How To- Splunk Log Management and Configuration

What Is Splunk?

Splunk is a software platform to search, analyze and visualize the machine-generated data gathered from the websites, applications, sensors, devices etc.

It can monitor and read different type of log files and stores data as events in indexers.

This tool allows you to visualize data in various forms of dashboards.

Features of Splunk

Important features of Splunk are:

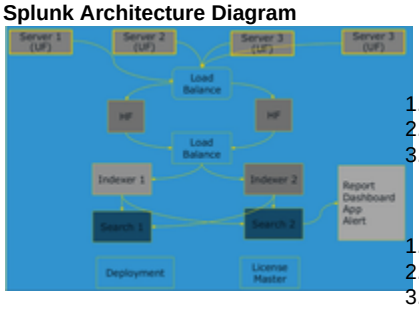
- Accelerate Development & Testing
- Allows you to build Real-time Data Applications
- Generate ROI faster
- Agile statistics and reporting with Real-time architecture
- Offers search, analysis and visualization capabilities to empower users of all types

Splunk Products

Splunk is available in three different versions.

- **Splunk Enterprise** :- Splunk Enterprise edition is used by large IT business. It helps you to gather and analyze the data from applications, websites, applications, etc.
- **Splunk Light** :- Splunk Light is a free version. It allows search, report and alter your log data. It has limited functionalities and feature compared to other versions.
- **Splunk Cloud** : Splunk Cloud is a hosted platform. It has the same features as the enterprise version. It can be availed from Splunk or using AWS cloud platform.

Splunk Architecture Diagram



The main components of Splunk Architecture are:

1. Indexers
2. Search Heads
3. Forwarders(Universal/Heavy)

The other components are:

1. Deployment Server
2. License Master
3. Master Cluster

Indexer:

Indexer process the incoming data in real-time. It also stores & Indexes the data on disk.

Indexer (LB):

Indexer helps you to store and index the data. It improves Splunk search performance. By default, Splunk automatically performs the indexing. For example, host, source, and date & time.

Search head (SH):

Search head is used to gain intelligence and perform reporting.

Search Head:

End users interact with Splunk through Search Head. It allows users to do search, analysis & Visualization.

Universal Forward (UF):

Universal forward or UF is a lightweight component which pushes the data to the heavy Splunk forwarder. You can install Universal Forward at client side or application server. The job of this component is only to forward the log data.

Heavy forward (HF):

Heavy forward is a heavy component. This Splunk component allows you to filter the data. Example: collecting only error logs.

Load Balancer (LB):

Load balancer is default Splunk load balancer. However, it also allows you to use your personalized load balancer.

Deployment Server(DS):

Deployment server helps to deploy the configuration. For example, update the UF configuration file. We can use a deployment server to share between the component we can use the deployment server.

License manager (LM):

The license is based on volume & usage — for example, 50 GB per day. Splunk regular checks the licensing details.

Forwarder:

Forwarder collect the data from remote machines then forwards data to the Index in real-time

Deploy Splunk Enterprise Docker containers

1. `docker pull splunk/splunk:latest`
2. docker run -d -p 8000:8000 -e 'SPLUNK_START_ARGS=--accept-license' -e 'SPLUNK_PASSWORD=password123' splunk/splunk:latest
3. `docker ps -a -f id=<container_id>`
4. When the status of the container becomes healthy, it means the container is already up and running. Open an Internet browser and access Splunk Enterprise inside the container through Splunk Web:

```
http://3.89.201.137:8000
```

username `admin` `Password password123`

Integrate Splunk with Jenkins

We need two components:

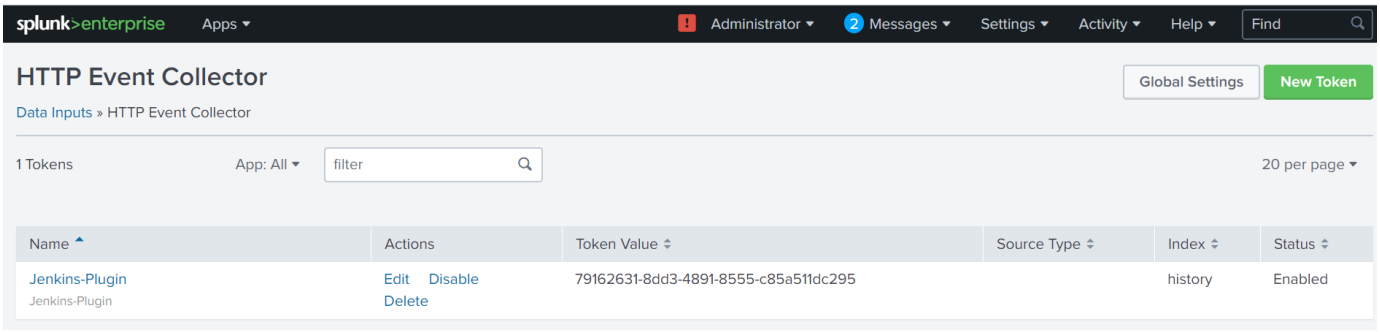
Splunk Plugin for Jenkins :- The Plugin automatically monitors, collects, sends Jenkins data to Splunk. It uses HTTP Event Collector (HEC) to send data to Splunk software eliminating the need for installing Splunk forwarders.

Splunk App for Jenkins. The App provides a set of pre-built dashboards for quick analysis of Jenkins data.

Splunk Plugin for Jenkins

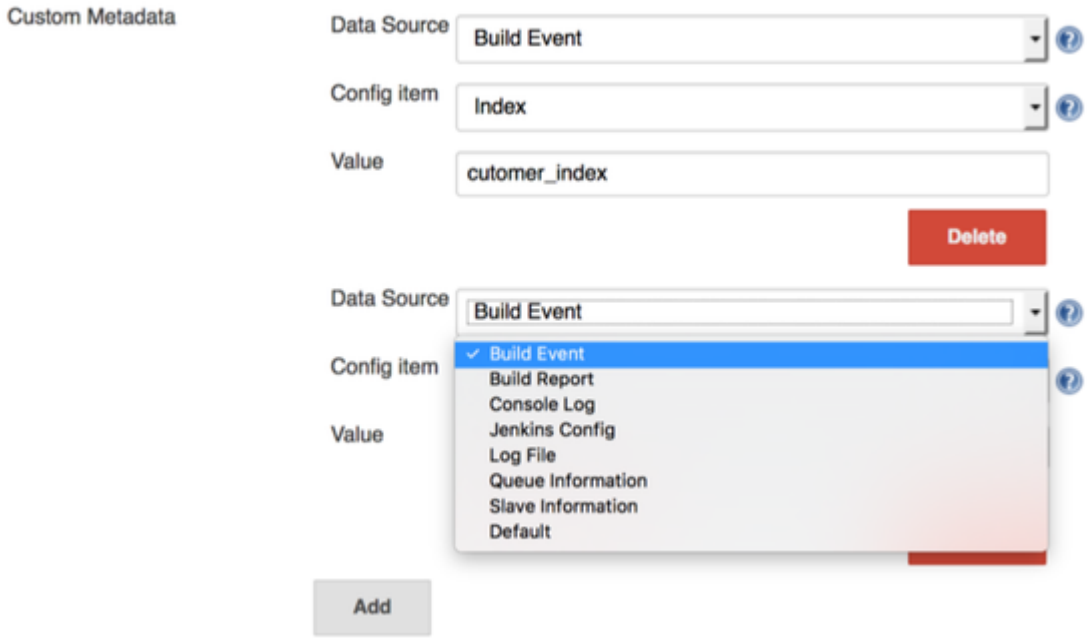
- Go to http://novartis.devops.altimetrik.io:8084/
- To install Splunk Plugins Go to >Manage Jenkins > Manage Plugins

- Go to Manage Jenkins > Configure System > Search for Splunk > Splunk for Jenkins Configuration
- Enter Hostname, Port, and Token > For Splunk enterprise user, the host name is the indexer host name, and port is 8088 by default
- SSL is enabled by default in Splunk, it will protect the data transferred on network.
- Click "Test Connection" to verify the setup
- Check "Enable" and Save



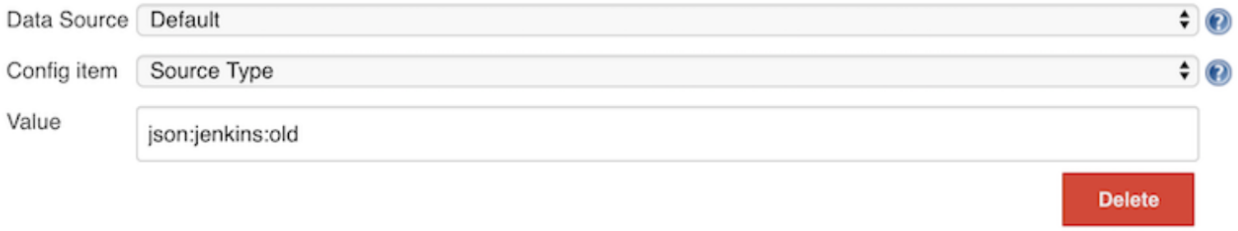
Configure Metadata

you can customize the index, sourcetype in the "Custom Metadata" section.

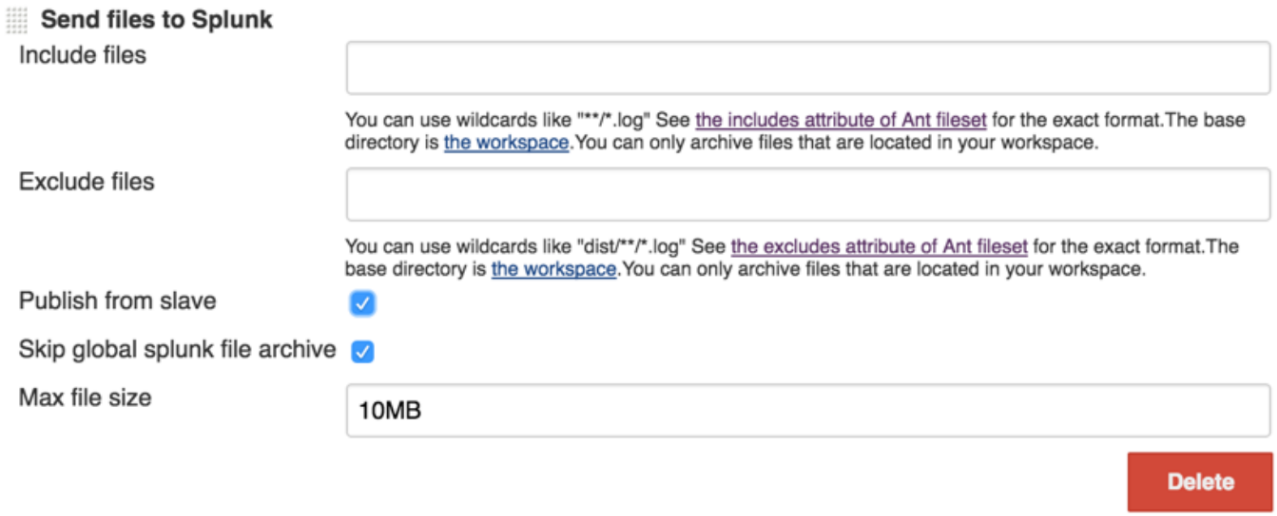


Metadata configuration for Splunk App for Jenkins

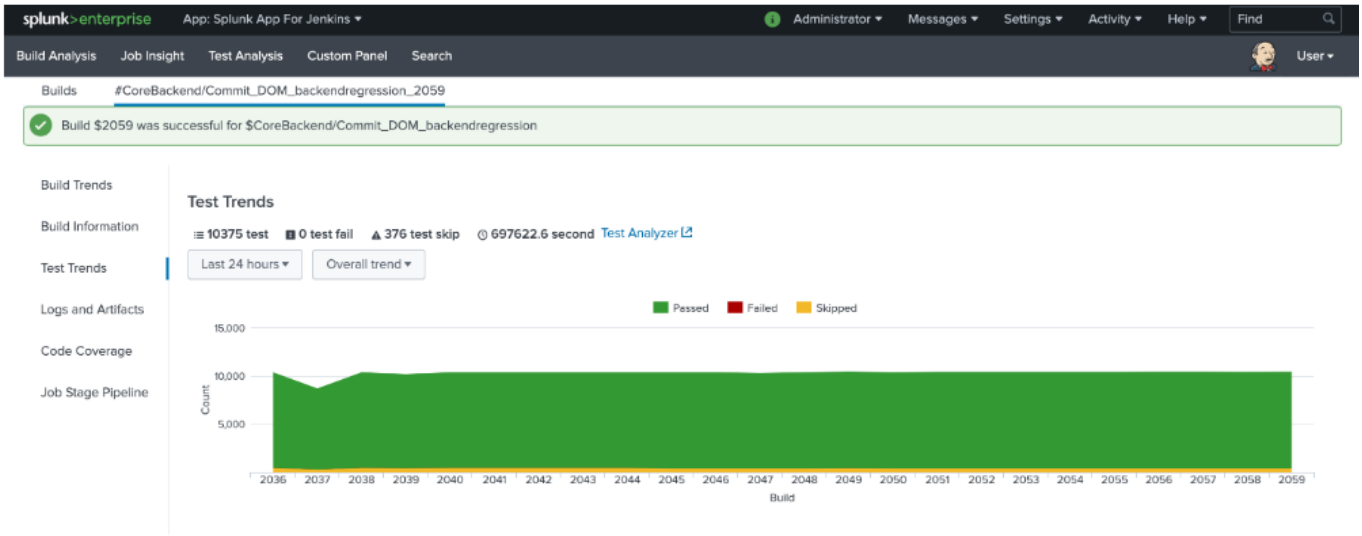
please adjust the default sourcetype to `json:jenkins:old` (please remove it if Splunk get upgraded to latest version, otherwise data will be extracted twice)

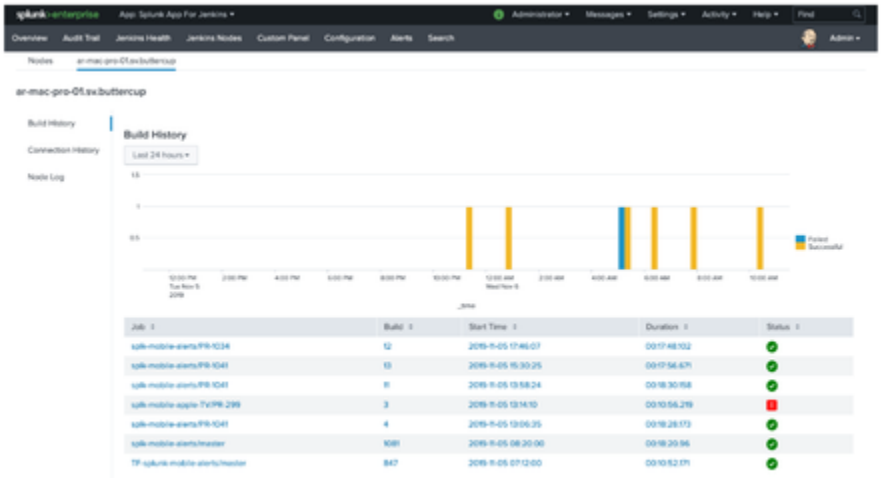


Customize log files at job level



Splunk Dashboards for Jenkins





GitLab Add-On

Providing a way to import your event data from GitLab into Splunk, this add-on gives you the ability to visualise and alert on your GitLab activity using Splunk. Need to see how many times a file has been modified, how many branches exist for a project, the amount of changes a particular developer has merged? This add-on gives you the ability to do just that.

Installation

Installation is as standard for a Splunk Add-On

Set-up

This input requires your GitLab token (Obtained via GitLab > Profile Settings > Access Tokens) for authentiction. When setting up inputs, you have the option to enter a project ID or not.

PLEASE NOTE

If you do not specify a specific project ID, you will only get event data associated with the account the token is associated with. This will also result in no CI data

Using Data

The add-on will collect any new information gathered since last run. This is done by storing the datetime of last run in the KV Store and then passing it into GitLab as a HTTP Param. The add-on focusses on the events, initially retrieving these before getting records associated with the event (Merge Requests, Project Info, Commits, Milestones...). During retrieval, it will then store everything in Splunk in JSON format.