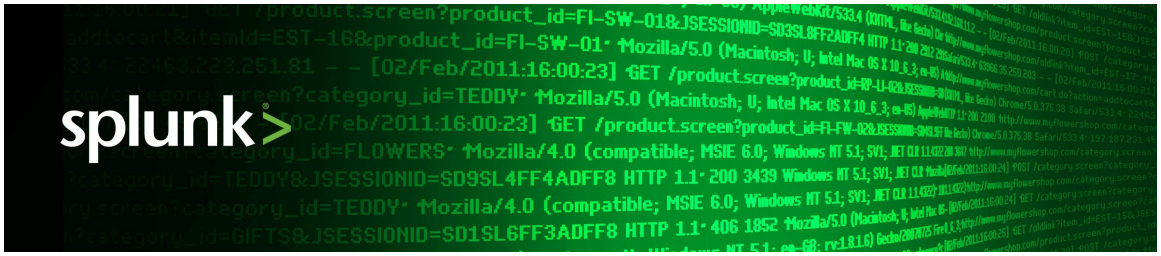

Splunk® Enterprise Capacity Planning Manual 6.5.0

Dimensions of a Splunk Enterprise deployment

Generated: 11/03/2016 10:54 pm

Copyright (c) 2016 Splunk Inc. All Rights Reserved

Dimensions of a Splunk Enterprise deployment

A Splunk Enterprise deployment has many dimensions. These scenarios determine whether a single reference machine can handle indexing and search load.

In some cases, a single reference machine can collect, store, and search data efficiently. In other cases, consider adding machines to your Splunk Enterprise deployment to increase performance. Below is a list of items that can have a significant impact on Splunk Enterprise performance.

- **Amount of incoming data.** The more data you send to Splunk Enterprise, the more time it needs to process the data into events that you can search, report, and generate alerts on.

- **Amount of indexed data.** As the amount of data stored in a Splunk Enterprise index increases, so does the I/O bandwidth needed to store data and provide results for searches.

- **Number of concurrent users.** If more than one person at a time uses an instance of Splunk Enterprise, that instance requires more resources for those users to perform searches and create reports and dashboards.

- **Number of saved searches.** If you plan to invoke a lot of saved searches, Splunk Enterprise needs capacity to perform those searches promptly and efficiently. A higher search count over a given period of time requires more resources.

- **Types of search you use.** Almost as important as the number of saved searches is the types of search that you run against a Splunk Enterprise instance. There are several types of search, each of which affects how the indexer responds to search requests.

- **Whether or not you run Splunk apps.** Splunk **apps** and solutions can have unique performance, deployment, and configuration considerations. If you plan to run apps, consider the resource requirements of the apps the you are using. See the documentation for the app for more information.

How do these dimensions impact overall performance?

While these factors have an impact on the basic sizing requirements of your Splunk Enterprise deployment, addressing each of them individually does not guarantee peak performance gain for the deployment. You must discover

through trial how these factors correlate with one another in your specific application.

For example, if your Splunk Enterprise deployment calls for a low amount of indexing but has a high number of concurrent users, it has significantly different resource needs than a setup with a low number of concurrent users and a high amount of daily indexing volume. Additionally, as both user count and amount of indexed data rise, you must distribute the environment across multiple servers to maintain a similar performance level. Search types complicate matters, because some searches strain available CPU resources, while others depend on the speed of the disk subsystem.

When should I scale my Splunk Enterprise deployment?

You must understand how the deployment dimensions described in this topic apply to your specific use case. Answer the following questions, and then refer to the performance checklist in this manual to determine when you should add more hardware resources:

- How much data do you expect to index daily?
- How much data do you need to retain and for how long?
- How many users do you expect to search through the data at any one time?
- Do you plan to use certain specific searches more than once?
- Do you want or need to use a Splunk app to present or manipulate your data?

The key to a well-performing installation is to develop a plan early in the deployment cycle to account for both your initial outlay of hardware resources and the addition of resources when the deployment scales up.