Guidelines for Government Departments On Service Level Agreement For Procuring Cloud Services

Ministry of Electronics and Information Technology (MeitY)
Electronics Niketan, 6,
CGO Complex New Delhi-110 003

March 31, 2017

This page is intentionally left blank

Table of Contents

This page is intentionally left blank

1 Background

MeitY has announced MeghRaj Policy to provide strategic direction for adoption of cloud services by the Government (http://meity.gov.in/content/gi-cloud-initiative-meghraj). The aim of the cloud policy is to realize a comprehensive vision of a government cloud (GI Cloud) environment available for use by central and state government line departments, districts and municipalities to accelerate their ICT-enabled service improvements. MeghRaj policy of MeitY states that "Government departments at the Centre and States to first evaluate the option of using the GI Cloud for implementation of all new projects funded by the government. Existing applications, services and projects may be evaluated to assess whether they should migrate to the GI Cloud."

1.1 Provisional Empanelment of Cloud Service Offerings

Taking demand into consideration, MeitY has initiated Provisional Empanelment of the cloud service offerings of Service providers that the end-user departments can leverage in addition to the National Cloud services offered by NIC for their e-governance solutions.

The following cloud service offerings offered by the Cloud Service Providers (http://meity.gov.in) for a combination of the Deployment Models (Public Cloud, Virtual Private Cloud and Government Community Cloud) have been provisionally empanelled by MeitY.

1. Infrastructure as a Service (IaaS)
2. Platform as a Service (PaaS)
3. Disaster Recovery as a Service (DRaaS)
4. Dev / Test Environment as a Service (DevOps)
5. Virtual Desktops as a Service (VDaaS)

The provisional empanelment shall be initially for two years from the date of accepting the terms and conditions by the empanelled cloud service providers. The provisionally empanelled cloud service providers will have the option to comply with the full-fledged guidelines / standards to get the certification for their offerings as and when the new guidelines / standards are published by MeitY.

STQC is in the process of auditing and certification of the empanelled services. MeitY has opened up the process of empanelment and henceforth the potential Cloud Service Providers can apply for empanelment and certification by MeitY. The audited/certified service offerings of various CSPs are available at www://meity.gov.in

1.2 Guidelines for Procurement of Cloud Services

To further facilitate the end user departments in procuring/ adopting cloud computing services, MeitY has prepared broad guidelines highlighting key considerations that Government Departments need to be aware of when procuring cloud services. The guidelines also detail out the services and requirements that can be used when preparing the Request for Proposal (RFP) document.

In addition to the above, as Cloud Service Level Agreements (Cloud SLAs) form an important component of the contractual relationship between a cloud service customer and a cloud service provider of a cloud service, following guidelines are provided for Departments to identify critical Service Levels for cloud and standardize the SLA terminologies across CSPs.

2 Service Level Agreement Guidelines for Procuring Cloud Services

It is best if the service level objectives offered by each cloud service provider for similar services can be easily compared. But SLA terminology, measurement methodology, metrics often differs from one cloud service provider to another, making it difficult for Government Departments/ Agencies to compare cloud service offerings and the Service Levels. In order to facilitate Government Departments, the key service level objectives are identified, described the SLAs terminologies, indicated the measurement methodology to be adopted for measuring the services, defined the metrics/target levels and penalties to be levied are indicated in case of non-performance.

However, the set of Service Level Objectives (SLOs) defined below is not exhaustive and other additional SLOs may be required specific projects. Also certain requirements such as Availability of Patches, Timely Cloud Service Change Notifications etc. are not mentioned in the SLAs, but CSPs are required to comply with the requirements mentioned in the empanelment RFP.

The key service level objectives that relate to the cloud service and the related aspects of the interface between the department and the cloud service provider are indicated below:

2.1 Availability

Availability is a key service level objective, since it describes whether the cloud service can actually be used, and it is typically necessary to specify numeric values for availability for Government Department/ Agency to identify which services/resources are to be monitored under availability and indicate the measure of availability. The indicative parameters for designing Availability linked SLAs are enlisted below:

a. Availability of Provisioned Resources for Pre-Production/Production environment - could include VMs, Storage, OS, VLB, Security Components..,

b. Availability of Critical Services- Some of the services such as Provisioning / De-Provisioning; User Activation / De-Activation; User Profile Management; Register Support Request or Incident; Access Utilization Monitoring Reports are critical for cloud services and hence the service levels needs to be monitored.

2.2 Support Channels - Incident and Helpdesk

Support is an interface made available by the cloud service provider to handle issues and queries raised by the Government Department/ Agency. The following are key parameters to be factored when gauging the support requirements from CSP.

a. Support hours: The hours during which the cloud service provider provides a support interface that accepts general inquiries and requests from the Government Department/ Agency.

b. Support responsiveness: The maximum time the cloud service provider will take to acknowledge an inquiry or request raised by the Department. It is typical for responsiveness to vary depending on a severity level which is attached to the customer request, with a shorter response time associated with higher severity levels. (Indicative Severity Levels enclosed in Annexure 3)

c. Resolution time: The time taken to complete any necessary actions as a result of the request.

2.3 Response Time

Response time is the time interval between an initiated event (stimulus) by the Department/ Agency and a cloud service provider initiated event in response to that stimulus. Response time can be a highly significant aspect of the user experience of a cloud service – for some critical requests such as (Service with significant number of end users or public users), response times that are greater than some threshold are regarded as unacceptable and can make the cloud service effectively unusable.

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

2.4 Performance

Service Level Objectives for Performance deal with the actual mechanisms used to guarantee that the Department's data is available (online or offline) in case of failures forbidding access to it. Proposed SLOs allow customers e.g., to fine-tune their risk assessment and business continuity procedures.

a. Latency: Latency may address the storage and the time when the data is placed on mirrored storage.

b. Maximum Data Restoration Time refers to the committed time taken to restore cloud service customer data from a backup. Concrete details related with to the frequency and method used by the cloud service provider's backup and recovery mechanism(s).

2.5 Security Incident and Management Reporting

Specifying measurable security level objectives in SLAs is useful to improve both assurance and transparency.

a. Percentage of timely incident reports: Describes the defined incidents to the cloud service which are reported to the Department in a timely fashion. This is represented as a percentage by the number of defined incidents reported within a predefined time limit after discovery, over the total number of defined incidents to the cloud service which are reported within a predefined period (i.e. month, week, year, etc.).

b. Percentage of timely incident responses: Describes the defined incidents that are assessed and acknowledged by the cloud service provider in a timely fashion. This is represented as a percentage by the number of defined incidents assessed and acknowledged by the cloud service provider within a predefined time limit after discovery, over the total number of defined incidents to the cloud service within a predefined period. (i.e. month, week, year, etc.).

c. Percentage of timely incident resolutions: describes the percentage of defined incidents against the cloud service that are resolved within a predefined time limit after discovery.

2.6 Vulnerability Management

Vulnerability Management refers to managing a weakness of an asset or group of assets, e.g. software or hardware related, that can be exploited by one or more threats. The set of service level objectives is not exhaustive, and not all the service level objectives are

applicable to all cloud services. The Information about technical vulnerabilities should be obtained in a timely fashion, the organization's exposure to such vulnerabilities evaluated and appropriate measures taken to address the associated risk.

a. Percentage of timely vulnerability corrections describes the number of vulnerability corrections performed by the cloud service provider, and is represented as a percentage by the number of vulnerability corrections performed within a predefined time limit, over the total number of vulnerability corrections to the cloud service which are reported within a predefined period (i.e. month, week, year, etc.).

b. Percentage of timely vulnerability reports describes the number of vulnerability reports by the cloud service provider to the Department, and is represented as a percentage by the number of vulnerability reports within a predefined time limit, over the total number of vulnerability reports to the cloud service which are reported within a predefined period (i.e. month, week, year, etc.).

c. Reports of vulnerability corrections is a description of the mechanism by which the cloud service provider informs the Department of vulnerability corrections applied to the provider's systems, including the frequency of the reports.

Apart from the above, some key measurements required for monitoring are also indicated in Annexure 2

2.7 Indicative SLOs for MSP/SI:

The following SLOs are pertinent if the Department engages a Managed Service Provider or a System Integrator for managing the cloud services.

a. Recovery Point Objective is the maximum allowable time between recovery points. RPO does not specify the amount of acceptable data loss, only the acceptable time window. In particular, RPO affects data redundancy and backup. A small RPO suggests mirrored storage of both transient and persistent data while a larger window allows for a periodic backup approach.

b. Recovery Time Objective is the maximum amount of time a business process may be disrupted, after a disaster, without suffering unacceptable business consequences. Cloud services can be critical components of business processes.

c. Availability of Reports (Reports such as Provisioning, Utilization Monitoring Reports, User Profile Management etc.)

d. Request – Response: Percentage of successful requests

e. Availability of the network links at DC and DR (links at DC / DRC, DC-DRC link)

The Definitions of the Key terms used in the Section are provided in Annexure 1. The indicative Penalties against breach in Service Levels are mentioned in Annexure 2 and the severity levels for the SLAs are defined in Annexure 3.

3 Measurement and Monitoring

The following clauses related to measurement and monitoring may be included in the RFP:

a. The SLA parameters shall be monitored on a <<periodic basis/ monthly basis>> as per the individual SLA parameter requirements. However, if the performance of the system/services is degraded significantly at any given point in time during the contract and if the immediate measures are not implemented and issues are not rectified to the complete satisfaction of Department or an agency designated by them, then the Department will have the right to take appropriate disciplinary actions including termination of the contract.

b. The full set of service level reports should be available to the Department on a monthly basis or based on the project requirements.

c. The Monitoring Tools shall play a critical role in monitoring the SLA compliance and hence will have to be customized accordingly. The CSP shall make available the Monitoring tools for measuring and monitoring the SLAs. The MSP may deploy additional tools and develop additional scripts (if required) for capturing the required data for SLA report generation in automated way. The tools should generate the SLA Monitoring report in the end of every month which is to be shared with the Department on a monthly basis. The Department or its nominated agency shall have full access to the Monitoring Tools/portal (and any other tools / solutions deployed for SLA measurement and monitoring) to extract data (raw, intermediate as well as reports) as required during the project. The Department or its nominated agency will also audit the tool and the scripts on a regular basis.

d. The measurement methodology / criteria / logic will be reviewed by the Department.

e. In case of default on any of the service level metric, the SP shall submit performance improvement plan along with the root cause analysis for the Department 's approval.

4 Periodic Reviews

The SLAs might require to be modified based on the project needs. Therefore Departments need to ensure that the relevant clauses are included in the Agreement that would allow the Departments to modify the SLAs. The following clauses are provided as guidance to the departments while preparing the Service Level Agreement.

a. During the contract period, it is envisaged that there could be changes to the SLA, in terms of measurement methodology / logic / criteria, addition, alteration or deletion of certain parameters, based on mutual consent of both the parties, i.e. the Department and SP.

b. The Department and SP shall each ensure that the range of the Services under the SLA shall not be varied, reduced or increased except by the prior written agreement of the Department and SP in accordance with the Change Control Schedule.

c. The SLAs may be reviewed on an annual basis by the Department in consultation with the SP and other agencies.

5 Penalties

For the Departments to ensure that the Cloud Service Providers adhere to the Service Level Agreements, this section describes the Penalties which may be imposed on CSPs. In case these service levels cannot be achieved at service levels defined in the agreement, the departments should invoke the performance related penalties. Payments to the SP to be linked to the compliance with the SLA metrics laid down in the agreement. To illustrate calculation of penalties, an indicative example is provided below.

a. The payment should be linked to the compliance with the SLA metrics.

b. The penalty in percentage of the <<Periodic Payment>>) is indicated against each SLA parameter in the table.

    I. For ex: For SLA1 if the penalty to be levied is 7% then 7% of the <<Periodic Payment (monthly)>>) is deducted from the total of the <<periodic/monthly> bill and the balance paid to the SP.

II. If the penalties are to be levied in more than one SLA then the total applicable penalties are calculated and deducted from the total of the <<periodic/monthly> bill and the balance paid to the SP.

For ex: SLA1 =7% of the <<Periodic Payment (monthly)>>, SLA12=10% of the <<Periodic Payment (monthly)>>, SLA19=2% of the <<Periodic Payment (monthly)>> then

Amount to be paid = Total <<periodic/monthly> bill – {(19% of the <<Periodic Payment (monthly)>>)}

c. In case multiple SLA violations occur due to the same root cause or incident then the SLA that incurs the maximum penalty may be considered for penalty calculation rather than a sum of penalties for the applicable SLA violations

d. Penalties shall not exceed 100% of the <<periodic/monthly>> bill. If the penalties exceed more than 50% of the total <<periodic/monthly> bill, it will result in a material breach. In case of a material breach, the operator will be given a cure period of one month to rectify the breach failing which a notice to terminate may be issued by the Department

Annexure 1: Definitions

i. **Cloud "Service Level Objective" (SLO)** means the target for a given attribute of a cloud service that can be expressed quantitatively or qualitatively.

ii. **Cloud SLAs** means documented agreement between the cloud service provider and the Department that identifies services and cloud service level objectives (SLOs).

iii. **Response time** is the time interval between a cloud service customer initiated event (e.g., logging of the request) and a cloud service provider initiated event in response to that stimulus.

iv. "**Scheduled Maintenance Time**" shall mean the time that the System is not in service due to a scheduled activity. Scheduled maintenance time is planned downtime with the prior permission of the Department, during non-business hours. The Scheduled Maintenance time <<within 10 hours a month>> as agreed shall not be considered for SLA Calculation.

v. "**Scheduled operation time**" means the scheduled operating hours of the System for the month. All scheduled maintenance time on the system would be deducted from the total operation time for the month to give the scheduled operation time.

vi. **"Availability"** means the time for which the cloud services and facilities are available for conducting operations on the the Department system.

Availability is defined as:

{(Scheduled Operation Time – System Downtime) / (Scheduled Operation Time)} * 100%

vii. "**Incident**" refers to any event/issue that affects the normal functioning of the services / infrastructure, reported by the cloud consumer to the Cloud Service provider (CSP) can be termed as an Incident.

This page is intentionally left blank

Annexure 2: Service Levels

S. No.	Service Level Objective	Measurement Methodology /	Target/Service Level	Penalty (Indicative)
Service Levels for CSP				
Availability/Uptime				
1.	Availability/Uptime of cloud services Resources for Production environment (VMs, Storage, OS, VLB, Security Components,)	Availability (as per the definition in the SLA) will be measured for each of the underlying components (e.g., VM, Storage, OS, VLB, Security Components) provisioned in the cloud. Measured with the help of SLA reports provided by CSP	Availability for each of the provisioned resources: >=99.5%	Default on any one or more of the provisioned resource will attract penalty as indicated below. <99.5% & >=99% (10% of the <<Periodic Payment>>) < 99% (30% of the <<Periodic Payment>>)
2.	Availability of Critical Services (e.g., Register Support Request or Incident; Provisioning / De-Provisioning; User Activation / De-Activation; User Profile Management; Access Utilization Monitoring Reports)	Availability (as per the definition in the SLA) will be measured for each of the critical services over both the User / Admin Portal and APIs (where applicable)	Availability for each of the critical services over both the User / Admin Portal and APIs (where applicable) >= 99.5%	Default on any one or more of the services on either of the portal or APIs will attract penalty as indicated below. <99.5% and >= 99% (10% of the <<Periodic Payment>>) <99% (20% of the <<Periodic Payment>>)

S. No.	Service Level Objective	Measurement Methodology /	Target/Service Level	Penalty (Indicative)
	over User / Admin Portal and APIs (where applicable)			
3.	Availability of the network links at DC and DR (links at DC / DRC, DC-DRC link)	Availability (as per the definition in the SLA) will be measured for each of the network links provisioned in the cloud.	Availability for each of the network links: >= 99.5%	Default on any one or more of the provisioned network links will attract penalty as indicated below. <99.5% & >=99% (10% of the <<Periodic Payment>>) < 99% (30% of the <<Periodic Payment>>)
4.	Availability of Regular Reports (e.g., Audit, Certifications,) indicating the compliance to the Provisional Empanelment Requirements.		15 working days from the end of the quarter. If STQC issues a certificate based on the audit then this SLA is not required.	5% of <<periodic Payment>>
The following SLAs apply both for CSP and MSP/SI. While the CSP will be responsible for maintaining the SLAs pertaining to the cloud infrastructure, network, controls etc., the MSP will be responsible for the SLAs related to managing and monitoring the cloud services.				
Support Channels - Incident and Helpdesk				

S. No.	Service Level Objective	Measurement Methodology /	Target/Service Level	Penalty (Indicative)
5.	Response Time	Average Time taken to acknowledge and respond, once a ticket/incident is logged through one of the agreed channels. This is calculated for all tickets/incidents reported within the reporting month.	95% within 15minutes	<95% & >=90% (5% of the <<Periodic Payment>>) < 90% & >= 85% (7% of the <<Periodic Payment>>) < 85% & >= 80% (9% of the <<Periodic Payment>>)
6.	Time to Resolve - Severity 1	Time taken to resolve the reported ticket/incident from the time of logging.	For Severity 1, 98% of the incidents should be resolved within 30 minutes of problem reporting	<98% & >=90% (5% of the <<Periodic Payment>>) < 90% & >= 85% (10% of the <<Periodic Payment>>) < 85% & >= 80% (20% of the <<Periodic Payment>>)
7.	Time to Resolve - Severity 2,3	Time taken to resolve the reported ticket/incident from the time of logging.	95% of Severity 2 within 4 hours of problem reporting AND 95% of Severity 3 within 16 hours of problem reporting	<95% & >=90% (2% of the <<Periodic Payment>>) < 90% & >= 85% (4% of the <<Periodic Payment>>) < 85% & >= 80% (6% of the <<Periodic Payment>>)
Security Incident and Management Reporting				

S. No.	Service Level Objective	Measurement Methodology /	Target/Service Level	Penalty (Indicative)
8.	Percentage of timely incident report	Measured as a percentage by the number of defined incidents reported within a predefined time (1 hour) limit after discovery, over the total number of defined incidents to the cloud service which are reported within a predefined period (i.e. month). Incident Response - CSP shall assess and acknowledge the defined incidents within 1 hour after discovery.	95% within 1 hour	<95% & >=90% (5% of the <<Periodic Payment>>) < 90% & >= 85% (10% of the <<Periodic Payment>>) < 85% & >= 80% (15% of the <<Periodic Payment>>)
9.	Percentage of timely incident resolutions	Measured as a percentage of defined incidents against the cloud service that are resolved within a predefined time limit (month) over the total number of defined incidents to the cloud service within a predefined period. (Month). Measured from Incident Reports	95% to be resolved within 1 hour	<95% & >=90% (5% of the <<Periodic Payment>>) < 90% & >= 85% (10% of the <<Periodic Payment>>) < 85% & >= 80% (15% of the <<Periodic Payment>>)
Vulnerability Management				
10.	Percentage of timely vulnerability corrections	The number of vulnerability corrections performed by the cloud service provider - Measured as a percentage by the number of vulnerability corrections performed	99.95%	>=99% to <99.95% [10% of Periodic Payment] >=98% to <99% [20% of Periodic Payment] <98% [30% of Periodic

S. No.	Service Level Objective	Measurement Methodology /	Target/Service Level	Penalty (Indicative)
		within a predefined time limit, over the total number of vulnerability corrections to the cloud service which are reported within a predefined period (i.e. month, week, year, etc.). • High Severity Vulnerabilities – 30 days - Maintain 99.95% service level • Medium Severity Vulnerabilities – 90 days - Maintain 99.95% service level		Payment]
11.	Percentage of timely vulnerability reports	Measured as a percentage by the number of vulnerability reports within a predefined time limit, over the total number of vulnerability reports to the cloud service which are reported within a predefined period (i.e. month, week, year, etc.).	99.95%	>=99% to <99.95% [10% of Periodic Payment] >=98% to <99% [20% of Periodic Payment] <98% [30% of Periodic Payment]
12.	Security breach including Data Theft/Loss/Corruption	Any incident where in system compromised or any case wherein data theft occurs (including internal incidents)	No breach	For each breach/data theft, penalty will be levied as per following criteria. Any security incident detected INR << 5 Lakhs>>.This penalty is applicable per incident. These penalties will not be part of

S. No.	Service Level Objective	Measurement Methodology /	Target/Service Level	Penalty (Indicative)
				overall SLA penalties cap per month. In case of serious breach of security wherein the data is stolen or corrupted, << Government Department / Agency>> reserves the right to terminate the contract.
13.	Availability of SLA reports covering all parameters required for SLA monitoring within the defined time		(e.g., 3 working days from the end of the month)	5% of <<periodic Payment>>
Service levels for MSP/SI				
14.	Recovery Time Objective (RTO) (Applicable when taking Disaster Recovery as a Service from the Service Provider)	Measured during the regular planned or unplanned (outage) changeover from DC to DR or vice versa.	<<RTO <= 4 hours>> [Government Department / Agency to indicate based on the application requirements]	10% of <<Periodic Payment>> per every additional 4 (four) hours of downtime
15.	RPO (Applicable when taking Disaster Recovery as a Service from the Service Provider)	Measured during the regular planned or unplanned (outage) changeover from DC to DR or vice versa.	<= 2 hours [Government Department / Agency to indicate based on the application requirements]	10% of <<Periodic Payment>> per every additional 2 (two) hours of downtime

S. No.	Service Level Objective	Measurement Methodology /	Target/Service Level	Penalty (Indicative)
16.	Availability of Root Cause Analysis (RCA) reports for Severity 1 & 2		Average within 5 Working days	5% of <<periodic Payment>>

Note:

1. Periodic Payment means Monthly Payment

This page is intentionally left blank

Annexure 3: Severity Levels

Below severity definition provide indicative scenarios for defining incidents severity. However <<Government Department/Agency>> will define / change severity at the time of the incident or any time before the closure of the ticket based on the business and compliance impacts.

Severity Level	Description	Examples
Severity 1	Environment is down or major malfunction resulting in an inoperative condition or disrupts critical business functions and requires immediate attention. A significant number of end users (includes public users) are unable to reasonably perform their normal activities as essential functions and critical programs are either not working or are not available	• Non-availability of VM. • No access to Storage, software or application
Severity 2	Loss of performance resulting in users (includes public users) being unable to perform their normal activities as essential functions and critical programs are partially available or severely restricted. Inconvenient workaround or no workaround exists. The environment is usable but severely limited.	• Intermittent network connectivity
Severity 3	Moderate loss of performance resulting in multiple users (includes public users) impacted in their normal functions.	