

# Service Level Agreement for Microsoft Azure operated by 21Vianet

Last updated: Feb 2017

## 1. Introduction

This Service Level Agreement for Azure (this “SLA” ) is made by 21Vianet in connection with, and is a part of, the agreement under which Customer has purchased Azure Services from 21Vianet (the “Agreement” ). This SLA applies to the following Azure Services:

- Application Gateway
- Automation
- Azure IoT Hub
- Backup
- CDN
- Cloud Services
- DocumentDB
- Event Hubs
- ExpressRoute
- HDInsight
- Media Services
- MySQL Database on Azure
- Multi-Factor Authentication
- Notification Hubs
- Power BI Embedded
- Redis Cache
- Stream Analytics
- Scheduler
- Service Bus
- Site Recovery
- SQL Database
- SQL Data Warehouse
- Storage
- StorSimple
- Service Fabric
- SQL Server Stretch Database
- Traffic Manager
- Key Vault
- Virtual Machines
- VPN Gateway
- App Service
- Virtual Machine Scale Sets

We provide financial backing to our commitment to achieve and maintain Service Levels for our Services. If we do not achieve and maintain the Service Levels for each Service as described in this SLA, then you may be eligible for a credit towards a portion of your monthly service fees. These terms will be fixed for term of your Agreement. If a subscription is renewed, the version of this SLA that is current at the time the renewal term commences will apply throughout the renewal term. We will provide at least 90

days' notice for adverse material changes to this SLA. You can review the most current version of this SLA at any time by visiting <http://www.windowsazure.cn/support/legal/sla>.

## **2. General Terms**

### **a. Definitions**

- i. "Claim" means a claim submitted by Customer to 21Vianet pursuant to this SLA that a Service Level has not been met and that a Service Credit may be due to Customer.
- ii. "Customer" refers to the organization that has entered into the Agreement.
- iii. "Customer Support" means the services by which 21Vianet may provide assistance to Customer to resolve issues with the Services.
- iv. "Error Code" means an indication that an operation has failed, such as an HTTP status code in the 5xx range.
- v. "External Connectivity" is bi-directional network traffic over supported protocols such as HTTP and HTTPS that can be sent and received from a public IP address.
- vi. "Incident" means any set of circumstances resulting in a failure to meet a Service Level.
- vii. "Management Portal" means the web interface, provided by 21Vianet, through which customers may manage the Service.
- viii. "21Vianet" means the 21Vianet entity that appears on Customer' s Agreement.
- ix. "Preview" refers to a preview, beta, or other pre-release version of a service or software offered to obtain customer feedback.
- x. "Service" or "Services" refers to a Azure service provided to Customer pursuant to the Agreement for which an SLA is provided below.
- xi. "Service Credit" is the percentage of the monthly service fees for the affected Service or Service Resource that is credited to Customer for a validated Claim.

- xii. "Service Level" means standards 21Vianet chooses to adhere to and by which it measures the level of service it provides for each Service as specifically set forth below.
- xiii. "Service Resource" means an individual resource available for use within a Service.
- xiv. "Success Code" means an indication that an operation has succeeded, such as an HTTP status code in the 2xx range.
- xv. "Support Window" refers to the period of time during which a Service feature or compatibility with a separate product or service is supported.
- xvi. "Virtual Network" refers to a virtual private network that includes a collection of user-defined IP addresses and subnets that form a network boundary within Azure.
- xvii. "Virtual Network Gateway" refers to a gateway that facilitates cross-premises connectivity between a Virtual Network and a customer on-premises network.

**b. Service Credit Claims**

- i. In order for 21Vianet to consider a Claim, Customer must submit the Claim to Customer Support within two months of the end of the billing month in which the Incident that is the subject of the Claim occurs. Customer must provide to Customer Support all information necessary for 21Vianet to validate the Claim, including but not limited to detailed descriptions of the Incident, the time and duration of the Incident, the affected resources or operations, and any attempts made by Customer to resolve the Incident
- ii. 21Vianet will use all information reasonably available to it to validate the Claim and to determine whether any Service Credits are due.
- iii. In the event that more than one Service Level for a particular Service is not met because of the same Incident, Customer must choose only one Service Level under which a Claim may be made based on the Incident.
- iv. Service Credits apply only to fees paid for the particular Service, Service Resource, or Service tier for which a Service Level has not been met. In cases where Service Levels apply to individual Service Resources or to separate

Service tiers, Service Credits apply only to fees paid for the affected Service Resource or Service tier, as applicable.

- c. SLA Exclusions.** This SLA and any applicable Service Levels do not apply to any performance or availability issues:
- i. Due to factors outside 21Vianet' s reasonable control (for example, a network or device failure external to 21Vianet' s data centers, including at Customer' s site or between Customer' s site and 21Vianet' s data center);
  - ii. That resulted from Customer' s use of hardware, software, or services not provided by 21Vianet as part of the Services (for example, third-party software or services purchased from the Azure Store or other non-Azure services provided by 21Vianet);
  - iii. Due to Customer' s use of the Service in a manner inconsistent with the features and functionality of the Service (for example, attempts to perform operations that are not supported) or inconsistent with published documentation or guidance;
  - iv. That resulted from faulty input, instructions, or arguments (for example, requests to access files that do not exist);
  - v. Caused by Customer' s use of the Service after 21Vianet advised Customer to modify its use of the Service, if Customer did not modify its use as advised;
  - vi. During or with respect to Previews or to purchases made using 21Vianet subscription credits;
  - vii. That resulted from Customer' s attempts to perform operations that exceed prescribed quotas or that resulted from throttling of suspected abusive behavior;
  - viii. Due to Customer' s use Service features that are outside of associated Support Windows; or
  - ix. Attributable to acts by persons gaining unauthorized access to 21Vianet' s Service by means of Customer' s passwords or equipment or otherwise resulting from Customer' s failure to follow appropriate security practices.

#### **d. Service Credits**

- i. The amount and method of calculation of Service Credits is described below in connection with each Service.
- ii. Service Credits are Customer' s sole and exclusive remedy for any failure to meet any Service Level.
- iii. The Service Credits awarded in any billing month for a particular Service or Service Resource will not, under any circumstance, exceed Customer' s monthly service fees that Service or Service Resource, as applicable, in the billing month.
- iv. For Services purchased as part of a suite, the Service Credit will be based on the pro-rata portion of the cost of the Service, as determined by 21Vianet in its reasonable discretion. In cases where Customer has purchased Services from a reseller, the Service Credit will be based on the estimated retail price for the applicable Service, as determined by 21Vianet in its reasonable discretion.

### **3. Azure Services Subject to SLA**

#### Application Gateway

##### **i. Additional Definitions**

- A. "Application Gateway Cloud Service" refers to a collection of one or more Application Gateway instances configured to perform HTTP load balancing services.

##### **ii. Monthly Uptime Calculation and Service Levels for Application Gateway**

- A. "Maximum Available Minutes" is the total accumulated minutes during a billing month during which an Application Gateway Cloud Service comprising two or more medium or larger Application Gateway instances has been deployed in a Azure subscription.
- B. "Downtime" is the total accumulated Maximum Available Minutes during a billing month for a given Application Gateway Cloud Service during which the Application Gateway Cloud Service is unavailable. A

given minute is considered unavailable if all attempts to connect to the Application Gateway Cloud Service throughout the minute are unsuccessful.

C. "Monthly Uptime Percentage" for a given Application Gateway Cloud Service is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for the Application Gateway Cloud Service. Monthly Uptime Percentage is represented by the following formula:

D. The following Service Levels and Service Credits are applicable to Customer' s use of each Application Gateway Cloud Service.

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

#### Automation Service

i. **Additional Definitions**

- A. "Job" means the execution of a Runbook.
- B. "Planned Start Time" is a time at which a Job is scheduled to begin executing.
- C. "Runbook" means a set of actions specified by Customer to execute within Azure.

ii. **Monthly Uptime Calculation and Service Levels for the Automation Service**

- A. "Delayed Jobs" is the total number of Jobs that fail to start within thirty (30) minutes of their Planned Start Times.
- B. "Total Jobs" is the total number of Jobs scheduled for execution during a given billing month.
- C. "Monthly Uptime Percentage" for the Automation Service is calculated as Total Jobs less Delayed Jobs divided by Total Jobs in a billing month for

a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

D. The following Service Levels and Service Credits are applicable to Customer' s use of the Automation Service:

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

### **Azure IoT Hub**

- i. For IoT Hub, we promise that at least 99.9% of the time deployed IoT hubs will be able to send messages to and receive messages from registered devices and the Service will be able to perform create, read, update, and delete operations on IoT hubs.
- ii. No SLA is provided for the Free Tier of IoT Hub.

### **iii. Additional Definitions**

- A. "Deployment Minutes" is the total number of minutes that a given IoT hub has been deployed in Azure during a billing month.
- B. "Maximum Available Minutes" is the sum of all Deployment Minutes across all IoT hubs deployed in a given Azure subscription during a billing month.
- C. "Message" refers to any content sent by a deployed IoT hub to a device registered to the IoT hub or received by the IoT hub from a registered device, using any protocol supported by the Service.
- D. "Device Identity Operations" refers to create, read, update, and delete operations performed on the device identity registry of an IoT hub.

- E. Downtime : The total accumulated Deployment Minutes, across all IoT hubs deployed in a given Azure subscription, during which the IoT hub is unavailable. A minute is considered unavailable for a given IoT hub if all continuous attempts to send or receive Messages or perform Device Identity Operations on the IoT hub throughout the minute either return an Error Code or do not result in a Success Code within five minutes.
- F. Monthly Uptime Percentage: The Monthly Uptime Percentage is calculated using the following formula:

Monthly Uptime % = (Maximum Available Minutes - Downtime) / Maximum Available Minutes

- G. Service Credit:

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

- H. Service Level Exceptions: The Free Tier of the IoT Hub Service is not covered by this SLA.

## Backup Service

### i. **Additional Definitions**

- A. "Backup" or "Back Up" is the process of copying computer data from a registered server to a Backup Vault.
- B. "Backup Agent" refers to the software installed on a registered server that enables the registered server to Back Up or Restore one or more Protected Items.
- C. "Backup Vault" refers to a container in which Customer may register one or more Protected Items for Backup.



- D. "Failure" means that either the Backup Agent or the Service fails to fully complete a properly configured Backup or Recovery operation due to unavailability of the Backup Service.
  - E. "Protected Item" refers to a collection of data, such as a volume, database, or virtual machine that has been scheduled for Backup to the Backup Service such that it is enumerated as a Protected Item in the Protected Items tab in the Recovery Services section of the Management Portal.
  - F. "Recovery" or "Restore" is the process of restoring computer data from a Backup Vault to a registered server.
- ii. **Monthly Uptime Calculation and Service Levels for Backup Service**
- A. "Deployment Minutes" is the total number of minutes during which a Protected Item has been scheduled for Backup to a Backup Vault.
  - B. "Maximum Available Minutes" is the sum of all Deployment Minutes across all Protected Items for a given Azure subscription during a billing month.
  - C. "Downtime" is the total accumulated Deployment Minutes across all Protected Items scheduled for Backup by Customer in a given Azure subscription during which the Backup Service is unavailable for the Protected Item. The Backup Service is considered unavailable for a given Protected Item from the first Failure to Back Up or Restore the Protected Item until the initiation of a successful Backup or Recovery of a Protected Item, provided that retries are continually attempted no less frequently than once every thirty minutes.
  - D. "Monthly Uptime Percentage" for the Backup Service is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

- E. The following Service Levels and Service Credits are applicable to Customer' s use of the Backup Service:

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

### CDN Service

i. **Monthly Uptime Calculation and Service Levels for CDN Service**

- A. 21Vianet is not responsible for comprehensive monitoring of customer content; this responsibility lies with customer. 21Vianet will review data from any commercially reasonable independent measurement system used by Customer.
- B. Customer must select a set of agents from the measurement system' s list of standard agents that are generally available and represent at least five major metropolitan areas in the People' s Republic of China.
1. Measurement System tests (frequency of at least one test per hour per agent) will be configured to perform one HTTP(S) GET operation according to the model below:
  2. A test file will be placed on Customer' s origin (e.g., Azure Storage account).
  3. The GET operation will retrieve the file through the CDN Service, by requesting the object from the appropriate Azure domain name hostname.
  4. The test file will meet the following criteria:
    - i. The test object will meet CDN caching specifications. For example, the response header should not include 'set-cookie' ; cache-control attributes cannot be set to 'private' or 'no-cache' ; the 'expires' value cannot be set at a past time value.

- ii. The test object will be a file at least 50KB in size and no larger than 1MB.
- iii. Raw data will be trimmed to eliminate any measurements that came from an agent experiencing technical problems during the measurement period.

Remarks:

1. In the event test results should reveal issues, please contact us as soon as possible so that we may verify said test results at first opportunity in order to locate the causes for these issues.
2. If the chosen test source itself contains problematic links, this will cause anomalies in the test results; these anomalies should be excluded when calculating the final availability value.

C. "Monthly Uptime Percentage" is the percentage of HTTP transactions in which the CDN responds to client requests and delivers the requested content without error. Monthly Uptime Percentage of the CDN Service is calculated as the number of times the object was delivered successfully divided by the total number of requests (after removing erroneous data).

D. The following Service Levels and Service Credits are applicable to Customer' s use of the CDN Service:

Monthly Uptime Percentage	Service Credit
<99.9%	5%
<99%	20%

Cloud Services

i. **Additional Definitions**

- A. "Cloud Services" refers to a set of compute resources utilized for Web and Worker Roles.
- B. "Web Role" is a Cloud Services component run in the Azure execution environment that is customized for web application programming as supported by IIS and ASP.NET.

C. "Worker Role" is a Cloud Services component run in the Azure execution environment that is useful for generalized development, and may perform background processing for a Web Role.

ii. **Monthly Uptime Calculation and Service Levels for Cloud Services**

A. "Maximum Available Minutes" is the total accumulated minutes during a billing month for all Internet facing roles that have two or more instances deployed in different Update Domains. Maximum Available Minutes is measured from when the Tenant has been deployed and its associated roles have been started resultant from action initiated by Customer to the time Customer has initiated an action that would result in stopping or deleting the Tenant.

B. "Downtime" is the total accumulated minutes that are part of Maximum Available Minutes that have no External Connectivity.

C. "Monthly Uptime Percentage" for Cloud Services is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

D. The following Service Levels and Service Credits are applicable to Customer' s use of Cloud Services:

Monthly Uptime Percentage	Service Credit
<99.95%	10%
<99%	25%

DocumentDB

We guarantee at least 99.99% of the time we will successfully process requests to perform operations against DocumentDB Resources.

"Average Error Rate" for a billing month is the sum of Error Rates for each hour in the billing month divided by the total number of hours in the billing month.

"Database Account" is a DocumentDB account containing one or more databases.

"Error Rate" is the total number of Failed Requests divided by Total Requests, across all Resources in a given Azure subscription, during a given one-hour interval. If the Total Requests in a given one-hour interval is zero, the Error Rate for that interval is 0%.

"Excluded Requests" are requests within Total Requests that result in an HTTP 4xx status code, other than an HTTP 408 status code.

"Failed Requests" is the set of all requests within Total Requests that either return an Error Code or an HTTP 408 status code or fail to return a Success Code within 5 minutes for Database Account creation and deletion, 3 minutes for updating the Performance/Offer level, and 5 seconds for all other operations.

"Resource" is a set of URI addressable entities associated with a Database Account.

"Total Request" is the set of all requests, other than Excluded Requests, to perform operations issued against Resources attempted within a one-hour interval within a given Azure subscription during a billing month.

"Monthly Uptime Percentage" for the DocumentDB Service is calculated by subtracting from 100% the Average Error Rate for a given Azure subscription in a billing month. The "Average Error Rate" for a billing month is the sum of Error Rates for each hour in the billing month divided by the total number of hours in the billing month. Monthly Uptime Percentage is represented by the following formula:

Monthly Uptime % = 100% - Average Error Rate

**Service Credit:**

---

<b>Monthly Uptime Percentage</b>	<b>Service Credit</b>
< 99.99%	10%
< 99%	25%

## Event Hubs

### i. **Monthly Uptime Calculation and Service Levels for Event Hubs**

- A. "Deployment Minutes" is the total number of minutes that a given Event Hub has been deployed in Azure during a billing month.
- B. "Maximum Available Minutes" is the sum of all Deployment Minutes across all Event Hubs deployed by Customer in a given Azure subscription under the Basic or Standard Event Hubs tiers during a billing month.
- C. "Downtime" is the total accumulated Deployment Minutes, across all Event Hubs deployed by Customer in a given Azure subscription under the Basic or Standard Event Hubs tiers, during which the Event Hub is unavailable. A minute is considered unavailable for a given Event Hub if all continuous attempts to send or receive Messages or perform other operations on the Event Hub throughout the minute either return an Error Code or do not result in a Success Code within five minutes.
- D. "Monthly Uptime Percentage" for Event Hubs is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:
- E. The following Service Levels and Service Credits are applicable to Customer' s use of the Basic and Standard Event Hubs tiers. The Free Event Hubs tier is not covered by this SLA.

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

## ExpressRoute

i. We guarantee a minimum of 99.95% ExpressRoute dedicated circuit availability.

ii. **Additional Definitions**

"Dedicated Circuit" means a logical representation of connectivity offered through the ExpressRoute Service between Customer's premises and Azure through an exchange provider or a network service provider, where such connectivity does not traverse the public Internet.

iii. **Monthly Uptime Calculation and Service Levels for ExpressRoute**

"Maximum Available Minutes" is the total number of minutes that a given Dedicated Circuit is linked to one or more Virtual Networks in Azure during a billing month in a given Azure subscription.

"Downtime" is the total accumulated minutes during a billing month for a given Azure subscription during which the Dedicated Circuit is unavailable. A minute is considered unavailable for a given Dedicated Circuit if all attempts by Customer within the minute to establish IP-level connectivity to the Virtual Network Gateway associated with the Virtual Network fail for longer than thirty seconds.

"Monthly Uptime Percentage" for a given Dedicated Circuit in a given billing month is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes. Monthly Uptime Percentage is represented by the following formula:

$$\text{Monthly Uptime \%} = (\text{Maximum Available Minutes} - \text{Downtime}) / \text{Maximum Available Minutes}$$

The following Service Levels and Service Credits are applicable to Customer's use of each Dedicated Circuit within the ExpressRoute Service:

Monthly Uptime Percentage	Service Credit
<99.95%	10%
<99%	25%

## HDInsight

### i. **Additional Definitions**

- A. "Cluster Internet Gateway" means a set of virtual machines within an HDInsight Cluster that proxy all connectivity requests to the Cluster.
- B. "HDInsight Cluster" or "Cluster" means a collection of virtual machines running a single instance of the HDInsight Service.

### ii. **Monthly Uptime Calculation**

- A. "Deployment Minutes" is the total number of minutes that a given HDInsight Cluster has been deployed in Azure.
- B. "Maximum Available Minutes" is the sum of all Deployment Minutes across all Clusters deployed by Customer in a given Azure subscription during a billing month.
- C. "Downtime" is the total accumulated Deployment Minutes when the HDInsight Service is unavailable. A minute is considered unavailable for a given Cluster if all continual attempts within the minute to establish a connection to the Cluster Internet Gateway fail.
- D. "Monthly Uptime Percentage" for the HDInsight Service is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:
- E. The following Service Levels and Service Credits are applicable to Customer' s use of the HDInsight Service:

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%



## Media Services

- For Media Services Encoding, we guarantee 99.9% availability of REST API transactions.
- For Streaming, we will successfully service requests with a 99.9% availability guarantee for existing media content when at least one Streaming Unit is purchased.
- For Live Channels, we guarantee that running Channels will have external connectivity at least 99.9% of the time.
- For Content Protection, we guarantee that we will successfully fulfill key requests at least 99.9% of the time.
- For Indexer, we will successfully service Indexer Task requests processed with an Encoding Reserved Unit 99.9% of the time.

### Additional Definitions

"Allocated Egress Bandwidth" is the amount of bandwidth configured by Customer in the Management Portal for a Media Service. Allocated Egress Bandwidth may be labeled "Streaming Units" or a similar name in the Management Portal.

"Channel" means an end point within a Media Service that is configured to receive media data.

"Encoding" means the processing of media files per subscription as configured in the Media Services Tasks.

"Encoding Reserved Unit" means encoding reserved units purchased by the customer in an Azure Media Services account.

"Indexer Task" means a Media Services Task that is configured to index an MP3 input file with a minimum five-minute duration.

"Media Service" means an Azure Media Services account, created in the Management Portal, associated with Customer's Azure subscription. Each Azure subscription may have more than one associated Media Service.

"Media Service Request" means a request issued to Customer's Media Service.

"Media Services Task" means an individual operation of media processing work as configured by Customer. Media processing operations involve encoding, converting, or indexing media files.

"Streaming Unit" means a unit of reserved egress capacity purchased by Customer for a Media Service.

"Valid Key Requests" are all requests made to the Content Protection Service for existing content keys in a Customer's Media Service.

"Valid Media Services Requests" are all qualifying Media Service Requests for existing media content in a Customer's Azure Storage account associated with its Media Service when at least one Streaming Unit has been purchased and allocated to that Media Service. Valid Media Services Requests do not include Media Service Requests for which total throughput exceeds 80% of the Allocated Bandwidth.

#### Monthly Uptime Calculation and Service Levels for Encoding Service

"Total Transaction Attempts" is the total number of authenticated REST API requests with respect to a Media Service made by Customer during a billing month for a subscription. Total Transaction Attempts does not include REST API requests that return an Error Code that are continuously repeated within a five-minute window after the first Error Code is received.

"Failed Transactions" is the set of all requests within Total Transaction Attempts that do not return a Success Code within 30 seconds from 21Vianet's receipt of the request.

"Monthly Uptime Percentage" for the Azure Media Services Encoding Service is calculated as Total Transaction Attempts less Failed Transactions divided by Total Transaction Attempts in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

Monthly Uptime % = (Total Transaction Attempts - Failed Transactions) / Total Transaction Attempts

The following Service Levels and Service Credits are applicable to Customer' s use of the Azure Media Services Encoding Service:

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

#### Monthly Uptime Calculation and Service Levels for Indexer Service

"Total Transaction Attempts" is the total number of Indexer Tasks attempted to be executed using an available Encoding Reserved Unit by Customer during a billing month for a subscription.

"Failed Transactions" is the set of Indexer Tasks within Total Transaction Attempts that either

1. Do not complete within a time period that is 3 times the duration of the input file; or
2. Do not start processing within 5 minutes of the time that an Encoding Reserved Unit becomes available for use by the Indexer Task.

"Monthly Uptime Percentage" for the Azure Media Services Indexer Service is calculated as Total Transaction Attempts less Failed Transactions divided by Total Transaction Attempts in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

Monthly Uptime % = (Total Transaction Attempts - Failed Transactions) / Total Transaction Attempts

The following Service Levels and Service Credits are applicable to Customer' s use of the Azure Media Services Indexer Service:

Monthly Uptime Percentage	Service Credit
<99.9%	10%

<99%	25%
------	-----

### Monthly Uptime Calculation and Service Levels for Streaming Service

"Deployment Minutes" is the total number of minutes that a given Streaming Unit has been purchased and allocated to a Media Service during a billing month.

"Maximum Available Minutes" is the sum of all Deployment Minutes across all Streaming Units purchased and allocated to a Media Service during a billing month.

"Downtime" is the total accumulated Deployment Minutes when the Streaming Service is unavailable. A minute is considered unavailable for a given Streaming Unit if all continuous Valid Media Service Requests made to the Streaming Unit throughout the minute result in an Error Code.

"Monthly Uptime Percentage" for the Azure Media Services Streaming Service is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

Monthly Uptime % = (Maximum Available Minutes - Downtime) / Maximum Available Minutes

The following Service Levels and Service Credits are applicable to Customer' s use of the Azure Media Services On-Demand Streaming Service:

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

### Monthly Uptime Calculation and Service Levels for Live Channels

"Deployment Minutes" is the total number of minutes that a given Channel has been purchased and allocated to a Media Service and is in a running state during a billing month.

"Maximum Available Minutes" is the sum of all Deployment Minutes across all Channels purchased and allocated to a Media Service during a billing month.

"Downtime" is the total accumulated Deployment Minutes when the Live Channels Service is unavailable. A minute is considered unavailable for a given Channel if the Channel has no External Connectivity during the minute.

"Monthly Uptime Percentage" for the Live Channels Service is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

Monthly Uptime % = (Maximum Available Minutes - Downtime) / Maximum Available Minutes

The following Service Levels and Service Credits are applicable to Customer's use of the Azure Media Services Live Channels Service:

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

#### Monthly Uptime Calculation and Service Levels for Content Protection Service

"Total Transaction Attempts" are all Valid Key Requests made by Customer during a billing month for a given Azure subscription.

"Failed Transactions" are all Valid Key Requests included in Total Transaction Attempts that result in an Error Code or otherwise do not return a Success Code within 30 seconds after receipt by the Content Protection Service.

"Monthly Uptime Percentage" for Azure Media Services is calculated as Total Transaction Attempts less Failed Transactions divided by Total Transaction Attempts in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

Monthly Uptime % = (Total Transaction Attempts - Failed Transactions) / Total Transaction Attempts

The following Service Levels and Service Credits are applicable to Customer's use of the Azure Media Services Content Protection Service:

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

### MySQL Database on Azure Service

i. **Additional Definitions**

A. "Database Server" is a MySQL database server deployed by Customer within the MySQL Database on Azure Service.

ii. **Monthly Uptime Calculation and Service Levels for the MySQL Database on Azure Service**

A. "Deployment Minutes" is the total number of minutes that a given Database Server has been deployed in Azure during a billing month.

B. "Maximum Available Minutes" is the sum of all Deployment Minutes across all Database Servers for a given Azure subscription during a billing month.

C. "Downtime" is the total accumulated Deployment Minutes across all Database Servers deployed by Customer in a given Azure subscription during which the Database Server is unavailable. A minute is considered unavailable for a given Database Server if all continuous attempts by Customer to establish a connection to the Database Server within the minute fail.

D. "Monthly Uptime Percentage" for the MySQL Database on Azure Service is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

E. The following Service Levels and Service Credits are applicable to Customer' s use of the MySQL Database on Azure Service:

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

### Multi-Factor Authentication

- i. We guarantee 99.9% availability of Azure Multi-Factor Authentication.
- ii. No SLA is provided for the Free tier of Multi-Factor Authentication.
- iii. Monthly Uptime Calculation and Service Levels for Multi-Factor Authentication Service

"**Deployment Minutes**" is the total number of minutes that a given Multi-Factor Authentication provider has been deployed in Azure during a billing month.

"**Downtime**" is the total accumulated Deployment Minutes, across all Multi-Factor Authentication providers deployed by Customer in a given Azure subscription, during which the Multi-Factor Authentication Service is unable to receive or process authentication requests for the Multi-Factor Authentication provider.

"**Maximum Available Minutes**" is the sum of all Deployment Minutes across all Multi-Factor Authentication providers deployed by Customer in a given Azure subscription during a billing month.

"**Monthly Uptime Percentage**" for the Multi-Factor Authentication Service is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

Monthly Uptime % = (Maximum Available Minutes - Downtime) / Maximum Available Minutes

The following Service Levels and Service Credits are applicable to Customer' s use of the Multi-Factor Authentication Service:

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

### Notification Hubs

#### i. **Monthly Uptime Calculation and Service Levels for Notification Hubs**

- A. "Deployment Minutes" is the total number of minutes that a given Notification Hub has been deployed in Azure during a billing month.
- B. "Maximum Available Minutes" is the sum of all Deployment Minutes across all Notification Hubs deployed by Customer in a given Azure subscription under the Basic or Standard Notification Hubs tiers during a billing month.
- C. "Downtime" is the total accumulated Deployment Minutes, across all Notification Hubs deployed by Customer in a given Azure subscription under the Basic or Standard Notification Hubs tiers, during which the Notification Hub is unavailable. A minute is considered unavailable for a given Notification Hub if all continuous attempts to send notifications or perform registration management operations with respect to the Notification Hub throughout the minute either return an Error Code or do not result in a Success Code within five minutes.
- D. "Monthly Uptime Percentage" for Notification Hubs is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:



E. The following Service Levels and Service Credits are applicable to Customer' s use of the Basic and Standard Notification Hubs tiers. The Free Notification Hubs tier is not covered by this SLA.

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

Power BI Embedded

ii. We guarantee at least 99.9% availability of Power BI Embedded for users to execute API calls and embed reports.

iii. Additional Definitions

"Deployment Minutes" is the total number of minutes for which a given workspace collection has been provisioned during a billing month.

"Maximum Available Minutes" is the sum of all Deployment Minutes across all workspace collections provisioned by customer in a given Azure subscription during a billing month.

"Downtime" is the total accumulated Deployment Minutes, during which the workspace collection is unavailable. A minute is considered unavailable for a given workspace collection if all continuous attempts within the minute to read or write any portion of Power BI Embedded data result in an Error Code or do not return a response within five minutes.

"Monthly Uptime Percentage" The Monthly Uptime Percentage is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for the Azure subscription. Monthly Uptime Percentage is represented by the following formula:

Monthly Uptime % = (Maximum Available Minutes - Downtime) / Maximum Available Minutes

The following Service Levels and Service Credits are applicable to Customer' s use of the Power BI Embedded.

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

### Redis Cache Service

iv. We guarantee at least 99.9% of the time that customers will have connectivity between the Cache endpoints and our Internet gateway.

v. Additional Definitions

"**Cache**" refers to a deployment of the Cache Service created by Customer, such that its Cache Endpoints are enumerated in the Cache tab in the Management Portal.

"**Cache Endpoints**" refers to endpoints through which a Cache may be accessed.

vi. **Monthly Uptime Calculation and Service Levels for Cache Service**

"**Deployment Minutes**" is the total number of minutes that a given Cache has been deployed in Azure during a billing month.

"**Maximum Available Minutes**" is the sum of all Deployment Minutes across all Caches deployed by Customer in a given Azure subscription during a billing month.

"**Downtime**" is the total accumulated Deployment Minutes, across all Caches deployed by Customer in a given Azure subscription, during which the Cache is unavailable. A minute is considered unavailable for a given Cache when

there is no connectivity throughout the minute between one or more Cache Endpoints associated with the Cache and 21Vianet' s Internet gateway.

"

Monthly Uptime % = (Maximum Available Minutes - Downtime) / Maximum Available Minutes

The following Service Levels and Service Credits are applicable to Customer' s use of the Cache Service, which includes the Standard tier of the Azure Redis Cache Service. The Basic tier of the Azure Redis Cache Service is not covered by this SLA.

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

### Stream Analytics

- i. We guarantee at least 99.9% availability of the Stream Analytics API.
- ii. We guarantee that 99.9% of the time, deployed Stream Analytics jobs will be either processing data or available to process data.
- iii. **Monthly Uptime Calculation for Stream Analytics API Calls**

"Total Transaction Attempts" is the total number of authenticated REST API requests to manage a streaming job within the Stream Analytics Service by Customer during a billing month for a given Azure subscription.

"Failed Transactions" is the set of all requests within Total Transaction Attempts that return an Error Code or otherwise do not return a Success Code within five minutes from 21Vianet' s receipt of the request.

"Monthly Uptime Percentage" for API calls within the Stream Analytics Service is calculated as Total Transaction Attempts less Failed Transactions divided by

Total Transaction Attempts in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

Monthly Uptime % = (Total Transaction Attempts - Failed Transactions) / Total Transaction Attempts

The following Service Credits are applicable to Customer' s use of API calls within the Stream Analytics Service.

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

**iv. Monthly Uptime Calculation for Stream Analytics Jobs**

"Deployment Minutes" is the total number of minutes that a given job has been deployed within the Stream Analytics Service during a billing month.

"Maximum Available Minutes" is the sum of all Deployment Minutes across all jobs deployed by Customer in a given Azure subscription during a billing month.

"Downtime" is the total accumulated Deployment Minutes, across all jobs deployed by Customer in a given Azure subscription, during which the job is unavailable. A minute is considered unavailable for a deployed job if the job is neither processing data nor available to process data throughout the minute.

"Monthly Uptime Percentage" for jobs within the Stream Analytics Service is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

Monthly Uptime % = (Maximum Available Minutes - Downtime) / Maximum Available Minutes

The following Service Credits are applicable to Customer' s use of jobs within the Stream Analytics Service.

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

### Scheduler

#### i. **Additional Definitions**

- A. "Planned Execution Time" is a time at which a Scheduled Job is scheduled to begin executing.
- B. "Scheduled Job" means an action specified by Customer to execute within Azure according to a specified schedule.

#### ii. **Monthly Uptime C**

- A. "Downtime" is the total accumulated minutes in a billing month during which one or more of Customer' s Scheduled Jobs is in a state of delayed execution. A given Scheduled Job is in a state of delayed execution if it has not begun executing after a Planned Execution Time, provided that such delayed execution time shall not be considered Downtime if the Scheduled Job begins executing within thirty (30) minutes after a Planned Execution Time.
- B. "Maximum Available Minutes" is the total number of minutes in a billing month.
- C. "Monthly Uptime Percentage" for the Scheduler Service is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

D. The following Service Levels and Service Credits are applicable to Customer' s use of the Scheduler Service:

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

### Service Bus Service

i. **Additional Definitions**

A. "Message" refers to any user-defined content sent or received through Service Bus Relays, Queues, Topics, or Notification Hubs, using any protocol supported by Service Bus.

ii. **Monthly Uptime Calculation and Service Levels for Relays**

A. "Deployment Minutes" is the total number of minutes that a given Relay has been deployed in Azure during a billing month.

B. "Maximum Available Minutes" is the sum of all Deployment Minutes across all Relays deployed by Customer in a given Azure subscription during a billing month.

C. "Downtime" is the total accumulated Deployment Minutes, across all Relays deployed by Customer in a given Azure subscription, during which the Relay is unavailable. A minute is considered unavailable for a given Relay if all continuous attempts to establish a connection to the Relay throughout the minute either return an Error Code or do not result in a Success Code within five minutes.

D. "Monthly Uptime Percentage" for Relays is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

- E. The following Service Levels and Service Credits are applicable to Customer' s use of Relays:

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

iii. **Monthly Uptime Calculation and Service Levels for Queues and Topics**

- A. "Deployment Minutes" is the total number of minutes that a given Queue or Topic has been deployed in Azure during a billing month.
- B. "Maximum Available Minutes" is the sum of all Deployment Minutes across all Queues and Topics deployed by Customer in a given Azure subscription during a billing month.
- C. "Downtime" is the total accumulated Deployment Minutes, across all Queues and Topics deployed by Customer in a given Azure subscription, during which the Queue or Topic is unavailable. A minute is considered unavailable for a given Queue or Topic if all continuous attempts to send or receive Messages or perform other operations on the Queue or Topic throughout the minute either return an Error Code or do not result in a Success Code within five minutes.
- D. "Monthly Uptime Percentage" for Queues and Topics is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:
- E. The following Service Levels and Service Credits are applicable to Customer' s use of Queues and Topics:

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

## Site Recovery Service

### i. **Additional Definitions**

- A. "Failover" is the process of transferring control, either simulated or actual, of a Protected Instance from a primary site to a secondary site.
- B. "On-Premises-to-Azure Failover" is the Failover of a Protected Instance from a non-Azure primary site to an Azure secondary site. Customer may designate a particular Azure datacenter as a secondary site, provided that if Failover to the designated datacenter is not possible, 21Vianet may replicate to a different datacenter in the same region.
- C. "On-Premises-to-On-Premises Failover" is the Failover of a Protected Instance from a non-Azure primary site to a non-Azure secondary site.
- D. "Protected Instance" refers to a virtual or physical machine configured for replication by the Site Recovery Service from a primary site to a secondary site. Protected Instances are enumerated in the Protected Items tab in the Recovery Services section of the Management Portal.

### ii. **Monthly Uptime Calculation and Service Levels for On-Premises-to-On-Premises Failover**

- A. "Maximum Available Minutes" is the total number of minutes that a given Protected Instance has been configured for On-Premises-to-On-Premises replication by the Site Recovery Service during a billing month.
- B. "Failover Minutes" is the total number of minutes in a billing month during which a Failover of a Protected Instance configured for On-Premises-to-On-Premises replication has been attempted but not completed.
- C. "Downtime" is the total accumulated Failover Minutes in which the Failover of a Protected Instance is unsuccessful due to unavailability of the Site Recovery Service, provided that retries are continually attempted no less frequently than once every thirty minutes.



D. "Monthly Uptime Percentage" for On-Premises-to-On-Premises Failover of a specific Protected Instance in a given billing month is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes. Monthly Uptime Percentage is represented by the following formula:

E. The following Service Levels and Service Credits are applicable to Customer' s use of each Protected Instance within the Site Recovery Service for On-Premises-to-On-Premises Failover:

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

iii. **Monthly Recovery Time Objective and Service Levels for On-Premises-to-Azure Failover**

A. "Recovery Time Objective (RTO)" means the period of time beginning when Customer initiates a Failover of a Protected Instance experiencing either a planned or unplanned outage for On-Premises-to-Azure replication to the time when the Protected Instance is running as a virtual machine in Azure, excluding any time associated with manual action or the execution of Customer scripts.

B. "Monthly Recovery Time Objective" for a specific Protected Instance configured for On-Premises-to-Azure replication in a given billing month is four hours for an unencrypted Protected Instance and six hours for an encrypted Protected Instance. One hour will be added to the monthly Recovery Time Objective for each additional 25GB over the initial 100GB Protected Instance size.

C. The following Service Levels and Service Credits are applicable to Customer' s use of each Protected Instance within the Site Recovery

Service for On-Premises-to-Azure Failover, assuming Protected Instance sizes of 100GB or less:

Protected Instance	Monthly Recovery Time Objective	Service Credit
Unencrypted	> 4 hours	100%
Encrypted	> 6 hours	100%

### SQL Database Service

#### Basic, Standard, and Premium Tiers

We guarantee at least 99.99% of the time customers will have connectivity between their single or elastic Basic, Standard, or Premium Azure SQL Database and our Internet gateway.

#### Web and Business Tiers (retired)

We guarantee at least 99.9% of the time customers will have connectivity between their Web or Business Azure SQL Database and our Internet gateway.

### **Additional Definitions**

"**Database**" means any Web, Business, single or elastic Basic, Standard, or Premium Azure SQL Database.

Monthly Uptime Calculation and Service Levels for Azure SQL Database Service (single or elastic Basic, Standard, and Premium Tiers)

"**Maximum Available Minutes**" is the total number of minutes that a given single or elastic Basic, Standard, or Premium Database has been deployed in Azure during a billing month in a given Azure subscription.

"**Downtime**" is the total accumulated minutes during a billing month for a given Azure subscription during which a given single or elastic Basic, Standard, or Premium Database is unavailable. A minute is considered unavailable for a given Database if

all continuous attempts by Customer to establish a connection to the Database within the minute fail.

**"Monthly Uptime Percentage"** for a given single or elastic Basic, Standard, or Premium Tier of the SQL Database Service is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

$$\text{Monthly Uptime \%} = (\text{Maximum Available Minutes} - \text{Downtime}) / \text{Maximum Available Minutes}$$

The following Service Levels and Service Credits are applicable to Customer' s use of the Basic, Standard, and Premium Tiers of the Azure SQL Database Service:

---

<b>MONTHLY UPTIME PERCENTAGE</b>	<b>SERVICE CREDIT</b>
< 99.99%	10%
< 99%	25%

Monthly Uptime Calculation and Service Levels for Azure SQL Database Service (Web and Business Tiers)

**"Deployment Minutes"** is the total number of minutes that a given Web or Business Database has been deployed in Azure during a billing month.

**"Maximum Available Minutes"** is the sum of all Deployment Minutes across all Web and Business Databases for a given Azure subscription during a billing month.

**"Downtime"** is the total accumulated Deployment Minutes across all Web and Business Databases deployed by Customer in a given Azure subscription during which the Database is unavailable. A minute is considered unavailable for a given

Database if all continuous attempts by Customer to establish a connection to the Database within the minute fail.

**"Monthly Uptime Percentage"** for the Web and Business Tiers of the SQL Database Service is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

Monthly Uptime % = (Maximum Available Minutes-Downtime) / Maximum Available Minutes

The following Service Levels and Service Credits are applicable to Customer' s use of the Web and Business Tiers of the Azure SQL Database Service:

---

<b>MONTHLY UPTIME PERCENTAGE</b>	<b>SERVICE CREDIT</b>
< 99.9%	10%
< 99%	25%

#### SQL Data Warehouse Service

We guarantee that, at least 99.9% of the time client operations executed on a SQL Data Warehouse database will succeed.

#### **Additional Definitions**

"Database" means any SQL Data Warehouse Database.

" Maximum Available Minutes" is the total number of minutes that a given Database has been deployed in Azure during a billing month in a given Azure subscription.

“Client Operations” is the set of all documented operations supported by SQL Data Warehouse.

“Downtime” is the total accumulated minutes during a billing month for a given Azure subscription during which a given Database is unavailable. A minute is considered unavailable for a given Database if more than 1% of all Client Operations completed during the minute return an Error Code.

“Monthly Uptime Percentage” for a given Database is calculated as Maximum Available Minutes less Downtime, divided by Maximum Available Minutes in a billing month in a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

$$\text{Monthly Uptime \%} = (\text{Maximum Available Minutes} - \text{Downtime}) / \text{Maximum Available Minutes}$$

The following Service Levels and Service Credits are applicable to Customer’ s use of the Basic, Standard, and Premium Tiers of the Azure SQL Database Service:

---

<b>MONTHLY UPTIME PERCENTAGE</b>	<b>SERVICE CREDIT</b>
< 99.9%	10%
< 99%	25%

### Storage Service

- We guarantee that at least 99.99% (99.9% for Cool Access Tier) of the time, we will successfully process requests to read data from Read Access-Geo Redundant Storage (RA-GRS) Accounts, provided that failed attempts to read data from the primary region are retried on the secondary region.

- We guarantee that at least 99.9% (99% for Cool Access Tier) of the time, we will successfully process requests to read data from Locally Redundant Storage (LRS) and Geo Redundant Storage (GRS) Accounts.
- We guarantee that at least 99.9% (99% for Cool Access Tier) of the time, we will successfully process requests to write data to Locally Redundant Storage (LRS) and Geo Redundant Storage (GRS) Accounts and Read Access-Geo Redundant Storage (RA-GRS) Accounts.

#### Additional Definitions

"Average Error Rate" for a billing month is the sum of Error Rates for each hour in the billing month divided by the total number of hours in the billing month.

"Blob Storage Account" is a storage account specialized for storing data as blobs and provides the ability to specify an access tier indicating how frequently the data in that account is accessed.

"Cool Access Tier" is an attribute of a Blob Storage Account indicating that the data in the account is infrequently accessed and has a lower availability service level than data in other access tiers.

"Excluded Transactions" are storage transactions that do not count toward either Total Storage Transactions or Failed Storage Transactions. Excluded Transactions include pre-authentication failures; authentication failures; attempted transactions for storage accounts over their prescribed quotas; creation or deletion of containers, tables, or queues; clearing of queues; and copying blobs between storage accounts.

"Error Rate" is the total number of Failed Storage Transactions divided by the Total Storage Transactions during a set time interval (currently set at one hour). If the Total Storage Transactions in a given one-hour interval is zero, the error rate for that interval is 0%.

"Failed Storage Transactions" is the set of all storage transactions within Total Storage Transactions that are not completed within the Maximum Processing Time associated with their respective transaction type, as specified in the table below. Maximum Processing Time includes only the time spent processing a transaction request within the Storage Service and does not include any time spent transferring the request to or from the Storage Service.

---

#### **REQUEST TYPES**

#### **MAXIMUM PROCESSING TIME**

---

REQUEST TYPES	MAXIMUM PROCESSING TIME
PutBlob and GetBlob (includes blocks and pages) Get Valid Page Blob Ranges	Two (2) seconds multiplied by the number of MBs transferred in the course of processing the request
Copy Blob	Ninety (90) seconds (where the source and destination blobs are within the same storage account)
PutBlockList	Sixty (60) seconds
GetBlockList	
Table Query List Operations	Ten (10) seconds (to complete processing or return a continuation)
Batch Table Operations	Thirty (30) seconds
All Single Entity Table Operations All other Blob and Message Operations	Two (2) seconds

---

These figures represent maximum processing times. Actual and average times are expected to be much lower.

Failed Storage Transactions do not include:

- 1.Transaction requests that are throttled by the Storage Service due to a failure to obey appropriate back-off principles.

2.Transaction requests having timeouts set lower than the respective Maximum Processing Times specified above.

3.Read transactions requests to RA-GRS Accounts for which you did not attempt to execute the request against Secondary Region associated with the storage account if the request to the Primary Region was not successful.

4.Read transaction requests to RA-GRS Accounts that fail due to Geo-Replication Lag.

"Geo Replication Lag" for GRS and RA-GRS Accounts is the time it takes for data stored in the Primary Region of the storage account to replicate to the Secondary Region of the storage account. Because GRS and RA-GRS Accounts are replicated asynchronously to the Secondary Region, data written to the Primary Region of the storage account will not be immediately available in the Secondary Region. You can query the Geo Replication Lag for a storage account, but [Azure](#) does not provide any guarantees as to the length of any Geo Replication Lag under this SLA.

"Geographically Redundant Storage (GRS) Account" is a storage account for which data is replicated synchronously within a Primary Region and then replicated asynchronously to a Secondary Region. You cannot directly read data from or write data to the Secondary Region associated with GRS Accounts.

"Locally Redundant Storage (LRS) Account" is a storage account for which data is replicated synchronously only within a Primary Region.

"Primary Region" is a geographical region in which data within a storage account is located, as selected by you when creating the storage account. You may execute write requests only against data stored within the Primary Region associated with storage accounts.

"Read Access Geographically Redundant Storage (RA-GRS) Account" is a storage account for which data is replicated synchronously within a Primary Region and then replicated asynchronously to a Secondary Region. You can directly read data from, but cannot write data to, the Secondary Region associated with RA-GRS Accounts.

"Secondary Region" is a geographical region in which data within a GRS or RA-GRS Account is replicated and stored, as assigned by Azure based on the Primary Region associated with the storage account. You cannot specify the Secondary Region associated with storage accounts.

"Total Storage Transactions" is the set of all storage transactions, other than Excluded Transactions, attempted within a one-hour interval across all storage accounts in the Storage Service in a given subscription.



Monthly Uptime Percentage: Monthly Uptime Percentage is calculated using the following formula:

100% - Average Error Rate

Service Credit – LRS, GRS and RA-GRS (write requests) Accounts:

---

<b>MONTHLY UPTIME PERCENTAGE</b>	<b>SERVICE CREDIT</b>
< 99.9%	10%
< 99%	25%

---

Service Credit – RA-GRS (read requests) Accounts:

---

<b>MONTHLY UPTIME PERCENTAGE</b>	<b>SERVICE CREDIT</b>
< 99.99%	10%
< 99%	25%

---

Service Credit – LRS, GRS and RA-GRS (write requests) Blob Storage Accounts (Cool Access Tier):

---

<b>MONTHLY UPTIME PERCENTAGE</b>	<b>SERVICE CREDIT</b>
< 99%	10%

---

<b>MONTHLY UPTIME PERCENTAGE</b>	<b>SERVICE CREDIT</b>
< 98%	25%

---

Service Credit – RA-GRS (read requests) Blob Storage Accounts (Cool Access Tier):

---

<b>MONTHLY UPTIME PERCENTAGE</b>	<b>SERVICE CREDIT</b>
< 99.9%	10%
< 98%	25%

---

### StorSimple Service

#### i. **Additional Definitions**

- A. "Backup" is the process of backing up data stored on a registered StorSimple device to one or more associated cloud storage accounts within Azure.
- B. "Cloud Tiering" is the process of transferring data from a registered StorSimple device to one or more associated cloud storage accounts within Azure.
- C. "Failure" means the inability to fully complete a properly configured Backup, Tiering, or Restoring operation due to unavailability of the StorSimple Service.

- D. "Managed Item" refers to a volume that has been configured to Backup to the cloud storage accounts using the StorSimple Service.
- E. "Restoring" is the process of copying data to a registered StorSimple device from its associated cloud storage account(s).

ii. **Monthly Uptime Calculation and Service Levels for StorSimple Service**

- A. "Deployment Minutes" is the total number of minutes during which a Managed Item has been configured for Backup or Cloud Tiering to a StorSimple storage account in Azure.
- B. "Maximum Available Minutes" is the sum of all Deployment Minutes across all Managed Items for a given Azure subscription during a billing month.
- C. "Downtime" is the total accumulated Deployment Minutes across all Managed Items configured for Backup or Cloud Tiering by Customer in a given Azure subscription during which the StorSimple Service is unavailable for the Managed Item. The StorSimple Service is considered unavailable for a given Managed Item from the first Failure of a Backup, Cloud Tiering, or Restoring operation with respect to the Managed Item until the initiation of a successful Backup, Cloud Tiering, or Restoring operation of the Managed Item, provided that retries are continually attempted no less frequently than once every thirty minutes.
- D. "Monthly Uptime Percentage" for the StorSimple Service is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:
- E. The following Service Levels and Service Credits are applicable to Customer' s use of the StorSimple Service:

Monthly Uptime Percentage	Service Credit
<99.9%	10%

<99%	25%
------	-----

### Service Fabric

- Service Fabric is a free service, therefore, it does not have a financially backed SLA itself. The availability of your Service Fabric cluster is based on the SLA of the underlying virtual machines and storage resources used. Please see the [Virtual Machine SLA](#) and [Storage SLA](#) for more details.

### SQL Server Stretch Database

#### i. **SLA for SQL Server Stretch Database**

We guarantee at least 99.9% of the time customers will have connectivity between their SQL Server Stretch Database and our Internet gateway.

#### ii. **SLA details**

#### iii. **Additional Definitions**

- A. "Database" means one instance of SQL Server Stretch database.
- B. Monthly Uptime Calculation and Service Levels for SQL Server Stretch Database Service
- C. "Maximum Available Minutes" is the total number of minutes that a given Database has been deployed in Azure for a given Azure subscription during a billing month.
- D. "Downtime" is the total accumulated minutes across all Databases deployed by Customer in a given Azure subscription during which the Database is unavailable. A minute is considered unavailable for a given Database if all continuous attempts by Customer to establish a connection to the Database within the minute fail.

- E. "Monthly Uptime Percentage" for the SQL Server Stretch Database Service is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

$$\text{Monthly Uptime \%} = (\text{Maximum Available Minutes} - \text{Downtime}) / \text{Maximum Available Minutes}$$

- F. The following Service Levels and Service Credits are applicable to Customer' s use of the Azure SQL Stretch Database Service:

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

### Traffic Manager Service

#### iv. **Additional Definitions**

- A. "Traffic Manager Profile" or "Profile" refers to a deployment of the Traffic Manager Service created by Customer containing a domain name, endpoints, and other configuration settings, as represented in the Management Portal.
- B. "Valid DNS Response" means a DNS response, received from at least one of the Traffic Manager Service name server clusters, to a DNS request for the domain name specified for a given Traffic Manager Profile.

#### v. **Monthly Uptime Calculation and Service Levels for Traffic Manager Service**

- A. "Deployment Minutes" is the total number of minutes that a given Traffic Manager Profile has been deployed in Azure during a billing month.
- B. "Maximum Available Minutes" is the sum of all Deployment Minutes across all Traffic Manager Profiles deployed by Customer in a given Azure subscription during a billing month.

- C. "Downtime" is the total accumulated Deployment Minutes, across all Profiles deployed by Customer in a given Azure subscription, during which the Profile is unavailable. A minute is considered unavailable for a given Profile if all continual DNS queries for the DNS name specified in the Profile that are made throughout the minute do not result in a Valid DNS Response within two seconds.
- D. "Monthly Uptime Percentage" for the Traffic Manager Service is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:
- E. The following Service Levels and Service Credits are applicable to Customer' s use of the Traffic Manager Service:

Monthly Uptime Percentage	Service Credit
<99.99%	10%
<99%	25%

### Key Vault

- i. We guarantee that we will process Key Vault transactions within 5 seconds at least 99.9% of the time.
- ii. Monthly Uptime Calculation and Service Levels for Key Vault

"**Deployment Minutes**" is the total number of minutes that a given key vault has been deployed in Azure during a billing month.

"**Maximum Available Minutes**" is the sum of all Deployment Minutes across all key vaults deployed by Customer in a given Azure subscription during a billing month.

**"Excluded Transactions"** are transactions for creating, updating, or deleting key vaults, keys, or secrets.

**"Downtime"** is the total accumulated Deployment Minutes, across all key vaults deployed by Customer in a given Azure subscription, during which the key vault is unavailable. A minute is considered unavailable for a given key vault if all continuous attempts to perform transactions, other than Excluded Transactions, on the key vault throughout the minute either return an Error Code or do not result in a Success Code within 5 seconds from 21Vianet's receipt of the request.

**"Monthly Uptime Percentage"** for the Key Vault Service is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

Monthly Uptime % = (Maximum Available Minutes - Downtime) / Maximum Available Minutes

The following Service Levels and Service Credits are applicable to Customer's use of the Key Vault Service:

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

### Virtual Machines

For all Virtual Machines that have two or more instances deployed in the same Availability Set, we guarantee you will have Virtual Machine Connectivity to at least one instance at least 99.95% of the time.

For any Single Instance Virtual Machine using premium storage for all disks, we guarantee you will have Virtual Machine Connectivity of at least 99.9%. For all Internet facing Virtual Machines that have two or more instances deployed in the same

Availability Set, we guarantee you will have external connectivity at least 99.95% of the time.

i. **Additional Definitions**

- A. "Availability Set" refers to two or more Virtual Machines deployed across different Fault Domains to avoid a single point of failure.
- B. "Fault Domain" is a collection of servers that share common resources such as power and network connectivity.
- C. "Virtual Machine" refers to persistent instance types that can be deployed individually or as part of an Availability Set.
- D. "Virtual Machine Connectivity" is bi-directional network traffic between the virtual machine and other IP addresses using TCP or UDP network protocols in which the virtual machine is configured for allowed traffic. The IP addresses can be IP addresses in the same Cloud Service as the virtual machine, IP addresses within the same virtual network as the virtual machine or public, routable IP addresses.

ii. Monthly Uptime Calculation and Service Levels for Virtual Machines

- A. "Maximum Available Minutes" is the total accumulated minutes during a billing month for all Virtual Machines that have two or more instances deployed in the same Availability Set. Maximum Available Minutes is measured from when at least two Virtual Machines in the same Availability Set have both been started resultant from action initiated by Customer to the time Customer has initiated an action that would result in stopping or deleting the Virtual Machines.
- B. "Downtime" means periods of Downtime related to network, hardware, or Service maintenance or upgrades impacting Single Instances. We will publish notice or notify you at least five (5) days prior to the commencement of such Downtime. is the total accumulated minutes that are part of Maximum Available Minutes that have no External Connectivity.
- C. "Monthly Uptime Percentage" for Virtual Machines is calculated as Maximum Available Minutes less Downtime divided by Maximum Available



Minutes in a billing month for a given Azure subscription. Monthly Uptime Percentage is represented by the following formula:

$$\text{Monthly Uptime \%} = (\text{Maximum Available Minutes} - \text{Downtime}) / \text{Maximum Available Minutes} \times 100$$

D. The following Service Levels and Service Credits are applicable to Customer' s use of Virtual Machines:

Monthly Uptime Percentage	Service Credit
<99.95%	10%
<99%	25%

### iii. Monthly Uptime Calculation and Service Levels for Single-Instance Virtual Machines

A. "Minutes in the Month" is the total number of minutes in a given month.

B. "Downtime" is the total accumulated minutes that are part of Minutes in the Month that have no Virtual Machine Connectivity. Downtime excludes Announced Single Instance Maintenance.

C. "Monthly Uptime Percentage" is calculated by subtracting from 100% the percentage of Minutes in the Month in which any Single Instance Virtual Machine using premium storage for all disks had Downtime.

D. Monthly Uptime % = 100% - (Minutes in the Month - Downtime) / Minutes in the Month

E. The following Service Levels and Service Credits are applicable to Customer' s use of Single-Instance Virtual Machines:

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

## VPN Gateway

We guarantee 99.9% availability for each Basic Gateway for VPN or Basic Gateway for ExpressRoute.

We guarantee 99.95% availability for each Standard Gateway for VPN or Standard Gateway for ExpressRoute.

We guarantee 99.95% availability for each High Performance Gateway for VPN or High Performance Gateway for ExpressRoute.

### **i. Monthly Uptime Calculation and Service Levels for VPN Gateway**

A. "Maximum Available Minutes" is the total accumulated minutes during a billing month during which a given VPN Gateway has been deployed in an Azure subscription

B. "Downtime" is the total accumulated Maximum Available Minutes during which a VPN Gateway is unavailable. A minute is considered unavailable if all attempts to connect to the VPN Gateway within a thirty-second window within the minute are unsuccessful.

C. "Monthly Uptime Percentage" for a given VPN Gateway is calculated as Maximum Available Minutes less Downtime divided by Maximum Available Minutes in a billing month for the VPN Gateway. Monthly Uptime Percentage is represented by the following formula:

Monthly Uptime % = (Maximum Available Minutes - Downtime) / Maximum Available Minute

D. The following Service Levels and Service Credits are applicable to Customer' s use of each VPN Gateway:

#### **Basic Gateway for VPN or ExpressRoute**

Monthly Uptime Percentage	Service Credit
<99.9%	10%
<99%	25%

## Standard Gateway for VPN or ExpressRoute / High Performance Gateway for VPN or ExpressRoute

Monthly Uptime Percentage	Service Credit
<99.95%	10%
<99%	25%

### App Service

We guarantee that Apps running in a customer subscription will be available 99.95% of the time. No SLA is provided for Apps under either the Free or Shared tiers.

#### i. **Additional Definitions**

- A. **"Deployment Minutes"** is the total number of minutes that a given App has been set to running in Azure during a billing month. Deployment Minutes is measured from when the App was created or Customer initiated an action that would result in running the App to the time Customer initiated an action that would result in stopping or deleting the App.
- B. **"Maximum Available Minutes"** is the sum of all Deployment Minutes across all Apps deployed by Customer in a given Azure subscription during a billing month.
- C. **"App"** is a Web App, Mobile App, API App or Logic App deployed by Customer within the App Service, excluding apps in the Free and Shared tiers.
- D. **"Downtime"** the total accumulated Deployment Minutes, across all Apps deployed by Customer in a given Azure subscription, during which the App is unavailable. A minute is considered unavailable for a given App when there is no connectivity between the App and Microsoft's Internet gateway.

The Monthly Uptime Percentage is represented by the following formula:

E. The following Service Levels and Service Credits are applicable to Customer' s use of the Websites Service:

Monthly Uptime Percentage	Service Credit
<99.95%	10%
<99%	25%

**Additional Terms:** Service Credits are applicable only to fees attributable to your use of Web Apps, Mobile Apps, API apps or Logic Apps and not to fees attributable to other types of apps available through the App Service, which are not covered by this SLA.

#### Virtual Machine Scale Sets

Virtual Machine Scale Sets is a free service, therefore, it does not have a financially backed SLA itself. However, if the Virtual Machine Scale Sets includes Virtual Machines in at least 2 Fault Domains, the availability of the underlying Virtual Machines SLA applies. See the [Virtual Machines SLA](#) for more details.