Service terms and description

Keystone

NetApp
February 11, 2024

Table of Contents

Keystone service terms and descriptions

NetApp Keystone is guided by the following service terms and descriptions.

- Keystone services are available for a minimum of one year and up to three years. After the initial term, the service is renewable on an annual basis. Capacity can be increased in increments as small as 1 TiB.
- The minimum capacity is 100 TiB per site, and each site can have one or more clusters to meet the minimum capacity requirement. In a partner-operated model, subscriptions with flexible minimums are created for a customer, per site and across service levels.
- The 100 TiB capacity can be one single service level or a combination of levels.
- Tenant subscriptions are limited to service levels that partners are subscribed to.
- 20% of burst capacity is available at each site; any burst usage is billed only for that billing period. If you need an additional burst requirements that is greater than 20%, contact support.
- Committed capacity or service levels cannot be altered during a contract term.
- Increasing capacity or changing to higher service level during term is allowed; however, moving from a higher service level to a lower level is not permitted.
- Any change request in the last 90 days of the term requires the customer to renew the service for a minimum of one year.

Keystone service capacity definitions

The NetApp Keystone service capacities include:

Logical capacity

This is the data placed into the Keystone infrastructure by a customer. All Keystone capacities refer to a logical capacity.
For example, if a 1 TiB file is stored on the Keystone infrastructure, then at least 1 TiB of capacity must be purchased.

Committed capacity

The minimum logical capacity billed each month for the duration of the term:

- Capacity is committed to each service level.
- Committed capacity cannot be decreased during the term.
- Committed capacity and additional service levels can be added during the term.

Changes to committed capacity

During the tenure of a subscription, you can change the committed capacities. However, there are certain preconditions:

- The committed capacity cannot be decreased
- The committed capacity cannot be increased 90 days prior to the expiry of your subscription, unless the subscription is to be renewed for an additional 12 month term.
- You can request changes to committed capacity through Keystone Success Manager.

Burst capacity

This is the logical capacity that has exceeded the committed capacity. Note the following points:

- Keystone service provides 20% capacity in excess of the committed capacity.
- Burst capacity can be consumed on an elastic basis and is charged on a daily basis of the consumed average.
- Burst capacity up to 20% is charged at a same rate as the committed capacity.
- Burst capacity greater than 20% of committed is charged at a premium rate. Contact support for any additional burst requirements greater than 20%.

Consumed/provisioned capacity

Consumed capacity refers to the capacity in TiB of storage currently being consumed on the service. Keystone service considers the sum of the provisioned sizes (not the logical or physical capacity used) of all volumes on a particular service level to be considered as consumed capacity for that service level. This includes:

- The capacity that is provisioned through the creation, modification, deletion, or potential auto-growth of volumes.
- The Snapshot copies and clones.

The amount of data stored within provisioned capacity, or the amount of data actually written to disk is not considered.

Billed capacity

Monthly bill = (committed capacity [TiB] * committed rate [$/TiB]) + (daily average provisioned burst capacity [TiB] * burst rate [$/TiB]). The monthly bill contains a minimum charge based on the committed capacity.

The monthly bill varies beyond the minimum charge based on daily average burst capacity consumption. For more information on billing, see Keystone billing.

Service levels

Keystone offers data storage capacity at pre-defined performance service levels (service levels) or rate plans on a subscription basis. Each volume managed by the Keystone services is associated with a service level.

Each service level is defined by its I/O density, that is IOPS/TiB/volume. This is the ratio of performance (input/output operations per second [IOPS]) and allocated storage capacity (TiB) which is IOPS/TiB at average latency per volume.

The I/O density for a volume is calculated at an hourly interval to report the peak I/O density. I/O density reports for volumes are generated monthly.

You select service levels based on your storage environment, and storage and consumption needs. The base service levels are available to you by default. Other service levels are additionally available, if you have opted for specific services. For example, for the data protection service, the *Data-Protect <base-service-level>* service level is assigned to your subscription.

The base service levels for file, block, and object storage types are listed here.

Service levels for file storage

Supported protocols: NFS, CIFS, iSCSI, and FC

Service level	Extreme	Premium	Standard	Value
Workload type	Analytics, databases	VDI, virtualization apps, Software dev	File shares, web servers	Backup
Maximum IOPS/logical TiB allocated per volume	12,288	4,096	512	N/A
Target IOPS/logical TiB allocated per volume	6,144	2,048	128	N/A
Maximum MBps/logical TiB allocated per volume	384	128	16	N/A
Average Latency	<1 ms	<2 ms	<17 ms	N/A
Block size	32K			

More on service levels for file storage

The base service level metrics depend on the following conditions:

- The service levels for file and block storage support ONTAP 9.7 and later.
- IOPS/TiB/volume, MBps/TiB/volume, and latency values for service levels are based on the amount of logical data stored in the volume, 32KB block size, and a random combination of 70% read and 30% write IO operations.
- Actual IOPS/TiB/volume and MBps/TiB/volume may vary based on the actual or assumed block size, system workload concurrency, or input-output operations.
- Latency does not include the following:
  - application or host latency
  - customer network latency to or from the controller ports
  - overheads associated with the data transfer to the object store in case of FabricPool
  - latency automatically applied by QoS to keep IO within service level maximums
- Latency values are not applicable to MetroCluster write operations. These write operations are dependent on the distance of remote systems.
- If one or more volumes on a storage system do not have an AQoS policy assigned, then these volumes are considered as non-compliant volumes, and no target service levels are applicable for those systems.
- *Expected IOPS* is targeted for FabricPool only if the tiering policy is set to "none" and no blocks are in the cloud. *Expected IOPS* is targeted for volumes that are not in a SnapMirror synchronous relationship.
- Workload IO operations need to be balanced across all deployed controllers, as determined by the

Keystone order.

Service levels for block storage

Supported protocols: FC and iSCSI

Service level	Extreme	Premium	Standard
Workload type	HPC	Video surveillance	Backup
Target IOPS/TiB	N/A		
Max IOPS/TiB	5,500	4,000	N/A
Max throughput MBps (32KB/IOP)	43	31	N/A
Average Latency	<0.5 ms	<0.5 ms	N/A
Block size	32K		

Service levels for object storage

Supported protocol: S3

Service level	Object
Workload type	Media repository, archiving
Target IOPS/TiB	N/A
Max IOPS/TiB	N/A
Max throughput MBps (32KB/IOP)	N/A
Average Latency	N/A

Service level metrics and definitions

The following terms and definitions are used within the NetApp Keystone:

- **GiB, TiB, and PiB.** Measurements of data storage capacity using base of 1024 (1 GiB = 1024^3 bytes, 1 TiB = 1024^4 bytes, and 1PiB = 1024^5 bytes).
- **IOPS/TiB.** The protocol operations per second requested by the application divided by the allocated logical size of the volume.
- **Availability** is measured as a percentage of number of I/O requests successfully responded to by the service, divided by total number of I/O requests made of the service, measured at the service demarcation, in a given month, not including scheduled service downtime or unavailability of required facilities, network or other services to be provided by customer.
- **Durability** is the percentage of data accessed without loss of fidelity, excluding customer-caused deletion or corruption.
- **Target IOPS per TiB.** The guaranteed IOPS for all I/O requests made to a volume before the target IOPs per TiB threshold is reached. Performance on the volume is capped at the selected IOPS per TiB.

The target IOPS per TiB performance metric is calculated based on the logical consumed capacity in TiB.

- **Latency.** Time to service an I/O request received from a client, measured at the service demarcation (storage controller I/O port).

Adaptive Quality of Service in Keystone

Storage quality of service (QoS) is a critical technology that ensures that applications obtain consistent and predictable performance. Without QoS, certain workloads, such as those for booting of multiple systems, may consume most or all of of the resources for a period, and affect other workloads. For information about QoS, see Guarantee throughput with QoS overview.

Adaptive QoS

Adaptive QoS (AQoS) is used by Keystone services to dynamically maintain the IOPS/TiB ratio based on the volume size. For information about AQoS policies, see About adaptive QoS.

Keystone provides you with AQoS polices that you can set up once your cluster is in production. You should ensure that all your volumes are associated with the correct AQoS policies that are already created and available in your system.

An ONTAP volume is non-compliant if it does not have an AQoS policy applied. A volume without a QoS policy is the last on the list of priority for the system to provide any available input-output operations. However, if any input-output operations are available, then the volume could consume all available IOs.

If you have not applied AQoS policies to your volumes, those volumes will be measured and charged at the highest service level as per your subscription. This may result in unintended burst charges.

Adaptive QoS settings

The Adaptive QoS (AQoS) settings vary with service levels.

Policy name	**Extreme** and **Extreme Tiering**	**Premium** and **Premium Tiering**	**Standard**	**Value**
Expected IOPS	6,144	2,048	128	64
Expected IOPS Allocation	Allocated space			
Peak IOPS	12,288	4,096	512	128
Peak IOPS Allocation	Allocated space			
Block Size	32K			

Keystone billing

NetApp Keystone enables predictable and upfront pricing for your storage subscription.

If you prefer operational expenditures (OpEx) consumption model to capital expenditure (CapEx) or leasing, you can opt for the Keystone pay-as-you-grow model for your flexible and scalable consumption needs.

Keystone provides you with the following billing facilities:

- You can pay based on IOPS and latency committed capacity to meet various workload needs. The different performance service tiers - Extreme, Premium, Standard, and Value enable you to manage your storage based on your purchased service level for your Keystone services.
- It presents predictable billing for the committed capacity and pay-per-use for variable (burst) capacity usage.
- You can select a bundle price for hardware, core OS, and support for one $/TiB price. You have a single invoice for each storage type, file, block, object, or cloud storage services.
- Select a flexible term for the services and payment: You can opt for 12 months, 100TiB, or more per site. Thereafter you can auto renew for 12 months or go month-to-month.

Keystone billing is based on committed capacity and variable burst consumption.

For information about committed and burst capacity usage, see Keystone service capacity definitions.

Billing based on committed capacity

Committed capacity refers to the capacities for various services in a single subscription, agreed upon by the parties involved (NetApp/partner and customer). This capacity is stated on each Keystone order and is billed, regardless of the actual consumption.

Metering of consumed capacity

As a part of the Keystone service deployment, NetApp continuously monitors and measures the consumption of the service. At least once in every five minutes, a consumption record is generated by the system, detailing the current consumed capacity for the subscription. These records are aggregated over the billing period to generate invoices and usage reports.

Billing based on burst consumption

When the consumed capacity is greater than the committed capacity for a given service level, burst consumption is recorded, and charges are applied accordingly. This process occurs for each consumption record generated. Burst consumption, therefore, is a reflection of both the amount and tenure of your over-consumed capacities on top of your committed capacities.

Billing schedules

Keystone services are billed monthly and yearly.

Monthly billing

Invoices are sent monthly. For the month in which the services are availed, an invoice is sent in the next month. For example, the invoice for the services you have used in January is delivered at the beginning of February. This invoice includes the charges for the committed capacity and if applicable, any burst usage.

Annual billing

An invoice is generated at the beginning of each subscription year for the minimum payment of the committed capacity. It is generated on the start date of the subscription.
Another invoice is sent at the end of a subscription quarter, summing up the applicable charges of any burst usage accrued in that quarter.
If the committed capacity is changed during a subscription, then an invoice is sent on the same day the change in the committed capacity is effective, for the prorated minimum payments for the rest of that subscription year.

Grace period in burst billing

During the first 60 days of the activation of a subscription, any burst consumption that you accrue is recorded, but not charged on your Keystone service invoice. This grace period grants you the time to configure your storage with the appropriate Adaptive Quality of Service (AQoS) settings so that it can be properly monitored and billed. On configuring your storage correctly, you can avoid any unexpected burst charges. Any burst usage accrued after the initial grace period of 60 days will be billed on your next invoice. This grace period does not affect any modifications or renewals to your existing subscriptions.

Miscellaneous scenarios for Keystone billing

There are several scenarios for Keystone billing and you should be familiar with those scenarios.

Billing for cloned volumes

If volumes are cloned in ONTAP and you use them for backing up and restoring your data, you can continue using the clones without any additional payments. However, cloned volumes used for any other purpose in your business for an extensive duration are charged.

Note the following:

- Clone volumes are free from charging as long as their size is less than 10% of the parent volume (the physical capacity used in the clone volume compared to the physical capacity used in the parent volume).
- There is no 24-hour grace period for cloned volumes, only the size of the clone is considered.
- Once the clone volume exceeds 10% of the physical size of the parent, the clone is billed as a standard volume (logical used capacity).

Billing for MetroCluster

Advanced Data Protection uses NetApp MetroCluster to mirror data between two physically separated clusters. On MetroCluster mirrored aggregates, data is written twice, once on each cluster. Keystone service charges for consumption on each side independently, resulting in two identical consumption records.
If you monitor your clusters through ONTAP System Manager (System Manager) or Active IQ Unified Manager (Unified Manager), you might see a discrepancy between the consumption reported on these tools and Keystone services. System Manager and Unified Manager do not report volumes on the mirrored (remote) cluster, and in doing so, reports half the consumption metrics that the Keystone service reports.

For Example:
Site A and Site B are set up in a MetroCluster configuration. When a user creates a volume of 10TB in site A, an identical volume of 10TB is created in site B. Keystone distinguishes both the volumes and records an additional 10TB of consumption in each site, for a total increase of 20TB. System Manager and Unified Manager reports a 10TB volume created in site A.

Billing for temporary volumes

Occasionally, temporary (TMP) volumes are created by ONTAP when moving volumes. These temporary volumes are short-lived, and the consumption on these volumes is not measured for billing.

Billing and adaptive QoS policies

Keystone measures consumption based on service levels. Each service level is associated with a specific adaptive quality of service (QoS) policy. During deployment, you will be informed of the details of each QoS policy for your subscribed Keystone services. During storage management operations, ensure that your volumes have the appropriate QoS policies assigned as per your subscribed service levels, to avoid unexpected billing.
For more information about QoS policies in ONTAP, see Guarantee throughput with QoS overview.

Billing for SnapMirror destinations

The pricing for the SnapMirror destination volume governed by the QoS policy for the service level assigned on the source. However, if the source does not have an associated QoS policy, the destination is billed based on the lowest available service level.

Billing for FlexGroups

FlexGroups are billed based on the adaptive QoS policy of the FlexGroup. The QoS policies of its constituents are not considered.

Billing for LUNs

For LUNs, usually the same billing pattern is followed as for the volumes that are governed by QoS policies. If separate QoS policies are set on LUNs, then:

- The size of the LUN is counted for consumption according to the associated service level of that LUN.
- The remainder of the space in the volume, if any, is charged according to the QoS policy of the service level set on the volume.

Billing for FabricPool usage

If data is tiered from a Keystone system to ONTAP Simple Storage Service (S3) object storage or NetApp StorageGRID, then the consumed capacity on the hot tier (Keystone system) will be reduced by amount of data that has been tiered off, impacting the resultant billing. This is regardless of the fact whether the ONTAP S3 storage or StorageGRID system is covered by the Keystone subscription.

For tiering your data to any third party object storage, contact your Keystone Success Manager.

For information on the use of FabricPool technology for your Keystone subscriptions, see Tiering.

Billing for system and root volumes

System and root volumes are monitored as a part of the overall monitoring of the Keystone service but are not counted or billed. The consumption on these volumes is exempted for billing.

Data protection

NetApp Keystone data protection service can back up your data and is able to recover it if required. The available data protection services are:

- Snapshots of disks and shares
- Backups of disks and shares (requires data protection service as part of the subscription)

- Disaster recovery for disks and shares (requires data protection service as a part of the subscription)

Backup and disaster recovery services are available as add-on services, while snapshot is available as a part of the basic storage service.

If you have subscribed to data protection services, specific service levels are assigned to your subscriptions, for example *Data-Protect Extreme*, *Data-Protect Premium*, and so forth. For information, see Reference charts for data protection.

	Single Region snapshots (Available as a part of the basic storage service)	**Multi-region Backup (data protection add-on)**	**Multi-region Disaster Recovery (data protection add-on)**
Use case	Mitigate the risk of user or application data deletion or corruption, not against infrastructure loss or failure	Mitigate the risk of complete loss of data on the primary volume due to infrastructure loss or failure	Mitigate the risk of complete loss of data on the primary volume due to infrastructure loss or failure with a recovery time objective
Policy	Hourly, daily, weekly, and monthly	Number of backups to retain based on hourly, daily, weekly, and monthly snapshots	1 hour, 4 hours, and daily
Topology	Source only	Backup	Async replication target
Target replication service level[1]	n/a	Standard	Same as primary

[1]Additional storage capacity to be subscribed

Subscription to a basic Keystone service does not automatically back up your data. You should subscribe to add-on data protection services and configure your system for data backup and disaster recovery services. If your storage system is not managed by Keystone services, NetApp can still support protecting the data on your storage system and help in connecting it with your Keystone services. However, NetApp is not responsible for any backup failures.

Tiering

NetApp Keystone service includes a tiering capability that identifies less frequently used data and tiers it to a cold storage that is owned, deployed, and managed by NetApp.

The tiering capability leverages the NetApp FabricPool technology that enables automated tiering of data to low-cost object storage tiers either on or off premises. With this capability, infrequently accessed data is automatically tiered to a lower cost storage either on premises or in the cloud, based on the services agreed upon.

Partners and tenants can avail this capability easily by opting for the two preconfigured service levels, the Extreme-tiering and Premium-tiering service levels while provisioning their storage. The Extreme-tiering has the same QoS policies as the Standard, Extreme, and Premium service levels.

The add-on tiering capability is available only with Extreme and Premium service tiers. NetApp assumes 25%

of data is hot and 75% is less frequently used and can be moved to a cold storage. Billing is determined based on the duration per volume is in each service level.

The following features are enabled:

- You can create reports of the inactive data for your disks and file shares and decide upon whether to change the service level. On moving or changing the tiering policy, the latencies can be higher if data is accessed from cold tier.
- You can change the service level of the volumes from Extreme and Premium to Extreme-tiering and Premium tiering respectively, provided that the destination tiering is enabled on the cluster.
- Likewise, you can change the tiering service levels to non-tiering for your volumes.
- Enable and disable backups for a volume on a tiering service level.
- Enable and disable disaster recovery for a volume on a tiering service level.

For information on billing, see Billing for FabricPool usage.

Non-returnable disk offering

As a part of NetApp Keystone services, NetApp extends the non-returnable disk (NRD) offering.

If you purchase the NRD offering for Keystone, NetApp does not recover the physical storage media used during the entire service tenure because of support and maintenance activities, or at service termination when NetApp otherwise recovers all of its physical assets used in the delivery of the service.

If you have purchased this service, note the following:

- Even on purchasing this service, you can opt for NetApp to recover the physical storage media.
- In case NetApp is not responsible for recovering the media, you are entitled to destroy the storage media or disks used in the delivery of the Keystone service at the end of the service tenure.
- You can add, modify, or terminate the NRD offering during the renewal of the subscription and not in the middle of the tenure.
- The cost associated with the NRD offering changes based on the committed capacity of the subscription. That is, if you opt to increase your committed capacity in the middle of the subscription period, the cost of NRD is revised likewise. The increase will be proportional to the increase in the committed capacity.
- You can retain only the physical storage media used in your service. Controllers, shelves, cables, switches, network cards, and any other equipment owned by NetApp will be recovered by NetApp.

U.S. Citizen Support (USCS)

United States Citizen Support (USCS) is an add-on offering for NetApp Keystone Subscriptions. It entitles you to receive delivery and support of ongoing Keystone services from U.S. citizens on U.S. soil.

Read the following sections to understand which elements of your subscriptions are bound by this add-on service; and are provided under the terms of NetApp Keystone Agreement. [1]

NetApp Global Services Support Center monitoring

NetApp Global Services and Support Center (GSSC) monitors the health of your products and subscribed services, provides remote support, and collaborates with your Keystone Success Manager. All personnel monitoring the products associated with the relevant Keystone subscription orders are U.S citizens operating on U.S. soil.

Keystone Success Manager

The Keystone Success Manager is a U.S. citizen operating on U.S. soil. Their responsibilities are specified in your NetApp Keystone Agreement.

Deployment activities

Where available, onsite and remote deployment and installation activities are conducted by U.S. citizens on U.S. soil. [2]

Support

Where available, the necessary onsite troubleshooting and support activities are conducted by U.S. citizens on U.S. soil. [2]

Keystone services powered by Equinix

NetApp has partnered with Equinix for hosting NetApp Keystone services in an Equinix data center to ensure the delivery of a unified solution for you.

Keystone services powered by Equinix is unchanged from the standard Keystone services offering.

In addition to the standard Keystone services offering, you will need to select an Equinix datacenter to host your Keystone equipment.

Copyright information

Trademark information

[1] The services and offerings described here are subject to, and limited and governed by a fully-executed Keystone Agreement.

[2] Availability of appropriate personnel for onsite activities is dependent of the geographical location at which the Keystone systems are deployed.