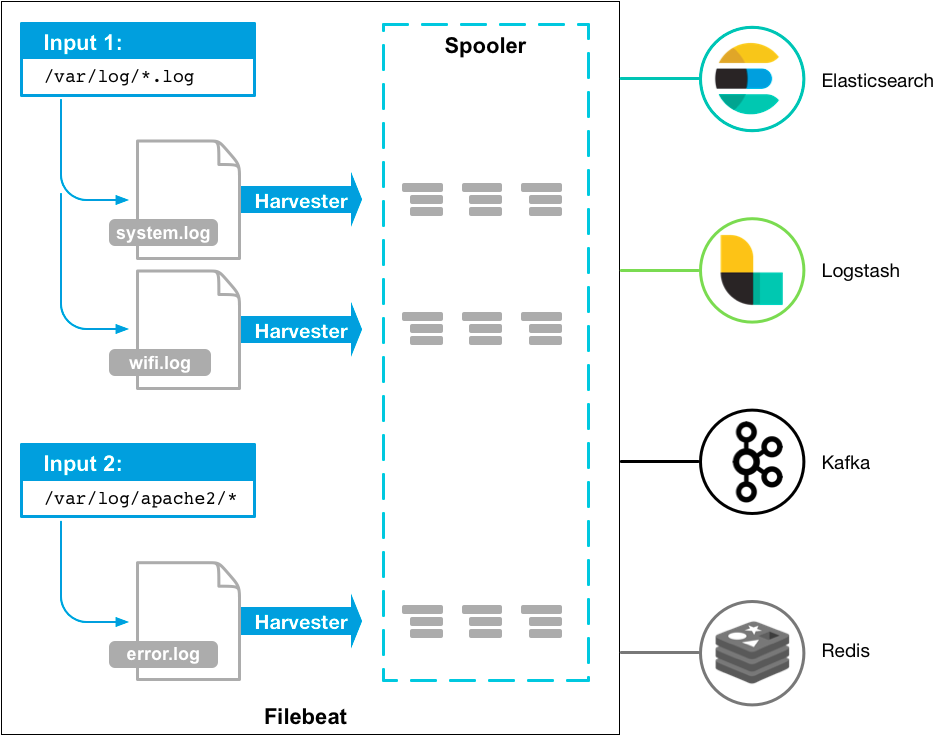
**What is Filebeat?**

Filebeat is a lightweight shipper for forwarding and centralising log data. Installed as an agent on your servers, filebeat monitors the log files or locations that you specify, collects log events, and forwards them either to Elasticsearch or logstash for indexing.

How filebeat works:

When we start filebeat, it starts inputs[one or more as per requirement] that look in the location that we specified for log data. For each log that filebeat locates, it starts a harvester[responsible for reading the content of a single file].

Each harvester reads a single log for new content and send the new log data to libbeat, which aggregates the events and sends the aggregated data to the output that we have configured to filebeat.



**Configure filebeat:**

To configure our filebeat, we have to edit the configuration file. The default configuration file is called filebeat.yml.

The location of the file varies by platform.

**Archive:**

Home: Home of the filebeat installation. -> path.home

Bin: The location for the binary files.

Config: The location for configuration files. -> path.config

Data: The location for persistent data files. -> path.data

Logs: The location for the logs created by filebeat. -> path.logs

**Default paths:**

Home: Home of the filebeat installation. -> /usr/share/filebeat

Bin: The location for the binary files. -> usr/share/filebeat/bin

Config: The location for configuration files. -> /etc/filebeat

Data: The location for persistent data files. -> /var/lib/filebeat

Logs: The location for the logs created by filebeat. -> /var/log/filebeat

**Docker:**

Home: Home of the filebeat installation. -> /usr/share/filebeat

Bin: The location for the binary files. -> usr/share/filebeat/bin

Config: The location for configuration files. -> /etc/filebeat

Data: The location for persistent data files. -> /var/lib/filebeat

Logs: The location for the logs created by filebeat. -> /var/log/filebeat

**1). Configure Inputs:**

To configure filebeat manually we have to specify the list of inputs in the filebeat.inputs section, in the filebeat.yml.

We can specify multiple inputs with a YAML array, so each input begins with a dash(-).

We can specify the multiple inputs and you can specify the same input type more than once.

filebeat.inputs:

- type: filestream

id: my-filestream-id

paths:

- /var/log/system.log

- /var/log/wifi.log

- type: filestream

id: apache-filestream-id

paths:

- "/var/log/apache2/\*"

fields:

apache: true

fields\_under\_root: true

Each filestream input has a unique ID to track the state of files.

Here, we can see that filebeat will harvest system.log & wifi.log.

There are a lot of input types in filebeat.

Fo example: AWS CloudWatch:

<https://www.elastic.co/guide/en/beats/filebeat/current/filebeat-input-aws-cloudwatch.html>

You can see the different inputs from documentation.

<https://www.elastic.co/guide/en/beats/filebeat/current/configuration-filebeat-options.html#filebeat-input-types>

**2). Config modules**

Filebeat modules provide a quick way to get started processing common log formats. They contain default configurations, Elasticsearch ingest pipeline definitions, and kibana dashboards to help you implement and deploy a log monitoring solution.

We can configure modules inthe modules.d directory, or in filebeat config file.

-> Configure modules in module.d directory.

We can see that in etc/filebeat directory the module.d directory is thare and consists various modules, we can enab; and disable specific module configuration using following command.

DEB: filebeat modules enable module\_name.

MacOS: ./filebeat modules enable module\_name

Linux: ./filebeat modules enable module\_name

Windows: ./filebeat.exe modules enable module\_name

-> Configure module in filebeat.yml

filebeat.modules:

- module: nginx

access:

error:

- module: mysql

slowlog:

- module: system

Auth:

**3). General settings**

We specify settings in the filebeat.yml config file to control the general behaviour of filebeat. This includes:

-Global options.

-General options.

**-> Global filebeat configuration options**

These options are in the filebeat namespace.

**Registry.path:** The root path of the registry. If a relative path is used, it is considered as a relative to the data path.

NOTE: The registry is only updated when new events are flushed and not on a predefined period. That means, in case TTL expires, they are only removed when new events are processed.

**registry.file\_permissions:** The permissions mask to apply on registry data file. The default value is 0600. The permissions options must be a valid UNIX-style file permissions mask expressed in octal notation. In Go, numbers in octal start with 0. This option is not available in windows.

0640: give read and write access to the file owner, and read access to members of the group associated with the file.

0600: give read and write access to all others.

Filebeat.registry.file\_permissions: 0600

**registry.flush:** This is the timeout value that controls when registry entries are written to disk (flushed). When an unwritten update exceeds this value, it triggers a write to disk. When registry.flush is set to 0s, the registry is written to disk after each batch of events has been published successfully. The default value is 1.

**-> General configuration options**

This options are supported in all beats in elasticsearch.

**Name:** It is used for a beat name. If this field is empty, then the name used is hostname.

**Tags:** Tags basically used to group servers by different properties.

**Fields:**  This is the optional part we can specify to add more information about the outcome.

**Processors:** This applies to the data generated by beat. We can define our events and return our events.

Example:[drop event]

processors:

- drop\_event:

when:

regexp:

message: "^DBG:"

**Max\_procs:**  We can set the max number of CPUS that can be executed simultaneously.

**Timestamp.precision:**  We can configure precision of all timestamps. We can specify it by microseconds,milliseconds,nanoseconds. By default it is set to microseconds.

**4). Configure project paths.**

In this section, we configure our project paths.

**5). Output:**  We configure filebeat to write a specific output by setting options in the outputs section of the filebeat.yml. Only a single output may be defined.

Example:

output.elasticsearch:

hosts: ["https://myEShost:9200"]

Accessing metadata fields:

@metadata: it is responsible to send metadata to logstash.

Beat: We can specify which index we have to choose for that.

Version: This specifies the version of our beat we specify in beat.

Example:

{

...

"@metadata": {

"beat": "filebeat",

"version": "8.3.3"

}

}

We specify the indices too. For specify the format of data which we want.

Example:

output.elasticsearch:

hosts: ['http://node01:9200']

indices:

- index: "filebeat-%{[agent.version]}-%{+yyyy.MM.dd}"

**6). Configure index lifecycle management:**

This is used to manage our filebeat to back indices to our data streams.

**Configuration options:**

**setup.ilm.enabled:** This is used to enable or disable ilm on any new indices create dby filebeat .

**setup.ilm.policy\_name:** This is used for lifecycle policy.

**7). Configure Elasticsearch index template loading**

This section is used for setting mapping in elasticsearch. If it is successfully connected then filebeat loads the index template.

**Settings:**

We can define the template name,fields,overwrite,settings, etc.

Example:

setup.template.name: "filebeat"

setup.template.fields: "fields.yml"

setup.template.overwrite: false

setup.template.settings:

index.number\_of\_shards: 1

index.number\_of\_replicas: 1

**8). Configure the kibana endpoint**

We can configure the endpoint in the section of filebeat.yml file.

**setup.kibana.host:**  This specifies the host where dashboards are loaded.

**setup.kibana.protocol:** It specifies the protocol where it is reachable or not.

Options: http or https

**setup.kibana.username:** This is used for authentication of users.

**setup.kibana.password:** This is used for authentication of passwords.

**setup.kibana.path:** This is the path prefix to the HTTP API.

**setup.kibana.ssl.enabled:**  This is using the SSL setting when we connect to kibana via HTTPS. If our requirement is to connect over https then we have to set this setting to true.

Example:

setup.kibana.host: "https://192.0.2.255:5601"

setup.kibana.ssl.enabled: true

setup.kibana.ssl.certificate\_authorities: ["/etc/client/ca.pem"]

setup.kibana.ssl.certificate: "/etc/client/cert.pem"

setup.kibana.ssl.key: "/etc/client/cert.key

**9). Configure kibana dashboard loading.**

This section is used for dashboards, visualisations and searches for data visualisation in kibana.

To enable the dashboard loading, configure the settings below in filebeat.yml file.

setup.dashboards.enabled: true

**setup.dashboards.enabled:**  This option is used to load the sample dashboard from the local kibana directory to the home path of the installation directory of filebeat.

**setup.dashboard.directory:** This is used for loading the dashboards where they are located. The default path is kibana.

**setup.dashboard.file:** This option is used when we are working with archives. We can specify the archive file instead of our local file.

**setup.dashboard.beat:** This option is used when we are working with multiple beats. So this setting lets us select which beat we want on the dashboard.

**10). Load balance the output host**

This section is used when we fetch the data in bulk and send the events to multiple hosts. So, it is hard to load such a large amount of data in a single worker.

So load balancer supports multiple workers per host[where we send the data].

**The total number of workers participating in load balancing:**

Number of hosts \* workers

**Example:**

filebeat.inputs:

- type: log

paths:

- /var/log/\*.log

output.logstash:

hosts: ["localhost:5044", "localhost:5045"]

loadbalance: true

worker: 2

In this case we can see that we have 2 hosts and 2 workers. So, there are total 4 worker participating in load balancing.

**11). Configure logging**

This section is used for logging the output.

logging.level: info

logging.to\_files: true

logging.files:

path: /var/log/filebeat

name: filebeat

keepfiles: 7

permissions: 0640

**12). Regular Expression Support**

This option supports regular expression support and is based on RE2.