APM性能日志传输解决方案

主讲： 鲁班

时间： 2018/3/13 8:20

地址： 腾讯课堂–图灵学院

概要：

1. 日志传输解决方案设计
2. filebeat 采集配置与实现
3. filebeat 集成 logstatsh & elasticsearch

讲师介绍：

一、日志传输解决方案设计

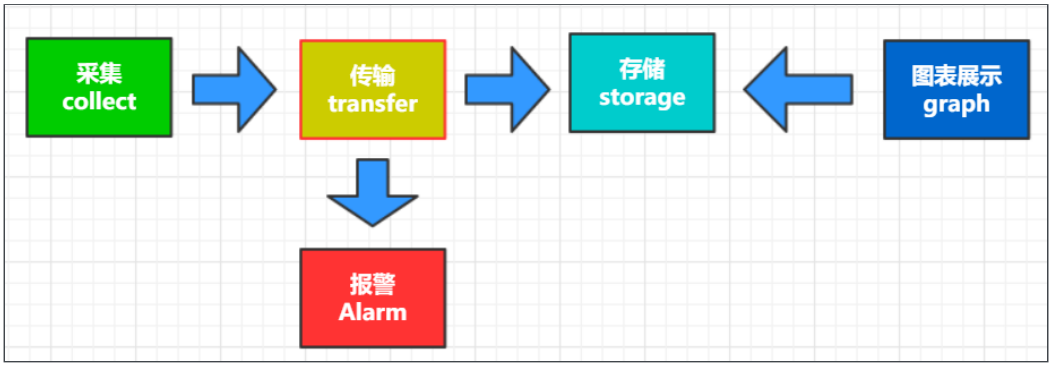
1. APM整体架构回顾
2. 传输模块解决方案

1、APM整体架构回顾：

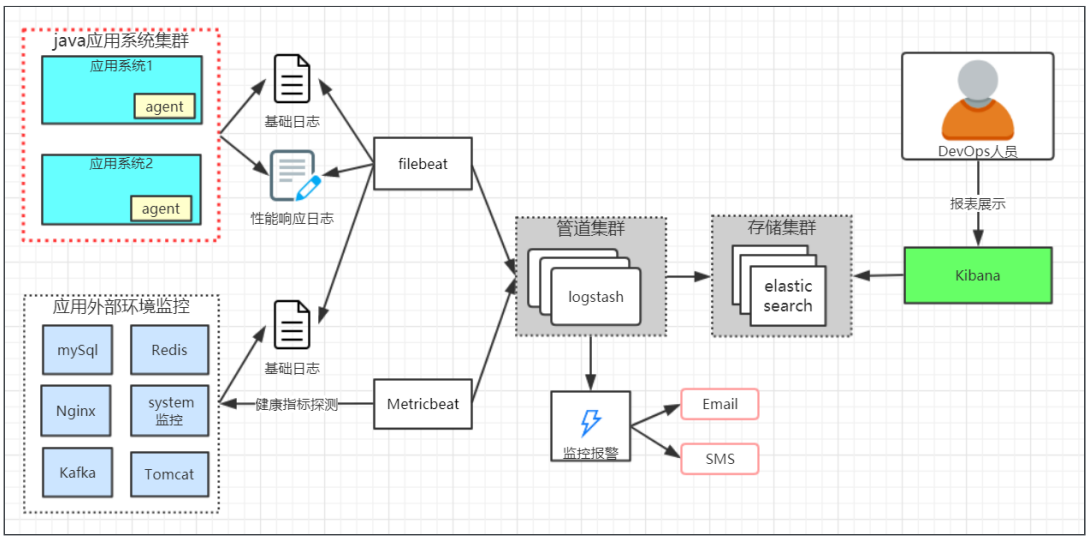
一个完整的APM系统，包含如下基本模块：

1. 性能日志采集
2. 数据传输
3. 异常判定与报警
4. 数据存储
5. 图表展示

各模块对应关系如下图：



各模块在我们的APM系统当中如何实现的?



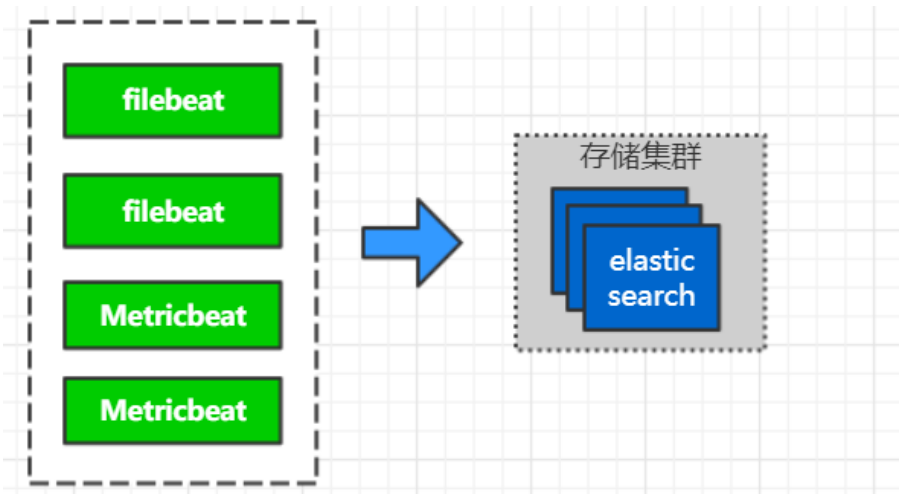
2、传输模块具体解决方案

传输模块的设计目标在于两点：

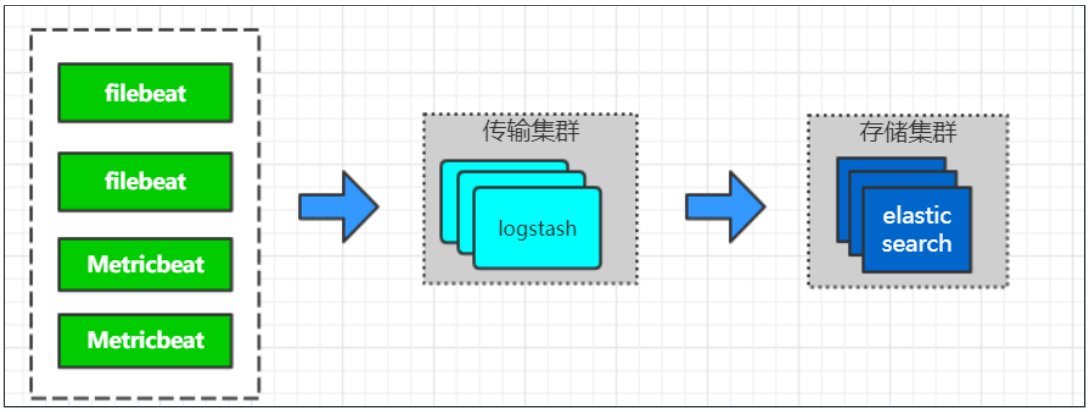
1. 保证信息能够顺利传输至报警判定中心与存储中心。
2. 可基于不同业务体量灵活调整架构规模，控制维护成本。

基于设计目标我们的传输方案分为以下三个档，以应对不同的业务体量。

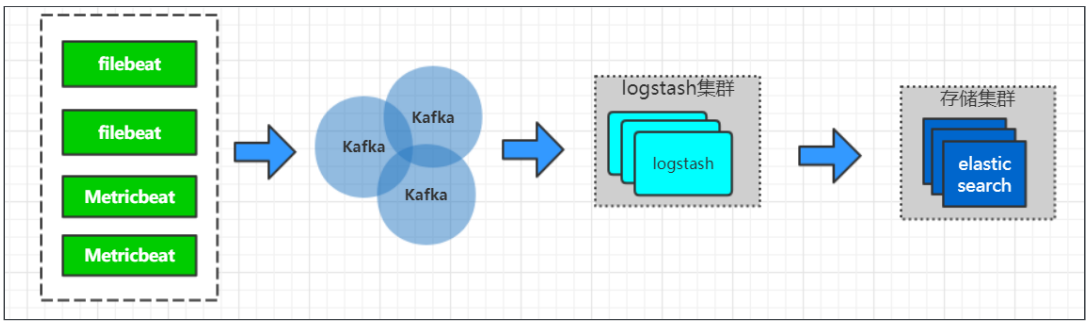
基本业务体量：



中等业务体量：



超大业务体量



二、filebeat 采集配置与实现

---

1. filebeat 简介
2. filebeat 安装与简单使用
3. filebeat 参数说明

1、filebeat 简介

filebeat 是elastic Beats 平台的一员，使用GO语言开发，期主用途为文件数据采集器。这些采集器安装后可用作轻量型代理，从成百上千或成千上万台机器向 Logstash 或 Elasticsearch 发送数据。其主要特点如下：

1. **轻量**：对使用环境无依懒。
2. **即插即用**：内部集成了一系列模块，用以简化收集、解析和可视化常见日志格式，诸如：NGINX、Apache
3. **可扩展**：可基于libbeat 自行构建采集器。

2、filebeat 安装 与使用

前往elastic 官网下载filebeat 指定版本 https://www.elastic.co/downloads/past-releases

```
#下载
wget https://artifacts.elastic.co/downloads/beats/filebeat/filebeat-5.6.6-
```

```
linux-x86_64.tar.gz
# 解压
tar -zxvf filebeat-5.6.6-linux-x86_64.tar.gz
#编辑配置文件
vim filebeat.yml
#启动
./filebeat
#后台启动
$nohup ./filebeat &
```

注：本次项目启用的Elastic系统版本统一为V5.6.6

filebeat 使用演示:

- [x] ~~编辑Filebeat配置文件，设置要采集的日志项与输出项~~
- [x] ~~重新启动Filebeat~~
- [x] ~~写入日志 查看Filebeat采集情况~~

3、filebeat参数说明

常用的filebeat 分为三个部分

1. prospectors 采集配置
2. Output 输出配置
3. global与 General 公共配置与一般配置

采用YAML 语言做为其配置语法。它的基本语法规则如下。

1. 大小写敏感
2. 使用缩进表示层级关系
3. 缩进时不允许使用Tab键，只允许使用空格。
4. 缩进的空格数目不重要，只要相同层级的元素左侧对齐即可

示例：

```
filebeat.prospectors:
- input_type: log
  paths:
    - /var/log/*.log
    - /var/log/nginx/*.access.log
- input_type: log
  paths:
    - /var/log/*.log
    - /var/log/nginx/*.access.log
output.elasticsearch:
  hosts: ["localhost:9200","192.168.0.01:9200"]
```

具体配置说明

模块	配置	说明
prospectors	input_type	可选值 log/stdin，分别代表日志输入和标准输入
	paths	数组：采集文本路径
	encoding	采集编码
	exclude_lines	排除指定符合正则表达示的行DEBUG TRACE INFO
	include_lines	包含除指定符合正则表达示的行
	exclude_files	排除非指定文件，正则表达示
	fields:	额外增加的字段和值，以键值对的形式配置level: debugreview: 1
	fields_under_root	true/false 额外字段是否存放至根目录
	document_type	指定 Elasticsearch 当中对应type 进行存储
	scan_frequency:10s	
	harvester_buffer_size:16384	
	max_bytes:10485760	
	json.message_key:	如果json 为多行格式时，指定一个root key，以进行标识
	json.overwrite_keys:false	false/true是否覆盖beat中的字段
	json.add_error_key:false	false/true 解析失败时是否存储解析失败信息
	json.keys_under_root:false	false/true 是否存放在根目录
	multiline.pattern: ^\[	多行文本 标识正则表达示
	multiline.match: after	after/before 文本匹配在 表达示之前还是之后
	multiline.max_lines: 500	
	output.elasticsearch: hosts:	
	output.logstash: enabled: true	
	output.console: enabled: true pretty: false	

三、filebeat集成logstatsh & elasticsearch

1. logstatsh & elasticsearch &Kibana 安装
2. filebeat 集成logstash
3. filebeat 集成elasticesarch

1、logstatsh & elasticsearch &Kibana 安装

前往 elastic 下载 3个软件的5.6.6 版本 https://www.elastic.co/downloads/past-releases

添另es用户（elasticsearch 不允许用root启动）

```
useradd es
```

2、filebeat 集成logstash

Nginx 日志配置

```
#配置json 日志格式
log_format json '{"@timestamp":"$time_iso8601",'
                  '"host":"$server_addr",'
                  '"clientip":"$remote_addr",'
                  '"remote_user":"$remote_user",'
                  '"request":"$request",'
                  '"http_user_agent":"$http_user_agent",'
                  '"size":$body_bytes_sent,'
                  '"responsetime":$request_time,'
                  '"upstreamtime":"$upstream_response_time",'
                  '"upstreamhost":"$upstream_addr",'
                  '"http_host":"$host",'
                  '"url":"$uri",'
                  '"domain":"$host",'
                  '"xff":"$http_x_forwarded_for",'
                  '"referer":"$http_referer",'
                  '"status":"$status"}';
#引入日志模板
 access_log logs/$server_name.access.log json;
```

FileBeat 配置

编辑filebeat.yml配置文件

```
vim filebeat.yml
```

```
filebeat.prospectors:
- input_type: log
  paths:
    - /root/svr/nginx/logs/access.log
    - /root/svr/nginx/logs/*.access.log
  #添加自定义字段
  fields:
    logIndex: nginx
    docType: nginx-access
  #自定义字段添加至根目录
  fields_under_root: true
#输出至logstash
output.logstash:
  # The Logstash hosts
  hosts: ["192.168.0.12:5044"]
```

logstash 配置

编写写 beats.conf

```
vim conf/beats.conf
```

在beats.conf 当中添加以下内容

```
input {
    beats {
        port => 5044
        codec => json
    }
}
filter {
    mutate {
        #删除filebeat自动添加的字段
        remove_field => ["tags", "beat"]
    }
}

output {
    stdout {
        codec => rubydebug
    }
}
```

3、filebeat 集成elasticesarch