Optimize Search Performance

Search optimization tools speed the search process, delivering query results in less time and improving productivity for forensic analysis and log management.

Search speed generally depends on the amount of data and the type of query run against the data. Search optimization tools segment the data and queue it up for quick results.

An index, or proper subset of the data, is central to search optimization. When you run a search against an index, search results are returned more quickly and efficiently because the search is run against a smaller data set.

Sumo Logic supports index-based and field-based methods for search optimization.

Index-based methods

Partitions route unstructured data into an index.

Scheduled Views pre-aggregate data and then index it.

Field methods

Field Extraction parses out fields and then routes the fields to an index.

Field Browser allows you to zero in on just the fields of interest in a search by displaying or hiding selected fields without having to parse them.

Search optimization process

When data enters Sumo Logic, search optimization is done in the following order:

1. Field Extraction Rules are applied.
2. Partitions and Scheduled Views are applied. If both Partitions and Scheduled Views are defined, the Partitions are applied first.
3. The data is indexed.
4. The optimized and indexed data is available for use with other Sumo Logic features.
Is there such a thing as creating too many indexes?

Yes. Indexes can be overused, and in some situations, they can even slow the search process. When designing your organization's indexes, think about the minimal amount of data it makes sense to index, regardless of the tool. When running a search on non-indexed data, Sumo Logic might need to process all indexed data as well, which can take a long time.

How do Partitions and Scheduled Views differ?

Partitions begin building a non-aggregate index from the date a Partition started, only indexing data moving forward.

Scheduled Views backfill, meaning that all data that extends back to the start date of the Scheduled View can be queried.

Choosing the right indexed search optimization tool

Here's a quick look at how to choose the right indexed search optimization tool.

<table>
<thead>
<tr>
<th>I want to...</th>
<th>Partition</th>
<th>Scheduled View</th>
</tr>
</thead>
<tbody>
<tr>
<td>Run queries against a certain set of data</td>
<td>Choose if the quantity of data to be indexed is more than 2% of the total data.</td>
<td>Choose if the quantity of data to be indexed is less than 2% of the total data.</td>
</tr>
<tr>
<td>Use data to identify long-term trends</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Segregate data by sourceCategory</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Have aggregate data ready to query</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Use RBAC to deny or grant access to the data set</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Re-use the fields that I'm parsing for other searches against this same sourceCategory</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

How is data added to Partitions and Scheduled Views?

As data enters Sumo Logic, it's first routed to any Partitions for indexing. It's then checked against Scheduled Views, and any data that matches the Scheduled Views is indexed.

Data can be in both a Partition and a Scheduled View because the two tools are used differently (and are indexed separately). Although Partitions are indexed first, the process doesn’t slow the indexing of Scheduled Views.

• Topic hierarchy