[VIP课程]

基于ELK搭建网站流量可视化监控平台

做技术人的指路明灯，职场生涯的精神导师！

➢ Tom老师QQ号：441221062

01

ElasticSearch与Spring集成

类比学习法

利用已有的知识和新的知识多维度对比
产生差异化的结果，达到加深印象的目的

关系型数据库和ES操作姿势对比

关系型数据库	ElasticSearch
建库(DB) 建表(Table) 键约束(Constraint)	建库（Index） 建表（Index Type） 主键(ID)

关系型数据库和ES操作姿势对比

JDBC操作

1、加载驱动类(JDBC驱动)
2、建立连接（Connection）
3、创建语句集（Statement）
4、执行语句集 execute()
5、获取结果集（ResultSet）
6、关闭结果、语句、连接

ES Client操作

1、建立连接（TransportClient）
2、条件构造（SearchRequestBuilder）
3、执行语句 execute()
4、获取结果（SearchResponse）
5、关闭以上操作

代码效果演示

我们来模拟微信摇一摇的功能

从10W条记录中快速搜索到附近的人

大数据　　< 0.1s　　排序算法

代码效果演示

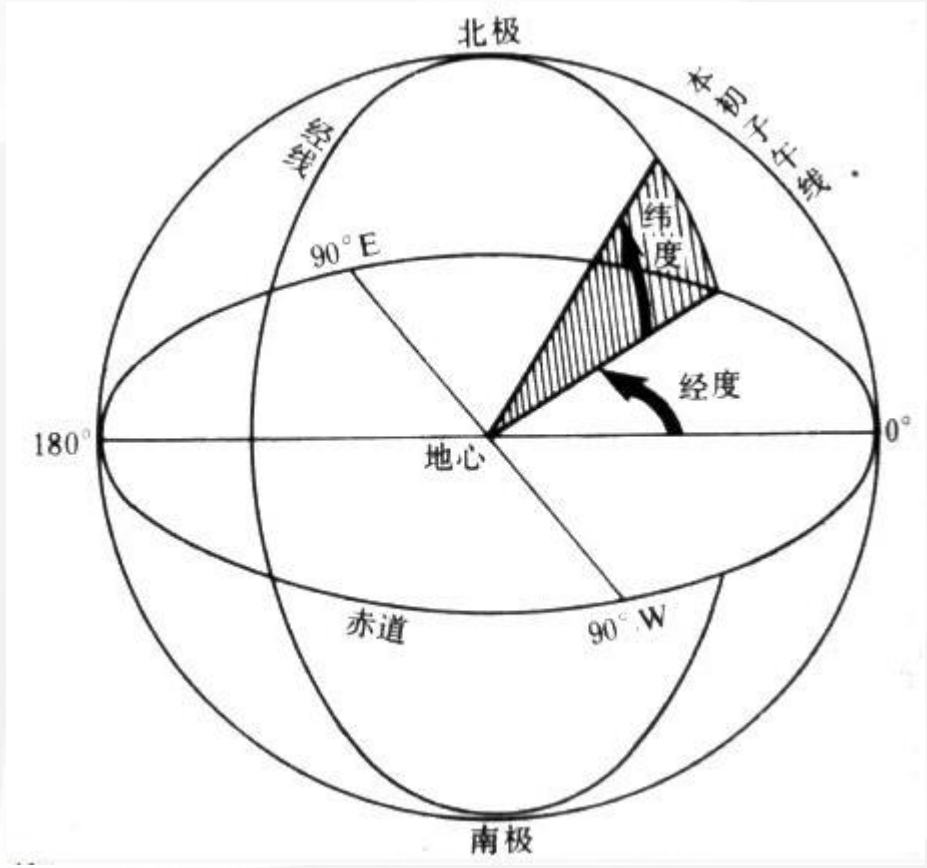
操作流程：

1、利用GPS设备获取每个人的位置（经纬度坐标）然后上传到服务器
2、根据Tom老师所在的位置，利用大数据搜索引擎实时搜索出附近的人，列出其个人信息
3、实现条件筛选，只看女生或者只看男生

代码实现：

1、利用ES搜索引擎随机生成10W条模拟数据
2、设定Tom老师所在位置（假如Tom老师就在长沙某地）
3、从模拟数据中匹配出符合条件的人（只找200米以内的人）
4、开始加人，聊天。

GEO Hash的基本原理



02

分布式带来的变革

问题

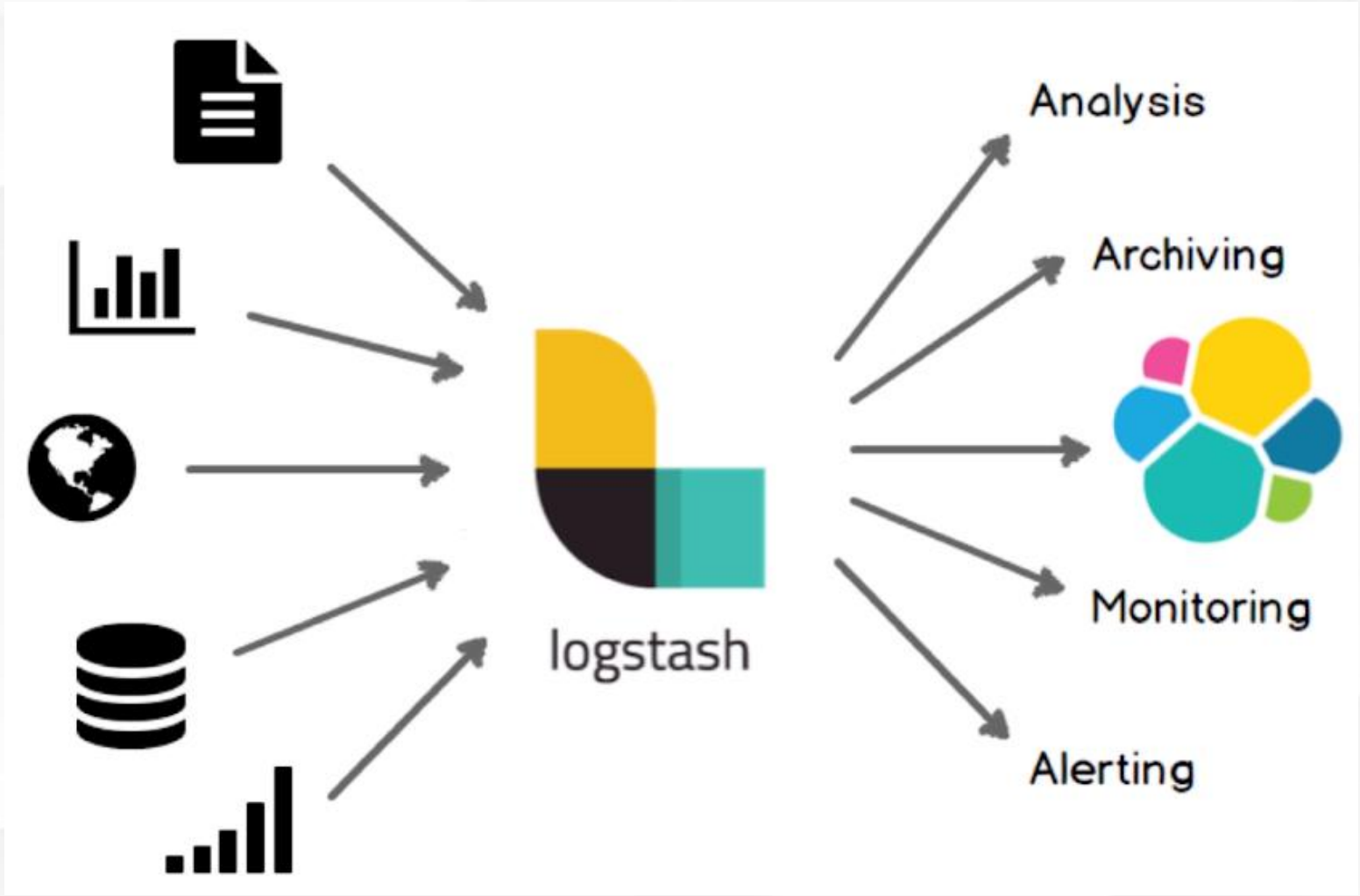
多节点

日志分散

运维成本高

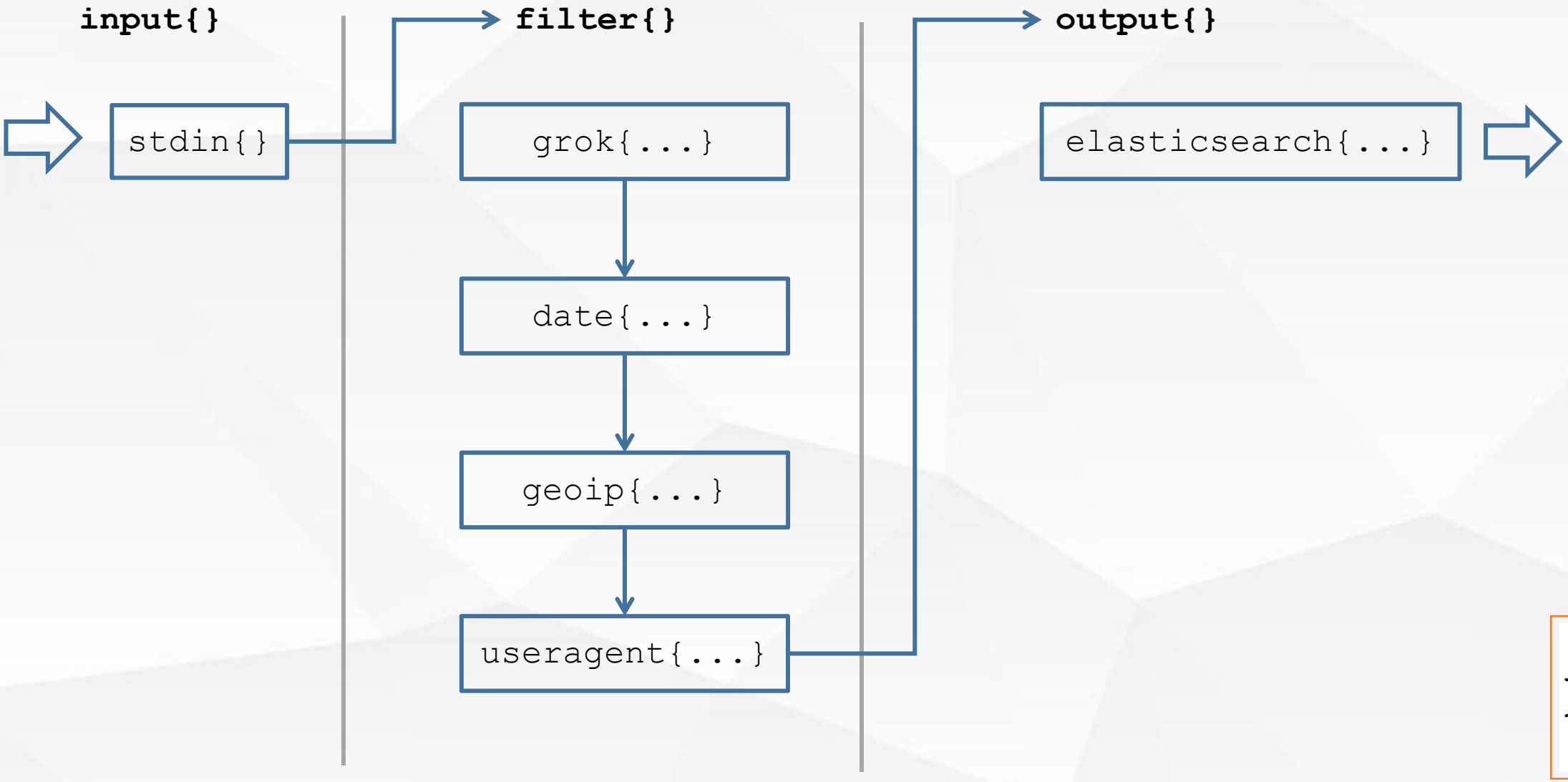
03 Logstash

架构简介



执行流程

执行流程



必须明白的概念

Pipeline

- input-filter-output的3个阶段处理流程
- 队列管理
- 插件生命周期管理

Logstash Event

- 内部流转的数据表现形式
- 原始数据在input被转换为Event，在output event被转换为目标格式数据
- 在配置文件中可以对Event中的属性进行增删改查

Input Plugin

input插件可以指定数据输入源，一个pipeline可以有多个input插件，我们简单几个常用插件的作用：

stdin　控制台标准输入
file　文件输入
tcp　TCP输入
http　通过HTTP协议输入

每个插件的属性都有差异，详情查看官方文档。

Codec Plugin

Codec（Code Decode）Plugin作用于input和output plugin，负责将数据在原始与Logstash之间转换，常见的codec有：

plain 读取原始内容

dots 将内容简化为点进行输出

rubydebug 将内容按照ruby格式输出，方便调试

line 处理带有换行符的内容

json 处理json格式的内容

multiline 处理多行数据的内容

Filter Plugin

Filter 是Logstash功能强大的主要原因，它可以对数据内容进行丰富的处理，比如解析数据、删除字段、类型转换等等，常见的有如下几个：

date 日期解析

grok 正则匹配解析

dissect 分隔符解析

mutate 对字段作处理，比如重命名、删除、替换等操作

json 按照json格式解析字段内容到指定字段中

geoip 增加地理位置数据

ruby 利用ruby代码来动态修改数据内容

04 现场搭建流量监控平台

代码演示

05

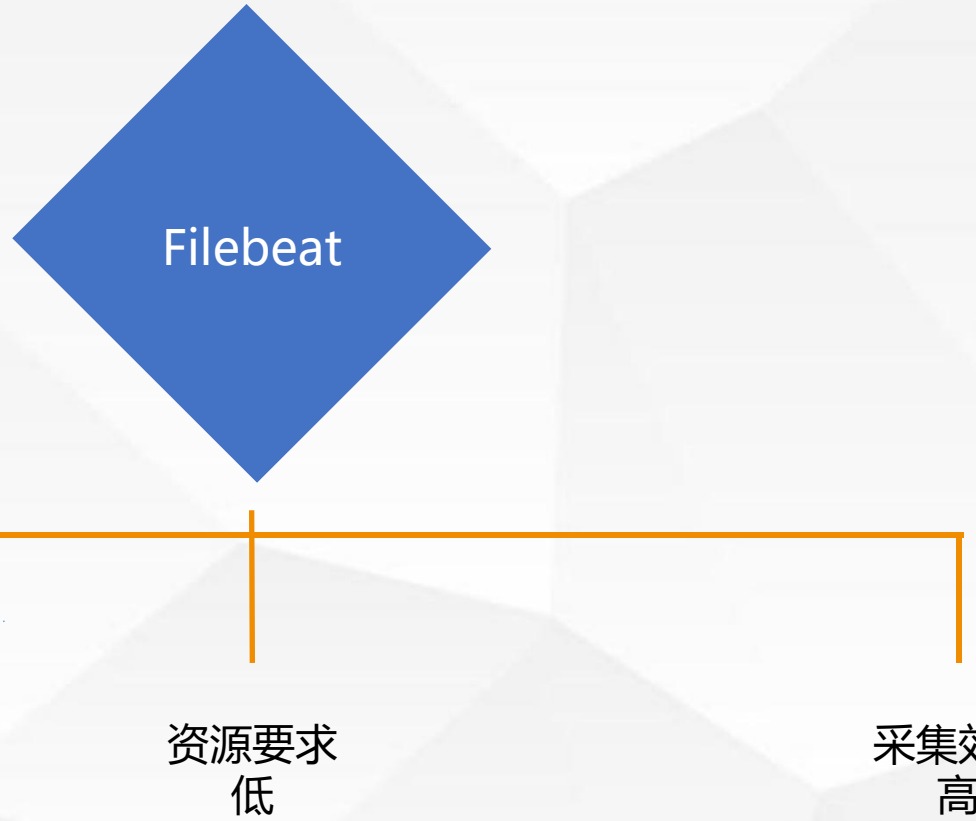
ELK的一般部署方案

ELK的一般部署

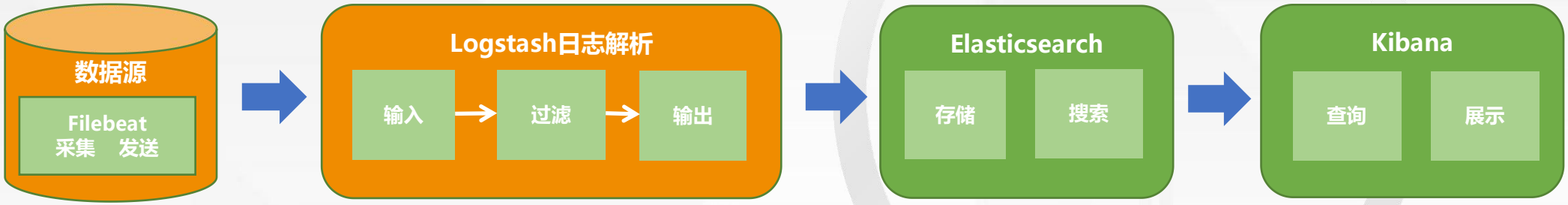
这种结构因为需要在各个服务器上部署Logstash，而它比较消耗CPU和内存资源，所以比较适合计算资源比较丰富的服务器，否则容易造成服务器性能下降，甚至可能导致无法正常工作，这是不可忍受的。

所以！我们需要一个资源消耗低，效率还不错的日志采集工具。

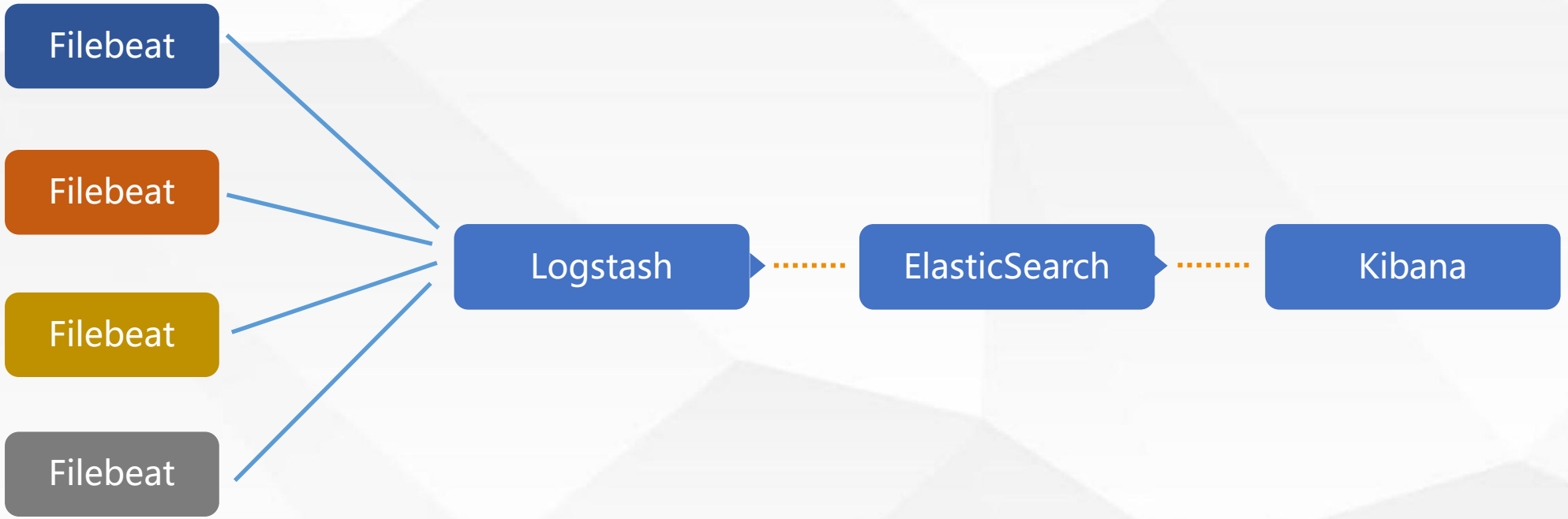
Filebeat介绍



ELK的一般部署

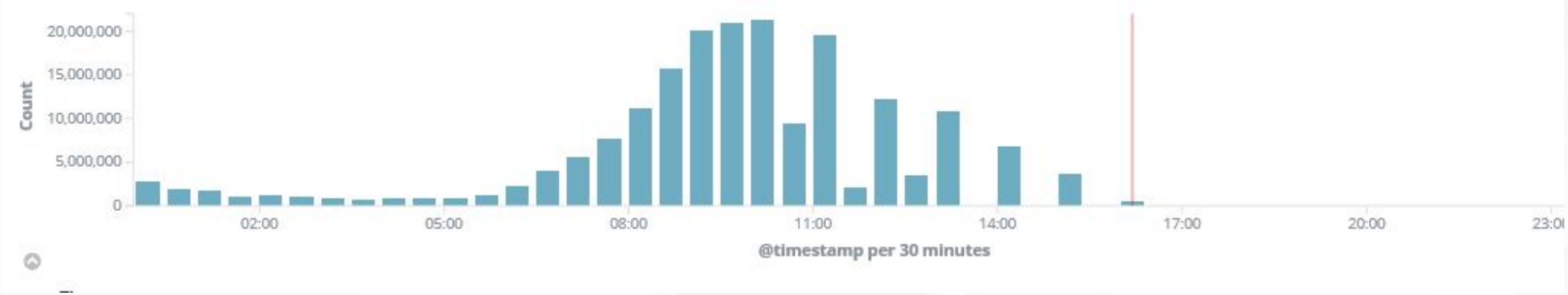


ELK的一般部署



ELK的一般部署

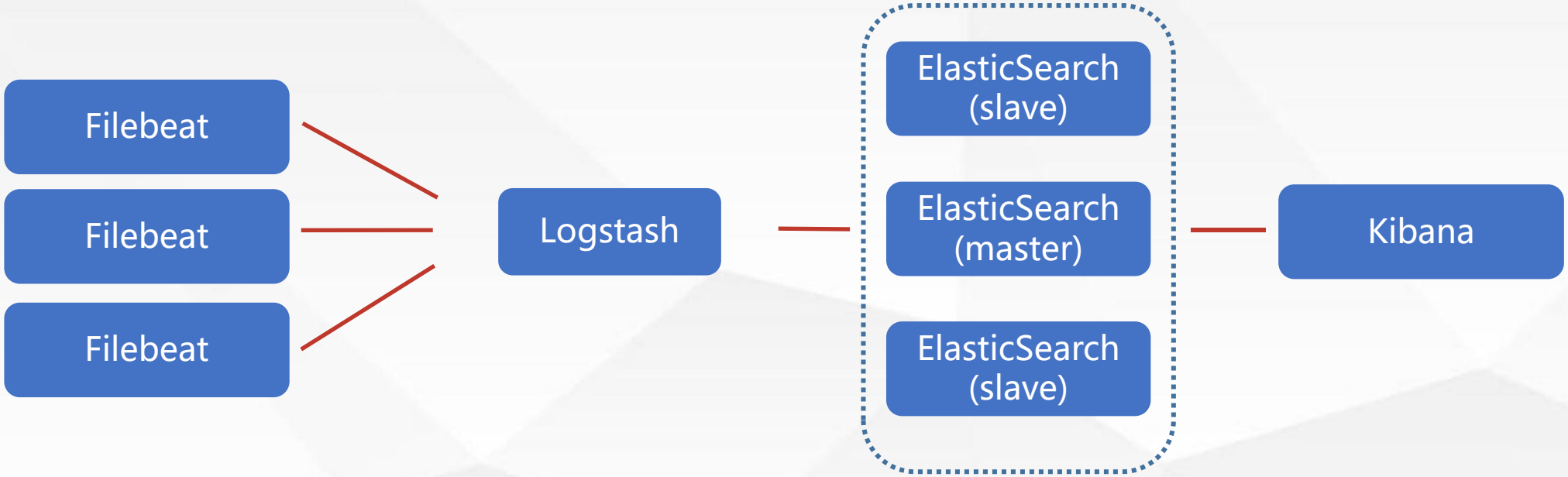
从Kibana上看发现：日志延时，缺失



如果解决这个问题？
提升采集效率才是王道

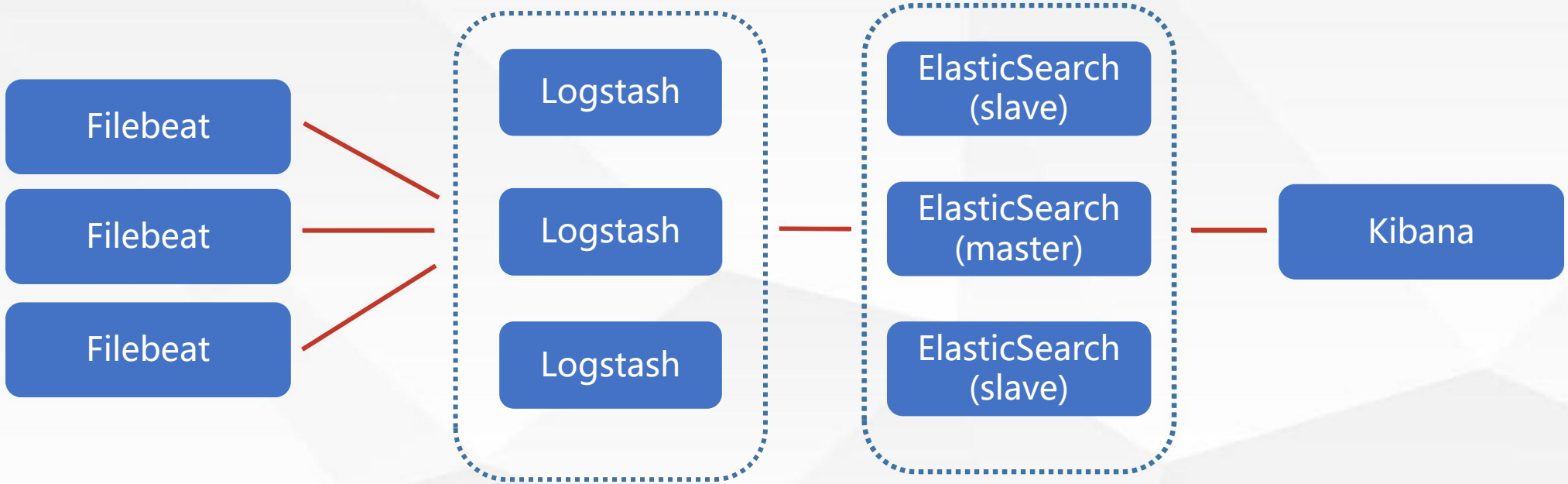
ELK的一般部署

架构继续调整，单个Filebeat支持配置往多个Logstash，再加上ElasticSearch集群



ELK的一般部署

架构继续调整，单个Filebeat支持配置往多个Logstash，再加上ElasticSearch集群



数据吞吐量持续增加怎么办？

ELK海量日志解决方案

ELK同步的采集机制

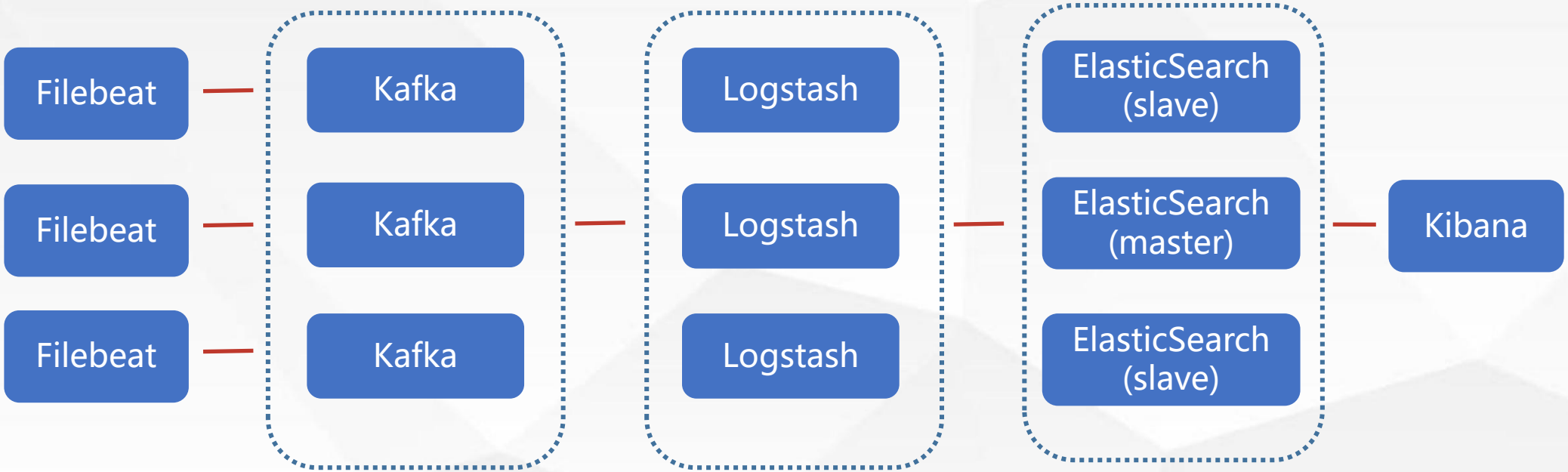
异步化

Filebeat支持异步化

引入消息队列机制　　如：Kafka

ELK海量日志解决方案

Filebeat与Logstash之间使用消息队列异步化



专业互联网IT教育服务平台

谢谢观看

咕泡学院，只为更好的你！

- Tom老师QQ号：441221062