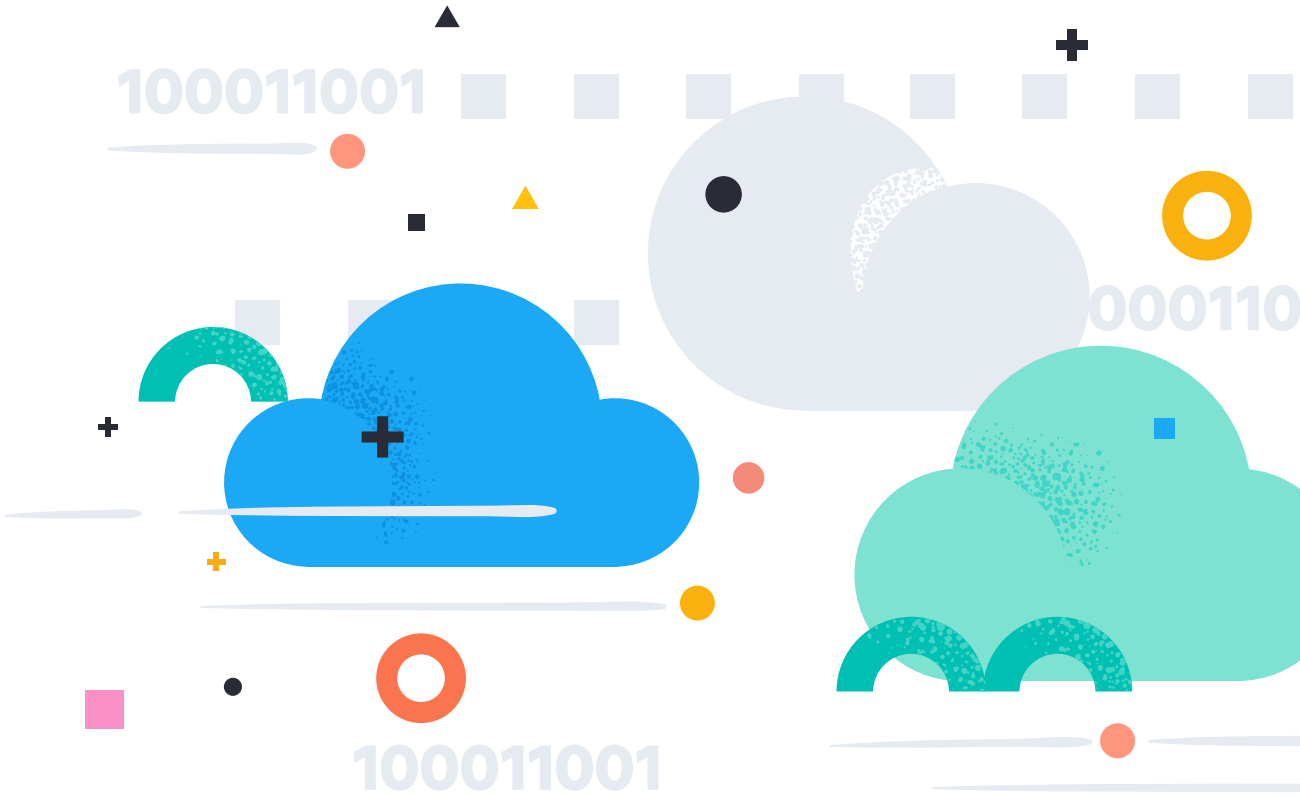
elastic

Search. Observe. Protect.

Tips and tricks for getting the most out of Elastic Cloud

elastic.co →

Table of Contents



Introduction

Elastic Cloud brings the power of the Elastic Stack — including Elastic Enterprise Search, Observability, and Security — to the cloud, allowing you to quickly and easily search your environment for information, analyze data to observe insights, and protect your technology investment.

Whether you're new to Elastic or you're an existing user, this guide will show you how to get the most out of Elastic Cloud, including:

- Tips that will make your life easier as you ramp up with Elastic Cloud
- Tricks from Elastic pros on setting up your Elastic Cloud deployment
- How to connect with the Elastic community

Getting started: Tips for new Elastic Cloud users

Choosing the right deployment configuration

Elastic Cloud supports a wide range of configurations. With this flexibility comes great freedom, but also the first rule of deployment planning: your deployment needs to match the workloads that you plan to run. Specifically, this means accounting for three things:

- What use case will your Elastic Cloud deployment be used for?
- Does your data need to be highly available?
- Do you know when to scale?

What use case will your Elastic Cloud deployment be used for?

We've streamlined the Elastic Cloud setup process by prepackaging resources such as VCPU, RAM, and storage into solution templates, taking the guesswork out of provisioning your deployment. The currently available solution templates are:

Elastic Stack

This solution template provides you with the flexibility to create a deployment using any of our **hardware profiles** so you can reliably and securely search, analyze, and visualize data in real time.

Elastic Enterprise Search

This solution template provides you with a **CPU-optimized deployment** to power modernized search experiences for your websites, applications, and workplaces.

Elastic Observability

This solution template provides you with an **I/O-optimized deployment** to run unified analysis across logs, metrics, APM, and uptime monitoring.

Elastic Security

This solution template provides you with an **I/O-optimized deployment** for threat prevention, detection, and response through a single UI.

You can start with one of the solution templates and then add more instances to your existing deployment at any point in time simply by moving a slider under deployment configurations within your cloud console. Keep in mind that Elastic Cloud trial users need to purchase an Elastic Cloud subscription to add instances.

Does your data need to be highly available?

With Elastic Cloud, your deployment can be spread across as many as three separate availability zones. Why this matters:

- Availability zones can and do encounter issues. There can be internet outages, earthquakes, floods, or other events that affect the availability of a single zone. With a single availability zone, you have a single point of failure that can bring down your deployment.

- Multiple availability zones help your deployment remain available. This includes your Elasticsearch cluster, provided that your cluster is sized so that it can sustain your workload on the remaining availability zones and that your indices are configured to have at least one replica.

- Multiple availability zones enable you to perform changes to resize your deployment with zero downtime.

Elastic Cloud trial users can access up to two zones. To gain access to three zones, you must purchase an Elastic Cloud subscription.

Do you know when to scale?

Knowing how to scale your deployment is critical, especially when unexpected workloads hit. Scaling with Elastic Cloud is easy: simply log in, click on your deployment, navigate to Edit to visit the configuration page, and drag the memory sliders to the desired levels. CPU, disk I/O, and storage are scaled up proportionally with memory as your cluster is resized.

Elastic Cloud trial users will need to purchase an Elastic Cloud subscription to scale a deployment.

Migrating existing Elastic data to Elastic Cloud

You're likely wondering how to get your existing Elasticsearch data into your new Elastic Cloud infrastructure. In addition to easily creating as many new deployments with Elasticsearch clusters as you need, you have several options for moving your data over. Choose the option that works best for you:

- Index your data from the original source if you still have a copy of the original data.
- Reindex from a remote cluster, which rebuilds the index from scratch.
- Restore from a snapshot, which copies the existing indices.
- Migrate from AWS Elasticsearch Service.

One of the many advantages of Elastic Cloud is that you can spin up a deployment quickly, try something out, and then delete it if you don't like it. This flexibility provides the freedom to experiment while your existing production cluster continues to work.

For detailed step-by-step migration instructions, please refer to our **migration guide.** If you're migrating from AWS Elasticsearch Service, please refer to **our documentation on migrating from Amazon Elasticsearch Service.**

Please visit our subscription comparison page for further information on which features are available at each subscription level.

If, at any time during your monthly subscription with Elastic Cloud, you decide you need features at a higher subscription level, you can easily make changes to your subscription. You can either upgrade to a higher subscription level or downgrade to a lower one.

Which features are available under each subscription level?

Some features are only available for particular Elastic Cloud subscription levels:

Gold

Gold subscription and above on Elastic Cloud provides access to:

- ✔ **Alerts with actions** via email, PagerDuty, ServiceNow®, Slack, Jira, IBM Resilient, webhooks
- ✔ Configurable retention policy for stack monitoring
- ✔ **Custom plugins**

Platinum

Platinum subscription and above on Elastic Cloud provides access to:

- ✔ **Machine learning** and anomaly detection on time series
- ✔ **Single sign-on** (SAML, OpenID Connect, Kerberos)
- ✔ JDBC, ODBC, and Tableau connector

Enterprise

Enterprise subscription on Elastic Cloud provides access to:

- ✔ **Searchable snapshots**

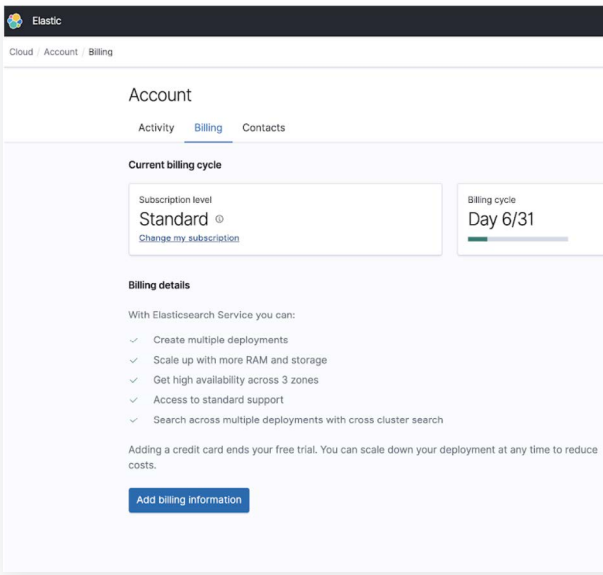
Please visit our **subscription comparison page** for further information on which features are available at each subscription level.

If, at any time during your monthly subscription with Elastic Cloud, you decide you need features at a higher subscription level, you can easily make changes to your subscription. You can either upgrade to a higher subscription level or downgrade to a lower one.

HOW TO

To change your subscription level:

1. Log in to the Elastic Cloud console
2. Click on your user icon on the top right corner
3. Select Account & Billing
4. Click the Billing tab
5. Click Add Billing Information
6. Under Subscription level box select Change my subcription
7. Save your changes



Adding a credit card ends your free trial, taking effect immediately, so to fully utilize the 14-day trial, we recommend entering your credit card information near the end of the 14 days. Becoming a paying customer also unlocks the trial memory restrictions. When you switch to a higher subscription level, the change takes effect immediately. However, changes to a lower subscription level take effect in the next billing cycle, with the exception of customers in their first billing cycle, who may immediately change to a lower subscription level.

Choosing the right subscription level for your required support service-level agreement (SLA)

When you decide to add your credit card and become a paying Elastic Cloud customer, be sure to choose a subscription level that includes the features you are going to use as well as the level of support you need.

Standard

Standard subscription comes with email-based support from the Elastic team to help keep your cluster green and available, from analysis of your cluster state to actions you can perform to stabilize your cluster. Elastic Cloud Standard support does not have a guaranteed initial or ongoing response time, but we do strive to engage on every issue within three (3) business days. We do not offer weekend coverage, so we are able to respond Monday through Friday only.

Gold

Gold subscription comes with both phone- and web-based support from the Elastic team to get help with break/fix, technical guidance, and best practices. Elastic Cloud Gold support has a guaranteed initial or ongoing response time of four (4) business hours for tickets with critical severity, with coverage between 8:00 a.m. and 6:00 p.m. in the time zone shown on the order form.

Platinum and Enterprise

Platinum and Enterprise subscriptions come with both phone- and web-based support from the Elastic team to get help with memory usage, architecture advice, scaling, upgrading, and more. Elastic Cloud Platinum support has a guaranteed initial or ongoing response time of one (1) business hour for tickets with critical severity, with 24/7/365 coverage.

Adding monitoring, operational, and billing contacts to receive alerts and more

If different people from your organization are involved in billing and operations, you can specify their relevant contacts in addition to the primary email address of your account.

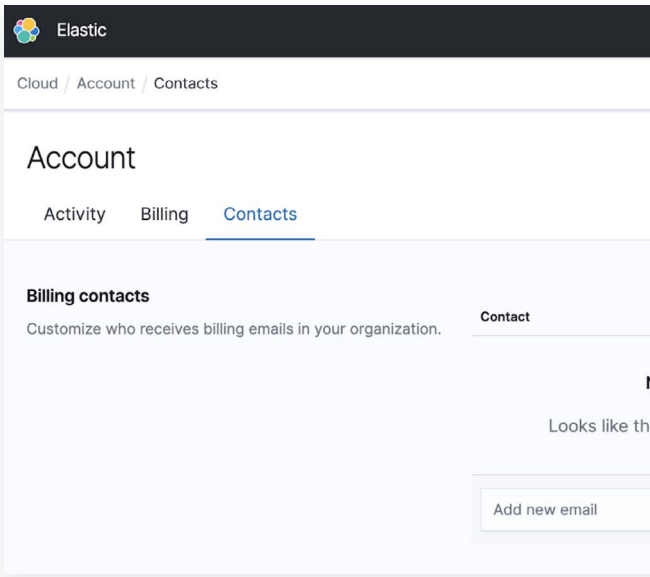
- Operational contacts can only receive operational notifications, such as out-of-memory alerts.
- Billing contacts can only receive receipts and billing notifications.
- Monitoring whitelist email addresses can only receive monitoring alerts sent from Elasticsearch Watcher as well as Kibana Alerts & Actions.

All operational, billing, and monitoring whitelist email contacts do not have access to the Elastic Cloud console. Currently, only the primary email address associated with your account is able to log in to the Elastic Cloud console. By default, the primary email address is used to sign up for Elasticsearch Cloud and to log in.

HOW TO

To update billing and operational contacts or whitelist an email address:

1. Log in to the Elastic Cloud console
2. Click on your user icon on the top right corner
3. Select Account & Billing
4. Click the Contacts tab



Multiple email addresses can be specified for each category. Operational and billing contacts become effective immediately and no further confirmation of the email addresses is required. Monitoring whitelist emails require verification via a link sent to the email address before being able to receive monitoring alerts.

Moving to marketplace billing

Cloud provider marketplace billing is a convenient way to subscribe to Elastics Cloud if you already use GCP, AWS, or Azure. With marketplace billing, you can add an Elastic Cloud subscription to your regular cloud provider bill and don't have to supply any additional credit card information to Elastic.

A few things to keep in mind when using marketplace billing:

- There is no trial offered with marketplace billing. Billing starts when you subscribe to Elastic Cloud.

- Previous Elastic Cloud accounts cannot be converted to marketplace billing. If you already have an account, you must use a different email address when you sign up for a subscription through the marketplace or change the email address of the existing Elastic Cloud account.

- Pricing for an Elastic Cloud subscription through a cloud provider marketplace differs from our direct billing model and follows the pricing outlined on the Elastic Cloud page in each cloud provider marketplace.

If you want to convert an existing Elastic Cloud account to marketplace billing, you must migrate your deployments over to the new account. To migrate, use a **custom repository** to take a snapshot and then restore that snapshot to a new deployment under your marketplace account. If you also want to use the same account email with marketplace billing, you must first change the primary email on your existing account before setting up your new marketplace account:

HOW TO

1. Log in to the Elastic Cloud console
2. Click on your user icon on the top right corner, select User Profile
3. Click the Edit button next to the Email Address input box
4. Enter a new email and your current password
5. Click Save

Getting help via documentation, free training, and support

When you're stuck, you have many options to get help from directly in the Elastic Cloud console, from our comprehensive documentation and free online training videos to online support! Our documentation is split into a number of different categories:

Elastic Stack

Elastic Enterprise Search

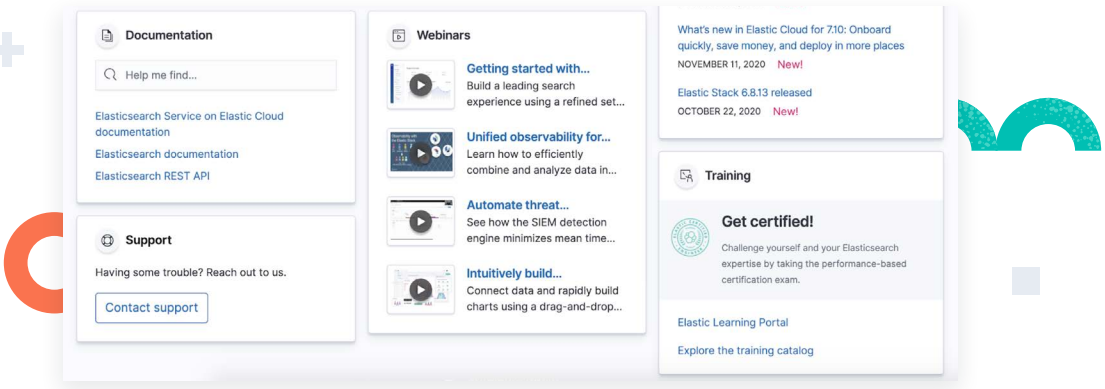
Elasticsearch

Elastic Observability

Elastic Cloud

Elastic Security

Kibana



From the training catalog, you can see that we have numerous free on-demand courses available:

- Quick Start guides to help you create your own Elastic Cloud cluster and explore data sets in less than 15 minutes

- Fundamentals training, which includes expertly designed materials, engaging demos, and hands-on lab exercises

With your Elastic Cloud subscription, you get access to support from the creators of Elasticsearch, Kibana, Beats, Logstash, and our exclusive features such as machine learning, Canvas, Elastic APM, index lifecycle management, Elastic App Search, Elastic Workplace Search, and more. Critical severity issues must be filed via the web-based support portal. Non-emergent issues may be filed via email or the web-based support portal. All email-based production tickets will be treated as non-emergent issues.

- **Elastic Cloud Standard** support is provided by email or through the Elastic Support Portal. The main focus of support is to ensure your Elasticsearch Service deployment shows a green status and is available.

- **Elastic Cloud Gold, Platinum, and Enterprise** support is provided by email or through the Elastic Support Portal. You get support not only for making sure your deployment is green and available, but also for how-to and development questions.

Getting the most out of Elastic Cloud: Tricks from Elastic pros

What is the Data Transfer Fee and how can I reduce it?

AWS Data Transfer in (per GB)	Free		$0.00
AWS Data Transfer Inter-Node (per GB)	$0.0160	1047	$16.75
AWS Data Transfer Inter-Node (per GB)	$0.0160	-1047	-$16.75
AWS Data Transfer Out (per GB)	$0.0320	47	$1.50
AWS Data Transfer Out (per GB)	$0.0320	-47	-$1.50

Data transfer costs account for the volume of data (payload) going into, out of, and within the Elastic Cloud deployments.

We are set to meter and bill data transfer using three dimensions:

- Data in (free)
- Data out
- Data intra-deployment

Data in accounts for all of the traffic going into the cluster. That includes index requests with data payload, as well as queries sent to the cluster (although the byte size of the latter is typically much smaller).

Data out accounts for all of the traffic coming out of the cluster. That includes search results as well as monitoring data sent out of the cluster. The same rate applies regardless of the destination of the data, whether to another region, to the Internet, or to the same region but a different account.

Data intra-deployment accounts for all of the traffic sent between the components of the deployment. That mostly includes the data sync between nodes of a cluster spread across different availability zones, which is managed automatically by Elastic Cloud cluster sharding. It also includes data related to search queries executed across multiple nodes of a cluster. Note that single-node Elastic Cloud clusters may also incur intra-cluster charges accounting for data exchanged with Kibana nodes or other nodes such as machine learning or APM. These charges are expected to be lower in such cases.

The data transfer usage is calculated similarly to the storage API requests, in that it is summed up to a cumulative amount within a billing cycle.

Currently, the rates for these three dimensions are:

- Data in: $0 per GB transferred (free)
- Data out: $0.032 per GB transferred
- Data Intra-deployment: $0.016 per GB transferred

We also provide a free allowance of 100GB/month, which covers data out and data intra-deployment separately, and across all the deployments of the account. Once this threshold is passed, a charge will apply for any data transfer used in excess of the 100GB/month free allowance.

Data transfer out of deployments and between nodes of the cluster is harder to control, as it is a function of the use case employed and cannot always be tuned. Some cases of batch queries executed at frequent intervals may be revisited, if applicable. Also, oversharding may cause high data transfer fees due to excessive shard balancing and replication.

Given the varied ways of using Elasticsearch in different use cases, it is hard to predict the exact data transfer costs for your account. For this reason, we have implemented two “grace” bills that will show the amount that would have been charged, but with the charges zeroed out. We hope that this helps you better estimate your costs and prepare for the change.

How can I back up my data in the cloud?

Snapshots provide a way to restore your Elasticsearch indices. They can be used to copy indices for testing, to recover from failures or accidental deletions, or to migrate data to other deployments.

By default, Elastic Cloud takes a snapshot of all the indices in your deployment every 30 minutes. You can set a different snapshot interval if needed for your environment. You can also take snapshots on demand without having to wait for the next interval. Taking a snapshot on demand does not affect the retention schedule for existing snapshots — it just adds an additional snapshot to the repository. This might be helpful if you are about to make a deployment change and you don’t have a current snapshot.

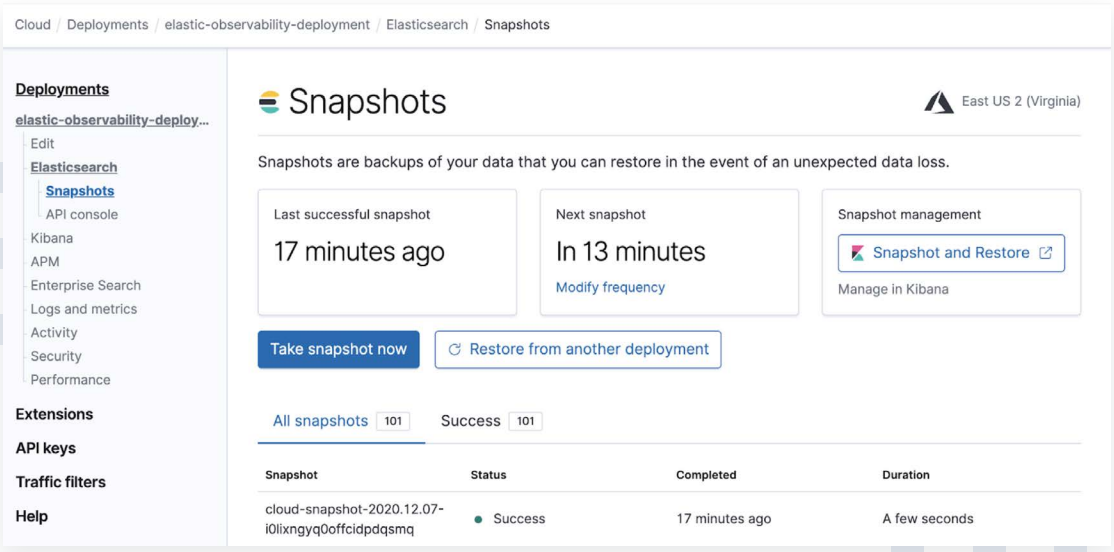
Snapshots only back up open indices. If you close an index, it is not included in snapshots and you will not be able to restore the data.

When you create a cluster on Elastic Cloud, a default repository called found-snapshots is automatically added to the cluster. This repository is specific to that cluster: the deployment ID is part of the repository's base_path, that is, /snapshots/[cluster-id].

Other repository options are available. You can:

- **Share a repository**, so that you can restore a snapshot from one cluster to another
- **Add your own custom repositories** to snapshot to and restore from

To configure your cluster snapshot settings or perform a restore, refer to **our documentation on working with snapshots.**



What is the Snapshot Fee and how can I reduce it?

AWS Snapshot Storage (per GB-month)	$0.0330	54	$1.78
AWS Snapshot Storage (per GB-month)	$0.0330	-54	-$1.78
AWS Snapshot Storage API (1K Requests)	$0.0018	5000	$9.00
AWS Snapshot Storage API (1K Requests)	$0.0018	-5000	-$9.00

Snapshot storage costs are tied to the cost of storing the backup snapshots in the underlying IaaS object store (for example, S3 on AWS or GCS on Google Cloud). These storage costs are not for the disk storage that persists the Elasticsearch indices, as these are already included in the cost of the Elastic Cloud deployment.

As is common with all cloud providers, we meter and bill snapshot storage using two dimensions:

- Storage size (GB/month)
- Storage API requests (1,000 requests/month)

Storage size is calculated by metering the storage space (GB) occupied by all the snapshots of all the deployments tied to an account. The same unit price applies to all regions. To calculate the charges due, we meter the amount of storage on an hourly basis and produce an average size (in GB) for a given month. The average amount is then used to bill the account for the GB used within a billing cycle (a calendar month).

For example, if the storage used in April 2019 was 100GB for 10 days, and then 130GB for the remaining 20 days of the month, the average storage would be 120 GB/month, calculated as (100*10 + 130*20)/30.

Storage API call costs are calculated by counting the total number of calls to backup

or restore snapshots made by all deployments associated with an account. Unlike storage size, this dimension is cumulative, summed up across the billing cycle, and is billed at a price of 1,000 requests.

Currently, the rates for these two dimensions are:

- Storage size: $0.033 per GB
- Storage API requests: $0.0018 per 1,000 API calls

We provide a free allowance of 100GB/month to all accounts across all the account deployments. Any metered storage usage below that amount will not be billed. Whenever the 100GB/month threshold is crossed, we bill for the storage used in excess of the 100GB/month free allowance.

We also provide a free allowance of 100,000 API requests to all accounts each month across all the account deployments. Once this threshold is passed, a charge will occur for the use of API requests in excess of the free allowance only.

> **NOTE**
>
> A single snapshot operation does not equal a single API call. There could be thousands of API calls associated with a single snapshot operation, as different files are written, deleted, or modified. The price we list is in 1,000s of API calls, meaning $0.0018 for 1000 API calls or $1.8 for a million calls.

The way snapshots work in Elasticsearch is by saving data incrementally at each snapshot event. This means the effective snapshot size may be larger than the size of the current indices. The size gets bigger as more data is used in the cluster, as well as when the data is changed frequently (added/deleted/modified records). In order to allow control beyond adapting the changes to the data (which is not always practical), we included an advanced parameter in the Elastic Cloud console under the snapshots sub menu, called Snapshot count. We have kept the current default of 100 snapshots (rolled over), but this can be changed to any value between [2 and 100].

> **WARNING**
>
> Reducing the number of snapshots effectively reduces the retention period of indices. This means that only a recent restore point will exist and may expire quickly.

As for API requests, these are executed every time a snapshot is taken or restored. While restore is not typically a frequent option, snapshots are taken by default every 30 minute to maintain a recent and fresh restore point. We have included a new parameter named Snapshot interval, which can be changed to up to 24 hours, resulting in fewer API calls.

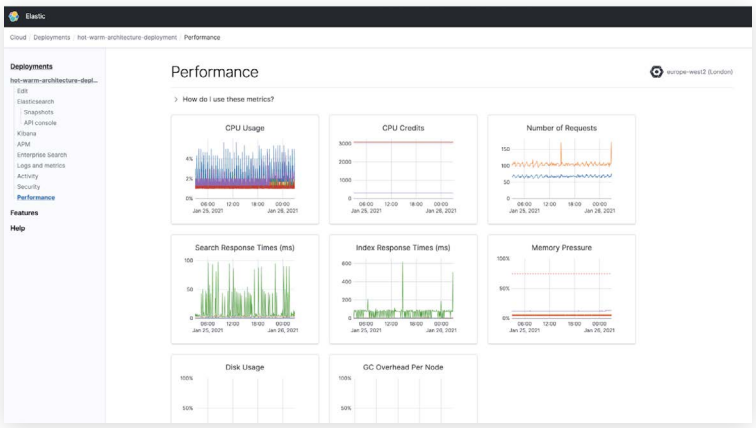
> **WARNING**
>
> Reducing the snapshot interval may result in partial data loss, as a restore operation of an older snapshot will not account for all the data that has changed since the last snapshot.

Finally, in case of any implemented logic using the Elasticsearch API to create/restore snapshots, it is advisable to revisit that process to avoid excess charges.

Given the varied ways of using Elasticsearch in different use cases, it is hard to predict the exact snapshot storage costs for your account. For this reason, we have implemented two “grace” bills that will show the amount that would have been charged, but with the charges zeroed out. We hope that this helps you better estimate your costs and prepare for the change.

Get a snapshot of your performance metrics to monitor your cluster's health

Cluster performance metrics available directly in the Elastic Cloud console give you a **quick overview** of the health of your deployments. This allows you to view the last 24 hours of the logs and metrics for your deployment.



To access basic cluster performance metrics such as CPU usage and credits, number of requests, search and index response, memory pressure per node, and garbage collection overhead per node:

1. Log in to the Elastic Cloud console.
2. Select your deployment.
3. From your deployment menu, go to the Performance page.

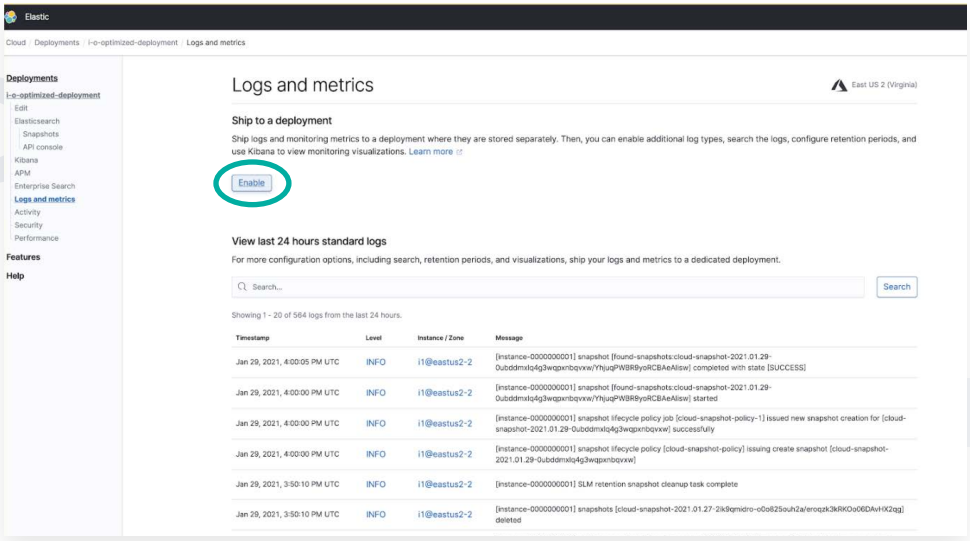
To view your Elasticsearch server logs, which are basic cluster logs:

1. Log in to the Elastic Cloud console.
2. Select your deployment.
3. From your deployment menu, go to the Logs and metrics page.

For additional tips on working with basic performance metrics as well as what to do when your cluster restarts after an out-of-memory failure, refer to our **documentation.**

Advanced monitoring for a deeper understanding of your deployment's health

Stack monitoring gives you greater insight into your deployment's performance. Just navigate to the Logs and metrics page and click Enable to take advantage of advanced views for Elasticsearch and Java virtual machine (JVM) metrics, including configurable retention periods that exceed 24 hours.



The Elastic Cloud deployment logging and monitoring feature lets you monitor your deployment in Kibana by shipping logs and metrics to a monitoring deployment. You can:

- View your deployment's health and performance in real time and analyze past cluster, index, and node metrics.
- View your deployment's logs to debug issues, discover slow queries, surface deprecations, and analyze access to your deployment.

To enable a cloud deployment to send its logs and metrics to a specific cloud deployment, go to the Logs and metrics page of the deployment you want to send

logs and metrics from, click Enable, and select any available deployment where you want to send your logs and metrics. You can opt to ship just deployment logs, just metrics, or both to any deployment in the region with the same major version.

Generally, we recommend sending your deployment logs and metrics to a separate dedicated monitoring deployment so that your production deployments are not impacted by the overhead of indexing and storing monitoring data. In the case of self monitoring, if the deployment is experiencing issues it won't be able to send out alerts. A dedicated monitoring deployment also gives you more control over the retention period for monitoring data.

If you want to adjust data retention for logs, you can adjust the default index lifecycle management policy created for the index, such as the amount of time you want logs to be retained for or rolled over. In addition to using the Logs app and default dashboards in Kibana, you can also create an elastic-cloud-logs-* index pattern to view your deployment's logs in Discover.

By default, the following logs will be sent when you enable this functionality:

- Elasticsearch logs
- Kibana server log
- APM logs
- Enterprise Search logs

You can also opt for more granular logging levels such as:

- **Audit logging**: logs security-related events on your deployment
- **Slow query and index logging:** helps find and debug slow queries and indexing
- **Verbose logging:** helps debug stack issues by increasing component logs like debug or trace

Keeping your Elastic Cloud deployment secure

At Elastic, we know that security is everyone's responsibility. That's why we bake security into the development of our products and into the foundation of Elastic Cloud. The security and privacy of your Elastic Cloud SaaS data also depend on keeping your Elasticsearch cluster configured securely and maintaining the confidentiality of your Elastic Cloud login credentials.

Here's a quick checklist:

- ✔ Don't share your credentials with others.
- ✔ Update your account profile to make sure information is correct and current.
- ✔ Add operational contacts as appropriate.
- ✔ Ensure that you've set secure passwords.
- ✔ Use caution when enabling custom plugins on your Elastic Cloud deployments.
- ✔ Consider setting the option to require index names when initiating destructive actions.

If you need to make changes that are not offered in the Elastic Cloud console, please create an Elastic Support case. If you believe an account has been compromised, please email **security@elastic.co**. If you need to make an erasure request, please email **privacy@elastic.co.**

To learn more about how to secure your Elastic Cloud deployment, refer to **securing your deployment.**

To learn more about how we provide security and privacy for your data on Elastic Cloud, refer to **Elastic Cloud Security.**

What you need to monitor vs what Elastic takes care of

Benefits of a managed service

When you choose our managed Elasticsearch offering, we handle the maintenance and upkeep so you can focus on moving your project forward. You'll get the latest version and security updates — including exclusive features and access to Elastic Support. Plus, Elastic Cloud grows with you as your requirements change over time. Let's take a look at some of the features and benefits of Elastic Cloud.

Benefits	Features
 Deployments made easy	• Deploy globally in minutes • Optimized compute, memory, and storage ratios based on the deployment size you select
 Security by default	• OS hardening and logical separation of discrete containers • Data is encrypted by default, whether in motion or in transit • We preconfigure secure communications between components and manage the encryption keys
 Hands-free maintenance	• Save time while we apply operating system updates and security patches in the background • Save money you would have spent on maintenance and focus on initiatives that drive greater business value
 We wear the pager	• We architect for service redundancy and resiliency to help prevent against unplanned downtime. • Rest easy knowing we are periodically backing up your data
 Future-proof	• Take advantage of the latest features and always be on the current version • Benefit from the latest hardware innovations • Easily upgrade to a Gold or Platinum subscription for access to additional features

Our shared responsibility

We think of a managed service as a shared responsibility. At a high level, Elastic is responsible for the security and operation of Elastic Cloud, including the underlying infrastructure and the Elastic software that runs on top of it. You make decisions about your specific configuration to ensure your deployment size meets your business needs. Let's take a closer look at what this means.

Our responsibilities

Infrastructure

- Building out infrastructure in Amazon Web Services, Microsoft Azure, and Google Cloud regions and availability zones
- Optimizing cloud storage and locally attached disk ratios for your deployment
- Managing availability, so you won't see insufficient capacity or stock out errors when you create a deployment
- Ensuring the underlying infrastructure receives operating system updates and security patches

Platform and network security

- Encrypting your cluster data and snapshots and managing your keys
- Encrypting communications between nodes and components such as Elasticsearch, Kibana, and APM

Orchestration of the Elastic Stack

- Ensuring connectivity between components
- Automatically applying your configuration changes and plugins
- Applying your specified capacity parameters
- Taking periodic snapshots (you set the cadence, we do the work)

Your responsibilities

Operations

- Giving careful consideration to the sizing requirements of your deployment
- Deciding when and how you want to restore snapshots
- Ensuring the health of your configurations, plugins, indexes, and shards

Security

- Securing communications between Elastic Cloud and other systems, such as applications, databases, and other cloud services
- Implementing configurations that adhere to security best practices
- Applying user- and account-based access rights

Data

- Managing the data you use in conjunction with the service

Troubleshooting from the trenches

I've changed the configuration of my cluster and it's taking too long to apply. Is it stuck?

Plan changes, either version upgrades or instance resize/scale-up, can take a significant amount of time depending on the amount of data stored and overall load of the deployment. Furthermore, hot-warm deployments with slower I/O high-storage nodes can take substantially longer during the resizing process. So don't worry if a plan change is taking too long — it's not stuck! Plan changes may take from a few minutes to several hours, depending on how much data is stored. For instance, a cluster holding 5TB+ of data can take up to 6 hours to complete a plan change. With the exception of major version upgrades for Elastic Stack products, Elasticsearch Service can perform configuration changes without having to interrupt your deployment. You can continue searching and indexing. The changes can also be

done in bulk: in one action, you can add more memory, upgrade, adjust the number of Elasticsearch plugins, and adjust the number of availability zones, for example.

Why did I receive an email saying nodes have been restarted/CPU or memory usage is high?

Occasionally, you may receive email alerts from Elastic Cloud indicating that your cluster is running out of hardware resources. You may feel the urge to increase your cluster size immediately, but before taking actions that incur more costs, it is better to understand why your cluster increased its demand for resources. A quick investigation will give you better direction as to what sort of changes the cluster really needs. Depending on the email you receive, there will be different suggestions to help diagnose the underlying issues:

Node restarted due to running out of memory or high memory usage

High memory usage on nodes can cause severe service outages (like nodes being restarted), dashboard taking an increased time to load, and slower querying in general, so make sure you **understand what the memory pressure indicator means.**

Additionally, it is worth noting that a node may still restart due to running out of memory, even though it isn't close to the heap usage limit. This happens because nodes not only use JVM heap memory but also native OS memory, which is not collected by stack monitoring. Despite a different type of memory being used, the same recommendations for reducing JVM heap memory usage still apply.

Storage space is low

This happens more often with logging and metrics use cases where time-based indices are not automatically deleted. By default, an Elastic Cloud cluster will snapshot all indices every 30 minutes (**configurable** to your liking) so it is safe for you to automate the deletion of older indices.

- For Elastic Stack versions 7.6 and above, please refer to the Kibana **Snapshot and Restore** documentation to learn more about the snapshot management features.

- For Elastic Stack versions 7.5 and below, you can directly set the snapshot interval and total number of snapshots via Elasticsearch Service to control the snapshot retention period for your clusters. The calculation displayed is calculated based on the snapshot interval, expected snapshot count, and current cluster size. Changing any of these factors results in a different retention period.

We recommend configuring **index lifecycle management (ILM)** policies that have a delete phase to automatically manage indices according to your storage requirements.

High CPU usage

High CPU usage alerts do not necessarily indicate a problematic cluster and may just mean that the cluster is under normal usage. However, to be sure there are no underlying issues, the first thing you should do is verify whether there are issues with applications connected to the cluster such as dashboards timing out, slow queries, and/or indexing. If that is the case, your CPU is the reason your cluster is not performing smoothly, especially when the cluster's **CPU credits** run out. The high CPU usage alert may also be triggered by the JVM's garbage collection, heavy indexing, or a combination of both. Make sure to verify whether **memory pressure** is impacted, indicating that the JVM's garbage collection may be at fault. Either way, the **Elasticsearch hot threads API** is a great tool for identifying possible tasks using up high amounts of CPU.

Why is my ingestion rate low?

Sometimes you may notice that your dashboards in Kibana are falling behind (for example, the newest indexed document is noticeably far in the past) or, if you have a queueing system for indexing (such as Logstash or Apache Kafka), you may also notice that the indexing queues are building up. This is a sign of suboptimal ingestion rate. Diagnosing the root cause for this type of issue can be quite challenging because, most of the time, it is a combination of issues across various components in your ingestion pipeline such as Beats, Logstash, and Apache Kafka processes, network connections, and disk I/O. Here is a list of things that we recommend investigating:

Network connection and throughput issues can have a significant impact on ingestion rate. Before making any changes, make sure that all components — from the original data source such as hosts running Beats to the Elasticsearch cluster — are within the same cloud region. Next, try pinging (use the "ping" tool) a destination network address from within a data source. For instance, ping from a host running Beats that is sending data to a Logstash server and check the ping time. Times of around tens of milliseconds are OK, but there may be an issue if ping times are 1000 milliseconds or more. Also don't forget to check package loss (reported by the "ping" tool), which also indicates issues at the network level. Lastly, check processes logs to see if there are repeated/frequent network connection drops, which also severely impact ingestion rate.

If you have a custom-built indexer sending data to Elasticsearch, make sure your indexer is using the **Elasticsearch bulk API**, as this API is crucial for good indexing performance. Most of the official Elasticsearch client libraries provide a helper component that facilitates the usage of the bulk API; for instance, in the Python client there is the **bulk helper.**

Check CPU and memory usage of processes. For instance, both Logstash and Apache Kafka run in a JVM so they may also suffer from JVM garbage

collection issues. If you're using Logstash, make sure you enable its **monitoring features**. Additionally, Logstash has a **hot threads API** which helps diagnose high CPU usage.

Check the indexer's log (such as Beats, Logstash, or whatever is connected and sending data to Elasticsearch) for **Elasticsearch bulk rejections**. If you frequently see this sort of error, it means that you are trying to ingest data much faster than your current Elasticsearch cluster can handle. In this scenario, check the Elasticsearch cluster for performance issues (such as high CPU and memory usage). If you are seeing lots of bulk rejections in the logs while Elasticsearch hardware resources are still underutilized, this may be due to an increased number of primary shards being used (that is, you've switched to a different index with more primary shards).

Investigate long-running tasks using the **Elasticsearch task management API**. A bulk index API call will spawn bulk index tasks in Elasticsearch that can be identified using this API. Long-running bulk index tasks may impact ingestion rate. Check if there are other long-running tasks such as search tasks that may be causing a performance hit on the nodes and always remember that some **tasks can be cancelled** if running for too long.

Although this list is not exhaustive, as this sort of issue is highly dependent on the overall architecture of your ingestion pipeline, we have provided the most common issues that we recommend for your initial investigations.

My dashboards are taking too long to answer and often time out

Slow dashboards and time outs are a sign that the current group of connected dashboards and clients are concurrently querying data in a way that is complex enough (either by querying large data size, querying too many/deep aggregations, or a combination of both) to overwhelm the current Elasticsearch cluster. The challenge in identifying the root cause of this situation is that, in general, there isn't a single problematic query. One powerful tool you can use to try to track down

problematic queries is a combination of several Elasticsearch features including the `X-Opaque-ID` header and slow logs, which allows you to **identify what triggered a slow-running query**. For this to work, a few modifications are required to your configuration/software:

> 1. **Enable Elasticsearch slow logs.**
> 2. Modify any custom applications that query Elasticsearch to include the `X-Opaque-ID` header in their search requests.
> 3. In the case of slow Kibana dashboards, unfortunately Kibana does not currently support adding `X-Opaque-ID` to requests (**issue #16493**); however, **users have provided a workaround.**

With the above configuration in place, all you need to do is to identify the slowest queries in the search slow logs and you can start taking action. For example:

- Troubleshoot the query performance using the **Query Profiler.**
- Identify the source application of the offending queries.

What happens when I run out of disk space?

When you run out of disk space, you hit the floodstage watermark, and then the cluster switches to read and delete only. At this point, you will need to delete data or scale up manually.

Connecting with the Elastic community

Discussion forums

Find advice or lend a helping hand. Ask your most burning questions about all things Elastic and share your wisdom with fellow users on our **discussion forums**, which are also available in your native language.

Slack and local communities

Join our fast-growing **Elastic Slack** to chat with other users and ask for advice in various channels: #elasticsearch, #kubernetes, #kibana-development, or others. Additionally, many **other online communities** have sprung up all over the world! Join one in your region to share your Elastic story with the local community.

Keep learning

Getting started with the Elastic Stack? Looking for detailed deep dives? Get hands-on with the **Elastic examples repo** and explore curated datasets and step-by-step instructions. Plus, see what's making the rounds in our dev team through our **community newsletter**.

We'd love to hear from you

As technology evolves, so does Elastic. We really value hearing from our community. **Please reach out to us** for help or to share your thoughts about your Elastic experience.

Elastic makes data usable in real time and at scale for enterprise search, observability, and security. Elastic solutions are built on a single free and open technology stack that can be deployed anywhere to instantly find actionable insights from any type of data — from finding documents, to monitoring infrastructure, to hunting for threats. Thousands of organizations worldwide, including Cisco, Goldman Sachs, Microsoft, The Mayo Clinic, NASA, The New York Times, Wikipedia, and Verizon, use Elastic to power mission-critical systems. Founded in 2012, Elastic is publicly traded on the NYSE under the symbol ESTC. Learn more at elastic.co.

AMERICAS HQ
800 West El Camino Real, Suite 350, Mountain View, California 94040
General +1 650 458 2620, Sales +1 650 458 2625

info@elastic.co