

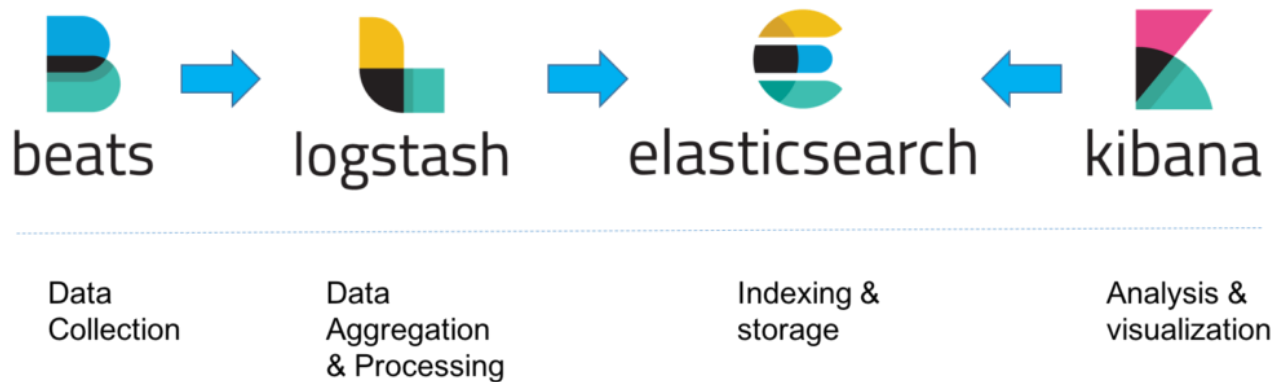
FluentD and Logstash are both open-source data collectors used for Kubernetes logging.

Logstash is centralized while FluentD is decentralized.

FluentD offers better performance than Logstash.

Filebeat is **a lightweight shipper for forwarding and centralizing 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.**



Data aggregation is the process where data is collected and presented in a summarized format for statistical analysis and to effectively achieve business objectives.

Data aggregation is vital to data warehousing as it helps to make decisions based on vast amounts of raw data.

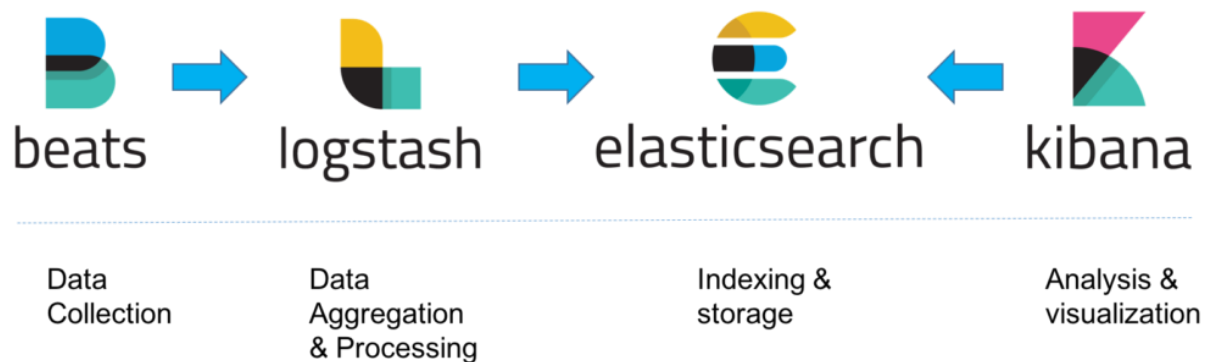
Jenkins Troubleshooting at Work.

We recently had one of our production Jenkins masters nearly run out of disk space due to an over-allocation of log data by just a couple of jobs running on our farm. In fact, 2 such builds consumed nearly half a terabyte for their build logs

alone! Further, closer examination of those log files revealed that the data therein was being excessively bloated by what appears to be some sort of markup or metadata attached to each line in the build log that looks something like this.

- Upon further investigation, it was discovered that Docker service was down.
- Docker was down because the disc was running out of space.
- The disc was full.
- It was full up to about 56 gigabytes out of 60 gigabytes

Solutions



- Deleted the log messages in the disc
- Killed file-beat (collect logs) and rsyslogd (log aggregation and processing) and this enabled the file to be released and the disc was free.
- Restarted Jenkins
- Docker was up again, and Jenkins was up.
- The company is using a container jenkins CI Server.