Splunk® Enterprise
Getting Data In 7.2.4

About event segmentation

Generated: 2/26/2023 3:13 am

About event segmentation

Event segmentation breaks events up into searchable **segments** at **index** time, and again at **search** time. Segments can be classified as major or minor. Minor segments are breaks within major segments. For example, the IP address `192.0.2.223` is a major segment. But this major segment can be broken down into minor segments, such as `192` or `0`, as well as groups of minor segments like `192.0.2`.

You can define how detailed the event segmentation is. This is important because index-time segmentation affects indexing and search speed, storage size, and the ability to use typeahead functionality, where Splunk Web provides items that match text you type into the Search bar. Search-time segmentation, on the other hand, affects search speed and the ability to create searches by selecting items from the results displayed in Splunk Web.

For more information about the distinction between index time and search time, see Index time versus search time in the *Managing Indexers and Clusters of Indexers* manual.

You can assign segmentation to specific categories of events in props.conf, as described in Set the segmentation for event data.

If you use Splunk Cloud Platform, configure index-time segmentation on heavy forwarder machines. You must file a Support ticket to configure search-time segmentation.

If you use Splunk Enterprise, configure index-time segmentation on the indexer or heavy forwarder machines, and configure search-time segmentation on the search head.

Types of event segmentation

There are three main types, or levels, of segmentation, which you can configure either at index or search time. You can also disable segmentation. The segmenters.conf configuration file, located in $SPLUNK_HOME/etc/system/default, defines all available segmentation types. By default, index-time segmentation is set to the `indexing` type, which is a combination of inner and outer segmentation. Search-time segmentation is set to full segmentation.

Inner segmentation

Inner segmentation breaks events down into the smallest minor segments possible. For example, when an IP address such as `192.0.2.223` goes through inner segmentation, it is broken down into `192`, `0`, `2`, and `223`. Setting inner segmentation at index time leads to faster indexing and searching and reduced disk usage. However, it restricts the typeahead functionality, so that a user can only type ahead at the minor segment level.

Outer segmentation

Outer segmentation is the opposite of inner segmentation. Under outer segmentation, the Splunk platform only indexes major segments. For example, the IP address `192.0.2.223` gets indexed as `192.0.2.223`, which means that you cannot search on individual pieces of the phrase. You can still use wildcards, however, to search for pieces of a phrase. For example, you can search for `192.0*` and you will get any events that have IP addresses that start with `192.0`. Also, outer segmentation disables the ability to click on different segments of search results, such as the `192.0` segment of the same IP address. Outer segmentation tends to be marginally more efficient than full segmentation, while inner segmentation tends to be much more efficient.

Full segmentation

Full segmentation is a combination of inner and outer segmentation. Under full segmentation, the IP address is indexed both as a major segment and as a variety of minor segments, including minor segment combinations like `192.0` and `192.0.2`. This is the least efficient indexing option, but it provides the most versatility in terms of searching.

No segmentation

The most space-efficient segmentation setting is to disable segmentation completely. This has significant implications for search, however. By disabling segmentation, you restrict searches to indexed fields, such as time, source, host, and source type. Searches for keywords will return no results. You must pipe your searches through the search command to further restrict results. See search in the *Search Reference*. Use this setting only if you do not need any advanced search capability.

Configure segmentation types

segmenters.conf defines segmentation types. You can define custom segmentation types, if necessary.

For information on the types of segmentation available by default, see the segmenters.conf file in $SPLUNK_HOME/etc/system/default.

> Do not modify the default file. If you want to make changes to the existing segmentation stanzas or create new ones altogether, you can specify the settings you want to change in a file in the $SPLUNK_HOME/etc/system/local/ directory or to a custom app directory in $SPLUNK_HOME/etc/apps/.

Set segmentation types for specific hosts, sources, or source types

You can configure index-time and search-time segmentation to apply to specific hosts, sources, or source types. If you run searches that involve a particular source type on a regular basis, you can use this capability to improve the performance of those searches. Similarly, if you typically index a large number of `syslog` events, you can use this feature to help decrease the overall disk space that those events take up.

For details about how to apply segmentation types to specific event categories, see Set the segmentation for event data.

See also

Related information
    Event segmentation and searching in the *Search Manual*.