exabeam

Cloud Archive

Exabeam Cloud Platform - Version 2021.1

Publication date February 5, 2021

Exabeam
1051 E. Hillsdale Blvd., 4th Floor
Foster City, CA 94404

650.209.8599

Have feedback on this guide? We'd love to hear from you!
Email us at docs@exabeam.com

Table of Contents

1. About Exabeam Cloud Archive 4
2. Exabeam Cloud Archive Architecture 5
    2.1. Exabeam Cloud Archive Technical Specifications 5
3. Using Exabeam Cloud Archive To Search Your Data 6
    3.1. Navigating The Search Page 6
        3.1.1. Search Bar 6
        3.1.2. Search Results Timeline 6
        3.1.3. Search Results List 7
    3.2. Export The Cloud Archive Search Result 9
    3.3. Search Best Practices 10
        3.3.1. Narrowing Down The Time Range 10
        3.3.2. Using Field Names And Values 10
    3.4. Data Ingestion Statistics 11
4. Technical Support Information 12

1. About Exabeam Cloud Archive

Exabeam Cloud Archive is a cloud-based log storage service that offers long-term data retention along with search capabilities. Similar to Exabeam Data Lake , Cloud Archive give you access to your logs in a clear and consumable manner. With the reliability of the Exabeam Cloud Platform backbone, Cloud Archive excels in providing an affordable storage cost while preserving search capabilities.

You can use Cloud Archive to search for log events captured months or years ago, and to scan for new indicators such as IP addresses or domain names over extended periods of time. In addition, Cloud Archive makes it easier for your organization to meet multi-year log retention requirements imposed by regulations such as PCI-DSS, Sarbanes-Oxley or HIPPA.

The Cloud Archive service is automatically provisioned as part of your Exabeam SaaS Cloud deployments and no further configuration steps are required for log data to enter the archive.

2. Exabeam Cloud Archive Architecture

Cloud Archive is a cloud-native, multi-tenant, log aggregation service, designed to handle very large data volumes. As described in the diagram below, Cloud Archive directly integrates with the Exabeam SaaS Cloud infrastructure to import all received logs. Cloud Archive indexes and stores logs in a cloud-native object store, then makes those logs available through its search service. Logs in Cloud Archive are parsed using security content packages offered by Exabeam. To ensure parsing consistency, Cloud Archive synchronizes the parser configuration with Data Lake every 24 hours.

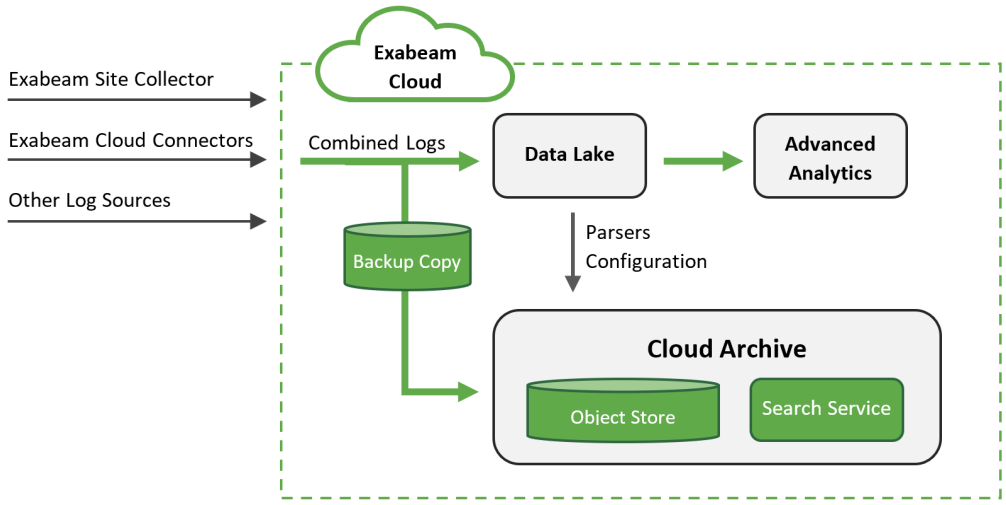


Figure 1. Cloud Archive architecture

2.1. Exabeam Cloud Archive Technical Specifications

Cloud Archive is an application deployed within the Exabeam Cloud Platform.

Specification	Value
Cloud Provider	Google Cloud
Geographical Regions	North America (us-west) Europe (europe-west3, Frankfurt)
Maximum events per Second (EPS)	200,000 EPS per tenant
Maximum retention	10 years
Maximum search query length	1 million characters
Maximum concurrent searches	5 per tenant

Log data received by Exabeam SaaS Cloud may take up to four hours to appear in Cloud Archive. Analysts should account for this possible delay when looking at recent events.

3. Using Exabeam Cloud Archive to Search Your Data

This chapter walks through the process of performing a search and understanding the results. The search capabilities in Cloud Archive are designed to match those of Exabeam Data Lake . While not identical, the search syntax and user interface of Cloud Archive will be familiar to any analyst who has previously used Data Lake .

3.1. Navigating the Search Page

3.1.1. SEARCH BAR

At the top of the page is a search bar where you can enter a simple text search or use the Lucene query syntax to search your data. Once you have entered a query, select an appropriate time range using the date and time selector located to the left of the **Search** button. Click **Search** to launch your query: the search will run and display results as soon as they are available.

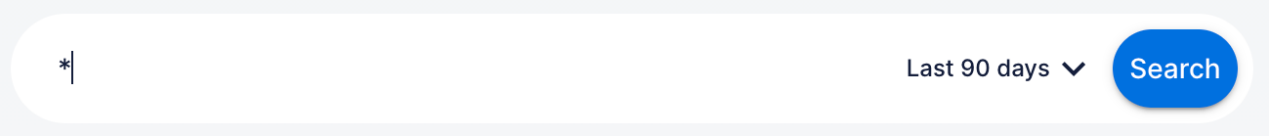


Figure 2. Search Bar before query input

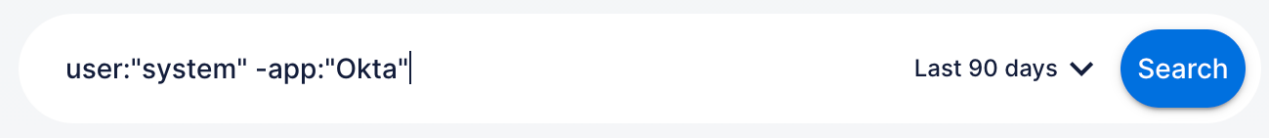


Figure 3. Search Bar before search action

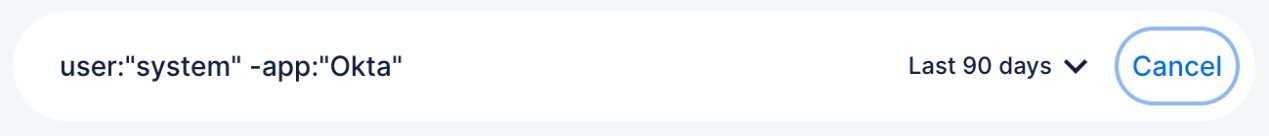


Figure 4. Search bar during search operation

Searches typically take several minutes to complete, depending on the data volume, the size of the time range, and the complexity or nature of the query. Click **Cancel** to edit your query or interrupt the search.

3.1.2. SEARCH RESULTS TIMELINE

Once the search has completed, Cloud Archive will display a histogram chart underneath the search bar. The timeline presents the count of events over the selected time range. Click any of the bars in the chart to further filter the search. After clicking a bar, or dragging over a group of bars, click the zoom-in **magnifier** icon to filter the results.

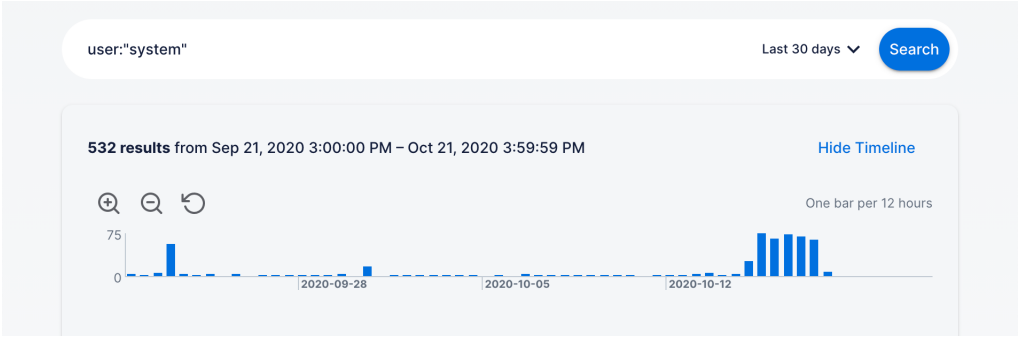


Figure 5. Search Results Timeline

The time range associated with each column dynamically adjusts based on the time range of your search. Each column may represent minutes, hours, or days in the timeline depending on how wide the search range.

3.1.3. SEARCH RESULTS LIST

When you submit a search request the **Timeline and Events List** are updated to reflect the search results. The most recent events that match the query are displayed in the Events List.

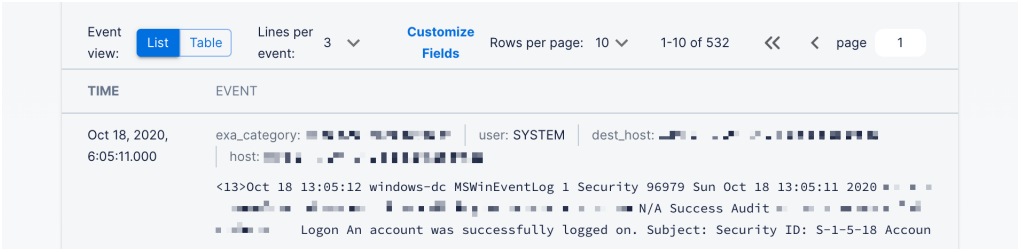


Figure 6. Search Results List

3.1.3.1. Lines Per Event

By default, the list displays the first three lines of each log event. Click the **Lines Per Event** control to change the number of lines displayed.

3.1.3.2. Rows Per Page and Pager

The list displays up to 100 events by default. Click **Rows Per Page** and select the number of lines that you want to display. Alternatively, use the control on the pager to navigate the search results pages.

3.1.3.3. Event Details

Click any event to reveal the full raw message of the event, along with the entire list of parsed fields for that event. The **Event Details** panel also offers controls to display or hide fields in the events list. Use the **Eye** icon to toggle whether a field is visible in the list of events.

Figure 7. Field visibility controls

3.1.3.4. Search Results Table

To display the search results in a tabular format, select the **Table** option in the **Event View** control. In **Table** view, a column is created for each visible field of the listed events. Use the **Customize Fields** dialog to control the visibility and order of the columns. While in **Table** view, click any event to display its full details in the **Event Details** panel.

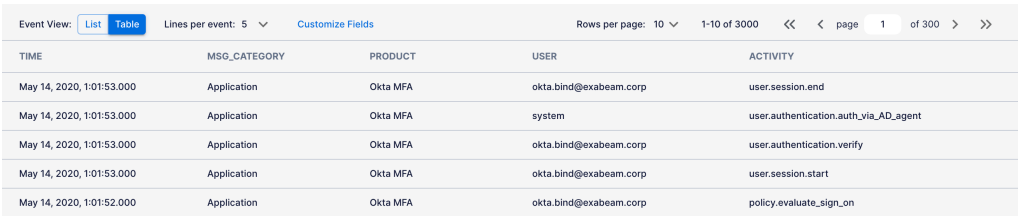


Figure 8. Search Results Table

3.1.3.5. Field Templates

The **List** and **Table** view present a default selection of fields per event that is curated based on the event's category. You can change what fields are made visible by selecting one of the templates listed in the **Field Template** picklist. Each template provides a different selection that is appropriate for the category. You can also create their own templates by clicking **Add New Field Template**.

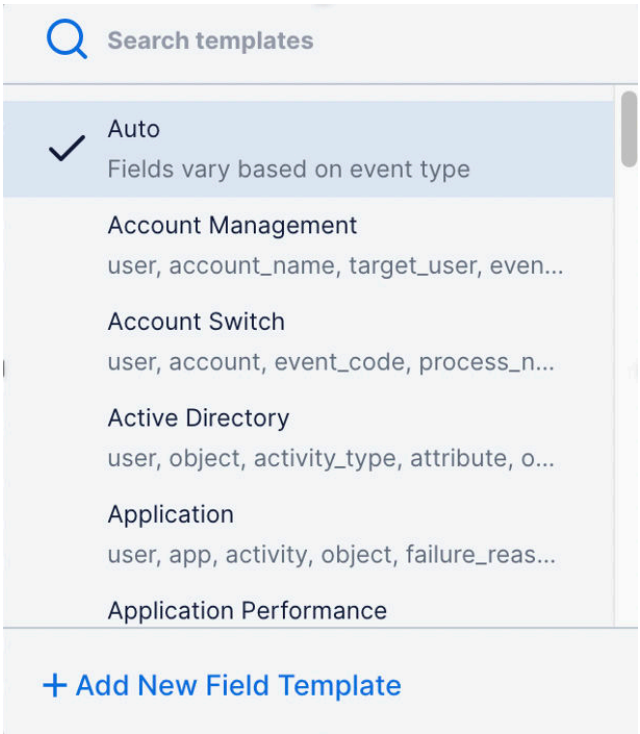
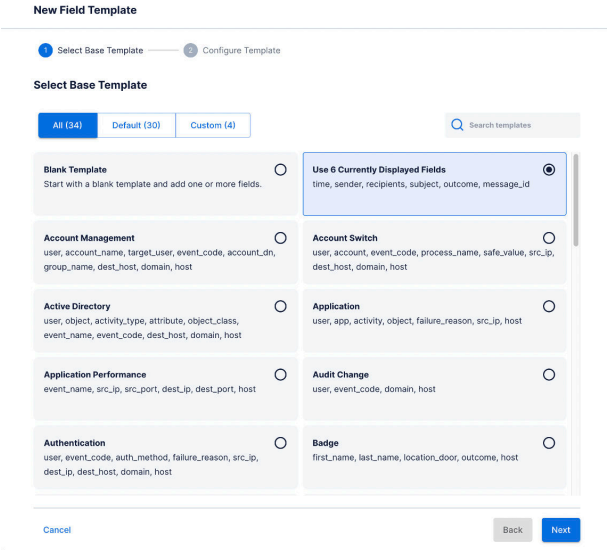


Figure 9. Field Template List

To add a new Field Template:

1. Click **Add New Field Template**.

2. In the **New Field Template** dialog, select whether you would like to start from a blank template, from the currently visible fields or from one of the existing templates, then click **Next**.
3. In the Configure Template step, enter a name for your template and select the fields to make visible from the **Available Fields** and **Displayed Fields** lists.



4. When your configuration if complete, click **Save**. The new template will automatically be selected under **Field Template**.

Keep in mind that Field Templates are shared across analysts, new field templates created by an analyst will be visible to all other analysts in the same environment.

3.2. Export the Cloud Archive Search Result

Use Cloud Archive's export functionality to download the search results to your local computer. You can capture the results and attach them to another system, for example a ticketing system, or when you need to work with the data outside of Cloud Archive's interface.

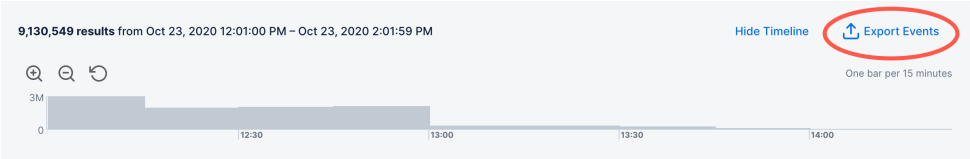
Cloud Archive Export allows output in the following formats:

- **Raw Log (txt)** --This format exports the events in a plain text file and separates each event by a carriage return line feed (CRLF). Use this format to attach events as evidence in a ticketing system, or to input the data into another tool such as the Exabeam Auto Parser Generator.
- **Time and Raw Log (csv)** -- This format exports the events in a comma separated value (CSV) file, where the first column includes the normalized ingestion time for the event, and the second column includes the raw message of the event. Use this format to import the search results into a spreadsheet, or into a tool that uses the time information present in the events.

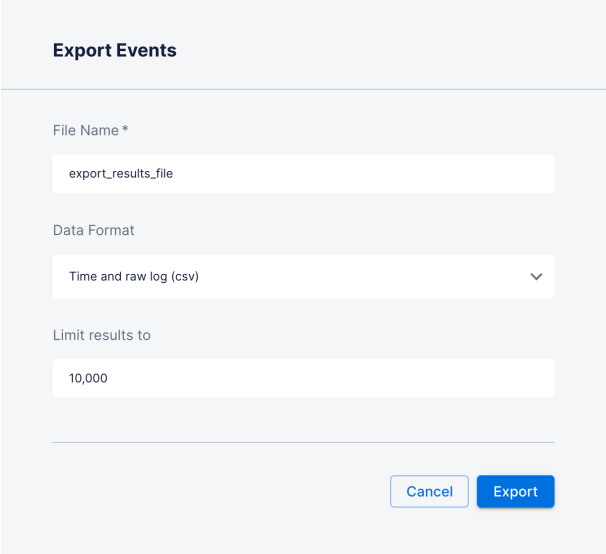
The export file is compressed in gzip format. Depending on the export format selected, the file's extension will be `.txt.gz` or `.csv.gz`.

To export events:

1. After or while a search is running, click **Export Events**.



2. Fill in the export parameters and then click **Export** to apply. The compressed exported events file will be downloaded to your local computer.



Exports include up to the most recent one million events.

3.3. Search Best Practices

Depending on the size of your environment, the logs stored in Cloud Archive may add up to petabytes of data. Searching through this much data can take a long time. However, there are several ways to speed up the search process.

3.3.1. NARROWING DOWN THE TIME RANGE

This may be the most straightforward approach: the longer the time range selected in the Search Bar, the more data Cloud Archive will have to search. In most cases, the search duration will grow linearly with the number of days or months Cloud Archive needs to scan.

3.3.2. USING FIELD NAMES AND VALUES

Cloud Archive stores parsed fields and their values in a separate data partition that is typically much smaller than the raw data it ingested. By using field names and values in the query string, you can help Cloud Archive find logs more efficiently.

The following query, looking for user John within the Okta logs, will run slowly because it forces Cloud Archive to look for these keywords anywhere they might exist in the raw logs.

```
Okta john
```

This query leverages field names and values to point Cloud Archive to the smaller partition and will complete faster than the previous one.

```
vendor:"Okta" AND user:"john"
```

3.4. Data Ingestion Statistics

Exabeam Cloud Archive stores logs beyond the 30-90 days standard retention in Exabeam Data Lake. Visibility into the usage of an archive helps gauge the current utilization and trend in order to plan for future expansions or system tuning.

The upper most bar shows the current day's volume of logs ingested out of the purchased allowance. The second bar displays the expended hours for the day in which the volume is measured.

Below the current day's data, the real-time data ingestion statistics is shown in 1-week, 1-month, 3-months, and 1-year collection windows. Daily ingestion totals are graphed in comparison to your purchased ingestion maximum, indicated by a blue line from the data volume axis. For days where the ingested data is less than the purchased ingestion maximum, the data point is presented in blue. When ingested data exceeds the purchased daily limit, the data point is presented in orange. The chart presents up to 12 months of daily ingested data.

Time on all bar charts and daily totals are referenced in UTC. For convenience, under the daily total, the 24 hour volume is shown in your local time zone (for example, 4PM 12/2 - 4PM 12/3).

To view the data ingestion statistics, navigate to **Settings** > **Data Ingestion** > **Overview**.

Figure 10. Data Ingestion Overview

If your usage is surpassing your purchased ingestion maximum (indicated by orange bars in the graph), please contact your account manager to discuss Exabeam account changes or ingestion increase.

4. Technical Support Information

To contact Exabeam Customer Success, please open a case via Community.Exabeam.com.