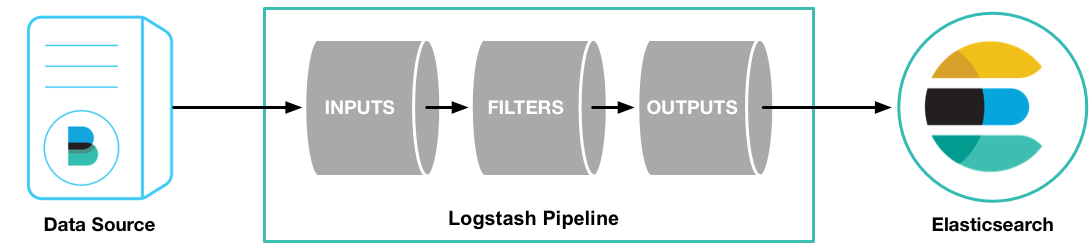
Logstash

提示：本节内容过于枯燥，请在儿童的陪伴下观看！

1. Logstash：

开源的流数据处理、转换（解析）和发送引擎，可以采集来自不同数据源的数据，并对数据进行处理后输出到多种输出源。Logstash是ELK Stash的重要组成部分

1. 工作原理：
   1. Logstash的数据处理过程主要包括：**Inputs**, **Filters**, **Outputs** 三部分，如图：



* + 1. **Inputs**：用于从数据源获取数据，常见的插件如beats、file、kafka、rabbitmq、log4j、redis等。

参考：https://www.elastic.co/guide/en/logstash/current/input-plugins.html

* + 1. **Filters**：筛选器是Logstash管道中的数据处理器，input时会触发事件，触发filter对数据进行transport，即转换解析各种格式的数据，常见的过滤器插件如下：
       1. **grok**：解析和构造任意文本。是Logstash过滤器的基础，广泛用于从非结构化数据中导出结构，当前，Grok是Logstash中将非结构化日志数据解析为结构化和可查询内容的最佳方法。
       2. **mutate**：对事件字段执行常规转换。支持对事件中字段进行重命名，删除，替换和修改。
       3. **date**：把字符串类型的时间字段转换成时间戳类型
       4. **drop**：完全删除事件，例如调试事件。
       5. **clone**：复制事件，可能会添加或删除字段。
       6. **geoip**：添加有关IP地址地理位置的信息。
    2. **Outputs**：用于数据输出，常见的插件如:
       1. elasticsearch：最高效、方便、易于查询的存储器，最有选择，官方推荐！
       2. **file**：将输出数据以文件的形式写入磁盘。[**stupid**](javascript:;)
       3. **graphite**：将事件数据发送到**graphite**，**graphite**是一种流行的开源工具，用于存储和绘制指标。<http://graphite.readthedocs.io/en/latest/>
       4. **statsd**：将事件数据发送到statsd，该服务“通过UDP侦听统计信息（如计数器和计时器），并将聚合发送到一个或多个可插拔后端服务”。

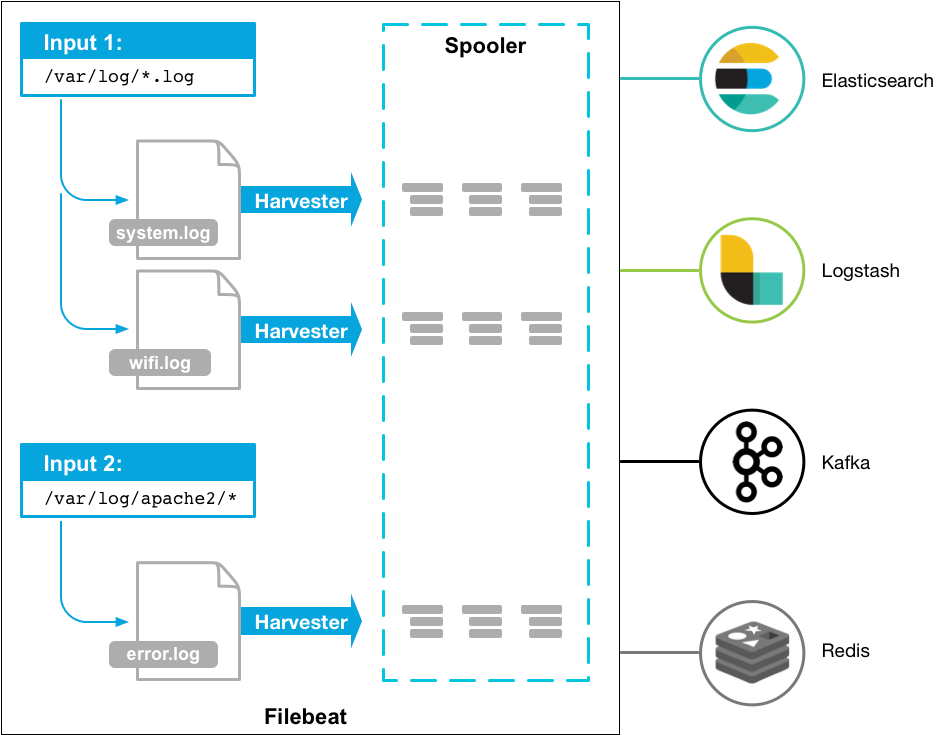
1. Beats：基于golang语言开发，开源的、轻量级的日志收集器的统称。
   1. 官方文档：

<https://www.elastic.co/guide/en/beats/libbeat/current/beats-reference.html>

* 1. 特点：
     1. 开源：社区中维护了上百个beat，社区地址：

<https://github.com/elastic/beats/blob/master/libbeat/docs/communitybeats.asciidoc>

* + 1. 轻量级：体积小，功能单一、基于go语言开发，具有先天性能优势，不依赖于Java环境。
    2. 高性能：对CPU、内存和IO的资源占用极小。
  1. 定位：就功能而言，Beats是弟弟，得益于Java生态优势，Logstash功能明显更加强大。但是Logstash在数据收集上的性能表现饱受诟病，Beats的诞生，其目的就是为了取代**Logstash Forwarder** 。
  2. 官方支持的Beats（下载地址：<https://www.elastic.co/cn/downloads/beats>）：
     1. Filebeat：文件日志监控采集 ，主要用于收集日志数据
        1. 结构:



* + - 1. 语法：yaml

filebeat.inputs:

-input\_type:log

paths:

-/var/log/apache/httpd-\*.log

-input\_type:log

paths:

-var/log/messages

-var/log/\*.log

* + - 1. 输入配置：
         1. 支持的输入类型

Azure Event Hub

Cloud Foundry

Container

Docker

Google Pub/Sub

HTTP JSON

Kafka

Log

MQTT

NetFlow

Office 365 Management Activity API

Redis

S3

Stdin

Syslog

TCP

UDP

* + - 1. 输出配置:
         1. 支持的输出类型：

Console

Elasticsearch

Logstash

Kafka

Redis

File

Elastic Cloud

* + - * 1. 输出到Elasticsearch:

output.elasticsearch:

hosts: ["myEShost:9200"]

username: "filebeat\_internal"

password: "YOUR\_PASSWORD"

setup.kibana:

host: "mykibanahost:5601"

username: "my\_kibana\_user"

password: "YOUR\_PASSWORD"

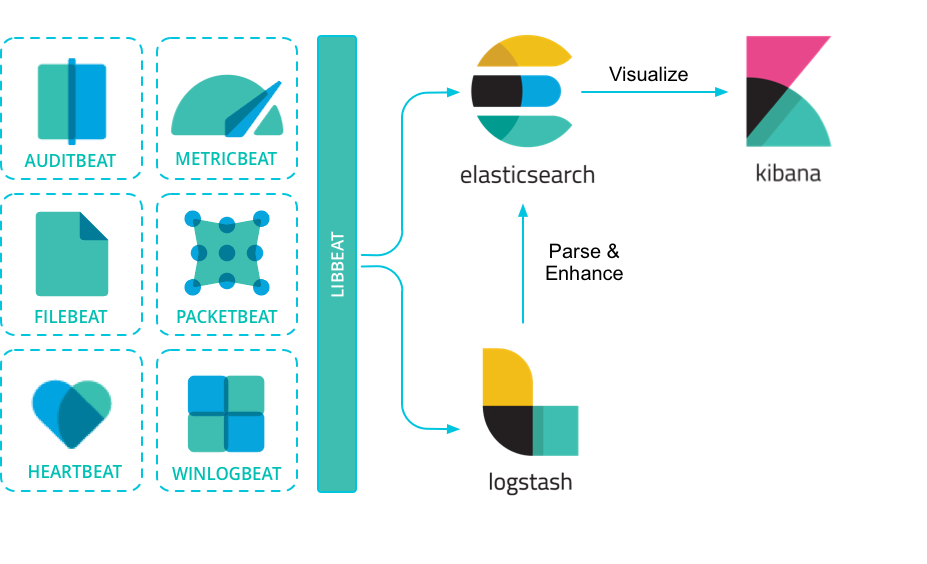
* + - 1. 执行：

head -n 1 ./logs/error.log |./filebeat -e -c ./filebeat.yml

* + 1. [Metricbeat](https://www.elastic.co/products/beats/metricbeat)：进行指标采集，指标可以是系统的，也可以是众多中间件产品的，主要用于监控系统和软件的性能
    2. [Packetbeat](https://www.elastic.co/products/beats/packetbeat)： 通过网络抓包、协议分析，基于协议和端口对一些系统通信进行监控和数据收集。
       1. 是一个实时网络数据包分析器，可以将其与Elasticsearch一起使用，以提供应用程序监视和性能分析系统
       2. 支持的协议：
          - ICMP (v4 and v6)
          - DHCP (v4)
          - DNS
          - HTTP
          - AMQP 0.9.1
          - Cassandra
          - Mysql
          - PostgreSQL
          - Redis
          - Thrift-RPC
          - MongoDB
          - Memcache
          - NFS
          - TLS
       3. 执行：

sudo ./packetbeat -e -c packetbeat.yml -strict.perms=false

* + 1. [Heartbeat](https://www.elastic.co/products/beats/heartbeat)：心跳检测(在配置的Url中喊一句：喂，有活着的么？有的话吱个声！)
    2. Winlogbeat：Windows事件日志
    3. [Auditbeat](https://www.elastic.co/products/beats/auditbeat)：审计数据（收集审计日志）
    4. Functionbeat：
    5. [Journalbeat](https://www.elastic.co/downloads/beats/journalbeat)：
  1. 架构：



1. 安装和部署
   1. Java依赖：Logstash要求java版本必须是java 8、java 11、java 14其中之一，并不是支持所有Java版本。配置以及环境要求，参考：<https://www.elastic.co/cn/support/matrix#matrix_jvm>

<https://www.elastic.co/cn/support/matrix>

* 1. 安装：参考ES安装
  2. 执行：

bin/logstash -f logstash-filter.conf