Table of Contents

1. About Exabeam Data Lake ................................................................. 5
  1.1. Features Of Data Lake ................................................................. 5
2. Exabeam Data Lake Search ................................................................. 6
  2.1. Exabeam Data Lake Search Page Overview ................................. 6
    2.1.1. Search Bar In Exabeam Data Lake ...................................... 6
    2.1.2. Events Table From Exabeam Data Lake Search Results .......... 7
    2.1.3. Timeline View From Exabeam Data Lake Search Results .......... 10
    2.1.4. Filtered Searches In Data Lake ............................................. 11
  2.2. Performing Searches In Exabeam Data Lake ................................. 13
    2.2.1. Types Of Exabeam Data Lake Queries .................................. 13
    2.2.2. Exabeam Data Lake Event Categorization ............................... 15
    2.2.3. Saved Searches In Exabeam Data Lake ................................. 16
    2.2.4. How To Run An Exabeam Data Lake Cross-Cluster Search ........ 17
3. Exabeam Data Lake Sort Logic .......................................................... 18
4. Exabeam Data Lake Search Fields ..................................................... 19
  4.1. Field Explorer ................................................................. 20
5. Visualize Results In Exabeam Data Lake ............................................. 21
  5.1. Creating A New Visualization In Exabeam Data Lake .................. 21
    5.1.1. Create Visualization From A New Search ............................... 22
    5.1.2. Create Visualization From A Saved Search ............................ 23
    5.1.3. Chart Builder In Exabeam Data Lake ................................. 23
    5.1.4. Aggregations In Exabeam Data Lake ..................................... 24
    5.1.5. Types Of Exabeam Data Lake Visualization Charts ................. 27
    5.1.6. Metric Visualization In Exabeam Data Lake ......................... 28
  5.2. Saving A Visualization In Exabeam Data Lake .............................. 28
6. Exabeam Data Lake Dashboard Setup ................................................ 30
  6.1. Building A Dashboard ........................................................... 30
  6.2. Saving A Dashboard ............................................................... 31
  6.3. Sharing A Dashboard ............................................................... 31
7. Exabeam Data Lake Reports .............................................................. 33
  7.1. Create An Exabeam Data Lake Report ........................................ 34
    7.1.1. Create A New Exabeam Data Lake Report From A Saved Search/Dashboard ........ 34
    7.1.2. Create An Exabeam Data Lake Report From New Search ............ 36
    7.1.3. Multi-Select Exabeam Data Lake Reports ............................... 38
  7.2. Compliance Reports In Exabeam Data Lake ................................. 39
    7.2.1. Out-Of-The-Box Reports: ................................................... 40
    7.2.2. Tags ............................................................................... 40
    7.2.3. Data Lake Event Categorization .......................................... 41
    7.2.4. Tags In Exabeam Data Lake ................................................. 41
  7.3. Import A Report ................................................................. 42
  7.4. How To Suppress Empty Exabeam Data Lake Reports .................... 42
8. Access Restrictions For Saved Objects In Exabeam Data Lake ................. 44
9. How To Forward Alerts Using Correlation Rules In Exabeam Data Lake ........... 46
  9.1. How Correlation Rules Work ................................................... 46
1. About Exabeam Data Lake

Data Lake is Exabeam's Log Management System. It is the data collection, indexing, and visualization engine. The goal of Data Lake is to present log data to the user in a clear and consumable manner. Log data isn't necessarily designed to be easily read by humans and finding what really matters can be time-consuming. That being said, effective log management is essential to both security and compliance. Monitoring, documenting, and analyzing events are a crucial component of an environment's security.

Data Lake allows large scale aggregation and storage of logs and provides access to those logs via a web interface. It enriches log events with contextual information. As data travels from the source Data Lake parses each event, identifies named fields to build structure, and transforms them to converge on a common format for easier, accelerated analysis and business value.

Data Lake is a scalable architecture that can be used as a solitary system or integrated with Exabeam's Advanced Analytics and products.

1.1. Features of Data Lake

Below, we give a short overview of the features of the Data Lake UI. They are described in more detail in the corresponding chapters. We recommend reading these in-depth parts in the order they are presented, since they build on each other.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indexing</td>
<td>Indexing essentially tokenizes fields within the logs, for example creating key-value pairs so that analysts can search for specific values and find the matching events.</td>
</tr>
<tr>
<td>Search</td>
<td>The Search page is where investigations begin and it is the primary way users navigate data in Data Lake. It displays all events in a selected time-span. You can query specific log events, search for specific conditions within a rolling time window, identify patterns in your data, and so on. Searches can be saved for future use, build visualizations, and power dashboard panels.</td>
</tr>
<tr>
<td>Visualize</td>
<td>Visualizations also begin with searches. In Data Lake a visualization is a graph, table, or other visual representation of an aspect of your data. On this page you will create or modify your visualizations.</td>
</tr>
<tr>
<td>Dashboard</td>
<td>Several Visualizations can be placed on one dashboard, as well as Saved Searches. They are made up of panels that contain modules such as search boxes, fields, charts, tables, and so on. This is typically something that an analyst would create for a business level or operational view.</td>
</tr>
<tr>
<td>Reports</td>
<td>Reports for firewall activity, account access, and authentication access attempts are available out-of-the-box. Customers can use the fully customizable reports to create unique views of their environments that can be sent to organizational stakeholders on a regular schedule.</td>
</tr>
<tr>
<td>Correlation Rules</td>
<td>Correlation rules are static rules on incoming logs that alert on known bad or non-compliant behaviors when specified conditions are met, such as service outages which have no security repercussions or any configuration changes on the firewall to meet compliance needs.</td>
</tr>
<tr>
<td>Settings</td>
<td>This is where you perform administrative actions on the Data Lake. You can add indices to, or remove them from, your Data Lake instance, manage your collectors remotely, and other admin tasks.</td>
</tr>
</tbody>
</table>

**NOTE**

If you have installed Exabeam Advanced Analytics, you can now directly access logs of interest in Data Lake. For more information on how to pivot from Advanced Analytics to Data Lake, see View a Data Lake Log from an Advanced Analytics Smart Timelines™ Event.
2. Exabeam Data Lake Search

One of the primary activities of an SOC is searching the log repository for specific events. For example, searching for the activities of a specific user in a given timeframe. Searching is the beginning of any investigation. It is where you access all your logs and filter through them, looking for events that match your criteria. Regardless of if you are building a visualization or a chart, you begin from a search.

You can interactively explore your data from the Search page. You have access to every event that matches the search query within the selected date and time range. You can submit search queries, filter the search results, and view event data. You can also see the number of events that match the search query and get field value statistics. The distribution of events over time is displayed in a histogram at the top of the page.

2.1. Exabeam Data Lake Search Page Overview

This section will isolate and identify the different sections of the Search page and give a brief overview of their functions.

2.1.1. Search Bar in Exabeam Data Lake

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. The total number of matching events are shown above the toolbar. To the right is a drop-down time filter that you can use to filter logs based on various relative and absolute time ranges.

Next to the time filter is a SAVE button that allows you to save searches - this function saves both the query string and the currently selected index pattern.

Clicking on LIBRARY opens a drop-down box that contains all of your saved searches, visualizations, and dashboards. Clicking on a Saved Search will populate the Search Bar with the saved query and launch the search.
For searches that have run a long period without timing out (for example, a search on a small data set does not complete in 5 minutes) or you have initiated in error, you can halt the query by clicking the cancel icon that appears when the search is started:

If your query string extends beyond the length of the search bar, the text will wrap to the next row. To view the whole expanded query, click ... located in the lower right corner of the search bar. Click elsewhere outside the search bar to collapse the expanded view.

To initiate a query with hot-keys, use Control-Enter.

**NOTE**
In a cluster deployment, shards that are not available (such as replicate shards) do not block searches on the remainder of the cluster. Partial results will be returned.

### 2.1.2. EVENTS TABLE FROM EXABEAM DATA LAKE SEARCH RESULTS

When you submit a search request the Timeline, Events Table, and Fields list are updated to reflect the search results. The most recent events that match the query are listed in the Events Table. You can use this to look at individual log messages and display log data filter by fields. If no fields are selected, then the entire log messages are displayed. By default, the table shows the localized version of the time field configured for the selected index pattern and the event source. You can add fields to the Events Table from the Fields list. You can sort the listed events by any indexed alpha column field that’s included in the table.

At the top of each event is a gray highlighted section. These are the fields within the event that Exabeam considers likely to be of interest to the analyst.
To view the event data as a separate page, click the event link. You can bookmark and share this link to provide direct access to a particular event.

The **Table** view allows you to create your own table with fields of your choosing, to be arranged as you see fit.

While viewing in **Table** view, sort the results according to a single column or multiple columns by clicking **SORT COLUMNS** or by clicking the arrow next to any column name.

The **SORT COLUMNS** menu lets you select the columns by which you want to sort data in the table. You can configure each column to sort by ascending or descending order. Additionally, you can sort search result tables on the **Search** page and data tables on the **Visualizations** by multiple columns.

In the **Raw** view, each event shows the first four lines of the event log. Generally, the first four lines include the timestamp and the actual payload of the event. You can see the full event payload by clicking **Show more** and then contract it back to a truncated view by clicking **Show less**.
The **Share** button allows you to create a shortened URL which links directly the current search results table. You can **COPY** and share the snapshot link with members of your organization who have access to the Data Lake UI.

The **Export** button allows you to save your search results as a PDF or a CSV file. For CSV exports, set the number of search results to export (up to 1 million). Large CSV exports are split into multiple CSV files and then exported to a ZIP file.

For each log event, you can select the **More Options** icon to open the submenu for export options:

- **Copy Raw Log** -- Cache the raw log text to your local buffer.
- **Copy Link** -- Cache a shareable link to your local buffer.
- **Open in New Tab** -- Present the parsed fields and raw log in a new tab in your web browser.
2.1.3. TIMELINE VIEW FROM EXABEAM DATA LAKE SEARCH RESULTS
The Timeline View is a date histogram bar graph that shows the count of logs over time, matched by the search and time filters. You can click on the bars to narrow the time filter.

To set a Time Filter from the histogram, do one of the following:

• Click the bar that represents the time interval you want to zoom in on.

• Click and drag to view a specific timespan. You must start the selection with the cursor over the background of the chart—the cursor changes to a plus sign when you hover over a valid start point.

The histogram lists the time range you’re currently exploring, as well as the interval range that is being used. To change the intervals, click the link and select an interval from the drop-down. The default behavior automatically sets an interval based on the time range.

To set a Time Filter from the histogram, do one of the following:

• Click the bar that represents the time interval you want to zoom in on.

• Click and drag to view a specific timespan. You must start the selection with the cursor over the background of the chart—the cursor changes to a plus sign when you hover over a valid start point.

The histogram lists the time range you’re currently exploring, as well as the interval range that is being used. To change the intervals, click the link and select an interval from the drop-down. The default behavior automatically sets an interval based on the time range.

You can narrow your search further using the Time Picker tool. See Selecting a Timeframe.

You can collapse or display the Timeline by clicking the Open/Close icon.

2.1.3.1. Selecting a Timeframe in Exabeam Data Lake Search Results
The time filter restricts the search results to a specific time period. You can set a time filter if your index contains time-based events and a time-field is configured for the selected index pattern. By default, the
time filter is set to the last 15 minutes. You can use the Time Picker to change the time filter or select a specific time interval or time range in the timeline view at the top of the page.

For all time-based data, you can select the time span that you want to analyze in the current view at the top right of the window. There are multiple ways to get to the events you are interested in: either use the Quick tab to select a date range like Today or Last 1 hour or use the Relative and Absolute tabs to specify more specific time spans you want to look at.

After you select a time range, you will see a timeline view at the top of the screen, which will show the distribution of events over time.

To use the Time Picker, either drag your mouse over a specific span of time.

Or, you can select a bar on the timeline that represents the time interval you want to zoom in on.

### 2.1.4. FILTERED SEARCHES IN DATA LAKE

In addition to using time constraints to narrow the amount of data to search, you can apply filters using context tables to optimize your queries. Filtered Searches can also be applied to dashboards, visualizations, and reports.
**NOTE**

Please consider the following:

- PDF export of filtered searches is currently not supported
- Filtered searches work only with key-only context tables of no greater than 10k records
- One context table at a time can be applied per filtered search
- Context table records must match the format of the field being queried
- Do not use string values with numeric characters
- Exact value matches will be applied

1. To narrow the data that you will run the query against, click **+ Context table** below the Search field to expand the menu to add a new filter to searches.

2. Select the **Field**, **is** or **is not** condition, and **In context table** that you want to apply. You can select from the drop down list or start typing in the fields to get possible matches displayed.

3. Click **ADD** to apply the filter. The filter will appear below the Search field.

4. Click **SAVE** to store the query to the library.
You can click the filter to edit it. If the parameter has already been applied, a check mark will appear next to the record. Click **UPDATE** to apply the filter.

For data formats supported in filtered searches, please see the [Exabeam Search Quick Reference Guide](#).

### 2.2. Performing Searches in Exabeam Data Lake

Data Lake is built on top of Elasticsearch, which uses the Lucene query language. For more detailed information on syntax and search options, see [Data Lake Search Quick Reference Guide](#).

**NOTE**

In a cluster deployment, shards that are not available (such as when a node goes down) do not block searches on the remainder of the cluster. Partial results will be returned.

#### 2.2.1. TYPES OF EXABEAM DATA LAKE QUERIES

Data Lake accepts searches in the Lucene query language. A query is broken up into terms and operators. There are two types of terms: Single Terms and Phrases. This section covers some of the basic operators for conducting searches.

- A Single Term is a single word such as "test" or "hello".
- A Phrase is a group of words surrounded by double quotes such as "hello world".
Multiple terms can be combined with Boolean operators to form a more complex query.

### 2.2.1.1. Text Searches

By default, the search box performs unstructured text searches. It searches for entries containing any of your search terms and a hyphen is considered a delimiter. This means that if no specific field is indicated in the search, the search will be done on all of the fields that are being analyzed. It will not tell you if your search has the wrong syntax.

[NOTE]

Text searches are not case sensitive. This means that 'category' and "CaTeGory" return the same results. Use double quotes ("") to search for an exact match.

### 2.2.1.2. Field Level Searches

The query language allows you to search inside any field, simply enter the name of the field and then a colon.

Some examples:

- To just search inside a field named “lang”
  
  `lang:en`

- To search for the language English or Spanish in the "lang" field
  
  `lang:(en OR es)`

Like the selected fields, the entered query will be persisted if you save your search.

You can search a range within a field. If you use brackets [], this means that the results are inclusive. If you use {}, this means that the results are exclusive.

Using the _exists_ prefix for a field will search the events to see if the field exists.

You cannot use wildcards inside of phrases.

For information on queries, see *Data Lake Search Quick Reference Guide*.

### 2.2.1.3. Logical Statements

Logical statements enable you to use more than one condition in a query. You can use parentheses to define complex logical statements and be sure that you use the proper format such as capital letters to define logical terms like AND or OR.

In some cases, you might want to compare the results of two separate queries. Data Lake can handle multiple queries by joining them with a logical OR.

To search for the logon event 4768 and the user Barbara Salazar:

`event_code:4768 AND user:bsalazar`
### Exabeam Data Lake Search

<table>
<thead>
<tr>
<th>Search Type</th>
<th>Operator(s)</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Search</td>
<td><code>*</code></td>
<td><code>*</code></td>
</tr>
<tr>
<td>Literal String</td>
<td><code>&quot;&quot;</code></td>
<td>&quot;geo-address&quot;</td>
</tr>
<tr>
<td>Single Field</td>
<td><code>&lt;Field name&gt;</code></td>
<td>country:</td>
</tr>
<tr>
<td>Missing Field</td>
<td><code>missing:</code></td>
<td>missing:vpn</td>
</tr>
<tr>
<td>Present Field</td>
<td><code>_exists_:</code></td>
<td><em>exists</em>.vpn</td>
</tr>
<tr>
<td>Wildcard</td>
<td><code>*</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td>? for any number of characters</td>
<td></td>
</tr>
<tr>
<td></td>
<td>for one character (Cannot be used in <code>_type</code> fields.)</td>
<td></td>
</tr>
<tr>
<td>Negative Terms</td>
<td><code>!</code>;<code>-</code>;<code>NOT</code></td>
<td><code>-VPN</code></td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>!VPN</code></td>
</tr>
<tr>
<td></td>
<td></td>
<td><code>NOT VPN</code></td>
</tr>
<tr>
<td>Range Search</td>
<td><code>[number TO number]</code></td>
<td>user.listed_count:[0 TO 10]</td>
</tr>
</tbody>
</table>

For more detailed information about running searches, see [Exabeam Search Quick Reference Guide](#).

### 2.2.2. EXABEAM DATA LAKE EVENT CATEGORIZATION

Data Lake supports multiple categorization attributes for each log or event type defined in the product. Different vendors use different fields and terms in their logs.

Categorizing events provides a consistent taxonomy for queries, reports, visualization, dashboard, search, and correlation rules. Our out-of-the-box compliance reports leverage this nomenclature.

For example, a log has the following value:

`exa_activity_type: authentication/local_logon`

This log will also be returned in the query:

`exa_activity_type=authentication`

Current categories are:

- `exa_category`
- `exa_device_type`
- `exa_activity_type`
- `exa_outcome`

Examples:

- `exa_activity_type = account-management/user/create`
- `exa_device_type = operating-system/network/firewall`
- `exa_outcome = success/allow`
For a complete list of Exabeam event categories, see How to Run Query Searches in Exabeam Data Lake > "Searches using Exabeam exa_category".

2.2.3. SAVED SEARCHES IN EXABEAM DATA LAKE

Saving searches enables you to reload query results quickly and use them as the basis for Visualizations, Dashboards, and Reports. A search can be saved into a library by clicking on the SAVE. You can access the Search Library at any time to get a list of all your saved searches. Selecting a saved search will populate the search box with the query and launch the search.

To Save Current Search:

- Click SAVE in the toolbar.
- Enter a name for the search and click SAVE TO LIBRARY.

To Open a Saved Search:

- Open the Search Library.
• Select the search you want to open.

To Edit or Delete a Saved Search:

The Search Library contains a list of all Saved Searches. To edit any Saved Search click on the vertical ellipse to the right of the date field. This opens a drop-down menu and from here you can edit the search or delete it.

2.2.4. HOW TO RUN AN EXABEAM DATA LAKE CROSS-CLUSTER SEARCH

In a multi-cluster deployment, you can perform searches simultaneously across all trusted clusters. Ensure you have permission to run a cross-cluster search and that the clusters of interest are available. You must have at least one remote cluster configured. For more information on setting up a searchable remote cluster, see Data Lake Administrator Guide > Configuring Data Lake > Cross-Cluster Search.

If you have permission to conduct cross-cluster searches, there will be a Local Cluster menu above the search field.

Select the clusters you want to apply the search to. Compose the query and its parameters as you would for typical searches, following prescribed syntaxes (see Type of Queries). To configure a remote cluster, see the Data Lake Administration Guide > Configuring Data Lake > Cross Cluster Searches.

⚠️ WARNING

Cross cluster search results export is limited to 10,000 events per search query. For local clusters, up to 1,000,000 events per search query can be exported in CSV format.
3. Exabeam Data Lake Sort Logic

The following is the sorting order using in Data Lake. Note that leading spaces are given greater weight than all other character or number when records are sorted.

Ascending order is:

1. Blank spaces (" ").
2. Underscore ('_').
4. Lower case alphabetic character.
5. Upper case alphabetic character.
6. Field values longer than maximum char limit of 256).

Descending order is:

The reverse of the above.

The following is an example of ascending order:

```
david
David
_david
_David
1david
2david
3david
david
daVid
daVid
David
```
4. Exabeam Data Lake Search Fields

Data Lake displays a list of fields that found in the events of the search results, at the left of the UI. You can click the field to add a column containing the contents of this field to the table. No matter what fields you have added as columns, you can always expand a row on the caret in the front. You can also remove fields that you don’t want to see as columns anymore in the section Selected Fields above the field list on the left.

You can expand any field in the fields list on the left by clicking on it. It will reveal the list of the most common values for that field. Use the – and + magnifier icons to quickly add a filter to show only events containing that value (+) or to exclude all events with that value (-).

If you add filters that way, this field will be added as a search term within the query.

Filters can also be set by expanding the table rows on the right which show the event contents and using the filter buttons which appear there. Note that events may contain fields which are not indexed and can thus not be used for filtering. You won’t find any filter buttons for those.

Additionally, click the View field visualization link to create a new visualization from a single selected field. Once the new visualization is created, you can further customize the view by adding or removing top terms you want to review.
4.1. Field Explorer

In addition to using manually created search strings, users have the option to filter data using out-of-the-box filters available in the Search UI.

The Field Explorer is the quick pick tool for viewing captured data in known categories (both out-of-the-box and custom filters). Click on the hyperlink for a given sub-category and menu of known values are listed to filter further. View field visualization can be selected to immediately visually organize data from the shown list.
5. Visualize Results in Exabeam Data Lake

Visualizations are graphic and interpretive representations of the results of one specific search. The Visualize page is where you create, modify, and view your own custom visualizations.

Visualizations are based on the filters that you set on your data. Data Lake then performs aggregations and creates distinct representations of the results. There are several different types of visualizations, ranging from Vertical Bar and Pie Charts. We will cover each of these chart types and their recommended uses. This chapter will also go into some depth on bucket and metric aggregations - concepts that are necessary in order to understand what Visualize is doing with your data.

Visualizations can be shared with other users as well as placed on dashboards in order to display trends you are interested in tracking.

5.1. Creating a New Visualization in Exabeam Data Lake

Visualizations are generated from the data queried during its creation or from saved searches. Exabeam offers an array of linear and spatial chart types to present search results.

From the Visualize Page, choose New Visualization at the top right.

You will then be asked to choose which type of chart you would like to create. This can be changed later. For more detailed information on each of the charts and their recommended uses, see Types of Charts.
1. Click to choose your chart type you need.

2. Decide whether your visualization is from either an existing search that already resides in your Search Library or an entirely new search.

5.1.1. CREATE VISUALIZATION FROM A NEW SEARCH

If you choose to create a Visualization from a New Search you will be taken to the Search Page, where you can enter queries using the Lucene query language and Data Lake will use this as a filter on the data. Visualize then runs aggregations on the results of this filtered data.

When your search results reflect the data you want to visualize, click the Visualize Search button.

Instead the query data will be stored with the visualization when you save it. This means that if you place the visualization on a Dashboard, the query that you stored with the visualization will still apply and be updated when you update the Dashboard.

After clicking the Visualization Search button, you will be taken to the Chart Builder Page.
5.1.2. CREATE VISUALIZATION FROM A SAVED SEARCH

If you choose to create a Visualization from a previously Saved Search you will be taken to your Search Library. Choose the Search you want to Visualize and click the Visualize button next to it. You will be taken to the Chart Builder page.

5.1.3. CHART BUILDER IN EXABEAM DATA LAKE

Chart Builder is your visualization workbench for creating charts from searches.

- Control the data being displayed using Builder Panel 1.
- Use the Styles Panel 2 to vary the look and feel of the chart.
- Select a color palette 3. Click this to see the different color palettes available.
- Create a new visualization 4.

5.1.3.1. Builder

The information in the Builder panel will change slightly depending on which chart type you've chosen. There are two aggregation types to configure - the Metric Aggregation which configures the Y-axis and the Bucket Aggregation which configures the X-axis.

For example, you could build a bar chart that shows the distribution of events by geographic location by specifying a terms aggregation on the src_country field.
See *Aggregations* on for more technical detail regarding how the different aggregations function.

### 5.1.3.2. Style
The information in the Styles panel will change slightly depending on which chart type you’ve chosen. Here you can choose the line types, the scale, the position of the legend, etc. After making any changes be sure to click the **Update** button at the bottom of the panel to see a preview of your chart.

### 5.1.4. AGGREGATIONS IN EXABEAM DATA LAKE
There are two separate types of aggregations - bucket and metric. Metrics aggregation happens on the y-axis, while buckets are on the x-axis. To understand how visualizations work it is essential to understand how the system aggregates your data.

In the visualizations the bucket aggregation will usually be used to determine the "first dimension" of the chart (e.g. for a pie chart, each bucket is one pie slice; for a bar chart each bucket will get its own bar). The value calculated by the metric aggregation will then be displayed as the "second dimension" (e.g. for a pie chart, the percentage it has in the whole pie; for a bar chart the actual high of the bar on the y-axis).

Because metric aggregations make more sense when they are run on buckets, we will cover Bucket Aggregations first.

#### 5.1.4.1. Exabeam Data Lake Bucket Aggregations
Bucket aggregations group all of your events into several buckets, each containing a subset of the indexed events. Which bucket a specific event is sorted into can be based on the value of a specific field, a custom filter, or other parameters.
**Date Histogram**
The Date Histogram aggregation requires a field of type date and an interval. It will then put all the events into one bucket, whose value of the specified data field lies within the same interval. Besides the easily recognized interval values (minutes, seconds, etc.) there is the value `auto`. When you select auto the actual time interval will be determined by Data Lake depending on how large you want the histogram to be, so that a reasonable number of buckets will be created.

*Example:* You can construct a Date Histogram on the `event_time` or `@timestamp` field of all messages with the interval minute. This will create a bucket for each minute and each bucket will hold all events that were generated in that minute.

**Histogram**
A Histogram is similar to a Date Histogram except that it can be used on any numerical field. As with the Date Histogram, you specify a number field and an interval. The aggregation then builds a bucket for each interval and includes all events whose value falls inside this interval.

*Example:* You can construct a Histogram on the `event_code` field with the interval of hour. This will create a bucket for each hour and each bucket will hold all events that were generated in that hour.

**Range**
The Range aggregation is like a manual histogram aggregation. You also need to specify a field of type number, but you have to specify each interval manually. This is useful if you either want differently sized intervals or intervals that overlap.

*Example:* Filtering on the `riskscore` field, you can plot the count of critical users (those with a risk score greater than 90), high risk users (those with a risk score between 80 and 90), and non-risky users (those below 80).

**Terms**
A Terms aggregation creates buckets based on the values of a field. You specify a field (which can be any type) and it will create a bucket for each of the disparate values that exist in that field.

*Example:* You can run a Terms aggregation on the field `src_country` which holds the country name of the user. Afterward you would have a bucket for each country and in each bucket the events of all users in the country.

**Filters**
A Filter aggregation is extremely flexible. You can specify a filter for each bucket and all events that match the filter will be in that bucket. Each filter is a query, as we discussed in the Search chapter of this document. The filter that you specify for each bucket is entirely at your discretion.

*Example:* You can create a filter aggregation with one query being `src_country: canada` to see all the events that originated in Canada.

**Significant Terms**
Significant Terms are "uncommonly common" terms in a set of events. In other words, this aggregation returns unusual or interesting terms in a given subset of events. Given a subset of events, this
aggregation finds all of the terms which appear in this subset more often that is expected from term occurrences in the entire events set. It then builds a bucket for each of the significant terms - each bucket contains all the events of the subset in which this term appears. The size parameter controls how many buckets are constructed, i.e. how many significant terms are calculated.

The subset on which to operate the significant terms aggregation can be constructed by a filter or you can use another bucket aggregation first on all events and then choose Significant Terms as a sub-aggregation which is computed for the events in each bucket.

Example: Using the search field at the top to filter events for `event_data.FailureReason` and setting the size parameter to 5 will return the top 5 most unusual failure reasons for a failed logon.

**Date Range**

Date Range aggregation is similar to Range aggregation, except that from and to values can be expressed. It is an aggregation dedicated to date values.

The date range expression itself starts with an anchor date, which can either be now, or a date string ending with `. This anchor date can optionally be followed by one or more expressions.

<table>
<thead>
<tr>
<th>y</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>Months</td>
</tr>
<tr>
<td>w</td>
<td>Weeks</td>
</tr>
<tr>
<td>d</td>
<td>Days</td>
</tr>
<tr>
<td>h</td>
<td>Hours (hour of day 1-12)</td>
</tr>
<tr>
<td>H</td>
<td>Hours (hour of day 0-23)</td>
</tr>
<tr>
<td>m</td>
<td>Minutes</td>
</tr>
<tr>
<td>s</td>
<td>Seconds</td>
</tr>
<tr>
<td>now</td>
<td>Current time</td>
</tr>
</tbody>
</table>

Examples of date expressions:

<table>
<thead>
<tr>
<th>Expression</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>now-1m/d</code></td>
<td>The current time minus one month, rounded down to the nearest day.</td>
</tr>
<tr>
<td>`2017-01-01</td>
<td></td>
</tr>
</tbody>
</table>

5.1.4.2. Exabeam Data Lake Metric Aggregations

After you have run a bucket aggregation on your data, you will have several buckets with events in them (the x-axis). You now specify one metric aggregation to calculate a single value for each bucket (the y-axis). The metric aggregation will be run on every bucket and result in one value per bucket.

**Count**

This is not really an aggregation. It just returns the number of events that are in each bucket as a value for that bucket.

Example: If you want to know how many events are from which country, you can use a term aggregation on the field `src_country` (which will create one bucket per source country) and afterward run a count metric aggregation. Every country bucket will have the number of events as a result.
**Average/Sum**
For the average and sum aggregations you need to specify a numeric field. The result for each bucket will be the sum of all values in that field or the average of all values in that field respectively.

*Example:* You can have the same country buckets as above again and use an average aggregation on the `user.logons` field to get a result of how many logons a user in that country has on average.

**Max/Min**
Like the average and sum aggregation, this aggregation needs a numeric field to run on. It will return the minimum value or maximum value that can be found in any event in the bucket for that field.

*Example:* If we use the country buckets again and run a maximum aggregation on the `logon_count` we would get for each country the highest number of logons a user had in the selected time period.

**Unique Count**
The unique count will require a field and count how many different/unique values exist in events for that bucket.

*Example:* This time we will use range buckets on the `user.logons` field, meaning we will have buckets for users with 1-100, 100-1000 and 1000- events. If we now run a unique count aggregation on the `src_country` field, we will get for each user's range the number of how many different countries users with x number of events would come. That could show us that there are users from 18 different countries with 1 to 100 events, from 30 different countries for 100 to 1000 events and from 4 different countries for 1000 events and above.

**Percentiles**
A percentiles aggregation is a bit different, since it won't result in one value for each bucket, but in multiple values per bucket. These can be shown as different colored lines in a line graph, for example.

When specifying a percentile aggregation, you have to specify a number value field and multiple percentage values. The result of the aggregation will be the value for which the specified percentage of events will be inside (lower) as this value.

*Example:* You specify a percentiles aggregation on the field `src.country_count` and specify the percentile values 1, 50 and 99. This will result in three aggregated values for each bucket. Let’s assume we have just one bucket with events in it. The 1 percentile result (and e.g. the line in a line graph) will have the value 7. This means that 1% of all the events in this bucket have a count of 7 or below. The 50-percentile result is 276, meaning that 50% of all the events in this bucket have a count of 276 or below. The 99-percentile have a value of 17000, meaning that 99% of the events in the bucket have a count of 17000 or below.

### 5.1.5. TYPES OF EXABEAM DATA LAKE VISUALIZATION CHARTS

**Area Chart**
An area chart displays a time series of logs as a set of points connected by a line, with all the area filled in below the line. This is a stacked chart that displays data contiguously. Commonly used to represent data that occurs over a period of time. Useful for showing proportional data that occurs over time. These charts are good for displaying totals for all series as well as the proportion that each series contributes to the total.
**Data Table**

A data table will display the raw data that in other visualizations would be rendered into graphs. Each column represents a data field and the rows hold the values for that field.

Data tables are useful for efficiently organizing large amounts of data.

**Line Chart**

A style of chart that is created by connecting a series of data points together with a line. A line chart can give you a fairly good idea of trends over time and are good for high density time series.

**Pie Chart**

The classic pie chart is perfect for displaying the parts of a whole. Ideally, keep the slices to 7 or fewer in order for the data to be clear and useful.

**Vertical Bar Chart**

The vertical bar visualization is best suited for visualizations where data on your x-axis is discrete, because each x-axis value will get its own bar(s) and there won’t be any interpolation done between the values of these bars.

You have three bar modes available:

- *Stacked*: this behaves like an area chart; the bars are stacked onto each other.
- *Percentage*: this uses 100% height bars, and only shows the distribution between the different buckets.
- *Grouped*: this places the bars for each x-axis value beside each other.

---

**5.1.6. METRIC VISUALIZATION IN EXABEAM DATA LAKE**

A metric visualization displays the result of a metrics aggregation. There is no bucket aggregation done; it will always apply to the whole data set that is currently selected (to change the data set, change the query in the filter box). The only view option that exists is the font size of the displayed number.

**5.2. Saving A Visualization in Exabeam Data Lake**

After you have created a visualization that you’re satisfied with you’ll need to save it. Next to the search bar is a Save button. Selecting this will open a drop-down panel that allows you to name your visualization and save it. You can then return to the Visualize Page to access all of your saved Visualizations in the future.

It is important to note that when switching between tabs (Visualize, Search, Dashboard) Data Lake will return you to the same visualization you’ve been editing. However, the edits are only temporarily retained while the session is open. Unless you save the visualization, any edits will not be persisted in Data Lake.
Saving visualizations is a necessary step for creating Dashboards. As you will learn in the next chapter, Dashboards are built using the Visualizations that you've created in the Visualize tab.
6. Exabeam Data Lake Dashboard Setup

The Data Lake Dashboard page is where you can create, modify, and view your own custom dashboards. Essentially, they are a place to consolidate visualizations and saved searches. With a dashboard, you can combine multiple visualizations onto a single page, then filter them by providing a search query or by selecting filters by clicking elements in the visualization. You can easily use one visualization on multiple dashboards and if you edit the visualization is will be updated automatically on every dashboard you use. Unsaved reports will be impacted by changes in searches and visualizations.

Dashboards are useful for when you want to get an overview of your logs and make correlations among various visualizations and events. They can be shared among your colleagues and used in reporting. Each chart will refresh itself with the most recent data, making them useful for performing recurrent tasks.

6.1. Building a Dashboard

To create a Dashboard, select the Dashboard tab at the top of the page.

If you have not created a dashboard before, you will see a mostly blank page that says Ready to get started?

To add visualizations to your Dashboard, tick the box(s) next to one of your Saved Visualizations listed to the left of the page. Click the Add to Dashboard button at the bottom right of the page.

**NOTE**

Having a saved Search or Visualization is a prerequisite for building a Dashboard.

Selecting any of your saved searches or visualizations will add them as a tile to the Dashboard. Use the filter at the top to search for a specific visualization or simply browse the list. Once you have added all the visualizations you want, collapse the Add menu. Now you can arrange each element to your satisfaction.
When you hover the mouse over a visualization a menu appears in the top right corner of the element. From this menu you can Edit, Move, or Delete the element respectively.

Selecting the **Edit** pencil will jump you to the Visualize page for that particular visualization. From here you can make any changes you like and save them.

To **Move** the visualization hover over the Move icon in the upper right corner and drag it to the desired location. The visualization can also be re-sized by hovering the mouse over any of the tile's corners and dragging.

**Deleting** will only remove it from this particular Dashboard, it will not delete the visualization itself.

Dashboards can be filtered further by entering a search query, changing the time filter, or clicking on the elements within the visualization. As with all the other pages, you can use the query language to enter queries in the top search box. This query will filter the data for all of the visualizations placed on the dashboard. If you have stored a query with a visualization, both queries will apply. This means you can use the dashboard search to filter out data, without compromising any logic inside the visualization.

You can also select particular elements within the visualization themselves. For example, if you click on a particular color segment in the histogram, Data Lake will allow you to filter on the significant term that the segment represents. Filters can be applied and removed as needed. The search and time filters work just like they do in the Search page, except they are only applied to the data subsets that are presented in the dashboard. Changing the time interval will apply to every visualization on the dashboard.

### 6.2. Saving a Dashboard

When you are pleased with the Dashboard's aesthetic, click the **Save** icon at the top right of the page. Name the Dashboard and Save it. As with visualizations and searches, when you make changes to the dashboard, you need to press save again to store these changes permanently. To store the time period specified in the time filter with the dashboard, select **Store time with dashboard**. **Save As** gives you the option to save the Visualization under a new name instead of overwriting the existing version.

All Dashboards are saved in your **Search Library** and can be accessed from anywhere in Data Lake.

### 6.3. Sharing a Dashboard

The **Share** button will open a drop-down menu that contains a link to your saved Dashboard. It can be accessed by anyone with valid Data Lake credentials. If you copy out the link written in the **src=" . . "**
attribute and share this, your users won’t have the option to modify the dashboard. This is not a security feature, since a user can just remove the embed from the URL, but it might be useful so people don’t modify the dashboards by accident.
7. Exabeam Data Lake Reports

Repetitive summaries and queries for known data of interest can be stored in reports.

The Reports page has two tabs:

- The My Reports tab lists only the Reports that you have created.
- The All tab lists all available Reports including those you have created as well as any Reports shared with you.

From here you can Edit or Delete reports by selecting the pencil icon or the trash icon that appears when you hover your mouse over individual reports.

In Data Lake, a Report can be created from either the enhanced view of a search’s results (the results of a search query), a table view (the results of a search query), or a dashboard (potentially multiple search results) or a brand new Search.

NOTE
Reports cannot be created from Visualizations alone. If you would like to create a Report with only a Visualization, save a Dashboard that includes the Visualization and create a Report from the Dashboard.

Reports have the following attributes:

- **Report Name**
- **Report Description**
- **Type**: What the Report is created from (Enhanced View, Table View, or Dashboard).
- **Creator**: Administrator who created the report.
- **Date**: When the report was last generated in the system.
- **Time Range**:
  - When Reports are scheduled, this parameter reflects a variable time range.
• When Reports are unscheduled, this parameter reflects whether queries in the report have a fixed or a variable time range.

7.1. Create an Exabeam Data Lake Report
You can create a Report from a Saved Search or Dashboard, or from a brand new Search.

A few notes about the time frame of reports:

• If a report is created from an enhanced view search, table view search, or dashboard of a newly saved search, the time frame in the report will be the same as the time frame of the saved search or dashboard. Note that the time frame applies to all searches and visualizations within the dashboard.

• Modification of the time frame must be performed in the Saved Search or Saved Dashboard.

Select the Reports Tab at the top right, which brings you to the Reports Page, click on the blue Add button at the bottom right. From here you can choose whether to create a Report from a Saved Search or Saved Dashboard or create a Report from a new search.

7.1.1. CREATE A NEW EXABEAM DATA LAKE REPORT FROM A SAVED SEARCH/DASHBOARD
If you choose to create a Report from a Saved Search or Dashboard, your Saved Library will be opened. Select the Saved Search or Saved Dashboard from which you would like to create a Report by selecting the Report.

The drop-down card gives you more information regarding the Dashboard or Search you have selected. Click Add to Report.
A preview of your report will appear on the next page. If you are satisfied with the preview, click **Add to Report** at the top right.

You will be taken to the Report Details page and asked to give the Report a name and description.

From here you can:

- Select **Save Report**, which saves the report to your Report Library without sending.
- Tick the **Send Now** box and then click **Save Report**, which immediately runs the report and saves the Report to your Reports library.
- Schedule your report by ticking the **Schedule a Report** box and enter the frequency with which you would like the report sent and well as all of the recipients who should receive it.
• Select **Save Report**. You will be returned to your **Reports Page**.

### 7.1.2. CREATE AN EXABEAM DATA LAKE REPORT FROM NEW SEARCH

If you choose to create a Report from a New Search, the Search landing page will be opened. Data Lake accepts searches in the Lucene query language. Input the search terms that you would like your Report to be based on.

Click the blue **Add to Report** button.
Your search will be run and the results displayed on the next page. This is a preview of what your Report would look like with those search terms.

If you are pleased with the Report preview, click **Add to Report**. If not, return to the previous screen and edit your search terms.

You will be taken to the Report Details page and asked to give the Report a name and description.
From here you can:

- Select **Save Report**, which saves the report to your Report Library without sending.
- Tick the **Send Now** box and then click **Save Report**, which immediately runs the report and saves the Report to your Reports library.
- Schedule your report by ticking the **Schedule a Report** box and enter the frequency with which you would like the report sent and well as all of the recipients who should receive it.

⚠️ **WARNING**

Scheduled reports with 3 billion or more logs can cause data outage in the PDF output.

- Select **Save Report**. You will be returned to your **Reports Page**.

### 7.1.3. MULTI-SELECT EXABEAM DATA LAKE REPORTS

Reports can be multi-selected in order for the user to perform mass operations on them. The following changes can be made:

- **Scheduled** - Reports can be scheduled. When multiple reports are selected and scheduled, they will be put on the same schedule and delivered to the same list of email recipients.
- **Export Template** - Reports can be exported. When multiple reports are selected, and the export button is clicked, all the reports (along with the underlying search, visualization, and schedule) are downloaded as a zip archive of JSON files.
- **Delete** - Reports can be deleted. When multiple reports are selected and the delete button is clicked, all reports will be deleted. This action cannot be undone.
7.2. Compliance Reports in Exabeam Data Lake

Exabeam offers compliance report templates for both U.S. and international regulations. Data Lake supports the following compliance reports out-of-the-box:

**NOTE**
Please contact your Data Lake administrator to enable/disable any out-of-the-box compliance reports listed below.

- **GDPR** – Protects the data and privacy for all individuals within the European Union (EU) and the European Economic Area (EEA). It also addresses the export of personal data outside the EU and EEA areas.

- **GPG** – Protects U.K. citizens by ensuring protective monitoring of business processes and technology. It provides visibility and understanding of who is accessing an organization’s sensitive data.

- **HIPAA** – Protects sensitive patient data. Any company that deals with protected health information (PHI) must ensure that all the required physical, network, and process security measures are in place and followed.

- **NIST** – Protects U.S.-based organizations in the science and technology industry by producing standards and guidelines to help these federal agencies meet the requirements of the Federal Information Security Management Act (FISMA).

- **PCI DSS** – A set of security standards designed to ensure that all companies that accept, process, store, or transmit credit card information maintain a secure environment.
- **SOX** – Protects shareholders and the general public from accounting errors and fraudulent practices in enterprises, and improves the accuracy of corporate disclosures.

Administrators have access to Exabeam-provided reports in Settings > Security Content > Exabeam Reports. Permission to access these reports must be granted by a user assigned the Administrator role.

### 7.2.1. OUT-OF-THE-BOX REPORTS:

- Access Granted/Revoked Activity
- Account Management Activity
- Successful Database Logon Activity
- Failed Database Logon Activity
- Audit Log Change Activity
- File Alert Activity
- Physical Access Activity
- Default Credential Usage/Change Activity
- Denied Web Access Activity
- Privileged Access
- Remote Session Overview
- Failed VPN Logons and Remote Session Timeouts
- Overall Log Monitor
- Protocols by Network Traffic
- Database Deletions
- Top Attackers
- Exabeam AA - Top Suspicious Users
- Windows User Privilege Elevation
- Unix User Privilege Elevation
- Vendor Authentication Activity
- Signature Update
- Successful Application Logon Activity
- Failed Application Logon Activity
- User Account Lockout Activity
- Disabled User Account Summary
- Deleted User Account Summary
- User Account Creation Summary
- Data Loss Prevention Activity Summary
- Object Access Summary
- Account Logout Summary
- System Startup and Shutdown Summary
- Security Alert Summary - Users
- Security Alert Summary - Impacted Hosts
- Security Alert Summary - Origin Hosts
- Windows Audit Failure Summary by Users
- Windows Audit Failure Summary by Hosts
- Vulnerabilities Detected
- System Critical and Error Activity Summary
- Policy Activity Summary

Only users with administrator privileges can view these reports unless they are shared amongst various roles.

These reports cannot be edited. However, you can make copies of reports to edit by you or roles you have shared the copy with.

### 7.2.2. TAGS

Data Lake Analysts are able to see which report is mapped to which regulation, sort and search by tag. Tags provide the ability to group different types of objects together. They can also be used in searching and filtering. Within Data Lake there are hundreds of out-of-the-box reports that could map to multiple regulations. Analysts can edit tags that Exabeam has added to a report as well as add tags to reports that they have created. When reports are exported, the tags will be included as part of the export.
7.2.3. DATA LAKE EVENT CATEGORIZATION

Data Lake supports multiple categorization attributes for each log or event type defined in the product. Different vendors use different fields and terms in their logs.

Categorizing events provides a consistent taxonomy for queries, reports, visualization, dashboard, search, and correlation rules. Our out-of-the-box compliance reports leverage this nomenclature.

For example, a log has the following value:

exa_activity_type: authentication/local_logon

This log will also be returned in the query:

exa_activity_type=authentication

Current categories are:

exa_category
exa_device_type
exa_activity_type
exa_outcome

Examples:

exa_activity_type = account-management/user/create
exa_device_type = operating-system/network/firewall
exa_outcome = success/allow

7.2.4. TAGS IN EXABEAM DATA LAKE

Data Lake Analysts are able to see which report is mapped to which regulation, sort and search by tag. Tags provide the ability to group different types of objects together. They can also be used in searching and filtering. Within Data Lake there are hundreds of out-of-the-box reports that could map to multiple
regulations. Analysts can edit tags that Exabeam has added to a report as well as add tags to reports that they have created. When reports are exported, the tags will be included as part of the export.

7.3. Import a Report

Import a report you manually created to move it between POC, UAT, and production clusters; or a report from a content pack to get data and dashboards about external partners and vendors.

If you move from a proof-of-concept (POC) or User Accepting Testing (UAT) cluster to a production cluster, and you also want to move any reports you manually created using searches or visualizations, you must export and import them to the new cluster. You can import the report only if the clusters are of the same version or adjacent versions.

1. If you’re moving between a POC, UAT, or production cluster, ensure that you exported the report(s) you’re moving. If you downloaded a content pack in a .tar.gz format, ensure that you untar it:

   ```
sos
tar -xvf [tarfile].tar.gz
   ```

2. Navigate to the REPORTS page.

3. Click IMPORT REPORT.

4. Select and upload the JSON file from your file system. You can only import a Data Lake report exported from another cluster, or JSON files provided in a security content package. The reports are sorted alphabetically.

   Data Lake automatically creates dashboards from these reports. To view these dashboards, navigate to the DASHBOARDS page, click Open Library, then click the Saved Dashboards tab.

7.4. How to Suppress Empty Exabeam Data Lake Reports

You can optimized Data Lake’s output by suppressing reports that have empty content.

1. To prevent empty reports from generating, navigate to Reports.

2. Find the report you want to suppress empty outputs and click the Edit icon.
3. In **Report Details**, click **Schedule Report** to expand the menu.
4. Click **Suppress empty reports** and then click **SAVE REPORT**.
8. Access Restrictions for Saved Objects in Exabeam Data Lake

Object-Based Access Control (OBAC) manages the viewing and editing of tangible output products such as visualizations, dashboards, and reports. Objects are subject to the following:

- Access to newly created objects are by default restricted to the user who created and saved them. New objects are created with default **Private** access configurations to the originating user.
- Permissions are granted or denied to user roles.
- Screenshot shares (such as search results page) are based on user roles.
- Access restrictions for **saved objects** are configured per object.

For more information about OBAC, see *Data Lake Administrator Guide > User Management > Object-Based Access Control*.

To configure the access restriction of a saved object:

1. Select the object and then click **Share**.
2. Select the **Access** type **Based on role**.
3. Select the type of permission for the role which will be granted or denied access, and then click **SAVE**.
Access Restrictions for Saved Objects in Exabeam Data Lake

Share settings

Shareable link
https://33.103.213.164:443/ceunlogs/enterprise/traceformdata/account?view=1

Access
Based on role
Address all selected roles can access and edit
Auditor
Data Privacy Officer
Test Analyst

Note:
Can view and edit
Can view
None

CANCEL  SAVE
9. How to Forward Alerts Using Correlation Rules in Exabeam Data Lake

Exabeam Data Lake provides search-based alerting by incorporating rules that trigger further actions like forwarding the found event logs to another analysis process. Correlation rules can be created by analysts or administrators who want alerting capabilities within Data Lake for known bad behaviors or non-compliant behaviors in their environment.

9.1. How Correlation Rules Work

Rules that correlate incoming event logs based on relationships that have been previously defined by analysts is the basis for triggering log forwarding. Each rule contains a search query, a rule type and condition that generates alerts. Data Lake is periodically queried with the defined search query, and the data is passed through the rule conditions, which determines when a match is found. When a match occurs, it triggers one or more alerts.

9.2. Correlation Rules in Data Lake vs Advanced Detection Rules in Advanced Analytics

Correlation rules are static rules on incoming logs that alert on known bad or non-compliant behaviors when specified conditions are met. Advanced Analytics rules, on the other hand, rely on understanding user and machine behaviors in an environment to identify the degree of anomaly and criticality of the anomalous behavior and allocate the relevant risk scores to the session associated with that anomaly.

One reason to leverage correlation rules is for alerting on the outage of a service which has no security repercussions. Sometimes, correlation rules are needed to meet compliance requirements. For example, an alert indicating a service failure does not necessarily have a security risk associated with it. This makes more sense to be monitored by a correlation rule within Data Lake. Another example is to alert on any configuration changes on the Firewall to meet compliance needs.

9.3. Auto Disable Correlation Rules during High Latency

The Correlation-Rule feature is a highly customizable tool to spotlight specific events when they are triggered. However, some correlation rules may be invalid or highly inefficient and cause performance latency and perhaps ultimately fail when triggered. Data Lake will automatically disable a correlation rule that has not successfully completed processing after 5 minutes 3 times in a row (default settings, can be tailored for your organization).

9.4. Rule Types in Exabeam Data Lake

The rule type is responsible for processing the data returned from Data Lake as follows:

- The rule is initialized with the rule configuration.
- It is then passed data that is returned from querying Data Lake with the rule’s filters.
- And finally the system outputs matches based on this data.

<table>
<thead>
<tr>
<th>Rule Type</th>
<th>Description</th>
<th>Required Fields</th>
<th>Example Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blacklist</td>
<td>Checks a certain field against a blacklist, and match if it is in the blacklist.</td>
<td>You must enter the name of the field to use to compare to the blacklist. You must enter a list of blacklisted values and/or select a context table that contain the blacklisted values.</td>
<td>Trigger when a user in the group &quot;Domain Admins&quot; is added, removed, or changed.</td>
</tr>
<tr>
<td>Rule Type</td>
<td>Description</td>
<td>Required Fields</td>
<td>Example Rule</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Frequency</strong></td>
<td>Matches when there are at least a certain number of events in a given time frame. This may be</td>
<td>You must enter the number of events which will trigger an alert.</td>
<td>Trigger if a user is created and deleted within 24 hours.</td>
</tr>
<tr>
<td></td>
<td>counted on a per-query key basis.</td>
<td>You must enter the time those events must occur within.</td>
<td></td>
</tr>
<tr>
<td><strong>Any</strong></td>
<td>Matches everything. Every hit that the query returns will generate an alert.</td>
<td>There are no additional required fields for this rule.</td>
<td>Trigger when a user is successfully authenticated, and logon type is 8.</td>
</tr>
<tr>
<td><strong>Cardinality</strong></td>
<td>Matches when the total number of unique values for a certain field within a time frame is higher or lower than a threshold.</td>
<td>You must enter the time period in which the number of unique values will be counted.</td>
<td>Trigger when a destination host has greater than 20 failed authentication events from different user accounts within 5 minutes.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>You must enter which field to count the cardinality for.</td>
<td></td>
</tr>
<tr>
<td><strong>Change</strong></td>
<td>Monitors a certain field and matches if that field changes. The field must change with respect to the last event with the same query key.</td>
<td>You must enter the names of the field to monitor for changes.</td>
<td>Trigger when a scheduled task is canceled, completed, created, started, from different countries within 24 hours.</td>
</tr>
<tr>
<td><strong>Flatline</strong></td>
<td>Matches when the total number of events is under a given threshold for a time period.</td>
<td>You must enter the minimum number of events for an alert not to be triggered.</td>
<td>Trigger if no events are received from a Critical System within the last 60 minutes.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>You must enter the time period that must contain less than the minimum number of events.</td>
<td></td>
</tr>
<tr>
<td><strong>Metric</strong></td>
<td>Matches when the value of a metric within the calculation window is higher or lower than a threshold.</td>
<td>You must enter the name of the field over which the metric value will be calculated.</td>
<td>Trigger when a user sends email volume greater than 100MB in a 1 hour timeframe.</td>
</tr>
<tr>
<td><strong>Aggregation</strong></td>
<td></td>
<td>You must enter the type of metric aggregation to perform on the field.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>This must be one of 'minimum', 'maximum', 'average', 'sum', 'cardinality', or 'value count'.</td>
<td></td>
</tr>
<tr>
<td><strong>Whitelist</strong></td>
<td>Will compare a certain field to a whitelist and match if the list does not contain the term.</td>
<td>You must enter the name of the field to use to compare to the whitelist.</td>
<td>Trigger when:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>You must enter a list of whitelisted values and/or a list of paths to files which contain the whitelisted values.</td>
<td>Direction is Local to Remote</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Connection is to Port 53</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The IP Address is not in the &quot;DNS Servers&quot; list.</td>
</tr>
</tbody>
</table>

**9.5. Creating a Correlation Rule in Exabeam Data Lake**

Correlation rules are used as parsing mechanisms to match known indicators that analysts consider to be signs of threat or events of interest. To create a correlation rule:

1. Log in to your instance of the Data Lake UI.
2. Click the settings icon at the top-right corner of any page.
3. Select **Settings**.

4. On the Settings page, select **Correlation Rules**.

5. Click the blue-and-white **Create** button.

   a. Select a rule type, and then click **Continue**.

   **NOTE**
   Refer to the **Rule Types** section for a full list of available rule types.
6. Build the rule conditions by selecting the events that this rule will evaluate against.
   a. You can either select an existing search or create a new search from scratch.
   
   b. Once you've selected or created a search, the **Build** box opens. Configure your rule according to your business needs.

   **NOTE**
   The build rule options (circled below) vary depending on the selected rule type. You can also change the rule type by clicking the **Build** dropdown menu.

For build types **FLATLINE** or **FREQUENCY**, you can create a terms query (aka aggregation query) whereby you can apply a **TERM SIZE** to aggregate and improve search performance. The default term size is 50.
7. Define the outcomes when the new correlation rule triggers.
   - **Alert** – Send an alert to Data Lake. The alert is also then parsed and indexed as an “Exabeam Alert” along with the parsed log and details regarding the rule that triggered the alert.
   - **Alert Severity** – Set the severity of the alert as shown in Data Lake (i.e., None, Low, Medium, High, or Critical).
   - **Suppress Repeating Alerts** – If the rule is particularly busy or triggers a high number of alerts, configure the rule to suppress alerts for set number of seconds, minutes, hours, days, or weeks.
   - **Add Risk to Entities** – Send an alert to Advanced Analytics and associate it with users and/or assets. Additionally, assign a risk score to the alert as it appears in Advanced Analytics.

**NOTE:** Add Risk to Entities only appears if you have deployed an instance of Advanced Analytics and configured log forwarding via syslog to Advanced Analytics ([Settings > Log Forwarding > Advanced Analytics Feed](#)). For more information on forwarding logs from Data Lake to Advanced Analytics, please refer to the [Data Lake Admin Guide](#).
• **Send an Email** – Send an email to specific individuals in your organization, include a subject line and description, and choose between one and four attachment options (e.g., link to alert, attach as CSV, attach as PDF, and search string (DL)).
IMPORTANT

You must have email notifications configured to use this option. To enable email notifications triggered by correlation rules, use the following steps to configure:

1. Navigate to **Settings > Notifications > Setup Notifications**.

2. Edit the existing email notification entry or click 🔄 to create a new email notification.

3. In the **Setup Email Notifications** menu, enter server and template information and then in the **Notifications by Product** section, ensure the **Scheduled Reports** checkbox is selected.

4. Click **SAVE** to apply the configuration.

- **Send a Syslog** – Send an alert to a Syslog server.

8. Click **Next**.

9. Save the rule.
   a. Select a **Rule Category**. You can choose multiple existing categories, or add a newly created custom category.
   b. Enter a **Rule Name**.
   c. Enter a **Rule Description**.
   d. Click **Save**.
The rule is now saved and appears within your correlation rule table.

9.6. Correlation Rules Table in Exabeam Data Lake

Now that you have created your correlation rules, use the rules table to edit, delete, and otherwise manage your correlation rules.

Located in the top-right corner of the correlation rules page, the search bar lets you search across rules by Categories, Type, Product, and Name.

The correlation rules table is organized into columns, including:

- **Title** – Name of the rule. Click the Title column header to sort the table alphabetically by title.
- **Product** – Exabeam product associated with the rule. Click the lined triangle icon next to the Product column header to filter your rules by specific products.
9.6.1. MANAGING CORRELATION RULES

The correlation rules table lets you sort your rules according to Title and Last Modified. It also lets you filter your rules according to Product, Type, Severity Level, and Categories.

Click the checkbox for any rule to select and either delete or edit a rule. Select multiple rules to perform bulk deletions or edits on all selected rules.

Deleting removes the selected rule(s) completely, while editing a rule lets you set the severity and status for the selected rule(s).
Hover over any single rule in the rules table to enable additional management actions.

Click the vertical ellipsis to edit, clone, disable, or delete the selected correlation rule. Disabled rules are displayed in gray.

**9.7. Blacklist/Whitelist Correlation Rules using Context Tables in Exabeam Data Lake**

Context table data can be used to build Blacklist and Whitelist Correlation Rules. Use pre-existing context or uploaded custom context tables (see section *Add a Context Table.*)

1. Navigate **Settings > Correlation Rules > Correlation Rules**.
2. In the upper right of the Correlation Rules UI, click **Create**.
3. Select the **Blacklist** or **Whitelist**.

4. Click **Add** or **Create a Search** to complete the query selection process. Click **Create Rule**.

5. On the **Build Rule Conditions** page, select the **Field** that you want to alert on, then click **Select** to choose a context table from the list.

6. Click **Next** to move to **Rules Outcomes**.

7. Define the rule outcomes.

8. Click **Next** to move to **Save Rule**.

9. Fill in the requisite fields and click **Save**.
Appendix A. Technical Support Information

To contact Exabeam Customer Success, please open a case via Community.Exabeam.com.
Appendix B. Supported Browsers

Exabeam products support the latest versions of:

- Chrome
- Internet Explorer
- Firefox
- Safari