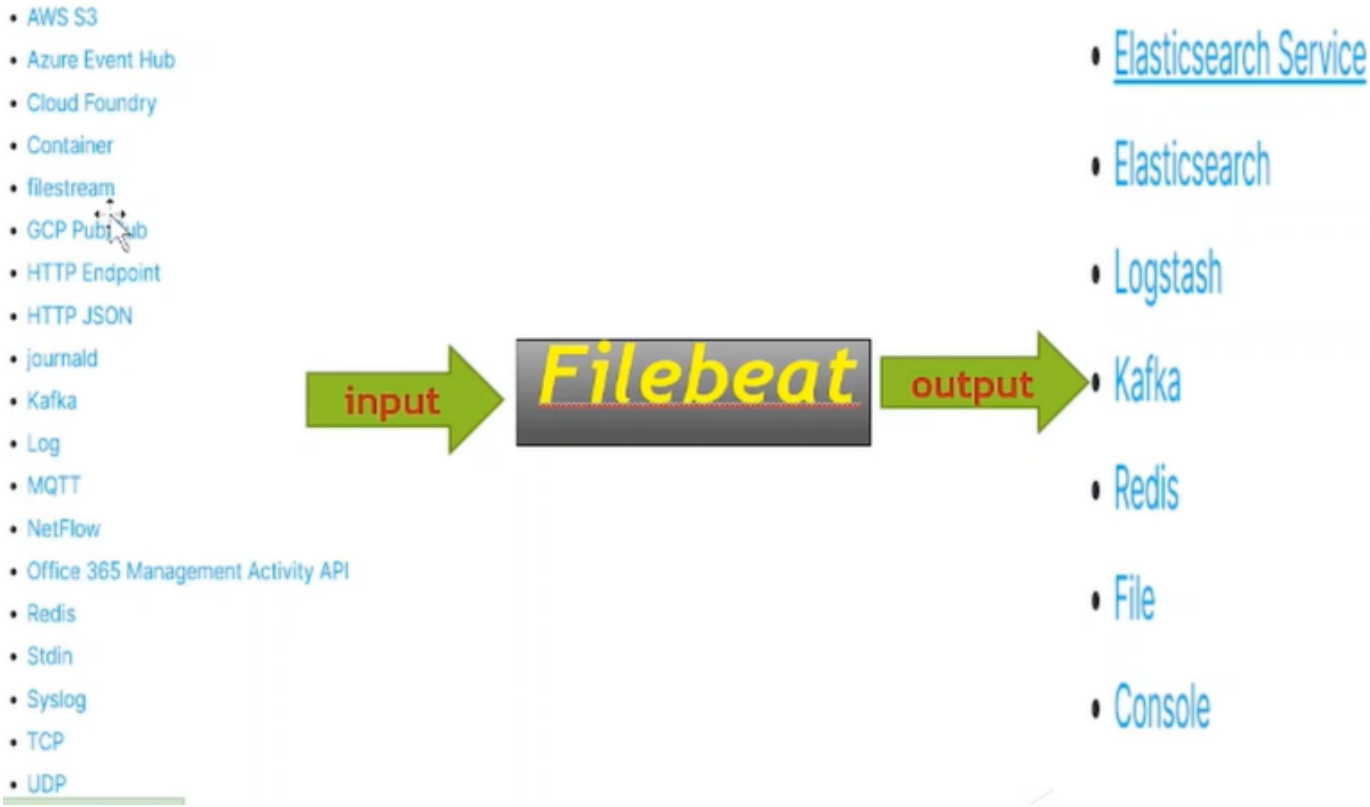
3. File beat

What is Filebeat?
Filebeat is part of the Elastic Stack and it works seamlessly with logstash, elasticsearch and kibana.
It is used as a data shipper/collector in the ELK stack application.

All the possible inputs and output file beat gives



Mainly there are two components for filebeat:;
Input and **Harvester**

1. **Input** : An input is responsible for managing the harvesters and finding all sources to read from. If the input type is log, the input finds all files on the drive that match the defined glob paths and starts a harvester for each file.
2. **Havester** :The harvester reads each file, line by line, and sends the content to the output. One harvester is started for each file. The harvested, Filebeat continues to read the file. This file, which means that the file descriptor remains open while the harvester is running.
   If a file is removed or renamed while it's being harvested, Filebeat continues to read the file. This has the side effect that the space on your dist is reserved ubtil the harvester closes. By default, Filebeat keeps the file open intil close_inactive is reached.

Configuration file

The file will have an extension as .yml. The default filebeat file name is **filebeat.yml**.

We can configure more than one inputs
But only single output can be configured.

Mainly the filebeat configuration will have two components i.e. **input** and the **output.**

The input component is where we mention the source of the filebeat(datasource) like kafka or logfile.

The output is the one where filebeat will send the event or the data to.

for more detail of how input configs whould be use you can visit::
https://www.elastic.co/guide/en/beats/filebeat/current/filebeat-input-log.html

Important properties of input configs for input type

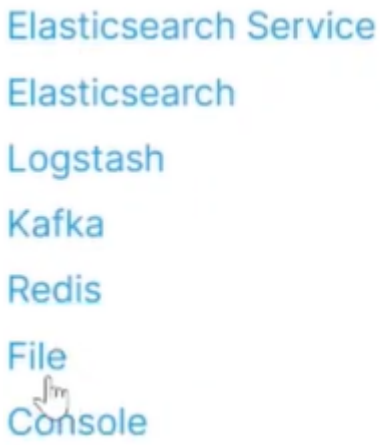
type - to mention the source of data like log, kafka etc
paths - to give the path of log files
Fields - to add the extra field in the filebeat output
exclude_lines - to remove the particular line from output if it matches the pattern
exclude_files - exclude the file from where filebeat need not to take data
include_lines - to add particular line in filebeat output if it matches the pattern.
enabled - to enbale particular input tupe. (Enabled is set to true.)
fields_under_root - to add the fields at the top in the output event so it is used with fields attribute.
encoding - to mention the encoding style.
scan_frequency - period after which filebeat should get the next line from source file.(The default setting is 10s.)
close_inactive - When this option is enabled, Filebeat closes the file handle if a file has not been harvested for the specified duration. Filebeat uses an internal timestamp that reflects when the file was last harvested, from that time the counter for close_inactive starts. The default is 5m.

Multiline - Options that control how filebeat deals with log messages that span multiple lines. you need to configure multiline settings in the filebeat.yml file to specify which lines are part of a single event. A feature which will combine the logfile lines and show in the single event,

multiple.type : pattern
multiple.pattern : regex pattern
multiple-negate : The default is false
multiple.after : the behavior of these setting depends on what you sepcify for negate.
mutiline.flush_pattern: specifies a regex at which the current multiline will be flushed, normally when log pattern is known.

```
2015-08-24 11:49:14,389] Start new event
2015-08-24 11:49:14,395] Content of processing something
2015-08-24 11:49:14,399] End event
2015-08-24 11:49:14,389] Start new event
2015-08-24 11:49:14,395] Content of processing something
2015-08-24 11:49:14,399] End event
```

Output config



available output of filebeat