Practical Guide to Cloud Service Level Agreements
Version 1.0
Contents
Practical Guide to Cloud Service Level Agreements Version 1.0 ................................................. 1
Acknowledgements .................................................................................................................. 4

  Workgroup Leaders ........................................................................................................ 4
  Extended Workgroup Members ....................................................................................... 4
  Additional Reviewers ..................................................................................................... 4

Introduction .......................................................................................................................... 5

Current SLA Landscape ........................................................................................................ 5

Guide for Evaluating Cloud Service Level Agreements ...................................................... 7
  Step 1: Understand Roles & Responsibilities ............................................................... 7
  Step 2: Evaluate Business Level Policies ..................................................................... 9
  Step 3: Understand Service and Deployment Model Differences .............................. 14
  Step 4: Identify Critical Performance Objectives ....................................................... 18
  Step 5: Evaluate Security and Privacy Requirements .................................................. 21
  Step 6: Identify Service Management Requirements .................................................. 28
  Step 7: Prepare for Service Failure Management .......................................................... 31
  Step 8: Understand the Disaster Recovery Plan ............................................................ 34
  Step 9: Develop an Effective Management Process ..................................................... 36
  Step 10: Understand the Exit Process .......................................................................... 39

Summary of Keys to Success ............................................................................................... 41

Works Cited .......................................................................................................................... 43

Additional References ......................................................................................................... 43
Acknowledgements

The Practical Guide to Cloud Service Level Agreements is a collaborative effort that brings together diverse customer-focused experiences and perspectives into a single guide for IT and business leaders who are considering cloud adoption. The following participants have provided their expertise and time to this effort.

Workgroup Leaders
John Meegan (IBM) – Lead Technical Editor; Introduction and Keys to Success Section Leader
Gurpreet Singh (Ekartha) – Current SLA Landscape and Disaster Recovery Section Leader
Steven Woodward (Cloud Perspectives) – Roles & Responsibilities; Performance Objectives Leader
Salvatore Venticinque (Second University of Naples) – Service & Deployment Model Section Leader
Massimiliano Rak (Second University of Naples) – Service & Deployment Model Section Leader
David Harris (Boeing) – Business Policies Section Leader
Gerry Murray (Fort Technologies) – Business Policies Section Leader
Beniamino Di Martino (Second University of Naples) – Business Policies Section Leader
Yves Le Roux (CA Technologies) – Security and Privacy Section Leader
John McDonald (CloudOne Corporation) – Service Management Section Leader
Ryan Kean (The Kroger Co.) – Service Failure Management Section Leader
Marlon Edwards (Hoboken Consulting Group, LLC) – Disaster Recovery Section Leader
Dave Russell (IBM) – Management Process Section Leader
George Malekkos (Powersoft Computer Solutions Ltd) – Exit Process Section Leader

Extended Workgroup Members
The workgroup leaders wish to recognize the following individuals for their outstanding efforts to provide content, share their expertise and ensure completeness of the Practical Guide to Service Level Agreements: Amy Wohl (Wohl Associates), Asher Bond (Elastic Provisioner, Inc.), Claude Baudoin (cebe IT & KM), Christopher Ferris (IBM), Melvin Greer (Lockheed Martin), Richard Miga (Synergistic Solutions), Thomas Somers (IBM).

Additional Reviewers
The following reviewers provided feedback on the Practical Guide to Cloud Service Level Agreements: Jenny Huang (AT&T), Karen Caraway (The MITRE Corporation), Kenneth Dilbeck (TMForum), Roopali Thapar (IBM), Tobias Kunze (Red Hat).
Introduction
This document is an extension of the Practical Guide to Cloud Computing white paper that was delivered by the Cloud Standards Customer Council (CSCC) in October, 2011.

The aim of this guide is to provide a practical reference to help enterprise information technology (IT) and business decision makers as they analyze and consider service level agreements (SLA) from different cloud service providers. The paper will give guidance to decision makers on what to expect and what to be aware of as they evaluate SLAs from their cloud computing providers. A checklist of key criteria for evaluating and comparing SLAs from different providers will be included. Additionally, this paper will highlight the role that standards play to improve interoperability and comparability across different cloud providers, and identify areas where future standardization could be effective.

SLAs are important to clearly set expectations for service between the cloud consumer (buyer) and the cloud provider (seller). Each cloud entity engaged by the enterprise should have a cloud SLA defined, including: cloud provider, cloud carrier, cloud broker and even cloud auditor. Consideration must also be given to the different models of service delivery: Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS)) as each model brings different requirements. This paper focuses primarily on the SLA details between the cloud consumer and cloud provider, and focuses on the requirements that are common across the various service models (emphasis is given to the IaaS service model since SLAs are more advanced in this area).

The Practical Guide to Cloud Service Level Agreements contains a set of guidelines and strategies to help decision makers in all major activities related to cloud SLAs.

The section entitled “Current SLA Landscape” explains the dynamics that currently exists between consumers and providers in the SLA space and the impact that company size has on the power to negotiate terms. This section also highlights the nuances of SLA development for different service models.

The section entitled “Guide for Evaluating Cloud Service Levels Agreements” is the heart of the guide. It provides a prescriptive series of steps that cloud consumers should take to evaluate cloud SLAs with the goal of comparing cloud service providers or negotiating terms with a provider. It provides guidance for the business and service level objectives of the SLA, highlighting what to be aware of and how to compare service levels across different cloud providers. This section takes into account the realities of today’s cloud computing ecosystem and postulates how it is likely evolve in the future, including the important role that standards will play to improve interoperability and consistency across providers.

Current SLA Landscape
The Service Level Agreement (SLA) serves as a means of formally documenting the service(s), performance expectations, responsibilities and limits between cloud service providers and their users. A typical SLA describes levels of service using various attributes such as: availability, serviceability,
performance, operations, billing, and penalties associated with violations of such attributes. Well-designed SLAs can significantly contribute to reducing causes of potential conflict and can facilitate issue resolution before a dispute materializes.

To guarantee an agreed service level, service providers must be capable of measuring and monitoring relevant metrics. But often there is a gap between correlating the metrics collected and monitored by service providers to higher-level functional guarantees which are of interest to consumers. This problem is quite challenging and development of SLAs by service providers has to take this into account. This is typical for all types of cloud services and is acute for SaaS providers which offer applications at higher levels of functionality. This is among the many reasons why SLAs for SaaS applications usually lack stringent service level guarantees. A vast majority of SaaS and PaaS providers simply offer no SLAs, although they strive to develop internal operations that contribute to increased reliability.

The situation for IaaS is better than SaaS and PaaS, but most public cloud infrastructure services are available only through non-negotiable standard contracts. They strictly limit the provider’s liability and the remedies do not provide significant benefit to consumers in case of service disruptions. Furthermore, most IaaS providers put the burden of SLA violation notification and credit request on their customers. Since a vast majority of the users of IaaS public clouds are small and medium businesses (SMB), the pressure on cloud providers to offer stringent SLAs is minimal.

In today’s nascent cloud services industry, SLAs provided by cloud vendors are increasingly being viewed by consumers as unsatisfactory forms of protection that weigh heavily in the provider’s favor. Disputes often arise over the structure of the agreements, monitoring of performance, and service unavailability. The reasons for this are numerous and cover both the challenges faced by cloud providers, and the lack of market power by consumers to potentially seek more stringent agreements.

Recent cases of cloud service outages highlight the dilemma for cloud consumers. In some cases, the SLAs offered by providers were loose enough to ensure that service providers were not violating the terms of the SLA despite a serious outage in mission critical areas of the service. In such cases, cloud consumers had no options to seek adequate penalties despite significant adverse effects. Cloud providers are taking account of such problems and are beginning to offer different service options that shield customers from such risks.

In general, the larger the customer deployment, which translates to higher subscription and upfront fees, the more power customers exert in negotiating stringent SLAs. Even in the case of SaaS providers, large customers are successful in negotiating a stronger agreement, where none may be offered to SMB customers. This does point to the trend that as cloud deployments proliferate to larger enterprise customers, the demand for stronger SLAs will intensify. As competition increases at all levels, better SLAs will inevitably become a competitive factor. At that time, large enterprises and SMBs, alike, will be able to choose based on more flexible and more favorable SLA terms and, in general, these terms will improve.
Guide for Evaluating Cloud Service Level Agreements

Before evaluating any cloud SLA, consumers must first develop a strong business case and strategy for their cloud computing environment. This includes identifying specific services that will be deployed in the cloud along with a clear understanding of the criticalness of these services to the business. A check on the exit clauses of current hosted services contracts is also important. Only after this strategic analysis has been completed can the consumer effectively evaluate and compare SLAs from different providers.

With the cloud business case and strategy as a prerequisite, this section provides a prescriptive series of steps that should be taken by cloud consumers to evaluate cloud SLAs with the goal of comparing cloud service providers or negotiating terms with a 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. Define an effective management 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

In order for consumers to understand specific roles and responsibilities explicitly or implicitly stated in a cloud SLA, it is important that they are aware of the various actors that can potentially participate in a cloud computing environment. The National Institute of Standards and Technology (NIST) Reference Architecture\(^1\) identifies 5 unique cloud actors:

- **Cloud Consumer.** The person or organization that maintains a business relationship with, and uses service from, cloud providers.
- **Cloud Provider.** The person, organization or entity responsible for making a service available to cloud consumers.
- **Cloud Carrier.** The intermediary that provides connectivity and transport of cloud services from cloud providers to cloud consumers.
- **Cloud Broker.** An organization that manages the use, performance and delivery of cloud services, and negotiates relationships between cloud providers and cloud consumers.
- **Cloud Auditor.** A party that can conduct independent assessments of cloud services, information system operations, performance and security of the cloud implementation.

---

\(^1\) Refer to [http://www.nist.gov/itl/cloud/refarch.cfm](http://www.nist.gov/itl/cloud/refarch.cfm) for more information on the NIST Reference Architecture.
The use of the term “broker” varies significantly and should be clarified with the various stakeholders, especially in context of a cloud SLA. An entity may provide broker services and functionality, but as a legal organizational entity not be recognized as a cloud broker. For example, an entity may perform research and negotiate on behalf of a consumer, but the actual SLA and contract terms are between the cloud consumer and cloud provider. The distinction of acting “broker like” vs. being an actual “broker” will evolve as the cloud computing industry matures and terminologies become more consistent. Due to these complexities this paper does not address all the SLA considerations for cloud brokering.

Consumers need to recognize the activities and responsibilities of each cloud actor that is engaged in delivering their cloud environment, and precisely define requirements and desired service levels for each actor. This paper focuses primarily on the cloud consumer/cloud provider SLA, although other SLAs may be addressed in a particular context.

In some cases, the consumer/provider relationship will indirectly include additional actors. Figure 1 below illustrates an environment where a cloud provider has established a SLA with two cloud carriers to establish service levels for communication and transport. In addition to cloud provider expectations, the consumer/provider SLA in this example may also include specific carrier and transport expectations. In this case, the cloud provider is also acting as a “broker” for the other two cloud carriers.

**Figure 1. Indirect Relationships**

![Diagram of indirect relationships between cloud provider, consumer, and carriers with SLAs](image)

Each cloud SLA will be unique based upon the consumers’ requirements and the cloud ecosystem under consideration. SLAs can contain various expectations between the actors 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 consumers gain a solid understanding of the
spectrum of SLAs that currently exist for cloud providers (and other actors as appropriate) in order to compare providers and assess tradeoffs between cost and service levels.\(^2\)

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

**Step 2: Evaluate Business Level Policies**

Consumers must consider key policy issues when reviewing a cloud SLA since there are interdependencies between the policies expressed in the SLA and the business strategy and policies developed in other aspects of the business. The data policies of the cloud provider, as expressed in the SLA, are perhaps the most critical business level policies that should be carefully evaluated.

The duty of care a cloud provider has to its clients and their data is partly governed by the data protection legislation applicable in the user’s local jurisdiction and also in those jurisdictions in which its data may reside or made available. Consumers should carefully consider these legal requirements and how the SLA their provider(s) offers deals with issues such as movement of data to offer multisite redundancy across several jurisdictions.

Table 1 highlights the critical data policies that need to be considered and included in the cloud SLA.

### Table 1. SLA data policies

<table>
<thead>
<tr>
<th>Data Policy</th>
<th>Description / Guidance</th>
</tr>
</thead>
</table>
| **Data Preservation** | • Timely and efficient capturing and preservation of data is critical to maintaining the organizational memory of a business or the general user.  
  • It enables the data controller to address operational, strategic and litigious situations from an informed perspective.  
  • Cloud users should ensure the service supports their data preservation strategy that includes sources, scheduling, backup, restore, integrity checks, etc. |
| **Data Redundancy**  | • Users should ensure they have an appropriate data preservation strategy that addresses redundancy within the system.  
  • This should be complemented by the SLA their cloud provider offers and can be tested to demonstrate service availability.  
  • The consumer should be concerned as to the protections offered or omitted by the provider. |

\(^2\) Refer to the Queen Mary School of Law Legal Studies Research Paper titled *Contracts for Clouds: Comparison and Analysis of the Terms and Conditions of Cloud Computing Services* at [http://ssrn.com/abstract=1662374](http://ssrn.com/abstract=1662374) for a comprehensive analysis of leading cloud service providers.
service provider.

| Data Location | • SLAs which cover single and multi-jurisdiction scenarios are challenging.  
|               | • Users should carefully consider as part of their data management strategy how the SLA will complement where their data will reside, where it is processed, and how this meets regulatory requirements.  
|               | • For example, can the provider truly deliver a sound technical solution when data spans several jurisdictions? Can the user trust that the data will be located where the provider commits it will be?  
| Verification of new data location | • Clients should ensure that when a provider elects to provide its service from another location it will be required to notify its clients of the new location and provide a means for the client to independently verify where the data will be relocated.  
|               | • Preferably the provider should be required to obtain the permission of the client to relocate the data before moving to a new location.  
| Data Seizure | • Legal powers enable law enforcement and other government agencies to seize data under certain circumstances.  
|               | • Consumers should also ensure there are arrangements in place to make a user’s data available in the event that their SLA provider goes out of business and the data center provider locks access to its systems pending payment of the outstanding account.  
| Data Privacy | • It is a requirement to conduct business in compliance with applicable laws on data privacy protection and data security.  
|               | • The declared data privacy policy from the provider should be included in the SLA.  
|               | • Examples of privacy terms that should be addressed in the SLA include the data sets gathered, data retention policies, how the data is communicated, how personal data is stored and used, etc.  
|               | • Refer to the Privacy section on page 24 for more information.  

In addition to data policies, there are a number of other business level policies expressed in the cloud SLA that require careful evaluation. All of these policies will impact and influence the consumer’s cloud strategy and business case. In many cases, these policies, as defined in the cloud SLA, 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 consumers.

Table 2 below highlights the critical business level policies that need to be considered and addressed in the cloud SLA.
Table 2. SLA business level policies

<table>
<thead>
<tr>
<th>Policy</th>
<th>Description / Guidance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guarantees</td>
<td>• SLA guarantees should be defined, objective and measurable with an appropriate scaled penalty matrix which complements the impact of non-performance by the provider. The SLA should clarify:</td>
</tr>
<tr>
<td></td>
<td>o What constitutes excused or excluded performance</td>
</tr>
<tr>
<td></td>
<td>o Escalation procedures</td>
</tr>
<tr>
<td></td>
<td>o How service-level bonuses and penalties are administered</td>
</tr>
<tr>
<td></td>
<td>o Remedy circumstances and mechanisms</td>
</tr>
<tr>
<td></td>
<td>• Such guarantees should include an availability measurement expressed as a percentage, e.g. 99.999%, which denotes the amount of time the service is guaranteed to be working.</td>
</tr>
<tr>
<td>Acceptable Use Policy</td>
<td>• Given the dislocated relationship that can exist between the provider and the consumer (the provider may not know the final end user of its service), the acceptable use policy will clearly describe how the consumer may use its service and the agreement generally will describe what actions the provider may take in the event of a breach.</td>
</tr>
<tr>
<td></td>
<td>• In today’s cloud environment, this policy is typically non-negotiable and the terms generally favor the cloud provider.</td>
</tr>
<tr>
<td>List of Services Not Covered</td>
<td>• The SLA will state under what conditions and with which described services the consumer is supported. The SLA may also state what is excluded and what constitutes illegal use.</td>
</tr>
<tr>
<td></td>
<td>• Consumers should look for explicitly stated exceptions and understand why the provider has excluded them.</td>
</tr>
<tr>
<td>Excess Usage</td>
<td>• Providers operate business models to drive revenue. Consumers may find that usage above their contracted thresholds may incur ‘premium’ rates which can be punitive and disrupt their budgets.</td>
</tr>
<tr>
<td></td>
<td>• Consumers should correctly size their usage requirements, reduce the opportunity for usage creep and consider and understand the ‘what-if’s’ of breach of their usage thresholds.</td>
</tr>
<tr>
<td></td>
<td>• The cloud is a great tool to vary the size of the resource pool your business buys but, as in any business model, buy only what you need and avoid excess usage charges.</td>
</tr>
</tbody>
</table>

3 Guarantees including measurable metrics will be covered in greater detail in the sections that follow.
### Activation

- Providers will delineate a time at which the service becomes active and the SLA commences operation. This time stamp can then be used to measure time and if an outage occurs may be used to establish the start of a penalty incurring ‘event’.

- Examples of activation include the time at which the consumer checks the ‘accept terms and conditions’ box or when the consumer acknowledges the secure URL providing login credentials to their new cloud service.

- From a SLA compliance perspective, it is important for consumers to understand the trigger points under the SLA so they can independently measure ‘event’ timing.

- It is also important to understand when the agreement is in operation as both the provider and the consumer have responsibilities such as Fair Use and Legal Use.

### Payment and penalty models

- The SLA should clarify when/how payment is to be made. Provider payment models vary. Monthly recurring or “pay as you use” models are typical.

- Associated with these payment models are credit terms which may 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 consumer needs to be diligent in obtaining service credit payments for outages.

### Governance / Versioning

- Providers’ 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 SLA was initially accepted are changed, the consumer should review the impact on their specific situation.

- A good provider will maintain a proactive policy of advising consumers of changes to their SLA and practice version control.

- Consumers should ensure that there is a mechanism in place to be informed of changes and, if not, amend their contract to put the onus on the provider to contact them at a designated contact point with updates within a reasonable timeframe.

### Renewals

- Renewals are an opportunity to bargain for better rates and relocate to another provider if necessary.

- Providers may operate auto renewals which state, for example, this contract will auto renew on its anniversary if the consumer does not give 90 days notice of their intention to cancel.

- Consumers should read the terms and conditions for the renewal arrangements, and consider the conditions under which a provider may vary the service terms at renewal. Provider systems are organic as are consumer requirements. More attractive service functionality may be available or more consumer flexibility required.
Transferability

- Consumers should consider the potential for needing to transfer an agreement in the event their business is sold.

- In addition, the provider’s business may be sold to a competitor and it may suit the consumer to relocate and discontinue the commercial agreement with the new owner.

- Consumers may operate several accounts with a provider and want to offset account credits between accounts. Is this supported in the provider’s contract terms?

Support

- The consumer should ensure they follow the reporting guidelines of the provider to ensure the support terms specified in the SLA are activated.

- An example of a support and escalation matrix is provided below. All three target times in the table are associated with the commencement ‘time stamp’ of the service or the notification of a service affecting event.

<table>
<thead>
<tr>
<th>Priority</th>
<th>Description</th>
<th>Target Response Time</th>
<th>Target Update Time</th>
<th>Target Fix Time</th>
</tr>
</thead>
<tbody>
<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>
</tbody>
</table>

Planned Maintenance

- All systems require maintenance. Complex systems may be designed to include sufficient resources such that maintenances can be carried out without affecting the service.

- The SLA may, however, describe ‘uptime’ as an availability percentage (e.g. 99.90%). This is the equivalent of 8.5 hours downtime per annum. SLAs may state that this does not include ‘planned maintenance’. Thus, the provider may have a service outage for 8.5 hours + maintenance time under the SLA and the consumer is not entitled to compensation.

Subcontracted Services

- Providers sometimes include in their SLAs indication that the SLA of an upstream (subcontracted) provider will be passed to the consumer and the only available penalties are those of the upstream provider even though these may be of a lesser quality than the consumer understands when reviewing the SLA of their immediate...
provider.

- Equally, the consumer should ensure that the immediate provider SLA is unambiguous in stating that their SLA applies to the complete service and not limited to the immediate provider’s component of the overall service.

## 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 ‘patching’ the licensed software over to the consumer once they have started to use the service. This can be preferable to the provider as later releases may negatively affect the consumer’s service.

- The provider may alternatively ‘push’ the update in which case the consumer, as specified in the SLA, may require notification of the update. Upon notification, the consumer can issue an ‘opt out’ option.

## Industry Specific Standards

- Regulated industries, like government, financial services, and healthcare, will have specific and usually quite onerous and costly standards which must be addressed in the cloud SLA and implementation.

- Consumers who operate in these regulated industries should ensure that the full resources of their legal team are brought to bear on the negotiation of the SLA.

## Additional terms for different geographic region or countries

- Consumers should consider the provider’s origins and primary market. Detailed refinements to the home market SLA may be required to properly cover consumers who are located in remote markets.

- Data protection legislation is one aspect of this. However, consumers should not limit their examination of the agreement to this solely.

### 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 cloud SLA. In addition, the level of clarity varies significantly for each service model. In general, PaaS and SaaS objectives are less precise than IaaS objectives since the variety of solutions and applications offered by providers is much broader for these service models.

Table 3 highlights the different SLA considerations for each of the cloud service models.

**Table 3. SLA considerations for Service Models**
Service Model | SLA Considerations
--- | ---
IaaS | • Cloud IaaS SLAs 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). Consumers should expect to find 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. Consumers are simply guaranteed to have the compute resources for which they paid with technical enforcement at the hypervisor level.
• 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 consumer.
• Denial-of-service (DoS) attacks may be explicitly excluded from the SLA even if the provider offers protection via firewalls and intrusion detection systems.
• There are several standardization efforts within the IaaS space which help describe and manage the services offered at this level. Whenever possible, consumers should ensure the cloud SLA 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 SLA, consumers should consider tradeoffs in flexibility, control, and ease of use to determine which approach best meets their business needs.
  o Integrated solutions, like Google App Engine (GAE), 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.

---

4 IaaS standards include: DMTF CIMI (Cloud Infrastructure Management Interface), DMTF OVF (Open Virtualization Format), SNIA CDMI (Cloud Data Management Interface), The Open Group’s SOCCI (Service-Oriented Cloud-Computing Infrastructure), OGF OCCI (Open Cloud Computing Interface), the ISO Study Group on Cloud Computing (SGCC), and de-facto industry alignment on IaaS service level objectives, warranties, guarantees, performance metrics, etc.

provider. Typically, service developers only have access to a provider-defined set of APIs (Task API in GAE) which offer limited control on the coordination of code execution.

- Deploy-based solutions enable deployment of middleware on top of resources acquired from an IaaS cloud provider, offering deployment services to the consumers which automate the process of installation and configuration of the middleware.\(^6\) 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.

- Consumers must distinguish between PaaS development environments and PaaS production environments when reviewing their cloud PaaS SLA. PaaS production environments will typically require more stringent service level objectives than PaaS development environments.

- The state of PaaS metrics is currently immature with metrics varying significantly across providers. Standards initiatives are just starting to emerge in this space. In the meantime, consumers are advised to identify the PaaS services that are critical to their business and ensure that their cloud SLAs contain clear and measurable metrics for these services.

- Standards that help identify PaaS services offered by cloud providers and standard interfaces for communicating with PaaS providers to provision or manage PaaS environments are also lacking. The lack of standards results in poor portability and interoperability across providers. Standards, like OASIS TOSCA\(^7\), are emerging which will help address this issue. Consumers should ensure their cloud SLA includes support for open standards, as they become available, to reduce vendor lock in.

---

**SaaS**

- 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 consumers to look out for in their cloud SLAs.

- Consumers should expect general SaaS service level objectives like *monthly cumulative application downtime, application response time, persistence of consumer information, and automatic scalability* to be included in their SLA.

- Consumers 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.

---

\(^6\) Deploy-based solutions are supported by commercial providers like IBM, Oracle and Microsoft as well government sponsored projects like OPTIMIS, CONTRAIL, Cloud4SOA and mOSAIC in Europe.

\(^7\) Refer to [http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=tosca](http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=tosca) for details on TOSCA.
In addition to service models, service deployment terms should be included in a cloud SLA. These terms should clarify to both parties signing the SLA 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 SLA should clearly specify one of the following options: Private, Community, Public, or Hybrid. Consumers must be well educated on the characteristics and differences in each of these deployment models since potential value and risk varies significantly.\(^8\)

Table 4 highlights the different SLA considerations across the deployment models.\(^9\)

**Table 4. SLA considerations for Deployment Models**

<table>
<thead>
<tr>
<th>Deployment Model</th>
<th>SLA Considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private (On-site)</td>
<td>• SLA 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, consumers must ensure that critical service objectives like availability and response time are met via ongoing measurement and tracking.</td>
</tr>
<tr>
<td>Private (Outsourced)</td>
<td>• SLA 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 consumer mitigates potential security and availability risks.</td>
</tr>
<tr>
<td></td>
<td>• Consumers should ensure the cloud SLA specifies security techniques for protecting the provider’s perimeter and the communications link with the provider.</td>
</tr>
<tr>
<td></td>
<td>• Consumers should consider the criticalness of the service being deployed to justify the added expense of this model over the Public model.</td>
</tr>
<tr>
<td>Public</td>
<td>• SLA considerations for the Public model are greater than the Private (Outsourced) model since the provider’s IT resources are now shared across multiple consumers.</td>
</tr>
<tr>
<td></td>
<td>• As a result, consumers should carefully review the cloud SLA to understand how the provider addresses the added security, availability, reliability and performance risks introduced by multi-tenancy.</td>
</tr>
</tbody>
</table>
|                          | • The ability to measure and track specific service level objectives becomes more

---

\(^8\) Refer to the Practical Guide to Cloud Computing for considerations on selecting a deployment model.

\(^9\) The Community deployment model is not called out explicitly in the table since it is very similar to the Public deployment model.
important in the Public deployment model. Consumers should also ensure the cloud SLA provides adequate methods and processes for ongoing measurement.

| Hybrid | • SLA considerations for the Hybrid model are similar to the Public model with the increased likelihood for unique integration requirements between cloud and enterprise services. As a result, consumers should ensure the cloud SLA adequately covers their service and data integration requirements. |

In addition to specifying the deployment model, the SLA 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 WAR file\(^ {10} \) (the application server enables WAR uploading).
- A grid application is deployed on a grid container as a GAR (Grid Archive) 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\(^ {11} \) (OVF) is recommended.

When SLAs are signed, a clear description of the technologies involved in the deployment of services may be specified (for example, WAR, GAR, OVF, etc.). 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 efficiency and accuracy of service delivery. Performance considerations often include: availability, response time, transaction rate, processing speed, but can include many other performance and system quality perspectives.\(^ {12} \) Consumers 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 consumer should be measureable and auditable, and documented in the SLA providing a comfort level to both parties. Performance considerations are dependent on the supported service model (IaaS, PaaS and SaaS) and the type of services provided within that model, for example, network, storage and computing services for IaaS.

\(^ {10} \) Refer to \url{http://java.sun.com/j2ee/tutorial/1_3-fcs/doc/WCC3.html} for more information on the WAR (Web application ARChive) file.

\(^ {11} \) Refer to \url{http://www.dmtf.org/standards/ovf} for more details.

\(^ {12} \) System quality measures that could be included in service performance include accuracy, portability, interoperability, standards compliance, reliability, scalability, agility, fault tolerance, serviceability, usability, durability, etc.
Of course, in order for performance objectives to be meaningful, measurement is a critical consideration where clarity and consistency are very important, if not crucial, to gain trust in the cloud. Measuring without context is useless and performance metrics are no different. It must be clear how the metric will be used and what decisions will be made from the metrics, continually evaluating and aligning with specific goals and objectives.

This section will focus on two performance metrics: availability and response time. The intention is to provide a basic framework to help identify and define cloud metrics that will be meaningful and consistent. This section is not meant to be exhaustive - it does not contain all the potential metrics and possible definitions. Many of the measurement and metric definitions may already be supported by your cloud provider; therefore the specific interpretation of the term in context of a specific solution is critical. Some calibration may be required between existing captured measures and those specifically requested as part of the SLA.

Industry standard measures with applicable definitions should be used to improve consistency, enabling meaningful comparative and trend analysis. For instance, IEEE also has good measurement definitions and categorizations for activities such as maintenance.13 Other organizations and private benchmarks exist. The key is to always calibrate values to get an “apple to apples” comparison to an appropriate level of accuracy.

To be effective, a performance metric must be clearly defined in the SLA and understood by both parties. 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 including delays (typically measured in milliseconds).

Table 5 describes three different example scenarios (network availability, storage availability, and service response time) and the specific performance information required for each.

**Table 5. Availability and Response Time metrics examples**

<table>
<thead>
<tr>
<th>Metric Name in SLA</th>
<th>Network Availability (example)</th>
<th>Storage Availability (example)</th>
<th>Service Response Time (example)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network Percentage Available Critical Business Hours</td>
<td>Storage Percentage Available</td>
<td>Service XXX Response Time in a Given Hour; Service YYY Response Time in a Given Hour.</td>
<td></td>
</tr>
</tbody>
</table>

13 Other standard measures include the International Software Measurement Association – International Function Point Users Group that retains several software measurement guidelines (such as ISO/IEC 20926) that are used for benchmarking and works closely with the International Software Benchmarking Standards Group.
### Constraints

| No constraints | Response times will only be evaluated for services XXX and YYY which are PaaS reuseable services that will be invoked by our applications. |

| Machine | Machine | Machine |

### Collection Method

| Machine | Machine | Machine |

### Collection Description

| Using the DMTF, OGF\(^{14}\), 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 will be collected. |

### Other Information

| Considered 60 seconds of uptime for each successful “ping”. | Considered 60 seconds of uptime 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 services reporting is needed (for example, listing of all services that exceeded SLA agreed response time). |

### Usage 1 in SLA

| Network availability will be 99.5% between 12AM GMT to 12PM GMT Monday thru Friday. | Storage availability will be 99.9%. | Response time for XXX service must 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 network charges. | For any day when storage availability is less than 99.9%, a 50% discount will be applied for the entire day 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. |

---

Cloud resources, both hardware and facilities, should also be considered when assessing critical performance objectives for cloud SLAs. Hardware includes: compute (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, communications, and other aspects of the physical plant.

Resources need to be clearly stated in the SLA to clarify scope, constraints and expectations for the cloud computing services of interest. To be successful, higher level business objectives and goals need to be understood such that critical resource metrics can be identified that address facility and hardware.

expectations. Some metrics may not be technical, but can include measures such as watts of power usage, cubic feet of network cabinets, or even revenue dollars from real estate sold. For example, one frequent objective for the IaaS service model is to reduce power usage, or perhaps reduce the footprint size of the data center.

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

- Understand the business level performance objectives for the cloud opportunity (for example, reduce cost and time to market per unit of software functionality).
- Identify the set of metrics that are critical to achieving and managing the business level performance objectives.
- Ensure these metrics are defined at the right level of granularity that can be monitored on a continuous basis cost effectively.
- Identify standards to provide consistency for cloud metrics in areas such as metric definitions and methods of collection.
- 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 than traditional IT solutions. At a basic level, assets supported by the cloud fall into two general categories:

- Data
- Applications/Functions/Processes

Information is either being moved into the cloud or applications are being executed in the cloud (from partial functions all the way up to full applications).

A critical initial step for ensuring sufficient cloud security is establishing a 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 a brief description of data retention and destruction requirements. The classification scheme should be used as the basis for applying controls such as access controls, archiving or encryption.

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:

1. How would the business be harmed if the asset became publically available and distributed?
2. How would the business be harmed if an employee of our cloud provider accessed the asset?
3. How would the business be harmed if the process or function were manipulated by an outsider?
4. How would the business be harmed if the process or function failed to provide expected results?
5. How would the business be harmed if the information/data were unexpectedly changed?
6. How would the business be harmed if the asset was unavailable for a period of time?

Table 6 below highlights the key steps consumers should take to ensure their cloud SLA sufficiently addresses their unique security requirements.

Table 6. Key security considerations for cloud SLAs

<table>
<thead>
<tr>
<th>SLA Security Considerations</th>
<th>Strategic Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assess asset sensitivity and application operational security requirements</td>
<td>• An assessment of the confidentiality, integrity, and availability requirements for the asset must be completed.</td>
</tr>
<tr>
<td></td>
<td>• Consumers must address application operational security and availability requirements in response to identified risks and in line with the organization’s data classification, information architecture, information security architecture, and risk tolerance.</td>
</tr>
<tr>
<td></td>
<td>• A common challenge organizations face with the cloud is managing data. Many organizations report individuals or business units moving sensitive data to cloud services without the approval or even notification of IT or security.</td>
</tr>
<tr>
<td></td>
<td>• Take steps to detect unapproved data moving to cloud services:</td>
</tr>
<tr>
<td></td>
<td>o Monitor for large internal data migrations with database activity monitoring (DAM) and file activity monitoring (FAM )</td>
</tr>
<tr>
<td></td>
<td>o Monitor for data moving to the cloud with URL filters and data loss prevention</td>
</tr>
<tr>
<td></td>
<td>• Protect data in transit. All sensitive data moving to or within the cloud at the network layer, or at nodes before network transmission should be encrypted, sensitive volumes should be encrypted to limit exposure to snapshots or unapproved administrator access, and sensitive data in object storage should be encrypted, usually with file/folder or client/agent encryption.</td>
</tr>
<tr>
<td>Understand Legal/Regulatory Requirements</td>
<td>• Due to potential regulatory, contractual and other jurisdictional issues it is extremely important to understand both the logical and physical locations of data.</td>
</tr>
<tr>
<td>Establish and track security</td>
<td>• Metrics and standards for measuring performance and effectiveness of information security management should be established prior to moving into</td>
</tr>
</tbody>
</table>
metrics the cloud.

- 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.
  - Refer to the following resources for specific information on security metrics: ISO 27004:2009\[^{16}\], NIST Special Publication (SP) 800-55 Rev.1, Performance Measurement Guide for Information Security\[^{17}\], and CIS Consensus Security Metrics v1.1.0\[^{18}\]

<table>
<thead>
<tr>
<th>Assess and compare cloud providers’ security capabilities</th>
<th>• Determine if the provider’s guarantees adequately address your security requirements.[^{19}]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• The provider’s security governance processes and capabilities should be assessed for sufficiency, maturity, and consistency with the user’s information security management processes.</td>
</tr>
<tr>
<td></td>
<td>• The provider’s information security controls should be demonstrably risk-based and clearly support these management processes.</td>
</tr>
<tr>
<td></td>
<td>• 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.</td>
</tr>
<tr>
<td></td>
<td>• Assess the cloud provider’s level of security and its maturity:</td>
</tr>
<tr>
<td></td>
<td>- If compliance to a normative standard (e.g. ISO 27001) is given:</td>
</tr>
<tr>
<td></td>
<td>- Verify the compliance certificate and its validity</td>
</tr>
<tr>
<td></td>
<td>- Look for verifiable evidence of resource allocation, such as budget and manpower to sustain the compliance program</td>
</tr>
<tr>
<td></td>
<td>- Verify internal audit reports and evidence of remedial actions for the findings</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Audit cloud provider’s security SLA compliance</th>
<th>• A right to audit clause in a cloud SLA gives customers the ability to audit the cloud provider, which supports traceability and transparency in the frequently evolving environments of cloud computing and regulation.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Use a normative specification in the right to audit clause to ensure mutual ...</td>
</tr>
</tbody>
</table>

---


\[^{19}\] The Cloud Security Alliance Consensus Assessments Initiative Questionnaire (CAIQ) provides a set of questions a cloud consumer and cloud auditor may wish to ask of a cloud provider. Refer to [http://cloudsecurityalliance.org/research/cai/](http://cloudsecurityalliance.org/research/cai/) for details.
Some of the security risks associated with cloud computing are unique, partly due to an extended data-centric chain of custody, and it is in this context that the business continuity, disaster recovery, and traditional security environments of a cloud service provider need to be assessed thoroughly and in reference to industry standards. Consumers should ensure that the provider is compliant with global security standards like ISO 27001 ISMS\(^{21}\) or other industry-standards such as TOGAF\(^{22}\), SABSA\(^{23}\), ITIL\(^{24}\), COSO\(^{25}\), or COBIT\(^{26}\).

Providers should notify consumers about 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 may want the provider to indemnify them if the breach was their fault.

**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 and the security of information and computer systems. Table 7 provides an overview of the worldwide privacy regulations that currently exist.

---


\(^{22}\) TOGAF® is an Open Group Standard see [http://www.opengroup.org/togaf/](http://www.opengroup.org/togaf/).


\(^{26}\) COBIT is an IT governance framework and supporting toolset - see [http://www.isaca.org/Knowledge-Center/COBIT/Pages/Overview.aspx](http://www.isaca.org/Knowledge-Center/COBIT/Pages/Overview.aspx) for details.
### Worldwide privacy regulations

<table>
<thead>
<tr>
<th>Region</th>
<th>Regulation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Asia Pacific region, Japan, Australia, New Zealand, and others</strong></td>
<td>These regions have adopted data protection laws that require the data controller to adopt reasonable technical, physical, and administrative measures in order to protect personal data from loss, misuse, or alteration, based on the Privacy and Security Guidelines of the Organization for Economic Cooperation and Development (OECD)(^{27}), and the Asia Pacific Economic Cooperation’s (APEC) Privacy Framework.(^ {28})</td>
</tr>
<tr>
<td><strong>Japan</strong></td>
<td>In Japan, the Personal Information Protection Act(^ {29}) requires the private sectors to protect personal information and data securely. In the healthcare industry, profession-specific laws, such as the Medical Practitioners' Law(^ {30}), the Law on Public Health Nurses, Midwives and Nurses(^ {31}), and the Dentist Law(^ {32}), require registered health professionals to protect the confidentiality of patient information.</td>
</tr>
<tr>
<td><strong>Europe, Africa, Middle East</strong></td>
<td>The European Economic Area (EEA) 30 Member States have enacted data protection laws that follow the principles set forth in the 1995 European Union (EU) Data Protection Directive and the 2002 ePrivacy Directive (as amended in 2009). These laws include a security component, and the obligation to provide adequate security must be passed down to subcontractors.</td>
</tr>
<tr>
<td><strong>Other countries that have close ties with the EEA, such as Morocco and Tunisia in Africa, Israel and Dubai in the Middle East have also adopted similar laws that follow the same principles.</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Americas</strong></td>
<td>North, Central, and South American countries are also adopting data protection laws at a rapid pace. Each of these laws includes a security requirement that</td>
</tr>
</tbody>
</table>

---

\(^{27}\) The OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data were adopted on 23 September 1980, see [http://www.oecd.org/document/18/0,3746,en_2649_34255_1815186_1_1_1_1,00.html](http://www.oecd.org/document/18/0,3746,en_2649_34255_1815186_1_1_1_1,00.html).  

\(^{28}\) In 2004, the APEC Privacy Framework was endorsed by APEC Ministers for more details see [http://www.worldlii.org/int/other/PrivLRes/2005/4.html](http://www.worldlii.org/int/other/PrivLRes/2005/4.html).  


places on the data custodian the burden of ensuring the protection and security of personal data wherever the data are located, and especially when transferring to a third party.

- In addition to the data protection laws of Canada\(^{33}\) and Argentina\(^{34}\) which have been in existence for several years, Colombia, Mexico, Uruguay, and Peru have recently passed data protection laws that are inspired mainly from the European model and may include references to the APEC Privacy Framework as well.

**United States**

- There is no single privacy law in the United States. A range of government agency and industry sector laws impose privacy obligations in specific circumstances. There are numerous gaps and overlaps in coverage.

- Current industry sector privacy laws include:
  - The Federal Trade Commission Act\(^ {35} \) which prohibits unfair or deceptive practices - this requirement has been applied to company privacy policies in several prominent cases.
  - The Electronic Communications Privacy Act of 1986\(^ {36} \) which protects consumers against interception of their electronic communication (with numerous exceptions).
  - The Health Insurance Portability and Accountability Act (HIPAA)\(^ {37} \) which contains privacy rules applying to certain categories of health and medical research data.
  - The Fair Credit Reporting Act\(^ {38} \) includes privacy rules for credit reporting and consumer reports.
  - The Gramm-Leach-Bliley Act (GLBA)\(^ {39} \) govern the collection,
disclosure, and protection of consumers’ nonpublic personal information for financial institutions

- These laws hold organizations responsible for the acts of their subcontractors. For example, the security and privacy rules under GLBA or HIPAA require that organizations compel their subcontractors, in written contracts, to use reasonable security measures and comply with data privacy provisions.

  - Government agencies, such as the Federal Trade Commission (FTC) or the State Attorneys General have consistently held organizations liable for the activities of their subcontractors.

  - The Payment Card Industry (PCI) Data Security Standards (DSS)\(^\text{40}\), which apply to credit card data anywhere in the world, including data processed by subcontractors has similar requirements.

When data is transferred to a cloud, the responsibility for protecting and securing the data typically remains with the collector 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 custodian of the data remains liable for any loss, damage, or misuse of the data. It is prudent, and may be legally required, that the data custodian and the cloud 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 SLA, the cloud consumer 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 consumer wishes to send HIPAA-covered information to the cloud, the consumer 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 cloud SLA? What happens if the preservation requirements outlast the terms of the SLA? If the consumer preserves the data in place; who pays for the extended storage and at what cost? Does the consumer have the storage capacity under its SLA? Can the consumer effectively download the data in a forensically sound manner?

\(^{39}\) See [http://www.gpo.gov/fdsys/pkg/PLAW-106publ102/content-detail.html](http://www.gpo.gov/fdsys/pkg/PLAW-106publ102/content-detail.html) for details.

\(^{40}\) PCI DSS provides an actionable framework for developing a robust payment card data security process -- including prevention, detection and appropriate reaction to security incidents. See [https://www.pcisecuritystandards.org/security_standards/](https://www.pcisecuritystandards.org/security_standards/) for details.
so it can preserve it off-line or near-line? These are some of the privacy related questions that need to be addressed in the cloud SLA.

Identity protection is another area that may need to be addressed in the cloud SLA. Almost all cloud-based storage systems require some authentication of participants (cloud consumer and/or cloud provider) to establish trust relations, either for only one endpoint of communication or for both. Although cryptographic certificates can offer sufficient security for many of these purposes, they do not typically cater to privacy because they are bound to the identity of a real person (cloud consumer). Any usage of such a certificate exposes the identity of the holder to the party requesting authentication. There are many scenarios (e.g., storage of electronic health records) where the use of such certificates unnecessarily reveals the identity of their holder.

Over the past 10-15 years, a number of technologies (e.g. Anonymous Credentials) have been developed to build systems in a way that they can be trusted, like normal cryptographic certificates, while at the same time protecting the privacy of their holder (i.e., hiding the real holder’s identity). Such attribute-based credentials are issued just like ordinary cryptographic credentials (e.g., X.509 credentials) using a digital (secret) signature key. However, attribute-based credentials (ABCs) allow their holder to transform them into a new credential that contains only a subset of the attributes contained in the original credential. These transformed credentials can be verified just like ordinary cryptographic credentials (using the public verification key of the issuer) and offer the same strong security.

**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 considerations in service management when entering into a service level 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, and the cloud is no different. Although the standards for SLA 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, organizations and people do what you inspect, not what you expect.

---

The objective of any SLA 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. 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. 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 existing 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), consumers should demand from any cloud SLA 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. **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 SLA includes the ability to see, assess and react to measurements in these areas will help keep your cloud infrastructure running smoothly.

**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 metering system employed by cloud providers, as embodied in the service level agreements you build. At a minimum, you must ensure that metering systems employed for your cloud providers include:

1. Assurance of accurate billing, and a methodology for handing 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.

**Rapid provisioning**

While auditing, monitoring, measuring and metering relate primarily to the cost savings features of the cloud, rapid provisioning is a key underlying quality of the improved flexibility that comes from the cloud. However, it’s not without its own unique qualities that must translate into your service level agreements with providers:

1. *Core provisioning speed.* As part of a cloud SLA, 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 SLA 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.
Resource change
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.

Upgrade to existing services
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 cloud SLA 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.

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, an SLA-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 services 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 service does not occur. Service capabilities and performance expectations should be explicitly documented in the cloud SLA. If not, the likelihood of misunderstandings between consumer and provider increases significantly. For example, a web service performing poorly in terms of response time may not be considered a service failure to the provider unless it is clearly called out in the SLA.
The level of service failure management will vary greatly dependent upon provider, and the ability to negotiate a greater level will vary upon size of the consumer. As a result, it is important for consumers to incorporate their own service failure management capabilities to ensure they are made aware of any issues in a timely fashion.

**Remedies**

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

It is also important to note that common service level agreements put the responsibility of reporting a service interruption on the consumer. The consumer will need to contact the provider and be prepared to show that their service has been impacted by a service interruption.

**Limitations**

Within each cloud provider’s agreement there are likely 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 providers who may also cite unscheduled downtime as being excluded from the SLA metrics. Consumers are strongly encouraged to fully understand all facets of their cloud SLA.

Given these limitations, it is important that consumers plan for unavailability. This may include keeping a separate, on-premise backup copy of the cloud data. While this may be more feasible for a large company, a small company may need to either ensure the provider is backing up data or contract a second provider to complete these backups. The frequency of the backups would be dependent upon the criticality of the data and the data’s rate of change.

**Roles / Responsibilities**

The roles of cloud computing service failure management are similar to the Information Technology Infrastructure Library (ITIL) incident management roles. The consumer incident manager would have responsibility to drive the incident and crisis management process. Assuming the failure is impactful to service delivery, the consumer incident manager should be plugged into the provider’s incident management process as well. This collaboration should be negotiated into the agreed upon SLA. This may also be more easily negotiated for large organizations. Smaller organizations may need to proactively reach out to the provider when an incident occurs and engage in a more manual fashion.
The end goal is to ensure that the consumer is knowledgeable of activity being taken to resolve provider incidents which impact service delivery.

On the consumer side, additional roles include the help desk, and development and engineering support teams. The help desk should be in communication with the incident manager so as to understand impact and estimated time to resolution. This is important as the help desk is likely receiving questions regarding the service availability from their customers. The development and engineering support teams would be engaged to triage and resolve applicable failure scenarios.

**Notification process of a failure (perceived or real)**

Notification of a process failure could be triggered by two paths. The first is the result of a synthetic transaction integrated into the consumer’s monitoring processes. A synthetic transaction is a “fake” transaction that flows through the entire system to test availability and response of all components. This path could alert on individual components throughout the transaction stream. An alert due to the execution failure of the SaaS component could be indicative of a corresponding issue at the provider. It is beneficial to build retry logic into the application to address items such as communication issues that would present themselves as a service call failure.

The capability of a synthetic transaction to identify issues is dependent upon the type of application transaction used (asynchronous or synchronous). It is recommended that calls to invoke a SaaS provided application component are routed through a service bus as this provides standard routing architecture with capabilities such as logging, queuing, retries, etc.

The second notification path would be triggered by the provider’s monitoring system. This alert would, ideally, be integrated into the consumer’s alerting system. This is not prevalent in today’s cloud service delivery, but it is an important point to try to negotiate into a cloud SLA. An alert would be sent to the consumer in the event that the issue causes a service interruption and, potentially, impacts agreed upon SLA terms. Upon receiving notification, the consumer should follow their established incident escalation process.

Problem identification should be a joint activity between the consumer and the service provider. This may be common place in the relationship between a large consumer and provider. For smaller consumers, the provider may provide problem information via an informational website rather than a personal interaction. It is critical that the consumer fully understand the provider’s monitoring and alerting processes. The monitoring and alerting processes should be called out in the agreed upon SLA. In the event that this is non-negotiable, the onus shifts to the consumer to provide monitoring and alerting. This is especially important in the instance of SaaS offerings that provide components to a larger solution.

As monitoring of problem identification should be a joint responsibility, the monitoring and alerting systems at the provider and consumer should have an integration point. This may include a statement in the SLA to send an e-mail or web service call when an incident occurs. This ensures that alerts created at the provider that impact the service delivery of the solution are captured within the consumer’s incident management process.
**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 consumers regarding disaster recovery planning. Just because businesses are outsourcing the infrastructure (IaaS), applications (SaaS), or platforms (PaaS) to the cloud 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 cloud SLAs provide inadequate guarantees in case of a service outage due to a disaster. Most cloud SLAs provide cursory treatment of disaster recovery issues, procedures and processes. That being said, it is rare for SMBs – the primary customer of clouds today, to internally develop the extensive disaster recovery infrastructure of large and established cloud providers.

Despite the limitations in cloud SLAs, 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 a 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 SLA. 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 SLAs.

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 consumers 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 vendor. 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 all the complexity, it also increases the difficulty of performing comprehensive due diligence. Lack of such diligence accompanied by weak SLAs represents a potential risk in the area of business continuity and disaster recovery. Thus, companies should view developing 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.

The cloud is growing at a rapid pace and cloud providers are facing a learning curve. In response to recent outages, providers are broadening the service options available to consumers. Recent announcements on increased interoperability among different providers for certain cloud services will be beneficial to customers. It also is an acknowledgement that industry players must work together in order to make the cloud more reliable for consumers.

**Step 9: Develop an Effective Management Process**
In the evolving world of cloud computing, there is a need for an effective management process for any problems that may arise. Today’s reality is that cloud SLAs contain very limited information on consumer-provider management processes except possibly for large enterprises that are capable of negotiating unique terms. Implementing an effective management process is an important step to ensuring internal and external user satisfaction with cloud based service(s).

Table 8 below highlights the key elements of a successful management process.

**Table 8. Management process**

<table>
<thead>
<tr>
<th>Element</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Establish monthly status</td>
<td>Establishes proactive management to quickly resolve SLA compliance issues.</td>
</tr>
<tr>
<td>meetings</td>
<td>Reduces the possibility of surprises and dissatisfaction with the service(s).</td>
</tr>
<tr>
<td></td>
<td>Led by cloud provider and attended by stakeholders from both the cloud consumer and the cloud provider.</td>
</tr>
</tbody>
</table>

- Meeting minutes should be documented.

**Ensure proper attendance**

It is important to have the right people in the status meetings. Both the cloud consumer and the cloud provider need to be represented and each party must have a stake in the overall success of the service(s). Examples of the attendees in the status meetings are as follows:

- **Cloud Consumer**
  - Customer Support management representative.
  - Technical Support management representative.

- **Cloud Provider**
  - Technical Support management representative.

**Topics of discussion**

While the primary purpose of the status meeting is to focus on ensuring all problems are resolved within the agreed to criteria, there are some additional items that can be discussed with the purpose of ensuring both parties are aligned with the continuing management of the service(s):

- Potential changes in the support processes or key personnel changes.
- Plans for enhancements to the existing services.
- Plans for adding new services to the existing environment.
- Other topics which may impact the level of service being provided or expected.

**Track key indicators**

Four key indicators should be tracked to ensure that the SLA 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.

**Produce reports**

In order to ensure SLA compliance, a set of reports, generated by the cloud provider, needs to be available to all parties.

- Reports that focus on the current reporting period addressing:
  - All problems reported (sorted by impact).
Problems closed (sorted by impact).

Duration of open problems (sorted by impact).

- A second set of reports which summarize the year to date (YTD) activity, by month and by impact, to identify trends.

It is the responsibility of the cloud provider to lead the status meetings and to provide a clear picture of the status of services, using the provided reports. To ensure that there is no confusion with agreed to actions and target dates, a set of minutes should be recorded. It is also very important that the person(s) responsible for the items be documented. Having a complete set of minutes from each status meeting accomplishes two objectives. Firstly, it eliminates any confusion in what and who is responsible for an action on a specific problem. Secondly, if there is ever a dispute over either a specific problem or the overall performance of the service, the minutes will serve as the record to resolve the dispute.

For problems which require management awareness, it is the responsibility of the attending stakeholders at the meeting 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 can not wait for the monthly status 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 9 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 9. Escalation considerations**

<table>
<thead>
<tr>
<th>Consideration</th>
<th>Description</th>
</tr>
</thead>
</table>
| **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.

**Types**

- **Immediate**
  - A critical business impact is identified.
  - Significant impact to a customer facing service.

- **As required.** Typically after the monthly status meeting 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 can not be reached through the escalation process then the terms of the SLA can be brought to bear to force resolution. Bringing the SLA terms into the discussion should be a last resort and only invoked should there be continual non-compliance of the SLA. One of the outcomes of continuous breaches to the SLA 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 will be 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 cloud SLA 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 consumer and producer 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 consumer data and applications is essential.

A consumer exit plan should always be prepared at the outset of the SLA 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 producer is effectively implementing these procedures.

By far, the most important aspect of any exit plan is the transmission and preservation of consumer data which is critical to achieving business continuity. In addition, consumers must ensure that their data is completely removed from the provider’s environment once the exit process is complete. Consumers should look out for and beware of the following details when they evaluate the exit clause included in a cloud SLA.

- The level of provider assistance in the exit process and associated fees should be clear in the SLA. In most cases, there should be no additional cost associated with the exit process.
- Providers should be responsible for extracting consumer data from their IT environments, or at least aid the consumer in extracting their data by providing clear and concise documentation.
- The format of the data transmitted from the provider to the consumer should be specified in the cloud SLA and should leverage standard data formats whenever possible to ease and enhance portability.
- The SLA should specify that all data and information belonging to the consumer be maintained for a specific time period after transition and then be completely removed immediately after.
  - The typical time period is 1-3 months which gives the consumer sufficient time to find a new provider and to continue receiving service from the current provider in the interim.
  - The time period should be explicitly documented in the cloud SLA and only with the consumer’s written notice should data be removed and destroyed before that time.
- Consumers should ensure that the SLA provides appropriate business continuity protection during the exit process.
- At the completion of the exit process, consumers should receive written confirmation from the provider that all of the consumer’s 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 consumer’s data for any reason in the future, including using the data for statistical purposes.

The bottom line is that consumers should undertake due diligence when evaluating and ultimately selecting a cloud provider. A trustworthy cloud provider should be prepared to work with consumers on a fair and effective exit strategy.
### Summary of Keys to Success

Table 10 summarizes the critical keys to success for any organization evaluating and comparing SLAs from different cloud providers.

#### Table 10. Summary of keys to success

<table>
<thead>
<tr>
<th>Key to Success</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Develop a strong business case and strategy for cloud computing environment</strong></td>
<td>• Assess criticalness of services being deployed in the cloud.</td>
</tr>
<tr>
<td></td>
<td>• Determine functional and non-functional requirements for each service (performance, availability, security, privacy, etc.).</td>
</tr>
<tr>
<td></td>
<td>• Understand legal and regulatory requirements concerning the data maintained in the cloud.</td>
</tr>
<tr>
<td></td>
<td>• Identify key performance metrics for each service.</td>
</tr>
<tr>
<td><strong>Assess provider’s SLA against functional and non-functional requirements</strong></td>
<td>• Based on the criticalness of the service being deployed in the cloud, determine if the cloud provider’s SLA is sufficient to address the functional, non-functional, legal, and regulatory requirements of the service.</td>
</tr>
<tr>
<td></td>
<td>• If not, determine if the cloud provider is willing to negotiate on the key aspects of the SLA that are not in line with your business strategy.</td>
</tr>
<tr>
<td></td>
<td>• If the cloud provider is not willing to negotiate on these critical points, seek alternative providers who more closely address your requirements.</td>
</tr>
<tr>
<td></td>
<td>• If a cloud provider who addresses your requirements cannot be found, strongly consider keeping the service within your enterprise IT environment.</td>
</tr>
<tr>
<td><strong>Determine how to monitor SLA performance</strong></td>
<td>• Assuming a cloud provider is found that meets your service requirements; understand the management process defined in the SLA.</td>
</tr>
<tr>
<td></td>
<td>• Ensure your SLA includes the ability to see, assess and react to key performance measurements that will help keep your cloud infrastructure running smoothly.</td>
</tr>
<tr>
<td></td>
<td>• Understand the notification process when service issues arise including method and timeliness of notifications along with prioritization and severity level assessment of issues.</td>
</tr>
<tr>
<td></td>
<td>• Be aware of remedies and liability limitations offered by the cloud provider when service issues arise.</td>
</tr>
<tr>
<td><strong>Ensure an adequate disaster recovery plan can be defined and executed</strong></td>
<td>• The cloud consumer bears the risk of disaster scenarios that severely limit the ability of their cloud provider to deliver service.</td>
</tr>
<tr>
<td></td>
<td>• Cloud consumers must understand the provider’s ability to support their data preservation strategy which includes criticalness of data, data sources,</td>
</tr>
</tbody>
</table>
scheduling, backup, restore, integrity checks, etc.

- Roles and responsibilities must be clearly documented in the SLA. In many cases, the cloud consumer may be responsible for implementing most of the data preservation strategy.

- Based on the criticalness of the data, cloud consumers should clearly define recovery time objectives.

- Consumers should test and verify the disaster recovery plan prior to production deployment.

- Cloud consumers 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 consumer should the relationship with the cloud provider terminate prematurely.

- The exit plan should be taken into account during the assessment phase of potential cloud providers.

- The provider’s SLA should be carefully reviewed to ensure the consumer defined exit plan is capable of being implemented.
  
  - The consumer 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 SLA. In many cases, the cloud consumer 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 consumers 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 consumers to effectively track SLA performance.
• Standardized security and regulatory compliance requirements to identify control points for risk management.

• Standards that enable coordinated end-to-end SLA management for both cloud consumers 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, consumers are able to effectively review and compare SLAs from different cloud providers to ensure the promise of the cloud is realized.

Works Cited


This guide is to provide a practical reference to help enterprise information technology (IT) and business decision makers adopt cloud computing to solve business challenges.

2010 International Conference on High Performance Computing and Simulation
Low level Metrics to High level SLAs - LoM2HiS framework: Bridging the gap between monitored metrics and SLA parameters in cloud environments. Emeakaroha, Vincent C.; Brandic, Ivona; Maurer, Michael; Dustdar, Schahram


This white paper defines cloud computing, the five essential characteristics, three service models, and four deployment models.

Queen Mary School of Law Legal Studies Research Paper No. 63/2010.

This document provides a comprehensive comparison of terms and conditions from the leading cloud service providers.


This document provides an actionable, practical road map to managers wanting to adopt the cloud paradigm safely and securely.

Additional References

CSMIC SMI – Service Measurement Index, Carnegie Mellon University, 2011

International Function Point Users Group, Software Non-Functional Assessment Process Release 1.0, International Function Point Users Group, September 2011


