Splunk Cloud™
Splunk Cloud User Manual 7.0.13

Monitor Splunk Cloud deployment health

Generated: 2/04/2021 3:51 pm

Monitor Splunk Cloud deployment health

The Cloud Monitoring Console lets Splunk Cloud administrators view information about the status of your Splunk Cloud deployment. Cloud Monitoring Console dashboards provide insight into how the following areas of your Splunk Cloud deployment are performing:

- search
- indexing
- forwarder connections
- indexer clustering and search head clustering, if applicable
- license usage, if applicable.

Locate the Cloud Monitoring Console

If you have a e-commerce Splunk Cloud deployment, to find the Cloud Monitoring Console:

1. Click **Settings**.
2. Click the Monitoring Console icon on the left.

If you have a managed Splunk Cloud deployment, the Cloud Monitoring Console is an app.

1. From anywhere in Splunk Web, click **Apps**.
2. Click **Cloud Monitoring Console**.

On the App Management page, the Cloud Monitoring Console is named splunk_instance_monitoring.

Enable Platform Alerts

The Cloud Monitoring Console provides preconfigured alerts that you can enable. If a platform alert is triggered, the Cloud Monitoring Console displays a notification. In addition, you can set up an alert action (for example, send an email) to be performed when a platform alert is triggered. See Set up alert actions in the *Alerting Manual* for more details.

- **E-commerce Splunk Cloud**: Go to **Cloud Monitoring Console > Settings > Alerts setup** and click **Advanced edit**. Notifications are displayed in the **Overview** dashboard.
- **Managed Splunk Cloud**: Go to **Settings > Searches, reports, and alerts** and select **Cloud Monitoring Console** in the App filter. Click the alert name and scroll down to **Alert actions**.

Dashboards

The Splunk Cloud Monitoring Console app provides information about your Splunk Cloud performance. The information is organized into several dashboards.

Dashboard	Description	For more information
Overview	Information about the performance of your Splunk Cloud deployment, including license usage (if applicable), indexing performance, and search performance information.	About license violations in the Splunk Enterprise *Admin Manual* What Splunk Cloud does with your data in *Getting Data In*

Dashboard	Description	For more information
		About jobs and job management
Search Usage Statistics	Information about how your users are running searches.	Write better searches in the Splunk Enterprise *Search Manual* Configure the priority of scheduled reports in the *Reporting Manual*
Scheduler Activity	Information about how search jobs (reports) are scheduled.	Configure the priority of scheduled reports in the *Reporting Manual*
HTTP Event Collector	Status of HTTP event collection, if you have enabled this feature.	Set up and use HTTP Event Collector in *Getting Data In*
Data Quality	Displays any line breaking, aggregation, or event breaking errors in your incoming data.	Resolve data quality issues in *Getting Data In*
Forwarders: Instance and **Forwarders: Deployment**	Information about forwarder connections and status. To get data to appear on the two forwarder dashboards, navigate to **Monitoring Console > Settings > Forwarder Monitoring Setup**, or click the setup link in either of the forwarder dashboards and follow the setup instructions.	Troubleshoot forwarder/receiver connection in *Forwarding Data*

Check your total data retention capacity

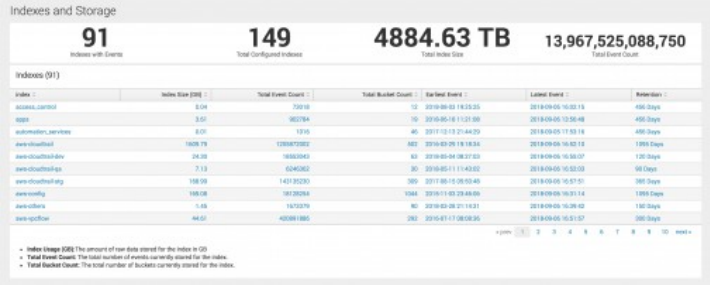
When you send data to Splunk Cloud, it is stored in indexes, and you can self-manage your Splunk Cloud index settings using the Indexes page in Splunk Web. Splunk Cloud retains data based on index settings that enable you to specify when data is to be deleted. Data retention capacity space in your Splunk Cloud service is based on the volume of uncompressed data that you want to index on a daily basis. Splunk Cloud provides 90 days worth data retention capacity with every subscription. For example, if your daily volume of uncompressed data is 100 GB your Splunk Cloud environment will have 9000 GB (9 TB) of data retention capacity. You can also purchase additional data retention capacity.

The Cloud Monitoring Console (CMC) Indexes and Storage dashboard provides insights into your data use so that you can better understand your current usage and predict future licensing needs.

In the Indexes and Storage dashboard, the CMC provides insights into your data retention based on the uncompressed data you have indexed.

Steps to find your retention usage and set an alert

1. Go to **CMC > Indexes and Storage**.



2. Take note of your total index size, which displays in the upper right. This represents your total uncompressed data that is currently retained.

3. Compare this value to your licensed entitlement amount to see if you need to update your license based on current usage. If you do not know your licensed entitlement, reach out to your Splunk sales representative.

4. Finally, create a query against CMC, and configure Splunk Cloud to generate an alert if the value exceeds your licensed usage. The following sample query shows the alert where `license_gb=10000000` should be replaced with your licensed data ingestion value (in GB):

```
 | dbinspect index=* cached=t
| where NOT match(index, "^_")
| stats max(rawSize) AS raw_size BY bucketId, index
| stats sum(raw_size) AS raw_size
| eval raw_size_gb = round(raw_size / 1024 / 1024 / 1024 , 2), license_gb = 10000000,
storage_usage_pct = round(raw_size_gb / license_gb * 100, 2)
| fields storage_usage_pct
```

5. Note that the query should be run against **All Time**.

More Information

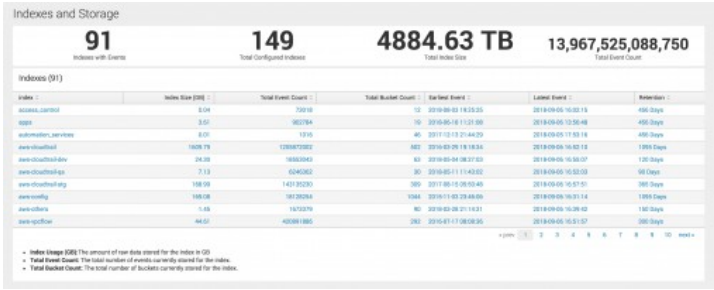
For more information about creating indexes, see Manage Indexes. For detailed instructions on creating alerts, see Alerts.

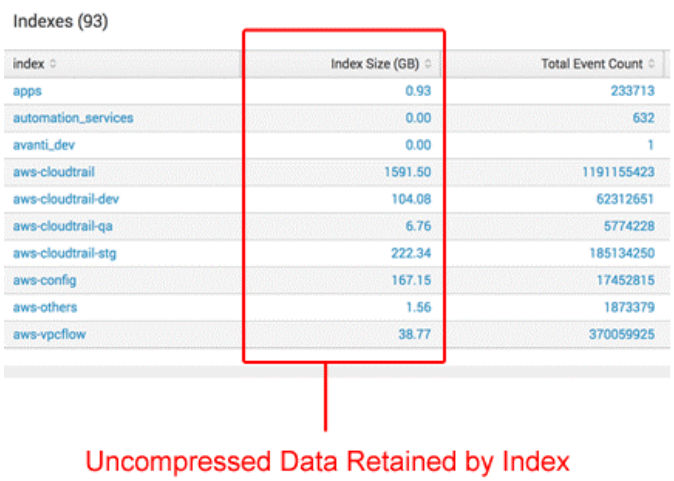
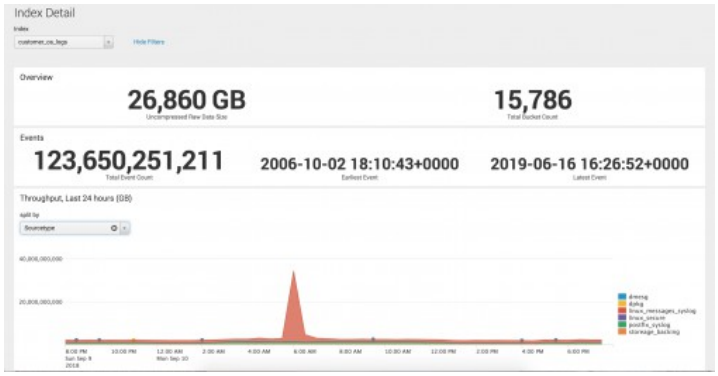
Understand your index data retention capacity

Your licensed data retention capacity is based on two variables: the daily licensed ingestion rate (e.g. 1 TB per day) and the amount of time Splunk is licensed to retain your data (e.g. 30 days). To understand how your data retention compares to your licensed retention, it's a good idea to view details about your index storage.

When you configure data retention for an index, you also configure two variables: the size of the index, and the number of days to retain the data. For example, you set data retention for 10 TB or 90 days, whichever comes first. If your data is retained for less time than you configured, it's likely that your ingestion rate is higher than expected. For example, if you configured your index to store data for 90 days or 10 TB, and you see that the data is being retained for 10 days, it's likely that you have hit the 10 TB threshold much sooner than expected, indicating a high ingestion rate. On the other hand, a longer retention than expected could indicate a misconfiguration of your index settings (i.e., you configured data retention for a time period that exceeds your licensed retention).

Steps to investigate your indexes

Steps	Detailed Image
Go to **CMC > Indexes and Storage**.	
A good method to determine if your data usage is running higher than expected is to check the date of the first event and the date of the last event and compare it to the retention setting for the individual index. For example, if the first event is 12/13/17, and the last event is 12/23/17 and the retention setting for the index is 90 days, then the data ingestion for the index was met long before the time retention setting was met. So, the data ingestion was greater than anticipated.	
Next, check to see which indexes are larger than others. You want to find which index is consuming the most storage and why. • To do this, check the index size, which shows the uncompressed data retained by the index. • Click the Index Size (GB) heading to sort the indexes by size. • Click the name of a larger index to open the Index Details page.	

Steps	Detailed Image
	
In the Index Details page, you can see if there's a spike or a higher trend line for an index. Both of these data points are clues that will tell you that you may need to adjust index settings or investigate further to determine what's causing the spike. • If you see a spike or rise in data, sort by source type or host to understand if there is a specific cause for the increase. • You may then need to investigate your host or source to determine if there is an issue. • If you don't see spikes or a higher trend line, you do not have an issue with ingestion.	

Check your data quality

This topic discusses how to check the quality of your data and how to repair issues you may encounter. However, the concept of data quality depends on what factors you use to judge quality. For the purposes of this document, data quality means that the data is correctly parsed.

Your data quality can have a great impact on both your system performance and your ability to achieve accurate results from your queries. If your data quality is degraded enough, it can slow down search performance and cause inaccurate search results. Therefore, it's important to take the time to check and repair any data quality issues before they become a problem.

Generally, data quality issues fall under three main categories:

- **Line breaks**. When there are problems with line breaks, the ability to parse your data into the correct separate events that it uses for searching is affected.
- **Time stamp parsing**. When there are timestamp parsing issues, the ability to determine the correct time stamp to use for the event is affected.
- **Aggregation**. When there are problems with aggregation, the ability to break out fields correctly is affected.

Guidelines

Finding and repairing data quality issues is unique to each environment. However, using these guidelines can help you address your data quality.

- It's a good idea to check your most important data sources first. Often, you can have the most impact by making a few changes to a critical data source.
- Data quality issues may generate hundreds or thousands of errors due to one root cause. Therefore, it is recommend that you sort by volume and work on repairing the source that generates the largest volume of errors first.
- Repairing data quality issues is an iterative process. Repair your most critical datasources first, and then run queries against the source again to see what problems remain.
- For your most critical source, you should ideally resolve all data quality issues. This helps to ensure that your searches are effective and your performance is optimal.
- Run these checks on a regular cadence to keep your system healthy.

Example

The following example shows the process of resolving a common data quality issue. The steps to resolve your data quality issues may differ, but you can use this example as a general template for resolving data quality issues.

1. Go to **Cloud Monitoring Console > Indexing > Data Quality** to see the Data Quality dashboard.

2. View the **Event Processing Issues by Source Type** dashboard. In this example, you can see that the greatest volume of issues are timestamp parsing issues in the `splunk_python` source. Since the `splunk_python` source has the most errors, and most are timestamp errors, we decide to work on timestamp errors. The steps below show you how to resolve timestamp errors.

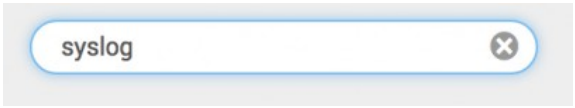
3. In this example, we are most concerned with timestamp errors in the syslog source, so we drill down into that source. Drilling down, we can see that the majority of issues are with the following source: `/var/log/suricata/stats.log`.

4. Clicking the source allows us to drill down further, where we can see searches against this source.

5. From here, we can look at a specific event. We can see that the issue is that Splunk was unable to parse the timestamp in the MAX_TIMESTAMP_LOOKAHEAD field.

6. To fix this, we go to **Settings > Source types**.

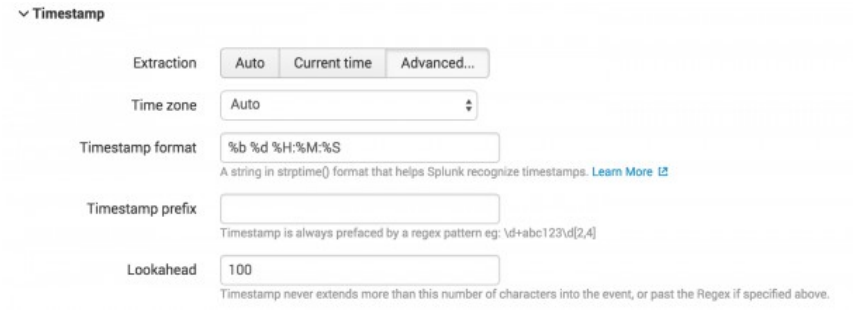
7. In the filter, enter syslog for the source type.



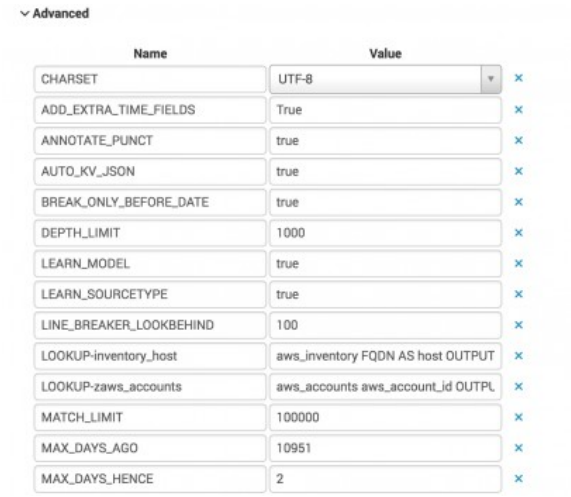
8. Under **Actions** select **Edit**. The **Edit Source Type** page opens.

9. Click **Timestamp > Advanced?** to open the Timestamp page for editing. Ensure you are satisfied with timestamp format and the Lookahead settings. In this case, we need to edit the Lookahead settings so that Splunk can parse

the timestamp correctly.



10. Returning to the main **Edit Source** Type page, go to the **Advanced** menu. From here you can make other changes if needed.



Understand your search performance

Healthy search loads are critical to the performance of your entire Splunk Cloud environment. Understanding search patterns can help you to determine if your search workload is aligned with best practices and optimized for the best performance. Often by looking more deeply into search patterns, you can see if a specific user, search, dashboard, or app is inhibiting your performance. If you encounter an issue, you can then work with users to improve performance. Search performance can be investigated by focusing on some key areas.

- Skipped Searches
- Search runtime

Skipped Searches

If you are skipping searches, it can be indicative of problems with your search scheduling or query formation. For example, maybe you have scheduled too many searches to run at the same time, and you can alleviate the problem by staggering the scheduled searches.

You may also find that you have a search that attempts to run before the previously scheduled search has completed. For example, if you schedule Search_A to run every five minutes, but the first instance of the search takes 10 minutes to complete, then the next time the search is scheduled to run, it will be skipped because the first search has not yet

completed. If this occurs, you may need to adjust the time range (set it to 10 minutes instead of 5), or you may need to optimize your search to align with search best practices to improve performance.

For more information about optimizing searches, see About Optimization.

Lastly, you may have skipped searches because your users have met the threshold for concurrency limits that you set in your Splunk System Limits. This is expected behavior, but it may also indicate that your users need help in optimizing their searches.

Search Runtime

When searches run for a long time, they may use too much compute and memory, causing an overall slowness of the Splunk instance, This commonly occurs when a few poorly formed searches are taking a large amount of resources. It can also occur if you have a dashboard that is being frequently used by multiple users concurrently. In each of these cases, investigating further can help you to pinpoint the searches that are long-running and determine if you can optimize them. Because each company's environment is different, it's not easy to set benchmarks for search performance. Generally, the best way to understand your search performance is to compare your historical search times with your current search times to see if there is a change. If search runtimes have slowed, review changes to your environment and new searches to determine if you need to optimize your searches or environment. For example, you may have added a poorly formed search, or you may have added a dashboard that has attracted a lot of traffic.

To check for skipped searches

1. Go to **Cloud Monitoring Console > Search > Skipped Scheduled Searches**.
2. In the **Time Range** field, select 24 hours to get a better picture of your searches historically.
3. In the **Count of Skipped Searches** pane, sort by **Reason**. Frequently, there are a number of skipped searches for the same reason. Take a note of the primary reason or reasons that searches are skipped.
4. Scroll down to see which report is generating the primary issues, and take note of the report name. If you determine that this is an expected behavior, you don't need to research any further. But, if the skipped searches are unexpected, continue to the next step.
5. Go to **Settings > Searches Reports and Alerts**.
6. If you know the application associated with the search or report, you can sort by the app. Otherwise, search by the report or search name.
7. Once you locate the search or report, click on it to open the search edit dialog box.
8. At this point, you may need to troubleshoot the formation of the search (look for wild cards, check to see if an index is specified, etc).
9. Or, if you found that scheduling is the problem, go to **Edit > Edit Schedule** to review the schedule for the search.
10. Verify that the schedule for the report or search is in line with how long the search should take to complete. For example, if the report is run every hour, but it takes 1.5 hours to run the search, the searches will be skipped.

To review searches by user

1. Go to **Cloud Monitoring Console > Search > Search Usage Statistics**.
2. Change the time frame to widen the scope. For example set it to week to date.
3. Split the search by users so that you can see if there are a few users that are typically running longer searches.
4. Sort by **Cumulative Runtime** to see which users have the most cumulative search time.
5. Sort by **Median Runtime** to see which users are running the median longest searches.
6. Click on the name of the user to drill down into more details about that user's searches.
7. If the user running the most or longest searches is the system user, you may want to review your applications to make sure that you have optimized them, and that they are providing the expected value. You may discover that some applications are not needed or are not used.

Reviewing this data will give you a better understanding of which users run a large number of searches (or run a few long-running searches). At this point, you may want to review the searches for that user in more detail so that you can better understand if they can be optimized.

For more information about optimizing searches, see About Optimization.

To review long-running searches

1. Go to **Cloud Monitoring Console > Search > Search Usage Statistics**.
2. Expand the time range to at least 24 hours. Searches are automatically sorted by long-running searches.
3. The **only ad-hoc Searches** toggle should be set to no. This ensures that you will see scheduled searches, which are more likely to be long-running searches than ad-hoc searches.
4. Scroll down to the **Search Details** pane where the searches are sorted by search runtime.
5. Click the search name to view more details, and scroll to the bottom of the screen. Two events are displayed. In the second event, you can see the search query.

If you discover a long-running query that runs frequently, you may want to expand the time range to a week or longer to see how commonly this search is run. If it is running frequently, consider optimizing the search.

For more information about optimizing searches, see About Optimization.