless of whether a cloud solution is used, is a *security classification scheme* that applies throughout the enterprise, based on the criticality and sensitivity of enterprise data. This scheme should include details about data ownership, definition of appropriate security levels and protection controls, and data retention and destruction requirements. The classification scheme should be used as the basis for applying access controls, archiving, and encryption methods.

In order to determine which level of security is required for a specific asset, a rough assessment of an asset's sensitivity and importance is required. For each asset, the following questions should be asked:

How would the business be harmed if…

1. The asset became publicly available and distributed?
2. An employee of our cloud provider accessed the asset?
3. The process or function was manipulated by an outsider?
4. The process or function failed to provide expected results?
5. The information was unexpectedly altered?
6. The asset was unavailable for a period of time?

Table 11 below highlights the key steps customers should take to ensure their CSA sufficiently addresses their unique security requirements.

Table 11: Key Security Considerations for CSAs

CSA Security Considerations	Strategic Activities
Assess asset sensitivity and the operational security requirements of applications	• Complete an assessment of the confidentiality, integrity, and availability requirements of the assets. • Complete a threat risk assessment and privacy risk assessment. • Address application operational security, availability requirements and privacy requirements in response to identified risks and in line with the organization's data classification, information architecture, information security architecture, and risk tolerance.
Understand legal/ regulatory data residency requirements	Understand the regulatory, contractual and other jurisdictional constraints about the logical and physical locations of data.

Put in place restrictions against unauthorized asset movement and accidental disclosure	• Establish policies to restrict the movement of sensitive data to cloud services by individuals or departments without the approval or, at minimum, notification of the Security/Privacy departments.
	• Take steps to detect such unapproved data moving to cloud services:
	  ○ Monitor for large internal data migrations with database activity monitoring (DAM) and file activity monitoring (FAM)
	  ○ Monitor for data moving to the cloud with URL filters and data loss prevention
	• Protect data in transit. All sensitive data moving to or within the cloud should be encrypted.
	• Protect data at rest. Sensitive volumes should be encrypted to limit exposure to snapshots or unapproved administrator access. Sensitive data in object storage should be encrypted, usually with file/folder or client/agent encryption.

Establish and track security metrics	• Metrics and standards for measuring performance and effectiveness of information security management should be established prior to moving to cloud computing. • At a minimum, organizations should understand and document their current metrics and how they will change when operations are moved into the cloud and where a provider may use different (potentially incompatible) metrics.[14]
Assess the cloud provider's security capabilities	• Assess the cloud provider's level of security and its maturity. • If compliance to a normative standard (e.g. ISO 27002/27017[15]) is asserted, then verify the compliance certificate and its validity. • Look for verifiable evidence of resource allocation, such as budget and manpower to sustain the compliance program • Verify internal audit reports and evidence of remedial actions for the findings

Assess the cloud provider's security governance	• Assess the provider's security governance processes and capabilities for sufficiency, maturity, and consistency with the customer's information security management processes.
	  ○ The provider's information security controls should be demonstrably risk-based and clearly support these management processes.
	  ○ Where a provider cannot demonstrate comprehensive and effective risk management processes in association with its services, customers should carefully evaluate use of the provider as well as the user's own abilities to compensate for the potential risk management gaps.
	• Determine if the provider's guarantees adequately address your security requirements.[16]

Audit the cloud provider's security CSA compliance	• A "right to audit" clause in a CSA gives customers the ability to audit the cloud provider, which supports traceability and transparency. • Use a normative specification in the "right to audit" clause to ensure mutual understanding of expectations. • In time, this right should be supplanted by third-party certifications (e.g., driven by ISO/IEC 27002/27017). If the cloud provider is not willing to subject itself to a customer audit, it should propose to use a trusted third-party that follows a normative standard.

Providers should notify consumers of the occurrence of any breach of its system, regardless of the parties or data directly impacted. The provider should include specific pertinent information in the notification, stop the data breach as quickly as possible, restore secure access to the service as soon as possible, apply best-practice forensics in investigating the circumstances and causes of the breach, and make long-term infrastructure changes to correct the root causes of the breach to ensure that it does not recur. Due to the high financial and reputational costs resulting from a breach, consumers should require the provider to indemnify them if the breach was their fault (the indemnification clauses contained in the provider's CSAs are often written the other way around: they are meant to protect the cloud provider from being sued for the consequence of customer actions).

Privacy

In many countries throughout the world, numerous laws, regulations, and other mandates require public and private organizations to protect the privacy of personal data stored in computer systems.

When data is transferred to a cloud, the responsibility for protecting and securing the data typically remains with the controller or custodian of that data, even if in some circumstances this responsibility may be shared with others. When it relies on a third party to host or process its data, the controller of the data remains liable for any loss, damage, or misuse of the data. It is prudent, and may be legally required, that the PII controller and the PII processor (i.e., the cloud service provider) enter into a written (legal) agreement that clearly defines the roles, expectations of the parties, and allocates between them the many responsibilities that are attached to the data at stake.

If privacy issues are not adequately addressed in the CSA, the cloud customer should consider alternate means of achieving their goals including seeking a different provider, or not sending sensitive data to the cloud. For example, if the customer wishes to send HIPAA-covered information to the cloud, the customer will need to find a cloud service provider that will sign a HIPAA business associate agreement or else not send that data to the cloud.

Preservation of information, included in some privacy regulations, can require that large volumes of data be retained for extended periods. What are the ramifications of this under the CSA? What happens if the preservation requirements outlast the terms of the CSA? If the customer preserves the data in place, who pays for the extended storage and at what cost? Does the customer have the storage capacity under its CSA? Can the customer effectively download the data in a forensically sound manner so it can preserve it offline or near-line? These are some of the privacy related questions that need to be addressed in the CSA.

The reverse risk may also exist: the backup and disaster recovery policies of a cloud provider may cause copies of data or code to be retained beyond the retention period intended by the customer. This may be a problem during the "discovery" phase of litigation. One party may claim that certain data (e.g., copies of old e-mails) has been erased, and the opposing party may discover that it still exists in the backups made by a cloud provider and subpoena the provider. In another scenario, a cloud customer may wish to implement an end user's "right to be forgotten," only to find that it has no ability to selectively delete the backup copy of the user's records.

Step 6: Identify Service Management Requirements

The fundamental goals of any cloud computing environment are to reduce cost, improve flexibility and increase reliability of the delivery of a service. Critical to meeting these goals

is a uniform, straightforward, transparent and extensible system for managing and monitoring cloud services. In this section we will outline some key things to consider in the area of service management when entering into a service agreement with a cloud computing provider.

Every computing system requires internal controls, management, automation, and self-healing in order to operate in today's interconnected world, an area commonly called Application Performance Management, or APM. A move to the cloud still requires these elements – perhaps even more so. Although the standards for CSA language for service management are evolving, it is of upmost importance to include provisions for the considerations outlined below in your agreements.

Auditing

First and foremost in ensuring manageability of cloud services is a methodology for auditing and reviewing those services. This helps discern between providers who are fully capable of deep manageability and those who provide only a simple veneer on someone else's offerings. As stated by many an experienced manager, people "do what you inspect, not what you expect."

The objective of any CSA terms in the area of auditing is multi-fold:

1. Provide you with an unbiased assessment of your ability to rely on the service provided

2. Assess the depth and effectiveness of the provider's internal systems and measures

3. Provide tools to compare quality levels with other competing providers

4. Ensure the openness needed to allow continuous review and improvement

5. Uncover issues in your own organization's ability to interface with the provider and provide uninterrupted services

This last objective is especially important. Many documented challenges have come not from a cloud provider's ability to service a customer, but the ability of the customer's systems to interface properly with the cloud. Therefore any audit scope should include both the provider and any internal systems exposed to the cloud to ensure a complete "envelope" of integrity.

When considering the scope of any auditing protocol, you must step beyond contract terms and conditions and ensure that you are addressing general issues of management and governance, including necessary resources to mitigate any risks found. For example, it's insufficient to include a provision to regularly audit security and encryption keys, only to neglect addressing any internal resource allocations, scheduling, review and approval processes needed to perform the audit and address any issues stemming from the audit. Consider carefully the importance of leveraging methods of audit and compliance that already exist in your organization, and look to extend those to the cloud vs. creating new ones.

Monitoring & Reporting

Transparency of the service level is extremely important to a successful service management protocol. While every cloud vendor offers different systems for visualizing data and its implications (web based, e-mail based, live, reactive, portal-based), customers should demand from any CSA a minimum set of capabilities:

1. *Cloud Performance Management*. This domain focuses on the response times for systems within the cloud architecture and between the cloud and the target user systems.

2. *Peak Load Performance*. This domain focuses on measurements and timings for when the cloud is under stress, either intentional or unintentional. As systems can perform differently when under different loads, and the interactions and dependencies of a complex cloud are often unknown in advance, it's important to visualize data both in a steady state as well as under load.

3. *Hybrid and Inter-cloud Performance*. As many clouds consist of different subsystems, often sourced from different cloud providers, it's critical to visualize data about the interactions between those hybrid cloud components.

4. *Application Performance*. This domain focuses on the applications executed from the cloud, particularly internal processing benchmarks as well as end-user experience measurement.

5. *Problem Notification*. This domain focuses on monitoring and reporting on failures and issues with the cloud system. Addressed are issues with prioritization, notification and severity level assessment.

Although the benchmarks in each of these areas are evolving, ensuring your CSA includes the ability to see, assess and react to measurements in these areas will help keep your cloud infrastructure running smoothly.

Measurement and Metering

A core characteristic of many cloud services is an on-demand model, where services used are billed as they are consumed, on a time or capacity basis. Therefore it is important to have confidence and transparency in the measurement and metering system employed by cloud providers, as embodied in the CSA you negotiate. At a minimum, you must ensure that metering systems employed by your cloud providers include:

1. Assurance of accurate billing, and a methodology for handling objections or challenges to any automated metered billing
2. The ability to segregate different services into different methods of billing: for example, performance testing, analytics, security scanning, backup, and virtual desktops might all be measured differently and metered separately.
3. Ability to handle taxation issues from geography to geography, and from user to user. As each country and municipality has implemented different approaches to taxation of online commerce, your provider must be able to discern between these sources of use and meter them independently.

Provisioning

While auditing, monitoring, measuring and metering relate primarily to the cost savings features of the cloud, provisioning is a key enabler of the improved flexibility that comes from the cloud. However, it's not without its own unique qualities that must translate into your CSA:

1. *Core provisioning speed*. As part of a CSA, there should be baseline expectations of the speed of deployment of new systems, new data, new users, new desktops or any function that's core to the service provided by the cloud vendor.

2. *Customization*. It's unusual that any templated method of rapid provisioning can be used "out of the box" without configuration and customization. Without careful management of the expectations and contractual levels for this function, any savings gained by automated rapid provisioning can evaporate in the face of delays in customizations post-deployment.

3. *Testing*. Important to any strong CSA are provisions for testing automated deployment and scaling prior to need. This is particularly acute in areas where provisioning is employed in disaster recovery or backup situations.

4. *Demand Flexibility*. It does no good to have a technical solution to rapid provisioning if the system is incapable of dynamic de-provisioning to match downturns in demand.

This is not an exhaustive list of considerations, only the basic requirements of any contractual definition of rapid provisioning. Each organization will need to add their own particular additional topics, particularly for different industries or IT applications running in the cloud.

Change Management

Change is an inevitable part of any IT system, and the cloud is no different. Fortunately, there is little that is special about the cloud in regards to considerations for change management. Procedures for requesting, reviewing, testing, and acceptance of changes differ little from those already in use with other IT subcontractor contracts and outsource agreements. The only unique issue is the sensitivity that many have to changes that have potentially radical implications, such as the cloud. In this case, extra care should be taken to manage the process carefully.

Upgrades & Patching

A subset of change management is upgrades or improvements in existing contracted services, such as when an upgrade or patch is needed, or when a new version of an underlying management system or SaaS application is rolled out. In these cases, it's important to outline in your CSA a set of basic steps for these inevitable needs.

1. **Responsibility to develop requested changes.** There should be a clearly defined responsibility set for which party is in the lead for different types of upgrades. For example, if the upgrade is dependent on many subsystems or people internal to an organization, not in the cloud, it might be advisable to center the responsibilities on the contracting organization vs. the cloud provider. On the other hand, if the majority of the upgrade happens with cloud-

provider personnel within the cloud space, it's likely the provider would assume primary responsibility.

2. **Process for identifying a timeline to develop, test and implement the change.** There must be a clearly defined "chain of command" and project plan for all changes made to the cloud environment, properly resourced and timed to ensure reasonable contingencies and problem resolution. Here too, little is different regarding a cloud solution vs. a traditional IT solution, with the exception of the increased anxiety and scrutiny that the cloud draws today. This is in many ways simply a special case extension of change management policies which should already be in place.

3. **Process for resolving problems resulting from change.** Since problems can often be compounded and result from multiple factors both within and outside the cloud, a CSA-based outline of upgrade procedures must include a clearly defined set of responsibilities and methods for resolving issues introduced by any upgrade.

4. **Back-out process if the changes cause major failures.** Even the best-laid plans often run aground on the rocks of reality. Cloud service providers should automatically embed rollback checkpoints throughout an upgrade plan in order to "pull the plug" and restore any upgrade to its initial state should an unexpected and unsolvable problem crop up during the upgrade procedure. Throughout the process, regular communication meetings should occur to keep both parties in sync.

Step 7: Prepare for Service Failure Management

Service failure management outlines what happens when the expected delivery of a cloud service does not occur. Cloud service capabilities and performance expectations should be explicitly documented in the CSA, as described in Step 4. It is important to note that the term "service failure" can cover a number of different things, from the complete unavailability of the cloud service, through response times that are longer than those promised in the SLA, through error responses to valid service requests made by the users. Service failure management covers activities both of the cloud service customer and also the cloud service provider.

Service failure management begins with the detection and alerting that a failure has occurred. The cloud service customer must ensure that cloud service failures can be detected. The cloud service provider may provide service monitoring capabilities to the cloud service customer and may in addition provide alerts to the customer when cloud service failures occur. However, the cloud service customer must establish whether the monitoring and alerting capabilities provided (if any) meet the customer's requirements. The cloud service customer may often need to put in place their own set of cloud service monitoring and alerting capabilities to ensure that all the potential cloud service failures of significance to the customer are detected.

Once a cloud service failure is detected, then the cloud service customer must ensure that a management system is in place to alert appropriate customer staff, to report the failure to the cloud service provider (assuming that the failure was not already detected and reported by the provider) and to put into action any processes to mitigate the failure. For some cloud services and for some types of service failure, the cloud service customer may need to provide suitable evidence to the cloud service provider that a failure has occurred. The cloud service customer must track the progress of each reported failure and if the failure is not rectified within stated timescales, an escalation process must be followed.

The cloud service customer must understand the cloud service provider's service failure management procedures:

- The process for reporting failures detected by the customer
- The process which the provider will follow to address a reported failure
- The timescales for remedial action
- The process that the cloud service provider will follow subsequent to a failure to improve the provider's operations to avoid the failure occurring again

Planning for cloud service failures on the part of the cloud service customer will also often involve having a disaster recovery plan, which will be brought into action if the cloud service failure is likely to have a significant impact on the business.

Remedies

The primary remedy for service failure is service credits. These are typically based upon a percentage of the fees paid by the cloud service customer during the billing cycle. The actual percentage will vary depending on the cloud service provider and on the nature of the cloud service itself. However, it is common that these service credits will not exceed 100% of the paid fees. This can result in service credits not being in proportion to business cost or risk to the cloud service customer.

Limitations

Within each cloud service provider's service agreement there may be liability limitations for certain types of service interruptions. While these may vary dependent upon the provider, a sampling of several major providers shared the following exclusions:

- Scheduled or emergency outages
- Acts of force majeure
- Suspension of service due to legal reasons
- Internet access issues outside the control of the provider

In addition to common, shared limitations, there are cloud service providers who may also cite scheduled downtime as being excluded from the CSA metrics.

Roles / Responsibilities

The roles of cloud computing service failure management are described in the ISO/IEC 17789 Cloud Computing Reference Architecture [4]. The cloud service administrator has the responsibility to drive the incident management process and so needs to receive an alert when a service failure is detected. Assuming the service failure is impacting the customer use of the service, the cloud service administrator will engage the cloud service provider's incident management process, as described in the service agreement.

On the cloud service customer side, additional roles may be involved, including the help desk and the cloud service integrator. The help desk should be aware of the service failure and the likely impact and estimated time to resolution so as to be able to answer questions about the cloud service from cloud service users. The cloud service integrator would be engaged to triage the service failure and potentially propose solutions or workarounds to reduce the impact on the customer's business.

Monitoring and Notification processes

Monitoring of a service failure can be done in one of two ways.

1. The cloud service customer puts in place system(s) which monitor the customer use of the cloud service. The concept is that the customer does not rely on any capabilities of the cloud service provider and instead places instrumentation of some form in the customer-side components that use the cloud service. This might, for example, involve routing all customer requests to the cloud service through an instrumented component such as an Enterprise Service Bus (ESB). Requests made to the cloud service can then be monitored for success or failure, for their response times and for any other characteristics of importance to the customer. A set of rules can be put in place which will determine if there is a service failure and an alerting process invoked when a service failure is detected.

2. The cloud service provider has in place a cloud service monitoring system which has an interface enabling the cloud service customer to monitor the behavior of the cloud service and to receive alerts in the case of a service failure. These alerts should be integrated into the cloud service customer's alerting system. An alert would be sent to the cloud service customer when a service failure occurs – but the cloud service customer must understand what type of failures are notified through this process and it may well be the case that not all service failures of importance to the customer will be notified. Upon receiving a notification, the cloud service customer should follow their established service failure management process.

For a typical cloud service, it is likely that the cloud service customer will use both approaches to the monitoring of the cloud service. In some cases, the cloud service provider monitoring will be absent or will be inadequate for the customer. In other cases, there may be factors that can only be monitored by the customer, such as the effect of the internet connection to and from the cloud service.

For the notification of a service failure, in the ideal situation there should be a two-way automated interface between the cloud service customer and the cloud service provider that is used to transmit notifications of a service failure in both directions. This caters for either party becoming aware of the service failure. In the case where two way notification is not provided, the cloud service customer should expect there to be a facility for the customer to report a service failure to the cloud service provider – and a process for the customer to track what is happening in regard to each reported service failure.

Step 8: Understand the Disaster Recovery Plan

Disaster recovery is a subset of business continuity and focuses on processes and technology for resumption of applications, data, hardware, communications (such as networking), and other IT infrastructure in case of a disaster. By the term disaster we mean either natural disaster or man-made events that have an impact of availability of IT infrastructure or software systems.

It is common to see a false sense of security among cloud customers regarding disaster recovery planning. Just because businesses are outsourcing the infrastructure (IaaS), applications (SaaS), or platforms (PaaS) to cloud service providers does not absolve them of the need for serious disaster planning. Every company is unique in the importance it assigns to specific infrastructure/ applications, thus, a cloud disaster recovery plan is specific to each

organization, and business objectives should play an important role in determining the specificity of disaster recovery planning.

The process of devising a disaster recovery plan starts with identifying and prioritizing applications, services and data, and determining for each one the amount of downtime that's acceptable before there is a significant business impact. Service priority, required recovery time objectives (RTOs), and recovery point objectives (RPO's) will determine the overall disaster recovery approach. For example, in some applications maintaining uptime may be more important than having the data precisely replicated as of the last time of failure. Further, while 99%+ uptime SLAs are common in cloud computing (approximately 4 days of down time a year), it may not be adequate for specific application and business needs.

In general, current CSAs provide inadequate guarantees in case of a service outage due to a disaster. Most CSAs provide cursory treatment of disaster recovery issues, procedures and processes. That being said, it is rare for SMBs to internally develop the extensive disaster recovery infrastructure of large and established cloud providers.

Despite the limitations in CSAs, cloud adopters should address key disaster recovery questions/issues with their service providers early in the process of cloud adoption. The key areas to address with cloud providers are:

- How is service outage defined?
- What level of redundancy is in place to minimize outages including co-location of services in different geographical regions?
- Will there be a need for scheduled down time?
- Who has the burden of proof to report outages? This can be difficult to prove in case of conflicts with the cloud providers.
- What is the process that will be followed to resolve unplanned incidents?
- How will unplanned incidents be prevented or reduced?
- When does the time clock start on lack of service availability in order to measure service credits?
- How will incidents be documented or logged?
- What actions will be taken in the event of a prolonged disruption or a disruption with a serious business impact?
- What is the process of performing disaster recovery testing, and how often are the tests conducted? Are the reports of the tests provided to clients and are the tests automated?
- What is the problem escalation process?
- Who are the key service provider and customer contacts (name, phone number, email address)?
- What is the contingency plan during a natural disaster?

• How is the customer compensated for an outage? It must be noted that cloud providers have limits on the maximum compensation provided in case of an outage, and the compensation is an insignificant remedy in case of serious outage.

• Does the cloud vendor provide cloud insurance to mitigate user losses in case of failure? Although this is a new concept, some major cloud vendors are already working with insurance providers.

Answers to the questions above will be highly specific to particular organizations, and their specific disaster recovery needs. For large enterprises the questions mentioned above can be used as a framework to seek a stronger disaster recovery component in a negotiated CSA. It is important to emphasize that this is only possible for large enterprises with large contracts. Established cloud vendors are quite resistant to altering existing CSAs.

There are large numbers of events that can have negative impact on the availability of cloud services provisioned by customers. Although, detailing all of them is out of the scope of this section, some of the important areas that cloud customers should consider are in areas of security/intrusion detection, denial of service, viability of a cloud provider, data ownership and recovery. As an example to highlight the above, consider a company using SaaS for critical applications, such as order management, billing, or ERP. The cloud user will face major technological hurdles in shifting to another provider in case of a disaster like a financial failure of the cloud provider. Cloud users should make it a priority to address key contingencies in case of such an event. Issues such as access to data and the application in a timely manner are critical to clarify.

While, in most cases, companies will be able to retrieve the application data from an established SaaS provider, the business logic and software systems will be left behind. One solution is to deploy the SaaS software onsite and run it internally – clearly a difficult and risky solution to implement. So, despite good planning, in some cases no easy solutions are available for negative events. Development of data and meta-data standards in specific application domains could provide a considerable benefit for customers and allow them to migrate to different SaaS solutions in the event of a disaster. The development of such standards though is in direct conflict with the interests of many providers, and will take time to materialize.

It is also important to understand that risk mitigation related to disaster recovery for cloud solutions will also depend upon the specific cloud type (IaaS, SaaS etc). Compared to the SaaS example above, in the case of a negative event for an application running on an IaaS, the client can implement a different set of solutions. One example solution would be to architect the application to continue performing in the face of individual resource failure (e.g., server failure, storage failure, network failure, etc), or in the case of a significant infrastructure failure use hot/warm sites in a different geographical zone or on a completely different cloud. The key point to understand is that risks and solutions associated with negative events will be different for SaaS, IaaS and PaaS.

When it comes to disaster recovery the public cloud presents a due-diligence paradox. While there are myriad options for implementing disaster recovery, and the cloud may simplify enterprise IT by abstracting away a lot of the complexity, it also increases the difficulty of performing comprehensive due diligence including testing of disaster recovery procedures. Lack of such diligence accompanied by weak CSAs represents a potential risk in the area of business continuity and disaster recovery. Thus, companies should view developing and testing a disaster recovery plan as an important part of moving to the cloud. Companies can consider using business continuity/disaster recovery standards as part of their planning

efforts. Existing standards such as BS 25999:2007, NFPA 1600:2010, NIST SP 800-34, ASIS SPC.1-2009, ISO 27031, and ISO 24762 can provide an effective starting point for planning disaster recovery.

Business Impact Analysis (BIA)

The fundamental task in business impact analysis (BIA) is understanding which processes in your business are vital to your on-going operations and to understand the impact the disruption of these processes would have on your business. From an IT perspective, as the National Institute of Standards and Technology (NIST) views it: "The BIA purpose is to correlate specific system components with the critical services that they provide, and based on that information, to characterize the consequences of a disruption to the system components."

According to the Business Continuity Institute (www.thebci.org), a recognized leader in business continuity management and certification, there are four primary purposes of the business impact analysis:

- Obtain an understanding of the organization's most critical objectives, the priority of each, and the time frame for resumption of these following an unscheduled interruption.
- Inform a management decision on Maximum Tolerable Outage (MTO) for each function.
- Provide the resource information from which an appropriate recovery strategy can be determined/recommended.
- Outline dependencies that exist both internally and externally to achieve critical objectives.

Business impact analysis is the process of figuring out which processes are critical to the company's on-going success, and understanding the impact of a disruption to those processes. Various criteria are used including customer service, internal operations, legal or regulatory, and financial. From an IT perspective, the goal is to understand the critical business functions and tie those to the various IT systems. As part of this assessment, the interdependencies need to be fully understood. Understanding these interdependencies is critical to both disaster recovery and business continuity, especially from an IT perspective.

Conducting a business impact analysis (BIA), we identify the criticality and the recovery time objective (RTO) for each service (that is, the maximum length of time the organization can afford to be without the service). We can also establish the recovery point objective (RPO) (that is, the point to which data must be recovered – e.g. start of day, end of day, or to a checkpoint). The results of this process will form the basis of the SLA requirements for availability and reliability (the number of incidents of outage) for each service. If a BIA has already been done for business continuity purposes, this needs to be retro-fitted into any existing SLAs so that they are made compatible with

the business continuity plan. The same applies to external suppliers: for instance, we may have a requirement for 99.5 percent availability (in a 24/7 operation, this equates to about four hours downtime a year). A maintenance contract for support of this activity which allows four hours to get on site is simply inadequate.

Business impact analysis includes the steps listed earlier, but we can break them out into a few more discrete activities or steps:

1. Identify key business processes and functions.
2. Establish requirements for business recovery.
3. Determine resource interdependencies.

4. Determine impact on operations.
5. Develop priorities and classification of business processes and functions.
6. Develop recovery time requirements.
7. Determine financial, operational, and legal impact of disruption.

The results can be sorted into tiers. A financial institution might, perhaps, define tiers as follows

Tier One: Continuous availability requirement: 99.999 percent availability, maximum of one outage and four minutes downtime per year.

Tier Two: High availability, maximum of one outage per year, maximum four hours outage per year.

Tier Three: Recovery essential within 24 hours; maximum three outages per year.

Tier Four: Recovery required within 3 days; maximum four outages per year

Tier Five: Delayed recovery – all other services.

Step 9: Develop an Effective Governance Process

The use of cloud services by a cloud service customer means that the customer organization is placing some parts of its IT operations – and hence part of it business processes - in the hands of outside suppliers in the form of one or more cloud service providers. As a result of the interface(s) between the customer and the provider, there is a need for strong and detailed governance of the use of the cloud services on the customer side.

The first part of the governance process involves the control and oversight of the previous steps outlined in this practical guide, which provide the necessary underpinnings for the selection and use of cloud services. The second part of the governance process is the regular ongoing review of the use of each cloud service, to ensure that it meets business requirements and to ensure both internal and external user satisfaction with the cloud services and the applications built on them. The governance process should also deal both with changing business and user requirements and also with any changes to the cloud service(s) that may be made by the cloud service provider.

Table 12 below highlights the key elements required to operate a successful governance process.

Table 12: Governance Process

Element	Description
Periodic assessment of achieved cloud service levels against agreed CSA	• Reports from cloud service provider of cloud service levels • Monitoring reports on cloud service usage created by customer cloud service administrators
Periodic assessment of compliance of cloud service	• Where the compliance of the cloud service to specific standards or regulations is important to the customer, it is necessary for the customer's governance process to periodically check that the cloud service still has valid proof of compliance.
Service failure reports	• Reports of any service failures or incidents which affect o Service availability o Security, particularly security breaches o Protection of personal data
Notification of changes from the cloud service provider	Any change notifications from the cloud service provider which relate to the cloud services being used (change of APIs, change of functionality, change of service level objectives, change or cloud service pricing, change of terms in the CSA)
Key indicator reports	Four key indicators should be tracked to ensure that the CSA criteria are being met and that the downstream users of the service (either internal or external to the enterprise) are experiencing the level of service that has been agreed to: • High impact problems and time to resolution • Number of open problems and their respective impact • Total view of problems not resolved within agreed to time frames • Trends of number of problems being reported with the resulting resolutions
Problem reports	In order to ensure CSA compliance, a set of reports needs to be produced: • Reports that focus on the current reporting period addressing: o All problems reported (sorted by impact) o Problems closed (sorted by impact) o Duration of open problems (sorted by impact)
Request reports	Reports on (non-problem) requests made by the cloud service customer to the cloud service provider: • All requests made • Number of open requests • Time to action requests
User satisfaction reports	Reports on user satisfaction with the cloud service(s)

The cloud service customer must periodically review the elements described in table 8 and decide on an appropriate course of action if the cloud services do not meet the terms of the agreement or do not meet business requirements. How the review is performed is a decision for the customer and it is likely to depend on the size and structure of the customer organization. A degree of formality and record keeping is advisable since in some cases, evidence may need to be prepared for presenting to the cloud service provider, especially if there are matters under dispute between the customer and the cloud service provider.

What constitutes an appropriate course of action will depend on the nature of the issue(s). Some breaches of the CSA terms may trigger remedy terms which imply some level of compensation to the customer – but it may often be the case that the customer must formally raise a request to the provider in order to trigger the terms of the remedy. More serious breaches or incidents are likely to require more significant action on the part of the customer. This may take the form of discussions between senior management from the customer with their counterparts from the cloud service provider. Alternatively, it may take the form of the customer deciding to switch their use of cloud services to another cloud service provider, triggering the termination process.

For problems that require higher management awareness, it is the responsibility of those involved in the governance process to advise their respective management chains on the status of a particular issue.

Escalation Process

Inevitably, there will be problems which fall outside the normal management process and will need additional focus to ensure a timely resolution. An example of the exceptional process is a major outage, i.e. loss of service, which cannot wait for a periodic meeting and requires an immediate notification of the management chain.

While we use the term escalation, the escalation process is really upward communication for awareness for a particular situation and not an upward delegation of responsibility for the resolution of the problem.

Table 13 below highlights the overall objectives of escalation, general guidelines for when to initiate an escalation, and the types of escalations that can be invoked.

Table 13: Escalation Considerations

Consideration	Description
Objectives	• Raise management awareness to avoid surprises (gives the perception that senior management is in control of the situation). • Gain agreement for action plans to resolve a problem. • Develop either a plan and gain agreement for additional resources, when required.
Guidelines	• Problem has a critical impact to the overall business to either an internal service or a customer facing service. • Service is still available but is significantly degraded; potential impact to a customer facing service. • Problem is of a significant impact and has missed the agreed to targets for resolution. • Independent of impact, problems are not being closed within the expected guidelines. • Number of problems is increasing with no agreed to resolution to reverse the trend. • Requests to the cloud provider to participate in root cause analysis or problem resolution in an associated system or tool are ignored.

Types	• **Immediate**
	  ○ A critical business impact is identified.
	  ○ Significant impact to a customer facing service.
	• **As required**. Typically after a review when:
	  ○ The duration of problem resolution is not being met.
	  ○ Number of open problems exceeds expectations.
	  ○ Trend for reported problems is increasing without a satisfactory resolution plan being offered.

Once an escalation has been initiated, the goal is to ensure that both chains of management understand the problem, its impact, and the currently agreed to action plan for resolution including containment of the problem, especially if the problem impacts an external customer service.

If a resolution of an escalated problem cannot be reached through the escalation process then the terms of the CSA can be brought to bear to force resolution. One of the outcomes of continuous breaches to the CSA can be termination of the agreement with the provider for the contracted service(s). It should be noted that the minutes generated from the management process is an important set of documentation to support the termination process.

Escalation should not be considered a last resort in the problem management process. Escalation should be used as an early warning activity to raise management awareness of a potential problem before it becomes critical. Escalation is a tool to manage the services and ultimately provide the best services to the users of the service(s), whether the users are internal or external to the organization.

Step 10: Understand the Exit Process

An exit clause should be part of every CSA and describes the details of the exit process including the responsibilities of the cloud provider and consumer in case the relationship terminates prematurely or otherwise.

There are numerous potential scenarios that could cause the termination of service between customer and provider which would result in the execution of the exit process. For example, a provider may be unable to deliver the required levels of performance and availability specified in the SLA, or it may be the case that the provider is going out of business. Regardless of the reason, a clearly defined exit process that ensures secure and speedy transfer of customer data and applications is essential.

A customer exit plan should always be prepared at the outset of the CSA and is an integral contractual annex. This plan should ensure minimal business disruption for the customer and ensure a smooth transition. The exit process should include detailed procedures for ensuring business continuity and it should specify measurable metrics to ensure the cloud provider is effectively implementing these procedures.

The most important aspect of any exit plan is the transmission and preservation of cloud service customer data, which is critical to achieving business continuity. In addition, customers must ensure that their data is completely removed from the provider's environment once the exit process is complete. Customers should look out for and be aware of the following details when they evaluate the exit clause included in a CSA.

- The level of provider assistance in the exit process and any associated fees should be clear in the CSA. In most cases, there should be no additional cost associated with the exit process.

- Providers should be responsible for removing customer data from their IT environments, or at least aid the customer in extracting and erasing their data by providing clear and concise documentation.

- The format of the data transmitted from the provider to the customer should be specified in the CSA and should leverage standard data formats whenever possible to ease and enhance portability.

• The CSA should specify that all data and information belonging to the customer is maintained for a specific time period after transition and then be completely removed after that time.

o The typical time period is 1-3 months which gives the customer sufficient time to find a new provider and to continue receiving service from the current provider in the interim.

o The time period should be explicitly documented in the CSA and only with the customer's written approval should data be removed and/or destroyed before that time.

• Customers should ensure that the CSA provides appropriate business continuity protection during the exit process.

• At the completion of the exit process, customer should receive written confirmation from the provider that all of the customer data has been completely removed from the provider's IT environment. The written confirmation should also state that the provider agrees not to use the customer data for any reason in the future, including using the data for statistical purposes.

The bottom line is that customers should undertake due diligence when evaluating and ultimately selecting a cloud provider. A trustworthy cloud provider should be prepared to provide customers on a fair and effective exit strategy.

4.3. Summary of Keys to Success

Table 14 summarizes the critical keys to success for any customer organization evaluating and comparing CSAs from different cloud providers.

Table 14: Summary of Keys to Success

Key to Success	Summary
Review internal policies and processes	• Identify key processes and policies that will be affected by a move to cloud services. • Purchasing and reporting are key areas for review.
Develop a strong business case and strategy for cloud computing environment	• Assess criticalness of services being deployed in the cloud. • Determine functional and non-functional requirements for each service (performance, availability, security, privacy, etc.). • Understand legal and regulatory requirements concerning the data maintained in the cloud. • Identify key performance metrics for each service.

Assess provider's CSA against functional and non-functional requirements	• Based on the criticalness of the service being deployed in the cloud, determine if the cloud provider's CSA is sufficient to address the functional, non-functional, legal, and regulatory requirements of the service. • If not, determine if the cloud provider is willing to negotiate on the key aspects of the CSA that are not in line with your business strategy. • If the cloud provider is not willing to negotiate on these critical points, seek alternative providers who more closely address your requirements. • If a cloud provider who addresses your requirements cannot be found, strongly consider keeping the service within your enterprise IT environment.
Determine how to monitor CSA performance	• Assuming a cloud provider is found that meets your service requirements; understand the management process defined in the CSA. • Ensure your CSA includes the ability to see, assess and react to key performance measurements that will help keep your cloud infrastructure running smoothly. • Understand the notification process when service issues arise including method and timliness of notifications along with prioritization and severity level assessment of issues. • Be aware of remedies and liability limitations offered by the cloud provider when service issues arise.
Ensure an adequate disaster recovery plan can be defined and executed	• The cloud customer bears the risk of disaster scenarios that severely limit the ability of their cloud provider to deliver service. • Cloud customers must understand the provider's ability to support their data preservation strategy which includes criticalness of data, data sources, scheduling, backup, restore, integrity checks, etc. • Roles and responsibilities must be clearly documented in the CSA. In many cases, the cloud customer may be responsible for implementing most of the data preservation strategy. • Based on the criticalness of the data, cloud customers should clearly define recovery time objectives. • Customers should test and verify the disaster recovery plan prior to production deployment. • Cloud customers should consider purchasing additional risk insurance if the costs associated with recovery are not covered under their organization's umbrella policy for IT services or operational risk riders.

Ensure support for an efficient exit process	• The goal of the exit plan is to ensure minimal business disruption for the customer should the relationship with the cloud provide terminate prematurely.
	• The exit plan should be taken into account during the assessment phase of potential cloud providers.
	• The provider's CSA should be carefully reviewed to ensure the customer defined exit plan is capable of being implemented.
	  ○ The customer should be able to terminate the agreement at any time, without penalty, provided sufficient notice is given to the provider.
	  ○ Data maintained on the provider's cloud resources should be stored using standard formats to ensure data portability.
	  ○ Transmission of data from the provider's cloud resources should leverage standard packaging and data transfer techniques.
	• Roles and responsibilities must be clearly documented in the CSA. In many cases, the cloud customer may be responsible for initiating most of the exit process steps.

In addition, emerging standards in the following areas will help improve the ability for customers to evaluate and compare the service levels offered by different providers:

- Standards that create consistent ways to describe services and associated terms including price.
- Standardized metrics that allow customers to effectively track and compare CSA performance.
- Standardized security and regulatory compliance requirements to identify control points for risk management.
- Standards that enable coordinated end-to-end CSA management for both cloud customers and cloud providers.

Cloud computing offers a value proposition that is different from traditional enterprise IT environments. With proper focus on the key success factors, customers are able to effectively review and compare CSAs from different cloud providers to ensure the promise of the cloud is realized.

5.CHAPTER V:

On service costs and metering of cloud services consumed: a perspective into cloud service billing

5.1. Cloud service system pricing and billing models based Cloud security as a service

When cloud services system is directly facing final consumers, the design of service pricing strategy is essential. Pricing strategy is directly related to the user experience and satisfaction, but also affects the cloud service provider's revenue. Cloud security is a factor that cloud service users are very concerned about, and directly determines the user of cloud services system availability and reliability. The cloud security as a service is the inevitable trend of cloud computing applications for users. To this end, the cloud service pricing and billing system should reflect the value of cloud security as a service for users. Cloud service system pricing system should be the clear, flexible, easy to understand and easy to select for individuals or enterprises users. and it's clear features means that the application can provide the functionality and the corresponding functional safety, and how each should charge is at a glance to cloud billing service users; flexible features means the different combinations of the functionality and the corresponding functional safety should be truthfully reflected in the price; Understanding is that price policy should have specify, scientific and reasonable framework; facilitate selection refers to the option of different types, different needs of users according to their situation. Based on this, the paper presents one reference model of the pricing and billing for cloud services system, as shown in Figure 9.

In this pricing model, according to the order of accounts, planning, feature pack type, feature pack level and cloud security level, the former and the latter is a one to many relationship, means that an account can have multiple plans, a plan may correspond to more than functional packet type, a function package types can correspond to multiple feature pack, a feature pack may correspond to multiple cloud security level. Along with this expansion-many relationship, the user selects more and more flexibility, but the choices are the lower the convenience, operability of the application weaker.

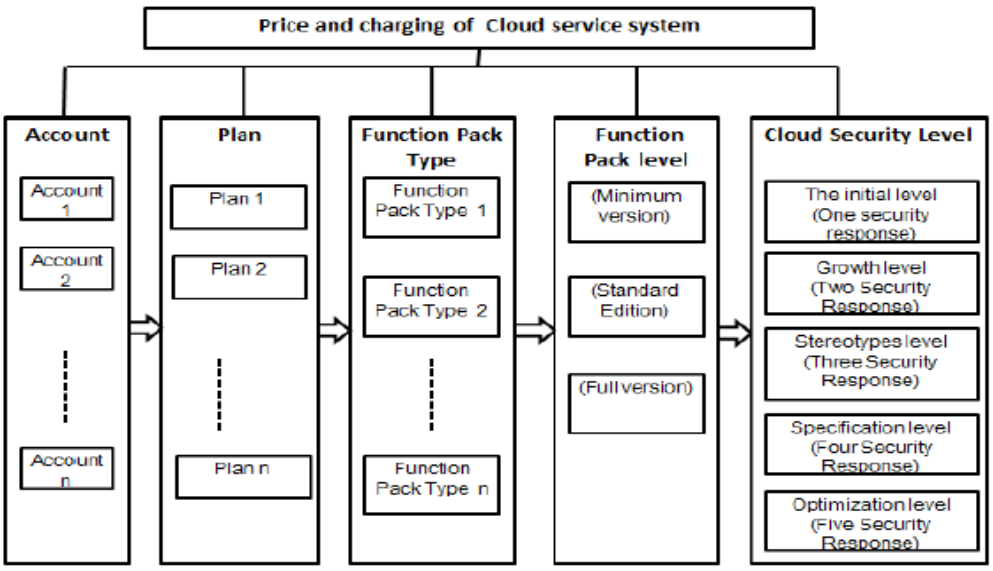

Figure 9: price and charging models of cloud service system

Billing plans added to the concept of time, it is different with the function package types, but can also come through differentiated market segments to user choice. Flexibility of billing account is at minimum, but provides the most convenient package solution to the cloud service users. An account is often of the users of multiple plans; provide a variety of planned combination according to account the different needs. Cloud service providers can refer to the above pricing model, select the appropriate option for pricing.

5.2. SLA Pricing

The service-level agreement between the consumer and the service provider states the lawful relationship between the two parties. This agreement protects the rights of both sides under any situation .SLA provides fairly knowledge about the service resources and features such as the quality of service and the price of providing the service .

The SLA Management part is the associated part of several resources, one of these resources is pricing. Pricing strategies are ways to determine the service price based on the service's demand and the service's equipment. Service providers use their rule to implement a smart pricing mechanism which will increase their profit. There are various charging strategies to set the service prices. For example, services price can be calculated built on the request peak/peak-off delivery time, service demand, service availability, service supply and charging rates even if it's static or dynamic. Most of the cloud computing providers use one of the three basic models for pricing which are; bid price model, static pricing, and dynamic pricing. Providers seem to prefer the dynamic pricing because they have to increase the service's price constantly, to maximize their incomes.

5.2.1. Dynamic Pricing:

Dynamic pricing means the continuous altering of the service price. The price changes continuously based on the service supply and service demand. If the service demand increased or the service supply decreased, the price will rise up. And if the service demand decreased or the service supply increased, the price will go down. Dynamic pricing has an influence on the SLA negotiation between the consumers and the providers because the price might change during the negotiation process. But, it is significant to understand that when an SLA has been approved, the price of that service level agreement must be fixed for the SLA lifetime rest. The formatting of an SLA contract and the particular price setting affect that specific provider-consumer collaboration only. The price might change if the same provider interacts with different consumers.

Price is calculated by certain functions that might be simple or complex depending on the variety and the quantity of its parameters. Simple functions will depend on few parameters. However, the complex functions will depend on many parameters. Those parameters can be measurements for the internal or the external state of the provider. Instances on internal state parameters are the service current, service loads, and historical data. External state parameters show the actual condition of the marketplace, but they are difficult to measure .

A critical parameter in service charging functions relates to the usage of the current service. The availability of resources is difficult to guarantee for future service supply and service

demand. That's why most of the biggest cloud providers such as Amazon EC2, Microsoft Azure, Dell Boomi, and Google Cloud state at their SLAs that the service availability is 99.9%.

Another main parameter in calculating the service price is the risk. In some situations the service level agreement in cloud has fixed deadlines or extreme obligation so the negotiators from both sides should find a solution to cover the liability. So, insurance premium must be involved in the price.

The base cost of service such as both hardware and software purchasing, storing cost, and maintaining cost is another major price parameter. Some service providers may face business's problems in their start so they offer their services in lower costs than the base cost. But that can't last in the long term because successful business model must be gainful.

5.2.2. Price Architecture SLA Negotiating:

SLAs negotiating price architecture consists of seven main functions which are; SLA template repository, resource capabilities, resource availability, business objectives, dynamic pricing component, SLA negotiator and the signer as shown in Figure 10. They are in details below;

- SLA template repository: The provider sends non-obligated SLA templates to the consumer as a procedure to announce the offering services by the cloud.
- Resource capabilities: Resource Capabilities presents the documents and the data that concerned with the capabilities of the service's resources.
- Resource availability: This function supplies the up-to date data about the system latest status, containing existent load, predicted request and upcoming reservations.
- Business objectives: This component is more related to the provider of the service. It is the Logical clarification of his business preferences, performance, behavior and management, *etc*.
- Dynamic pricing component: Dynamic pricing component calculates the service price based on the previous functions.
- SLA negotiator: This function is the core function that allows the consumer and the service provider to communicate and negotiate the SLA. The protocol of the negotiation process describes the messages which are sending to consumer's negotiators by the provider's negotiators and vice versa. Those messages might consist of quotes requests, the current quotes, deals, discount, and at last both approved and unapproved notifications.
- Signer: After the final agreement on the service's price, all parties including the provider and the consumers should present their approval digitally signing.

The architecture seems only a specific communication between a consumer and a provider. But in an actual system, several negotiations are rolling at the same period. Every single negotiation might be separate from the other negotiations.

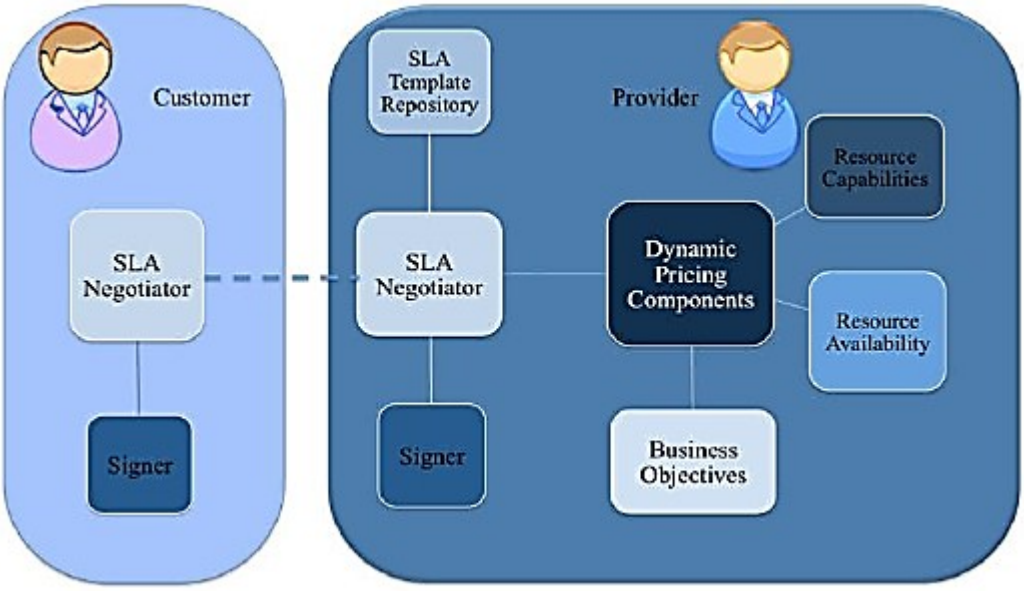

Figure 10: SLA Negotiating Price Architecture

5.3. Charges and Billing

5.3.1. Billing Options

The amount payable for the Cloud Service offerings are specified in the Order Document as follows:

- Entire commitment amount upfront
- Monthly (in arrears)
- Quarterly (upfront)
- Annually (upfront)

The selected billing option will be valid for the length of the term as specified in the Order Document. The amount payable per billing cycle will be based on the monthly or annual subscription fee and number of billing cycles in a year plus any overage charges.

5.3.2. Partial Month Charges

The Partial Month charge is a pro-rated daily rate. The Partial Month Charges are calculated based on the remaining days of the partial month starting on the date you are notified by CSP that your access to the Cloud Service offering is available.

5.3.3. Overages

If your actual usage of the Cloud Service during the measurement period exceeds the entitlement stated on the POE portion of the Order Document, then you will be invoiced for the overage, as set forth in the Order Document.

5.3.4. Term and Renewal Options

5.3.4.1. Term

The term of the Cloud Service will begin on the date that CSP notifies you that you have access to the portions of the Cloud Service that are described in the Order Document. The PoE portion of the Order Document will confirm the exact date of the start and end of the term, as well as how or whether the term will renew. You are

permitted to increase your level of use of the Cloud Service during the term by contacting CSP.

5.3.5. Cloud Services Term Renewal Options

Your Order Document will set forth whether the Cloud Service will renew at the end of the term, by designating the term as one of the following:

a. Automatic Renewal

If your Order Document states that your renewal is automatic, you may terminate the expiring Cloud Service term by written request, at least ninety (90) days prior to the expiration date of the term that is set forth in the Order Document. If CSP does not receive such termination notice by the expiration date, the expiring term will be automatically renewed for either a one year term or the same duration as the original term as set forth in the PoE portion of the Order Document.

b. Continuous Billing

When the Order Document notes that your billing is continuous, you will continue to have access to the Cloud Service and will be billed for the usage of the Cloud Service on a continuous billing basis. To discontinue use of the Cloud Service and stop the continuous billing process, you will need to provide CSP with ninety (90) days written notice requesting that your Cloud Service be cancelled. Upon cancellation of your access, you will be billed for any outstanding access charges through the month in which the cancellation took effect.

c. Renewal Required

When the Order Document notes that your renewal type is “terminate”, the Cloud Service will terminate at the end of the term and your access to the Cloud Service will be removed. To continue to use the Cloud Service beyond the end date, you will need to place an order with your CSP sales representative to purchase a new subscription term.

6.CHAPTER VI

Understanding the perspective of cloud services from the point of view of cloud service providers

6.1. Cloud Providers Considered

We briefly give an overview of cloud services offered by Amazon , Rackspace , Microsoft Azure , . These providers offer IaaS and PaaS compute and storage services. The compute service comprises of a virtual machine (or instance) or CPU cycles that a customer can purchase on an hourly, monthly, or yearly basis. The storage service allows storage and retrieval of blob or structured data. We interchangeably use customer and user to refer to the clients of cloud providers.

6.1.1. Amazon

Amazon is an IaaS provider and offers compute (EC2) and storage (S3) services. In EC2, a customer can obtain virtual machines (instances) by the hour or reserve them in advance for an entire year . In addition, EC2 offers spot instances where a customer can bid for compute capacity. EC2 SLA is applicable to hourly, spot, and reserved instances. The storage service S3 provides mechanism for storing and retrieving data objects using put(), get() operations with data size ranging from one byte to five tera bytes.

Amazon also provides a remote disk capability for its virtual machines, namely, Elastic Block Store (EBS). EBS volumes are replicated within an availability zone. A data center (or region) can contain multiple availability zones. The availability zones do not have power, networking, or hardware equipment in common. EBS volumes are not backed by any SLA; however, the snapshots of EBS volumes can be stored on S3, which as mentioned before is backed by a SLA. When creating an instance, the user must specify the region and availability zone in which she creates the instance.

Amazon also provides a Simple DB service which is a simplified relational database service. However, the service is still in beta at the time of writing of this paper. Among S3, EBS, and SimpleDB services, only S3 is backed by an SLA .

6.1.2. Windows Azure

Windows Azure is a PaaS and IaaS cloud provider that offers compute (Azure Compute) and storage (Azure Storage) services. Azure Compute comprises of three types of compute services (which it refers to as roles), namely, web, worker, and a VM. A web role provides a web based front end for an application and comprises of an IIS server . A worker role is useful for generalized development. It can run Apache Tomcat and Java Virtual Machines (JVMs) and can be used to perform background processing for a web role. A VM role is similar to instances in Amazon EC2, and gives user complete control over the virtual machine. However, at this time, VM roles are only available in beta and are not covered by Azure Compute SLA . The compute service can only be purchased on an hourly basis, and cannot be reserved in advance for the entire year. Azure Compute service defines the notion of a fault domain and an upgrade domain. Each compute role belongs to a fault domain and an upgrade domain. A fault domain comprises of a single point of failure and is at least a physical machine, but may also be a rack of machines; the precise details of what comprises a fault domain are not available. An upgrade domain defines which compute roles can simultaneously receive the software or operating system updates. A fault domain may span several update domains. Likewise, an update domain may also span several fault domains. Azure also provides Azure Storage , an S3 like storage service, which can be used for storing and retrieving blob and structured data. It also provides a queuing service and remote disks (known as Azure Drive). Azure storage service is backed by a SLA .

6.1.3. Rackspace

Rackspace is an IaaS provider that provides compute instances similar to Amazon EC2 and VM role of Azure, which it refers to as "Cloud Servers". A customer can obtain VMs on an hourly basis which are covered by an SLA . However, unlike EC2, Cloud Servers cannot be reserved in advance for the entire year. Rackspace also provides a managed service level for Cloud Servers. As part of the managed service, Rackspace is responsible for applying software and security patches for operating system and middleware. Rackspace provides a storage service called "Cloud Files" which allows a customer to store and retrieve files in the cloud and is covered by an SLA . The stored files are internally replicated by Rackspace.

6.2. Description of SLAs

We describe SLAs of compute and storage services offered by cloud providers considered in this chapter.

6.2.1. Amazon

Amazon EC2 and S3 services are backed by distinct SLAs. Below, we describe the SLAs of these services in detail.

6.2.1.1. EC2 SLAs

Amazon EC2 SLA is defined on a per data center (region in Amazon speak) basis instead of per instance. EC2 offers a 99.95% region availability rate (service guarantee). If a user is unable to access her instances in one region during a contiguous period of five minutes or launch replacement instances, the region is deemed to be unavailable during those five minutes. The burden of providing the evidence for region unavailability is on the user. Strictly speaking, if a user is running at least one VM which she cannot access during a five minute interval and cannot launch a replacement, she is eligible for a service credit if the credit value is above one dollar. A customer can claim a service credit anytime the service falls below the availability SLA in the last 365 days or since the last time a service credit claim was filed by the customer. The service credit is up to 10% of a customer's bill (excluding any one-time costs) for the instances affected by the outage. Service credits are typically only applicable towards future EC2 payments. Amazon requires that the service credit claim be received from the customer within 30 business days of the last reported incident in the filed claim.

Amazon does not provide any service credit for failures of individual instances not attributable to region unavailability. This clause means that even if a region (data center) is available, but some services in that region fail such as EBS on which an instance depends, Amazon is at least legally not bound to provide a service credit, although it may provide a credit at its own discretion. For example, Amazon provided a service credit for its April 2011 outage due to EBS failures. Further, Amazon does not provide any service credits if VMs suffer from any performance issues. A VM can suffer performance degradation due to co-location or hardware differences of the underlying physical machine .

Amazon EC2 SLA does not specify that scheduled and unscheduled maintenance are excluded from the service guarantee. EC2 SLA is defined on a data center basis, and, arguably, the data center being unavailable for scheduled maintenance is unlikely because it will impact all customers running their instances in that data center.

6.2.1.2. S3 SLAs

Amazon S3 SLA provides storage request completion guarantee of 99.9% over a billing month (service guarantee time period). A storage request is considered failed if S3

server returns an “Internal Error” or “Service Unavailable” response to a request. These responses correspond to HTTP response codes 500 and 503. The burden of reporting request failure and providing evidence is on the customer. S3 calculates failed requests over a five minute interval, which are then averaged over a month. The failed requests are calculated by dividing the number of requests generating an error response to the total number of requests in the five minute interval. The percentage of completed transactions in the billing month is calculated by subtracting from 100% the average of failed request rates from each five minute period.

The service credit is 10% of the customer bill if completion rate is below 99.9% and 25% of the customer bill if completion rate is less than 99%. Amazon must receive the claim within 10 business days after the billing month in which the incident occurred. Similar to EC2 SLA, Amazon S3 SLA does not exclude scheduled and unscheduled maintenance from service guarantee. Moreover, S3 service does not specify any performance guarantees on the storage requests.

6.2.2. Windows Azure

Azure compute and storage service are backed by separate

SLAs which are described below.

6.2.2.1. Azure Compute SLA

Azure Compute SLA provides connectivity and uptime service guarantees for its non-beta compute roles over a billing month (service guarantee time period). For Azure Compute SLA to be applicable, a customer must deploy at least two instances of a compute role type in different update domains.

Unlike Amazon EC2, which provides availability SLA on a per data center basis, Azure SLA is calculated as an aggregate over the deployed roles. Azure SLA defines two service guarantees, namely, external network connectivity and uptime which are calculated on a monthly basis. The connectivity service guarantee is defined as the aggregate time since all the Internet facing roles have been started minus the five minute intervals during which any role does not have connectivity, divided by the aggregate time since roles have been started. Like Amazon EC2, Azure calculates downtime for its compute roles in increments of five minute intervals.

The uptime service guarantee is defined as the aggregate time since roles have been deployed and started minus the time across all role instances which do not run for more than two minutes without corrective action being initiated, divided by the aggregate time since roles have been started. Any performance or availability issues due to regular platform upgrades and patches are excluded from the uptime service guarantee calculation. The service credit is 10% of the customer bill if connectivity and uptime percentage is below 99.95% and 99.9%, respectively, and 25% if less than 99.9%. The onus for reporting a SLA violation and providing evidence is on the customer. Microsoft requires that a customer notifies it of the incident within five business days following the incident in order to be eligible to file a claim. Then, Microsoft must receive the claim within a month of the billing month in which the incident occurred.

6.2.2.2. Azure Storage SLA

Azure Storage SLA defines service guarantee as percentage of completed transactions in a billing month. A request is considered failed if the maximum time to process the request exceeds the time specified in the service guarantee.

Azure Storage calculates failed requests over one hour interval by dividing the total number of failed requests to the total storage requests. The percentage of completed transactions within a billing month is calculated by subtracting from 100% the average of failed request rates from each one hour period in the billing month. Similar to Azure Compute, the onus for reporting an SLA violation is on the customer. Microsoft requires that a customer notifies it of the incident within five business days following the incident in order to be eligible to file a claim. Then, Microsoft must receive claim within a month of the billing month in which the incident occurred. The service credit is 10% of the customer bill if number of completed transactions are below 99.9% and 25% of the customer bill if less than 99%. Similar to S3 SLA, Azure Storage SLA excludes any transactions from SLA computation that are beyond its reasonable control, and that result from customer's fault or abuse of the system. Unlike S3 SLA, Azure storage SLA gives detailed examples of excluded transactions such as pre-authentication failures, abusive transactions, creation or deletion of containers, tables or queues, or flooding requests not obeying back off principles.

6.3. Comparison between SLAS of the Existing Cloud Service Providers

In this section, we explore how different cloud providers implement SLA. The characteristics chosen for the sake of comparison are selected based on similarities in attributes in the cloud SLAs we examined. The comparison outcomes can be found in Table 15. Furthermore, there is a number of steps developed by the cloud standards customer council that presents a series of ten steps for the consumer of the cloud service to evaluate and base its negotiation with the cloud vendor based on. The steps are explained briefly below:

1. To understand the roles and responsibilities: AUPs (acceptable use policies) are what cloud consumer is mainly concerned about. Reviewing them thoroughly and carefully allows the consumer to understand exactly what their roles and responsibilities are along with the cloud providers' roles and responsibilities.
2. Evaluate business level policies: When reviewing SLA the consumer should consider major policy issues because the SLA policies, the business strategy and policy are somewhat dependent.
3. Understand Service and Deployment Model Differences: This step is to make sure that the consumer understand what is the service model of the cloud (SaaS, PaaS, or IaaS), what are its characteristics? What are its objectives and KPI's? Furthermore, to understand the deployment model of the cloud presented in the service agreement (private, public, community, or hybrid). It is critical that the consumers understand the differences between those models to select the best to suit their requirements.
4. Identify Critical Performance Objectives: Four key components are considered in this step; service commitment, credits, credit process, and exclusions.
5. Evaluate Security & Privacy Requirements: Security and Privacy assurances should be obvious, distinct, and in clearly stated documents. Consideration should be taken for the consumer's data privacy.

6. Identify Service Management Requirements: Consumers should follow reasonable steps to guarantee that the provider is managing the level of service properly.
7. Prepare for Service Failure Management: Considering the offerings of public clouds, consumers must keep in mind the possible impact of service failure on their business operations.
8. Understand the Disaster Recovery Plan: The consumer should plan mainly for cases of disasters because the precautions taken by the cloud provider may not be sufficient to ensure the consumer satisfaction.
9. Define an Effective Management Process: Usually consumers expect good management from the cloud provider for any problem they might encounter. That is not the case actually. Concurrent cloud SLAs does not contain delivery of consumer-provider management process.
10. Understand the Exit Process: Every cloud SLA should contain an exit clause. An exit clause describes in details how the exit process is to be handled, what the provider and consumer has to do on contract termination.

SLA characteristic / cloud provider	**Amazon EC2 [48]**		**Microsoft Azure Storage [49]**		**Rackspace Cloud Servers [50]**	**Dell Boomi [51]**	**Google Cloud Storage [52]**	
Type of cloud service	IaaS		PaaS		IaaS	SaaS	PaaS	
Service Provider Discovery	Manual discovery		Manual discovery		Manual discovery	Manual discovery	Manual discovery	
Service Availability	<99.95%		<99.9%		100% excluding scheduled maintenance periods down time	<99.9% Service is down for 1 minute once a week for scheduled maintenance	>= 99.9%	
SLA Outlining	Predefined terms and QoS parameters		Predefined terms and QoS parameters		Predefined terms	Predefined terms and QoS parameters	Predefined terms and QoS parameters	
Agreement Establishment	SLA document provisioned by the provider		SLA document provisioned by the provider		SLA document provisioned by the provider	SLA document provisioned by the provider	SLA document provisioned by the provider	
Service Management & Monitoring	Third party monitoring systems can be used under the terms of Amazon's AWS Agreements		Management services are delivered by the provider. Consumer can use third party monitoring systems		Provider offers proactive infrastructure monitoring, operating sys maintenance and patching, application maintenance	Provider offers system upgrades and scheduled maintenance and emergency maintenance	Third party monitoring systems can be used	
SLA Violation Credits (provider penalty)	Monthly Uptime %	Credits	Monthly Uptime %	Credits	5% for each 30 minutes network or infrastructure downtime and 5% for each additional hour up to 100% of the fees	Users are not entitled to a credit if they are in breach of their services agreement with Rackspace	Monthly Uptime %	Credits
	99.0% – < 99.95%	10%	<99.9%	10%			99.0% – < 99.9%	10%
	<99%	30%	<99%	25%			95.0% – < 99.0%	25%
							< 95.0%	50%
SLA Exclusion	The service commitment does not apply to any unavailability, suspension or termination or performance issues that are due to the reasons stated in the SLA		SLA commitment doesn't apply in case any of the performance & availability issues stated in the SLA Exclusion section happens		Users are not entitled to a credit if they are in breach of their services agreement with Rackspace	Customer shall not receive any credits under this SLA in connection with any failure or deficiency of Service Availability caused by or associated with cases stated in it	The SLA does not apply to few exceptional cases stated in the document	

Table 15: Comparison between SLAS of the Existing Cloud Service Providers

7. CHAPTER VII

Examining security and its implications in the context of cloud service level agreements

7.1. Security in SLAs

Security still represents one of the main limits in the adoption of cloud computing. It is not rare the case where cloud service providers (CSPs) offer *non-transparent* security mechanisms, embedded in the systems, which are non-negotiable and, above all, vulnerable. The common approach followed by CSPs is a yes/no solution: they provide (or they declare that they provide) the higher security level available with their technological solutions.

Even though service availability and performance often are identified as critical issues, the number one barrier of adopting Cloud computing services is assurance : how can a potential customer be sure that it is safe to place data and applications in the Cloud? Since the SLA is used to explicitly state the obligations of the provider, the implemented security mechanisms, their effectiveness and the implications of possible mismanagement should be a part of this agreement. This concept is also known as Quality of Protection (QoP), which comprises the ability of a service provider to deliver service according to a set of specific security requirements. A standardized framework for constructing a SLA in the Cloud, based on guaranteed levels of these attributes and the consequences of mismanagement, is therefore of utmost importance for creating trustworthy and reliable Cloud computing services. This includes clarifying the consequences of a service provider's possible failure to deliver the service in accordance with the contract. Figure 11 outlines the basic structure for such a SLA.

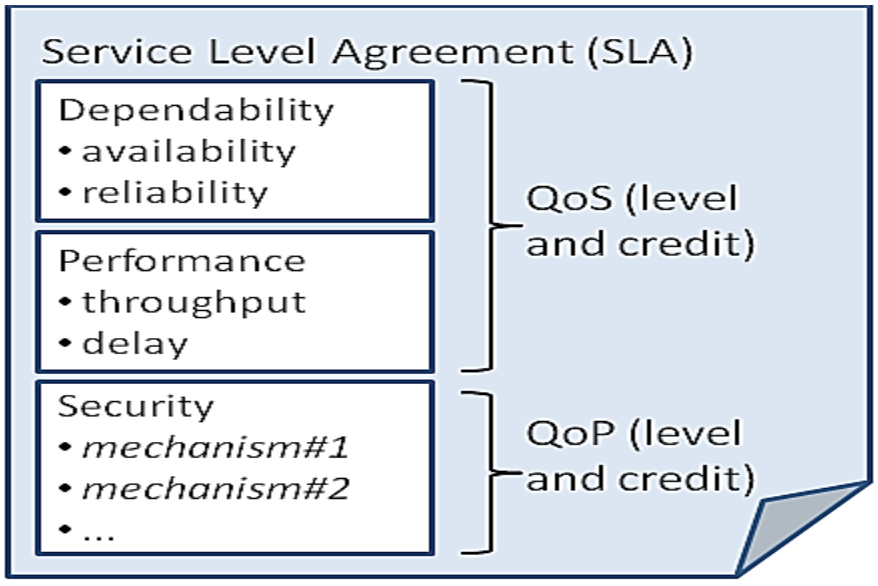

Figure 11: The basic structure of a SLA with dependability, performance and security guarantees

7.2. SECURITY MECHANISMS FOR CLOUD SLAs

Returning to the Cloud SLA outlined in **Figure 11**,in order to identify what security mechanisms to include in the contract we need to take a further look into the security related threats associated with the Cloud. As explained in the introduction, Cloud computing presents some fundamentally new security challenges in addition to the traditional ones. Cloud computing has five important characteristics, namely:

i) on-demand self-service, ii) broad network access, iii) resource pooling, iv)rapid elasticity, and v) measured service. The on-demand aspect must be taken into account when providing Cloud SLAs, since a SLA must also be allowed to be composed on-demand. Network access does not present any new security challenges but network Security becomes even more important in the Cloud, where large amounts of confidential data are regularly transmitted over the public Internet. Resource pooling and rapid elasticity presents some new challenges with respect to security, as will be described in the following. In addition, we have identified three categories of security mechanisms that need special attention in Cloud computing, and therefore should be a part of the SLA. These are access control, audit verification and compliance and incident management and response. Together with secure resource pooling and secure elasticity these five categories can be used in a structured approach to pick the right security mechanisms for a particular service.

7.2.1. Secure Resource Pooling

Resource pooling in Cloud computing to day is achieved by using virtualization either at the hardware level or at the application level. Both techniques enable multi-tenancy,i.e., different users share the same resources, and virtualization ensures the isolation of data and applications owned by different users. The sharing of physical resources in the Cloud gives rise to new security threats. One of the most imminent is unauthorized access to applications or data through the hypervisor, which may occur if proper isolation of applications and data is not achieved. It is therefore necessary to make sure that protection mechanisms exist and that they are stated in the SLA. In the framework outlined in Figure 2 this is illustrated as RP1: Data isolation and RP8: Application isolation", which are related to storage services and processing services, respectively. Moreover ,resource sharing implies that the customers need guarantees that their property remains confidential (RP3:Data encryption) and is integrity protected (RP6:Data integrity, RP12: Application integrity),that their data and applications are properly deleted from the physical hardware when requested (RP2: Data deletion),and that the data can be brought back in-house if necessary(RP5:Data portability, RP7: Data back-up).The customer should also have the possibility to put restrictions on the geographic location of storage and processing (RP4:Data location, RP9:Application location).Regarding network services (inside Clouds, between Cloud data centers and between the Cloud and the customer's premises), the customer should make sure that his traffic is properly protected (RP13:Network encryption, RP15:Integrity protection)and isolated from other customers traffic(RP14:Traffic isolation).

7.2.2. Secure Elasticity

Cloud computing promises rapid scalability of resources, scaling up and releasing resources as needed. This elasticity is also enabled by virtualization. Adding more virtual resources on the same physical machine does not in itself pose any new threats , but migrating virtual resources to new physical resources requires a secure migration process(E1:Secure data migration, E2:Secure virtual machine migration),including the actual network transfer. It must also be ensured that the new physical resource fulfils the same security requirements.

7.2.3. Access Control

Access control is especially important in the Cloud, where both competing customers sharing the same resources as well as insider personnel may try to gain unauthorized access to the customer's data. There source must also be protected from unauthorized remote access. It is therefore crucial to make sure that proper access control mechanisms are implemented (AC1:Identity management, AC2:Access management, AC3:Key management),and that there are strict restrictions on e.g. who may enter Cloud data centers (AC4:Internal security control).

7.2.4. Audit and Verification

The possibility to audit and verify the security of a service is very often crucial to the customer; however in the Cloud this is often not standard practice. Customers may require access to server logs, failed login attempts records or database change records (AU1: Logging) and sometimes also the possibility to audit the activity on specific Cloud resources (AU2: Auditing). In addition, the customers may want to make sure that a security certification scheme exists and is adapted to the

Cloud infrastructure (AU3: Certification). Customers may also have privacy concerns (AU4: Customer privacy).

7.2.5. Incident Management and Response

To make sure that the Cloud provider detects and responds to threats, the SLA may contain mechanisms for intrusion and malware detection (IM1:Intrusion detection, IM2:Malware detection),that security breaches are recorded and reported(IM3:Breach reporting), that data and applications can be reconstructed in the case of disasters(IM4:Recovery)and that mechanisms to prevent and mitigate DoS attacks are implemented(IM5:DoS mitigation). The framework in Figure 12 represents a first step towards a security-aware SLA. The purpose of the framework is to serve as a basis for constructing a SLA for a specific Cloud service, by identifying security mechanisms that should be stated in the SLA.

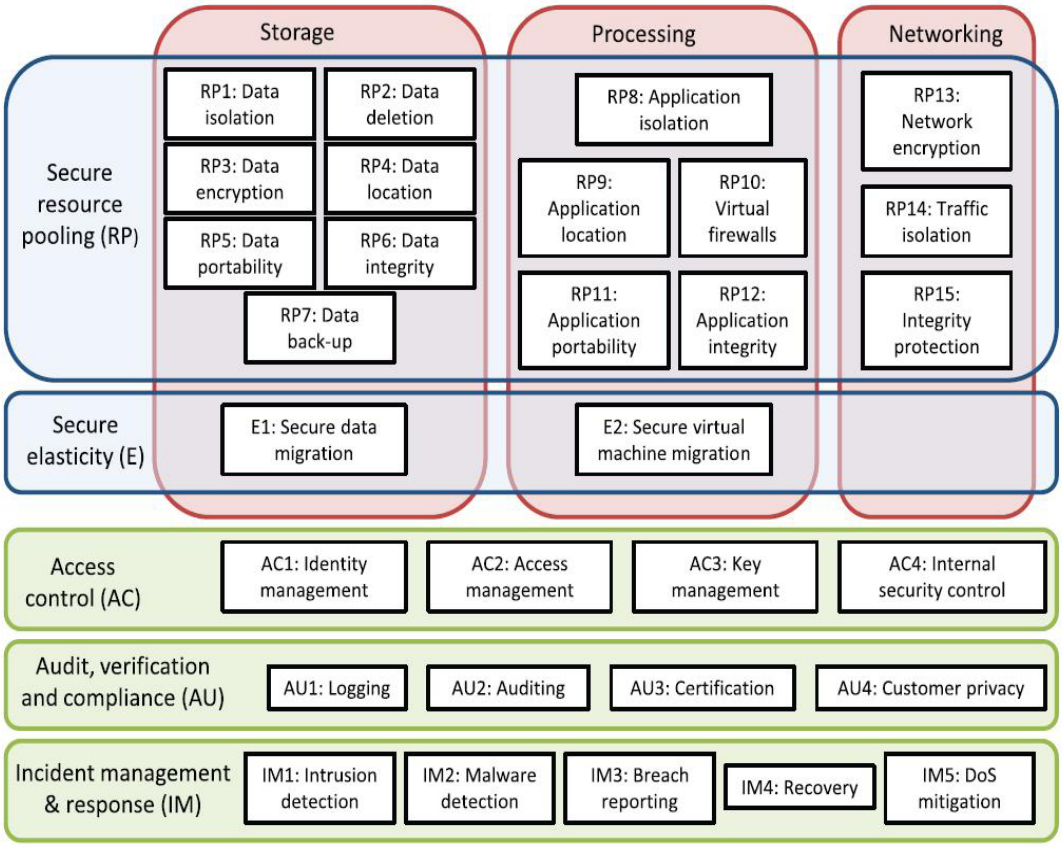

Figure 12: A framework for security mechanisms for Cloud SLAs

7.3. Legal Considerations for Security SLAs

Typically, the SLA is part of a Master Service Agreement (MSA). The MSA has many parts to it including Service Description, Confidentiality Requirements, Indemnification, Insurance Coverage, Business Continuity Commitments, Acceptable Use, Security and Privacy Policy and SLA. Metrics and objectives might be found within these other documents. These too should be carefully reviewed to determine if they provide sufficient risk mitigation and ensure there are not conflicting commitments (e.g., Security Policy commitments regarding patching conflicting with SLA commitments). In addition to providing measurable targets for service delivery, remedies are

typically identified when SLA obligations are not sustained. As mentioned earlier, not all metrics are equal. Therefore, weights are assigned to metrics as an incentive to assure vendor services are sustained to predetermined commitment levels. In addition to weight, some metrics might have a clause that recurring failures within a finite timeline have multipliers. For example, a service level objective failure in month 1 is 5%, subsequent failure in month 2 is 10% and subsequent failure in month 3 is 20%. Some Cloud services represent complex business processes and require significant investments to integrate before going live (e.g., ERP, CRM, etc.).

Recovering data and redeploying to another service provider (or internally) might require months or even years to execute. Therefore weights and multipliers for SLA metrics are very important for customers that are not prepared to uproot from a vendor with short notice. Vendor services evolve and improve over time. This is especially true for Cloud Computing related services. Because of this condition, new service objectives (e.g., recurring web application penetration testing every 90 days) and improved service quality commitments (e.g., account creation in less than 4 hours as compared to next day) will be offered by the vendor. This should be considered when negotiating the MSA. If the vendor offers these new commitments subsequent to the signing of the MSA, then the MSA should be written in a way that the customer automatically inherits these new service commitments without renegotiation of contract and pricing. Vendors frequently use business partners to help provide the total service offering. Rarely does a vendor use their own resources exclusively to provide services "top-to-bottom". This pool of resources is commonly referred to as Cloud Federation. For example, the vendor might provide the business application (e.g., Software-as-a-Service Model), however the foundational infrastructure might be provided and maintained by a third-party business partner (e.g., Infrastructure-as-a-Service Model). Cloud Federations work by distributing services and operational risk to providers that specialize in specific services (e.g., data center hosting, database administrator services, virtualization management, application integration, etc.). This approach can be beneficial for the vendor to maximize efficiency, increase scalability, and reduce cost. These are substantially the same benefits endpoint Cloud customers enjoy. With Cloud Federations, key customer security controls might be under the responsibility of a third-party that the customer has no direct, legal relationship with. Risk propagation (and liability assignment) therefore must be clearly understood by Cloud customer. "Weakness assessment and vulnerability analysis must be abstracted, i.e., not based on specific system details, and made relevant to the external black boxed cloud domain...to prevent violation scenarios and thus ensure security control compliance." (Hale and Gamble, 2012). Further, security certifications (e.g., ISO 27001, SSAE 16, PCI DSS, etc.) might be held by the third-party and not the vendor with which the customer has contracted. This creates a potential compliance issue as well as unintended liability. When negotiating Master Service Agreements and Service Level Objectives, the customer should have a clear understanding of the third-party involvement and risk propagation. There should be mutual understanding that SLAs apply to the vendor and vendor's agent(s). Timing for negotiating the SLA is important. The recommendation to codify security controls in a formal contract was a good idea in the past when negotiating traditional data center outsourcing and still good advice today with modern Cloud services. Waiting until after the contract is signed to establish a Service Level Agreement severely disadvantages the customer and presents an opportunity for unplanned risk. Further, contracts that reference vendor internal documents and marketing material might result in a moving target. Commitments that existed in marketing materials or vendor standards at the time of negotiation (e.g., encryption and data destruction) might not be sustained. The SLA is a critical part of the customer-vendor relationship and should be formally established early within the contract—not informally as a website URL that might change without notice. Cloud vendors provide a variety of ways to report on service level performance. In some cases, the reporting is similar to that of an outsourcing engagement in which the vendor presents reports monthly (pro-actively or upon customer request). This approach is beginning to fade away as the cost to produce these metrics can be significant. A growing trend is to present the evidence on-line for customer review. Evidence of commitment to security controls including Service Guarantees and SLAs that are common across multiple tenants are also being presented to Cloud security authorities that serve as informal certification and accreditation organizations. For example, the Cloud Security Alliance established Security, Trust, and Assurance Registry (CSA, 2014) in 2011. CSA STAR is a free, publicly accessible registry that documents the security controls provided by

various cloud computing offerings. Cloud service providers have the option to complete the CSA STAR control assertion questionnaire to demonstrate due care and satisfy customer evidentiary requirements. The program is based on an open certification framework that begins with vendor assertion, progressing through third-party certification (like that of SSAE 16 or PCI), to finally continuous monitoring-based certification. Lastly, formal audits of controls can be performed using the same approach as those used for data centers. For example, ISO 27001 registration, AICPA/ISAE SOC1/SOC2 report, and Payment Card Industry Data Security Standard (PCI DSS) Service Provider assessment are common considerations for Cloud vendors. In some cases these audit standards might be required to do business and will be a condition of the Master Services Agreement. Depending exclusively on an audit to confirm the effectiveness of the security controls once a year might be acceptable for some Cloud service offerings. Once a year assurance by an auditor might not be adequate for those customers intending to transport, store, or process confidential data. For these conditions, Service Level Agreements provide continuity of visibility into the operational effectiveness of controls and better risk response at or near the time a control fails.

7.3.1. Selection criteria for Security SLAs

The SLA standards proposed in this paper are focused on security controls for cloud computing. NIST SP800-53 proposes there are 18 families of security controls organized into 3 classes. Families are assigned to their respective classes based on the dominant characteristics of the controls in that family. The organization of these security controls can be helpful when determining how the Service Level Agreement conditions are to be presented in the Master Services Agreement. The table 16 provided below lists the NIST security controls into technical, operational, and management classes.

Technical	Operational	Management
(AC) Access Control (AU) Audit and Accountability (IA) Identification and Authentication (SC) System and Communication Protection	(AT) Awareness and Training (CM) Configuration Management (CP) Contingency Planning (IR) Incident Response (MA) Maintenance (MP) Media Protection (PE) Physical and Environmental Protection (PS) Personnel Security (SI) System and Information Integrity	(CA) Certification, Accreditation and Security Assessment (PL) Planning (RA) Risk Assessment (SA) System and Services Acquisition (PM) Program Management

Table16: NIST SP800-53 Control 18 Families and 3 Classes

For the technical class, security controls are generally architectural or policy based. Generally, these controls would be defined as part of an Information Security Policy. An approach might be to have the Master Services Agreement state that the vendor must have an Information Security Policy that requires this class of controls including automatically disabling inactive accounts after a finite time (AC-02), encryption to protect the confidentiality of remote access sessions (AC-17), regularly review/analyze audit records for indications of inappropriate or unusual activity (AU-06), time stamps for use in audit record generation (AU-08), multifactor authentication (AI-02), function isolation (SC-

03), and protects the confidentiality of transmitted information (SC-09). The examination of these controls is typically performed by a third-party security practitioner engaged by the cloud vendor yearly, and the audit approach can be based on a credible standard (e.g., ISO27001, SOC 2, etc.). The absence of these technical controls might be sufficient cause for contract termination. For this reason, annexing an Information Security Policy might be more effective than creating a unique service level metric for each of the security controls in this family. The management class of security controls is similar to the technical control family in that the obligations might be stated in the Master Service Agreement without complex metrics. Management security controls can be examined by a third-party auditor to confirm the controls are in-place and practiced. For example, the Master Service Agreement might state that a Risk Assessment must be performed yearly (RA-03), vulnerability scanning must be performed every month and immediately after every significant change (RA-05), and Acceptable Use Policy must be presented for review and signature yearly (PL-04). The absence of these management controls might be sufficient cause for cloud service provider contract termination. For this reason, annexing these requirements in the Master Service Agreement with a simple in-place obligation might be more effective than creating a unique service level metric for each of the security controls in this family. The operational class of security controls generally require frequent and recurring monitoring to demonstrate commercially reasonable due care. Relying on management attestation or audit once a year might not sufficiently demonstrate operational effectiveness nor manage risk. Therefore these controls lend themselves to Service Level Objectives (SLO). Based on asset valuation and cloud model, the customer might have specific requirements including configuration change control (CM-03), DR/BCP (CP-02), incident handling (IR-04), monitoring physical access intrusion (PE-06), patching (SI-02), malware prevention (SI-03), intrusion detection (SI-04), and error handling (SI-11). Therefore, the NIST operational class of security controls provides a useful reference for determining the SLA metrics. Since all industries and business requirements cannot be met with a single, universal collection of security control standards, service level obligations share the same limitations. This section proposes a minimum baseline that is intended to be broadly adoptable by all cloud service providers. Standards are categorized by Cloud Service Model (IaaS, PaaS, and SaaS). Since each cloud service model builds upon the underlying service model, the security SLA standards follow the same approach. PaaS proposed security SLA standards are intended to include IaaS. SaaS security SLA standards are intended to include IaaS and PaaS. Additional regulatory (e.g., Health Insurance Portability and Accountability Act), contracted (e.g., Payment Card Industry), and business requirements can supplement this baseline.

7.3.2. Key Metrics for IaaS SLAs

As stated earlier, Infrastructure-as-a-Service (IaaS) is one of three cloud service delivery models. IaaS is intended to provide basic computer infrastructure in a virtual environment so that the consumer does not have to purchase assets. The customer typically assumes substantially all data and application security risks. When progressing from IaaS to PaaS to SaaS, more technology abstraction is introduced reducing the customer direct visibility into and control over the environment. There are some infrastructure components such as networking that the customer will not have access to, but are critical for the security program. These require service level obligations for the vendor since the customer has no “hands-on” access to configure or examine the associated security controls. This section proposes key security SLA standards for these components common to IaaS as they serve an important role in the information security program. The authoritative sources that are identified and align with the proposed metrics are NIST SP800-53r4 (NIST), Cloud Security Alliance Cloud Control Matrix v3.01 (CSA), and ISO 27001-2013 (ISO).References to the specific section of each authoritative source are provided with each SLA recommendation for additional guidance.

#	Key Security SLAs	NIST	CSA	ISO
1	Change Control and Configuration Management	CM	CCC	A12.1.2
2	Data Center Asset Management	CM	DCS	A8.1.1
3	Disaster Recovery and Business Continuity Planning	CP	BCR	A.17.1.3
4	Secure Configuration and Server Hardening	CM	IVS	A.12.5.1
5	Malware and Intrusion Prevention	SI	TVM	A.12.2.1
6	Network Vulnerability and Penetration Testing	RA	IVS	A.14.2.3
7	Software Lifecycle and Patch Management	SA	TVM	A.12.6.1
8	Security Incident Handling	IR	SEF	A.16
9	Secure Network Protocols and Data Transport	SC	IPY	A.13
10	Security Event Logging	AU	IVS	A.12.4

Table 17 : Key Security Service Level Agreement Metrics for IaaS

7.3.3. Key Metrics for PaaS SLAs

This section proposes key security SLA standards for PaaS. These are in addition to the aforementioned IaaS SLA standards. Several new SLA metrics for PaaS are proposed (e.g., secure application and program interfaces). Some SLA standards are listed again reflecting the scope change between IaaS and PaaS. For example, Change Control and Configuration Management not only applies to infrastructure, but also middleware, databases, and messaging components introduced as part of PaaS. There are more "moving parts" to PaaS as compared to IaaS. To deliver the PaaS, multiple vendors might be collaborating creating a cloud federation. For example, separate vendors may be engaged for data center, network, systems, database and middleware services. Therefore, some clarification might be required during the SLA negotiation to understand actual metric source and reporting accountability.

#	Key Security SLAs	NIST	CSA	ISO
1	Change Control and Configuration Management	CM	CCC	A12.1.2
2	Secure Application and Program Interfaces	SC	AIS	A.14.1.3
3	Disaster Recovery and Business Continuity Planning	CP	BCR	A.17.1.3
4	Secure Configuration	CM	IVS	A.12.5.1
5	Intrusion Prevention	SI	TVM	A.14.1.2
6	Vulnerability and Penetration Testing	RA	IVS	A.14.2.3
7	Software Lifecycle and Patch Management	SA	TVM	A.12.6.1
8	Data Protection/Portability/Retention/Destruction	MP	DSI	A.8
9	Encryption and Key Management	SC	EKM	A.10.1.2
10	Application and Database Logging	AU	IVS	A.12.4

Table 18 : Key Security Service Level Agreement Metrics for PaaS

7.3.4. Key Metrics for SaaS SLAs

This section proposes key security SLA standards for SaaS. The IaaS and PaaS key Security SLAs are cumulative and would apply to a SaaS environment. In addition to this cumulative approach, it is worth noting that some of the same key security metrics are listed in all three cloud service models (e.g., Disaster Recovery, Intrusion Prevention, Software Lifecycle and Patch Management, etc.). These metrics remain relevant because the security requirements and operations duties to fulfil these requirements are substantially different for each cloud service model. Intrusion kill chain analysis demonstrates that cyber-attacks have pivoted using infrastructure, platform, and software to gain unauthorized access to confidential data (USSCCST, 2014). Each cloud service level introduces new surfaces of attack. Intrusion detection and prevention mechanisms are different for each cloud service level because the threats grow. Therefore, the security SLA metrics remain in all three service model recommendations.

#	Key Security SLAs	NIST	CSA	ISO
1	Change and Release Management	CM	CCC	A12.1.2
2	Secure Application and Program Interfaces	SC	AIS	A.14.1.2
3	Disaster Recovery and Business Continuity Planning	CP	BCR	A.17.1.3
4	Secure Configuration	CM	IVS	A.12.5.1
5	Intrusion Prevention	SI	TVM	A.14.1.2
6	Vulnerability and Penetration Testing	RA	IVS	A.14.2.3
7	Software Lifecycle and Patch Management	SI	TVM	A.12.6.1
8	Secure Coding Practices	AT	HRS	A.14.2
9	Identity Access Management	AC	IAM	A.9.2

Table 19 : Key Security Service Level Agreement Metrics for SaaS

7.4. Managed Security Services

Security is becoming increasingly important for companies, especially for the extension of networking to mission-critical environments, with new intranet and extranet and e-commerce applications. An increasing number of companies use the services of outsourcing in the security, with the aim to delegate effectively their security infrastructures management for focusing instead on their core-business. As computer attack patterns shift and threats to networks change and grow almost daily, it is critical that organizations achieve reliable information security. Investment decisions about information security are best considered in the context of managing business risk. Risks can be accepted, mitigated, avoided, or transferred. An organization needs to understand the level of information security risk in outsourcing any managed security service when developing the Request for Proposal (RFP). The costs to procure, operate, and manage provider service delivery, including review for compliance with the Service Level Agreement (SLA) and the overall contract, should not exceed the anticipated benefit.

The Managed Security Services suite includes:

- ➢ Managed services for Firewall
- ➢ Managed Virtual Private Network (VPNs)
- ➢ Managed Intrusion Detection Systems (IDSs)
- ➢ Managed anti-virus and content filtering services
- ➢ Managed information security risk assessments
- ➢ Vulnerability Assessment
- ➢ Vulnerability assessment and penetration testing

- Data archiving and restoration
- On-site Consulting
- Security monitoring (may be included in network boundary protection)
- Emergency response and forensic analysis (This service may be in addition to security monitoring)
- Information security risk assessments

The full suite of Managed Security Services is aimed at all those companies that, at low cost, wish to maintain high levels of security and control of their infrastructure, intending to focus internal resources on their core business activities. The main advantages of the adoption of services payable by the NOC (Network Operation Center) are:

- **Reducing costs:** Companies outsource activities that are not its core business and reduce the cost of operations related to the ongoing monitoring of their security infrastructure (firewall, IDS / IPS, Antivirus, etc.).
- **Business continuity:** The Managed Security Services suite helps you identify potential security issues and resolve them proactively. These results in improved continuity of service for the network, systems, applications: a proactive approach to safeguarding your data, productivity, service to the customer.
- **Optimizing resources:** Possibility for the customer to focus on their core business activities, with a focus on processes/services and technologies. The customer does not have the burden of managing complex technologies, in continues changing.

7.4.1. Benefits of Engaging an MSS Provider

The results from engaging a reputable, competent MSSP have the potential to be far superior to anything an organization can achieve on its own. Described in this section are reasons for contracting with a MSSP and some of the benefits that may result from the relationship. All of these factors can contribute to reducing the risks faced by the client through a combination of risk mitigation and risk/liability sharing between the client and the MSSP.

Cost: The cost of a managed security service is typically less than hiring in-house, full-time security experts. An MSSP is able to spread out the investment in analysts, hardware, software, and facilities over several clients, reducing the per client cost. As one example, an MSSP claims it can set up and monitor security on a 250-user network on a single T1 (1.5 Mbps) Internet gateway for about $75,000 a year, excluding hardware. Replicating these actions within the organization produces similar hardware costs, plus at least $240,000 in annual compensation to hire three full-time specialists, based on data from the magazine InformationWeek's most recent Salary Survey 2. A client organization can convert variable costs (when done in-house) to fixed costs (services), realize a tax advantage by deducting MSSP fee expenses from current year earnings versus depreciating internal assets, and experience cash flow improvements resulting from the transfer of software licenses (and possibly personnel) to the MSSP.

Staffing: A shortage of qualified information security personnel puts tremendous pressure on IT departments to recruit, train, compensate, and retain critical staff. The cost of in-house network security specialists can be prohibitive. When outsourcing, the costs to hire, trains, and retain highly skilled staff becomes an MSSP responsibility. An MSSP is likely to retain security experts by offering a range of career opportunities and positions from entry level to senior management, all within the information security field. In addition, if a client organization can outsource repetitive security monitoring and protection functions, then they can then focus internal resources on more critical business initiatives.

Skills: An in-house staff member who only deals with security on a part-time basis or only sees a limited number of security incidents is probably not as competent as someone who is doing the same work full-time, seeing security impacts across several different clients, and crafting security solutions with broader applicability.

MSSPs have insight into security situations based on extensive experience, dealing with hundreds or thousands of potentially threatening situations every day, and are some of the most aggressive and strenuous users of security software.

Facilities: MSSPs can also enhance security simply because of the facilities they offer. Many MSSPs have special security operations centers (SOCs) located in various parts of the country. These are physically hardened sites with state-of-the-art infrastructure managed by trained personnel.

Objectivity and Independence: An organization may have multiple, ad hoc solutions to handle the same types of security problems. There may be no enterprise-wide management of security or of strategy. Moving security to a capable security service provider may help simplify and strengthen the enterprise's security posture. An MSSP can provide an independent perspective on the security posture of an organization and help maintain a system of checks and balances with in-house personnel. An MSSP can often provide an integrated, more coherent solution, thereby eliminating redundant effort, hardware, and software.

Security Awareness: It is difficult for an organization to track and address all potential threats and vulnerabilities as well as attack patterns, intruder tools, and current best security practices. An MSSP is often able to obtain advance warning of new vulnerabilities and gain early access to information on countermeasures. An MSSP can advise on how other organizations handle the same types of security problems. An MSSP is likely to have contact with highly qualified and specialized international security experts as well as other MSSPs. These resources can be brought to bear to diagnose and resolve client issues.

Prosecution: The MSSP are often well connected to law enforcement agencies around the world and understands what forensic analysis and evidence are required to successfully support legal proceedings.

Service Performance: When an organization contracts for security monitoring services, the service can report near real-time results, 24 hours a day, 7 days a week, and 365 days a year. This is a large contrast with an in-house service that may only operate during normal business hours. MSSPs can be held accountable for the service standards they provide. They guarantee service levels and assure their availability; failing to do so can have financial repercussions. Their operational procedures are designed to ensure uninterrupted service availability. Also, if the MSSP is providing service systems, then it is their responsibility to upgrade software and hardware and to maintain a secure network configuration. Because MSSPs have strict contractual obligations to their clients and must maintain their reputation in the marketplace, their control procedures are generally both well documented and carefully enforced. In all instances, the client needs to verify these performance characteristics.

Service Security and Technology: Service security solutions and technologies such as firewalls, intrusion detection systems (IDSs), virtual private networks (VPNs), and vulnerability assessment tools are far more effective because they are managed and monitored by skilled security professionals. For example, when an intrusion is detected, MSSPs can use a remote monitoring connection to determine whether the alarm is justified and block further intruder actions. A managed service can protect the client’s network from unsecured VPN endpoints. For products developed by the MSSP and used in their services, the client organization receives an enhanced level of product support. The MSSP may use other third party provider products as the basis for providing service (such as firewalls and IDSs). Based on the size of the MSSP’s client base, the MSSP may be able to influence the product provider to improve the security of their products by, for example, addressing new attacks and vulnerabilities.

7.4.2. Risks in Engaging an MSS Provider

While an MSSP may have more competent staff to manage security services, they may not be as effective in applying remedies that meet the specific needs of the client. MSSPs sometimes run the risk of applying solutions that are too generic to benefit the client. Also, sometimes the client's staff is more adept at providing the best solution. In deciding to engage an MSSP, an organization needs to treat the potential action as a risk mitigation sharing decision. Regardless of an MSSP's role, the client is responsible for addressing the impact of a risk that has become a reality. The client must always be prepared to manage and respond to manifested risks.

There are counter arguments and issues to consider when weighing the risks against the benefits described above. Some of these include the following:

Trust: The challenge of establishing a good working relationship and building trust between a client and MSS provider remains as a significant hurdle in deciding to outsource security services. Any MSSP has access to sensitive client information and details about the client's security posture and vulnerabilities. The intentional or inadvertent public release of such information can be extremely damaging to the client. A signed confidentiality agreement enacted in the later stages of contract negotiations can help mitigate this risk.

Dependence: An organization can become operationally dependent on a single MSSP and be greatly affected by the MSSP's business viability (refer to Practice 1, P1.1 Business Attributes), other clients, and business partnerships. One risk mitigation approach is to outsource to multiple providers, but this comes with additional cost and management oversight responsibilities. An organization needs to carefully examine the provider's proposal to understand whether they use tiered providers and how they work. (Tiered providers are the subcontractors used by the MSSP and any other downstream subcontractors) Organizations must ensure that both the client and provider have the necessary and contractual checks and balances with respect to tiered provider performance.

Ownership: A client retains ownership and responsibility for the secure operation of its infrastructure and the protection of its critical assets regardless of the scope of services provided by an MSSP. An organization may start to ignore pressing security issues due to "out of sight, out of mind" thinking, having delegated this concern to the provider. The client must ensure that it retains sufficient competency to fulfil its responsibility and that contractual and service level agreement language supports this. Risk mitigation approaches include making information security the primary responsibility for one or more staff members and managers and conducting regular user security awareness and training sessions.

Shared Environment: The shared operational environment used by many MSSPs to service multiple clients poses more risks than an in-house environment. Sharing a data transmission capability (such as a common network) or a processing environment (such as a general purpose server) across multiple clients can increase the likelihood of one organization having access to the sensitive information of another.

Implementation: Initiating a managed security services relationship may require a complex transition of people, processes, hardware, software, and other assets from the client to the provider or from one provider to another, all of which may introduce new risks. IT and business environments may require new interfaces, approaches, and expectations for service delivery. Roles and responsibilities are often redefined.

Clients should ask for an implementation timeline and duration as well as a high-level implementation plan as part of a provider's proposal.

Partnership Failure: One of the greatest risks comes from inadequate, incomplete planning and infrequent communication and review between the provider and the client. This partnership can fail at any stage. Like any business relationship, it requires attention, care, and due diligence.

Hidden Costs and Impacts: Certain costs are overlooked or ignored because they are difficult to quantify. An organization needs to factor these into its risk analysis and decision-making processes before engaging an MSSP. Some of the hidden costs and areas where issues could arise are listed below.

- ➢ Costs associated with giving up control (experience, knowledge, skill development associated with) of critical assets and security technologies

- ➢ What happens at the end of the contract period? What happens if the original provider goes out of business, delivers poorly, or is more expensive when the contract is re competed? What is the cost of switching to a new provider?

- ➢ Would an MSSP do the job with the same quality and thoroughness that an organization would do for itself?

- ➢ How are needs met and services provided for multiple clients and how are they prioritized by the MSSP?

Legal Issues: An organization and an MSSP need to evaluate and discuss potential legal issues that could arise during a security incident involving both parties. The client needs to understand the jurisdiction under which the provider operates, the applicable laws and regulations, whether or not these laws apply to the client when engaging provider services, and if so, if these laws are compatible with the client's operation and acceptable to the client. This applies to tiered providers as well.

Questions for Cloud Service Provider - identifying strengths

1. Do you publish SLAs, and how are these documents accessed?

2. If you do not publish SLAs, do you publish service level objectives (SLOs)?

3. How do your SLA targets differ from your competitors? You may be surprised that SLAs do no vary that much.

4. Why were your SLA targets chosen? Targets are often defined competitively or based on the best or worst capability of the underlying products.

5. How often have you violated your SLAs in the last three months, six months, 12 months?

6. Do you publish your SLA results openly? How frequently?

7. Which SLA metrics do you fail at most often, even if it has no impact on your customers?

8. How often do you increase or decrease your SLA targets, and what has the trend been? Any reduction or removal of a target may mean scalability challenges.

9. What SLA metrics have been removed in the last 12 months?

10. How often do you test your own SLAs? You really want to hear that the metrics are continuously tested.

11. How are SLA claims validated? How am I compensated for an SLA violation? Your provider should be doing the work here, not requiring you to prove a failure.

12. Do I receive detailed incident response information? This is necessary to fully inform your organization or customers of the problem and the solution. Never waste a failure; make sure your provider is identifying the root cause and resolving it.

13. Do you use any third parties to monitor your SLAs? This can provide additional validation of the seriousness of SLA measurement.

14. Are the SLAs relevant to the areas that need alignment, such as availability, transaction time, storage, and performance?

15. How transparent is the cloud vendor in sharing SLA performance (daily, weekly, or monthly)? You need broad visibility into situations that may result in breaches of the SLA.

16. Are the SLAs results-oriented? You're in the business of creating value for your customers; the SLAs should help.

17. SLAs can be narrow or broad, simple or complex. Check with each cloud service provider (CSP) you are considering: what is included in each SLA? Which best meets your needs?

18. What happens if data is lost?

Cloud SLA Vocabulary

Definition	Description
Application programming interface (API)	The collection of invocation methods and associated parameters used by a certain (part of) cloud service or software component to request actions from and otherwise interact with another cloud service or software component.
Auditability	The capability of supporting a systematic, independent and documented process for obtaining audit evidence and evaluating it objectively to determine the extent to which audit criteria are fulfilled.
Availability	The property of being accessible and usable upon demand by an authorized entity.
Cloud computing	A paradigm for enabling network access to a scalable and elastic pool of shareable physical or virtual resources with self-service provisioning and administration on-demand.[11] Examples of resources include servers, operating systems, networks, software, applications, and storage equipment.
Cloud infrastructure	The collection of hardware, software and other related goods and resources that enables the provision of cloud services.
Cloud service	One or more capabilities offered via cloud computing invoked using a defined interface.
Cloud service customer	A party which is in a business relationship for the purpose of using cloud services, for this document not being consumers. NOTE – A business relationship may not necessarily imply financial agreements or similar arrangements.
Cloud service customer data	class of data objects under the control, by legal or other reasons, of the cloud service customer that were input to the cloud service, or resulted from exercising the capabilities of the cloud service by or on behalf of the

	cloud service customer via the published interface of the cloud service An example of legal controls is copyright.
Cloud service derived data	class of data objects under cloud service provider control that are derived as a result of interaction with the cloud service by the cloud service customer Cloud service derived data includes log data containing records of who used the service, at what times, which functions, types of data involved and so on. It can also include information about the numbers of authorized users and their identities. It can also include any configuration or customization data, where the cloud service has such configuration and customization capabilities.
Cloud service level objective (SLO)	Target for a given attribute of a cloud service that can be expressed quantitatively or qualitatively.
Cloud service provider (CSP)	A party which makes cloud services available.
Cloud service provider data	class of data objects, specific to the operation of the cloud service, under the control of the cloud service provider Cloud service provider data includes but is not limited to resource configuration and utilization information, cloud service specific virtual machine, storage and network resource allocations, overall data centre configuration and utilization, physical and virtual resource failure rates, operational costs and so on.
Cloud service user	natural person, or entity acting on their behalf, associated with a cloud service customer that uses cloud services Examples of such entities include devices and applications.
Cloud SLA life cycle	Service level agreements life cycle i.e.; assessment, negotiation, contracting, operation, amendment, escalation and termination, and other arrangements and matters.
Cloud SLAs	Documented agreement between the cloud service provider and cloud service customer that identifies services and cloud service level objectives (SLOs).
Cryptographic key management	Key management is the management of cryptographic keys in a cryptosystem. This includes dealing with the generation, exchange, storage, use, and replacement of keys, as well as cryptographic protocol. It includes cryptographic protocol design, key servers, user procedures, and other relevant protocols[12].
Data	Data of any form, nature or structure, that can be created, uploaded, inserted in, collected or derived from or with cloud services and/or cloud

	computing, including without limitation proprietary and non-proprietary data, confidential and non-confidential data, non-personal and personal data, as well as other human readable or machine readable data.
Data controller	The natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data.
Data format	One or more formats in which the data is in one or more phases of its data lifecycle.
Data integrity	The property of protecting the accuracy and completeness of assets.
Data intervenability	The capability of a cloud service provider to support the cloud service customer in facilitating exercise of data subjects' rights. Note: Data subjects' rights include without limitation access, rectification, erasure of the data subjects' personal data. They also include the objection to processing of the personal data when it is not carried out in compliance with the applicable legal requirements.
Data life cycle	The handling of data that commonly includes six (6) phases, (1) create/derive, (2) store, (3) use/process, (4) share, (5) archive, (6) destroy.[13]
Data location	The geographic location(s) where personal data may be stored or otherwise processed by the cloud service provider.
Data portability	Ability to easily transfer data from one system to another without being required to re-enter data.
Data processor	A natural or legal person, public authority, agency or any other body which processes Personal data on behalf of the Data controller.
Data protection	The employment of technical, organisational and legal measures in order to achieve the goals of data security (confidentiality, integrity and availability), transparency, intervenability and portability, as well as compliance with the relevant legal framework.
Data subject	An identified or identifiable natural person, being an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity.
Hybrid cloud	Deployment model of cloud computing using at least two different cloud deployment models.
Identity assurance	The ability of a relying party to determine, with some level of certainty, that a claim to a particular identity made by some entity can be trusted to actually be the claimant's true, accurate and correct identity.
Incident notification and transparency	Notifications and transparency about incidents under the SLA that may be required as per (a) mandatory law and legislation (such as under the EU

REST	Representational state transfer (REST) is a software architectural style consisting of a coordinated set of architectural constraints applied to components, connectors, and data elements, within a distributed hypermedia system.
Reversibility	Process for cloud service customers to retrieve their cloud service customer data and application artefacts and for the cloud service provider to delete all cloud service customer data as well as contractually specified cloud service derived data after an agreed period.
Sensitive data	Any classified, personal, proprietary or confidential information or data of any form, nature or structure, that can be created, uploaded, inserted in, collected or derived from or with cloud services and/or cloud computing whose access, use, disclosure or processing is subject to restriction either by applicable law or contact.[14]
Software as a services *(SaaS)*	The capability provided to the cloud service customer is to use the cloud service provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through either a thin client interface, such as a web browser (e.g. web-based email), or a program interface. The cloud service customer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.
Temporary data	Data or a data set that is created during the operation of the cloud service and becomes unused after a predefined period of time.
Vulnerability	A weakness of an asset or group of assets, e.g. software or hardware related, that can be exploited by one or more threats.
xaaS	A collective term of diverse but re-useable components, including without limitation infrastructure, platforms, data, software, middleware, hardware or other goods, made available as a service with some kind of use of cloud computing.

References

[1] Andrieux, A., Czajkowski, K., Dan, A., Keahey, K., Ludwig, H., Nakata, T., Pruyne, J., Rofrano, J., Tuecke, S., & Xu, M. (2007). Web Services Agreement Specification (WSAgreement). OGF proposed recommendation (GFD.107).

[2] AWS *EC2 Service Level Agreement*. Retrieved 03 28, 2010, from AWS: http://aws.amazon.com/ec2-sla/

[3] AWS S3 *Service Level Agreement*. Retrieved 03 28, 2010, from AWS: http://aws.amazon.com/s3-sla/

[4] Battre", D., Hovestadt, M., Kao, O., Keller, A., & Voss, K. (2007). Planning-based scheduling for SLA-awareness and grid integration. *PlanSIG,* (pp. 1).

[5] Blythe, J., Deelman, E., & Gil, Y. (2004). Automatically Composed Workflows for Grid Environments. *IEEE Intelligent Systems*, (pp. 16-23).

[6] Bonell, M. (1996). The UNIDROIT Principles of International Commercial Contracts and the Principles of European Contract Law: Similar Rules for the Same Purpose. *Uniform Law Review*, (pp. 229-246).

[7] Boniface, M., Phillips, S., Sanchez-Macian, A., & Surridge, M. (2009). Dynamic service provisioning using GRIA SLAs. *Service-Oriented Computing-ICSOC 2007 Workshops*, (pp. 56-67). Vienna, Austria.

[8] Cloud Standards Customer Council (2013). *Public Cloud Service Agreements: What to Expect and What to Negotiate* http://cloud-council.org/resource-hub.htm#migrating-applications-to-public-cloud-services

[9] Cloud Standards Customer Council (2014). *Practical Guide to Cloud Computing, Version 2.0.*

http://cloud-council.org/resource-hub.htm#practical-guide-cloud-computing-v2

[10] Cloud Standards Customer Council (2015). *Security for Cloud Computing: 10 Steps to Ensure Success, Version 2.0.* http://cloud-council.org/resource-hub.htm#security-for-cloud-computing-10-steps-to-ensure-success

[11] ISO/IEC 17789: Cloud Computing Reference Architecture http://www.iso.org/iso/catalogue_detail?csnumber=60545

[12] Amazon. https://www.amazon.com, 2011. [Online; accessed August 2011].

[13] Amazon EC2. https://aws.amazon.com/ec2/, 2011. [Online; accessed August 2011].

[14] Amazon EC2 Reserved Instances.https://aws.amazon.com/ec2/reserved-instances/, 2011. [Online;accessed August 2011].

[15] Amazon EC2 SLA. https://aws.amazon.com/ec2-sla/, 2011. [Online; accessed August2011].

[16] Amazon S3. https://aws.amazon.com/s3/, 2011. [Online; accessed August 2011].

[17] Amazon S3 SLA. https://aws.amazon.com/s3-sla/, 2011. [Online; accessed August 2011].

[18]. Wikipedia, http://en.wikipedia.org/wiki /Cloud_Computning

[19]. Rafael Moreno-Vozmediano,Rubén S. Montero, Ignacio M. Llorente," Key Challenges in Cloud Computing -*Enabling the Future Internet of Services*", Published by the IEEE Computer Society 1089-7801/13/ © 2013 IEEE, IEEE internet computing

[20]. M.Rajendra Prasad, R. Lakshman Naik, V.Bapuji," Cloud Computing : Research Issues and Implications ", International Journal of Cloud Computing and Services Science (IJ-CLOSER)

Vol.2, No.2, April 2013, pp. 134~140 ISSN: 2089-3337.

[21]. Francesco M.A and Gianni F. “An approach to a cloud Computing network”, IEEE, August 2008, pp113-118.

[22]. Huaglory Tianfield,”Cloud Computing Architectures”, 978-1-4577-0653-0/11/©2011 IEEE.

[23]. S. V. Kavitha, “A survey on security issues in service delivery models of cloud computing” Journal of Network and Computer Applications, Elsevier, vol. 34, (2011), pp. 1–11.

[24]. P. Adams, “Advantages and Disadvantages Of Cloud Computing System”, Advantages And Disadvantages Of Cloud Computing System, (2011) November 18.

[25] Gartner Group, “Cio-priorit¨aten und budgets 2011.” [Online]. Available: http://www.cio.de/strategien/analysen/2262709/ [retrirved: june, 2013].

[26] Distributed Management Task Force, “Architecture for managing clouds.” [Online]. Available: http://dmtf.org [retrieved: june, 2013].

[27] ISO/IEC SC 38 Study Group, “Jtc 1/sc 38 study group report on cloud computing,” International Organization for Standardization, Tech. Rep., 2011. [Online]. Available: http://isotc.iso.org [retrirved: june, 2013].

[28] A. Keller and H. Ludwig, “The wsla framework: Specifying and monitoring service level agreements for web services,” Journal of Network and Systems Management, vol. 11, no. 1, pp. 57–81, Mar. 2003.

[29] F. Liu, J. Tong, J. Mao, R. Bohn, J. Messina, L. Badger, and D. Leaf, “Nist cloud computing reference architecture,” NIST special publication, vol. 500, p. 292, 2011.

[30] J. Happe, W. Theilmann, A. Edmonds, and K. Kearney, Service Level Agreements for Cloud Computing. Springer-Verlag, 2011, ch. A Reference Architecture for Multi-Level SLA Management, pp. 13–26.

[31] H. Ludwig, A. Keller, A. Dan, R. P. King, and R. Franck, “Web Service Level Agreement (WSLA) Language Specification, v1.0,” Jan. 2003. [Online]. Available: http://www.research.ibm.com/wsla/WSLASpecV1-20030128.pdf [retrirved: may, 2013].

[32] K. T. Kearney, F. Torelli, and C. Kotsokalis, “Sla*: An abstract syntax for service level agreements,” 11th IEEE/ACM International Conference on Grid Computing, pp. 217–224, 2011.

[33] SLA@SOI. SLA@SOI projekt website. http://sla-at-soi.eu/. [retrirved: june, 2013].

[34] MIRROR-42, “Kpi library.” [Online]. Available: http://mirror42.com [retrieved: june, 2013].

[35] W. Sun, Y. Xu, and F. Liu, “The role of xml in service level agreements management,” in Services Systems and Services Management, 2005. Proceedings of ICSSSM ’05. 2005 International Conference on, vol. 2, 2005, pp. 1118–1120.

[36] Security Guidance for Critical Areas of Focus in Cloud Computing, https:// downloads.cloudsecurity alliance.org/initiatives/guidance/csaguide.v3.0.pdf.

[37] Marco Comuzzia, Constantinos Kotsokalisb, George Spanoudakisa,Ramin Yahyapour. Establishing and Monitoring SLAs in complex Service Based Systems, IEEE International Conference on Web Services, 2009: 783-790.

[38] FENG Deng-guo, ZHANG Min, ZHANG Yan，et al.Study on Cloud Computing Security［J］.Journal of Software, 2011, 22(1) : 71-83.

[39] WANG Lin-song，LIU De-shan，GUO Jin. Design of Public Cloud Security Architecture [J]. Journal of Jilin University (Information Science Edition), 2013, 31 (2):166-169

[40] Fu Yingxun,Luo Shengmei,Shu Jiwu. Survery of Secure Cloud Storage System and Key Technologies [J]. Journal of Computer Research and Development, 2013, 50 (1):136-145

[41] Xu Yingying, GaoFei,Shang Fengying.New Cloud Security solutions and its key technologies . Huazhong Univ.of Sci.&Tech. (Natural Science Edition),2012,40(Z1): 74-78

[42] Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J., & Ghalsasi, A. (2011). Cloud computing—The business perspective. Decision Support Systems, 51(1), 176-189.

[43] Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., & Zaharia, M. (2010). A view of cloud computing. Communications of the ACM, 53(4), 50-58.

[44] Motta, G., You, L., Sacco, D., & Sfondrini, N. (2013, May). Cloud computing: the issue of service quality: an overview of cloud service level management architectures. In Service Science and Innovation (ICSSI), 2013 Fifth International Conference on (pp. 230-233). IEEE.

[45] Garg, S. K., Versteeg, S., & Buyya, R. (2013). A framework for ranking of cloud computing services. Future Generation Computer Systems, 29(4), 1012-1023.

[46] Lee, S. Y., Tang, D., Chen, T., & Chu, W. C. (2012, July). A QoS Assurance middleware model for enterprise cloud computing. In Computer Software and Applications Conference Workshops (COMPSACW), 2012 IEEE 36th Annual (pp. 322-327). IEEE.

[47] Zhu, F., Li, H., & Lu, J. (2012, May). A service level agreement framework of cloud computing based on the Cloud Bank model. In Computer Science and Automation Engineering (CSAE), 2012 IEEE International Conference on (Vol. 1, pp. 255-259). IEEE.

[48] P. Mell and T. Grance, "The NIST Definition of Cloud Computing", http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf, **(2011)** September.

[49] R. Buyya, C. H. Yeo, S. Venugopal, J. Broberg and I. Brandic, "Cloud computing and emerging IT platforms, Vision, hype, and reality for delivering computing as the 5th utility", Future Generation Computer Systems, http://www.buyya.com/papers/Cloud-FGCS2009.pdf, **(2009)** June.

[50] C. S. Yeo DeAssuncao MD, Y. J. Sulistio, A. Venugopal, S. Placek and M. Buyya, "Utility computing on Global Grids", http://www.buyya.com/papers/HandbookCN_Utility_Grids.pdf, **(2006)**.

[51] "Sun Microsystems, Service Level Agreement in the Data Center", **(2010)** March. http://www.sun.com/blueprints

[52] L. Wu and R. Buyya, "Service Level Agreement (SLA) in Utility Computing Systems", The University of Melbourne. Australia. http://arxiv.org/ftp/arxiv/papers/1010/1010.2881.pdf.

[53] E. Marilly, O. Martinot, S. Betgé-Brezetz and G. Delègue, "Requirements for Service Level Agreement Management", France, http://www-rp.lip6.fr/adanets/PublicDoc/Papers/IPOM2002_SLA_corrected_modifencoursv2.pdf.

[54] M. Alhamad, T. Dillon and E. Chang, "Conceptual SLA Framework for Cloud Computing", Australia,http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5610586&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D5610586.

[55] A. Sahai, S. Graupner, V. Machiraju and V. Moorsel, "Specifying and Monitoring Guarantees in Commercial Grids through SLA", http://www.hpl.hp.com/techreports/2002/HPL-2002-324.pdf.

[56] A. Keller and H. Ludwig, "The WSLA Framework: Specifying and Monitoring Service Level Agreements for Web Services", http://clip.dia.fi.upm.es/Projects/S-CUBE/papers/keller03:wsla_framework.pdf, **(2003)**.

[57] P. Allen, "Service Level Agreements", , http://www.cbdiforum.com/report_summary.php3?page=/secure/interact/2006-12/service_level_agreements.php&area=silverH, **(2003)** .

[58] P. Bianco, G. A. Lewis and P. Merson, "Service Level Agreements in Service-Oriented Architecture Environments", http://www.sei.cmu.edu/reports/08tn021.pdf, **(2008)** September.

[59] "SLA Management Handbook, Concepts and Principles, TeleManagement Forum", http://www.afutt.org/Qostic/qostic1/SLA-DI-USG-TMF-060091-SLA_TMForum.pdf, vol. 2, **(2005)**.

[60] A. Pichot, "Dynamic SLA-negotiation based on WS Agreement", http://cui.unige.ch/~dimarzo/courses/services/Lecture6.pdf, **(2008)**.

[61] P. Rubach and M. Sobolewski, "Dynamic SLA Negotiation in Autonomic Federated Environments", http://sorcersoft.pl/publications/papers/2009/OTM-2009.pdf, **(2009)**.

[62] Cloud Security Alliance. (2014). *CSA Security, Trust and Assurance Registry.* Retrieved from

https://cloudsecurityalliance.org/star/#_overview

[63] Cloud Standards Customer Council. (2012). *Practical Guide to Cloud Service Level Agreements Version 1.0.* Retrieved from http://www.cloud-council.org/2012_Practical_Guide_to_Cloud_SLAs.pdf

[64] U.S. Senate Committee on Commerce, Science, and Transportation. (2014). *A Kill Chain Analysis of the 2013 Target Data Breach.* Retrieved from http://www.commerce.senate.gov/public/?a=Files.Serve&File_id=24d3c229-4f2f-405d-b8dba3a67f183883

[65] ENISA. (2012). *Benefits, risks, and recommendations for information security.* Retrieved from

http://www.enisa.europa.eu/events/speak/cloud.jpg/view

[66] European Commission. (2014). *Cloud Service Level Agreement Standardisation Guidelines.* Retrieve from http://ec.europa.eu/digital-agenda/en/news/cloud-service-level-agreement-standardisationguidelines

[67] Grance, Timothy and Jansen,Wayne J. (2011). *NIST SP800-144: Guidelines on Security and Privacy in Public Cloud Computing.* Retieved from http://www.nist.gov/customcf/get_pdf.cfm?pub_id=909494

[68] Hale, Matthew & Gamble, Rose. (2012). *Risk Propagation of Security SLAs in the Cloud.* Retrieved from http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=6477665&url=http%3A%2F%2Fieeexplore.ieee.org%2Fiel7%2F6470041%2F6477486%2F06477665.pdf%3Farnumber%3D6477665

[69] Henning, Ronda. (1999). *Security Service Level Agreements: Quantifiable Security for the Enterprise?* Retrieved from http://www.nspw.org/papers/1999/nspw1999-henning.pdf

[70] Hogben, G. & Dekker, M. (2012). *Procure Secure: A guide to monitoring of security service levels in cloudcontracts. Technical report, European Network and Information Security Agency* (ENISA). Retrieved from http://www.enisa.europa.eu/activities/Resilience-and-CIIP/cloudcomputing/ procure-secure-a-guide-to-monitoring-of-security-service-levels-in-cloud-contracts

[71] ISO. (2013). *Information Technology – Security techniques – Information security management systems – Requirements*. Retrieved from http://www.iso.org/iso/catalogue_detail?csnumber=54534

[72] Krutz, Ronald L. & Vines, Russell Dean. (2010). *Cloud Security: A Comprehensive Guide to Secure Cloud Computing.* Wiley Publishing:Indianapolis, IN.

[73] Butler, Brandon. (2012). *Nine security controls to look for in cloud contracts.* NetworkWorld. Retrieved from http://www.networkworld.com/article/2161443/cloud-computing/nine-security-controls-tolook-for-in-cloud-contracts.html

[74] Brodkin, J. (2008). *Gartner: Seven cloud-computing security risks*. Infoworld. Retrieved from

http://www.infoworld.com/article/2652198/security/gartner--seven-cloud-computing-securityrisks.

Html

[75] Chartered Institute of Purchasing and Supply. (2009). *How to prepare Service Level Agreements*. Retrieved from
http://www.cips.org/Documents/Resources/Knowledge%20How%20To/How%20to%20prepare%20Service%20Level%20Agreements.pdf?bcsi_scan_3F31264ACB0CFD71=hHhIS/

Acronyms

A

API-application programming interface
AUP -Acceptable Use Policy

B

BIA- Business impact analysis

C

CSCC -Cloud Standards Customer Council

CSLA -Cloud Service Level Agreement

CSP-cloud service provider

CSF-critical success factors
CSA - Cloud service agreement

D

DMTF -Distributed Management Task Force

E

EEA -European Economic Area
EBS- Elastic Block Store

G

GAR -Grid ARchive

H

HVAC - heating, ventilation and air conditioning

I

IaaS- infrastructure as-a-service
IDE-integrated development environment
IBM -International Business Machines
IDS- Intrusion Detection Systems

J

JVM- Java Virtual Machines

K

KPI-Key Performance Indicator

M

MSA -Master Service Agreement
MTO - Maximum Tolerable Outage

N

NPM- network performance metrics
NOC- Network Operation Center

O

OLA- Operational-level agreements
IOPS -Input / Outputs per second
OVF -Open Virtualization Format

P

PaaS- platform-as-a-service
PII -personally identifiable information

Q

QoS- Quality of Service
QoP - Quality of Protection

R

REST-Representational state transfer
RPO -Recovery Point Objective
RTO- Recovery Time Objective
RFP -Request for Proposal

S

aaS- software-as-a-service
SLA -service level agreement
SP -Service Provider
SLM- service level management
SDK-software development kit
SLO-service level objectives
SQP -Service Quality Plan
SOC- security operations center

T

TMF-Tele Management Forum

V

VM-Virtual Machine
VPN-Virtual Private Network

W

WAR- Web application ARchive

INDEX