
中生代技术
FRESHMAN TECHNOLOGY
IT大咖说
知识分享平台
ArchData
技术峰会上海站
主办方
中生代技术
快CTO
2017年9月9日 上海徐汇区田林路200号C座一楼COCOSPACE
ArchData技术峰会上海站
ArchData上海站

DerbySoft可伸缩的日志分析平台实践

付学良

日程

1. DerbySoft简介 & 系统背景
2. 日志分析
3. DerbySoft日志分析平台演进过程
4. 相关技术细节

德比软件全球酒店数据网络

GDN——Global Data Network

系统背景

- 典型分布式系统, 服务分布在全球十几个可用区
- 面向服务架构, 服务数量多
- 大部分服务部署在AWS

日志

- Apache HTTP Server

  2017-09-05T03:19:39 192.168.21.67 "GET /dsysmn/dsysmn-icinga.jsp HTTP/1.1" 200 19 rt=1093 0 cs=+ "-" "Zabbix" 127.0.0.1

- Tomcat

  07:01:52.813 INFO {main} [org.apache.catalina.startup.Catalina] : Server startup in 6288 ms

- ssh + cat / tail / less / grep
- cat * | grep kw1 | grep kw2 > abc.log
- python script

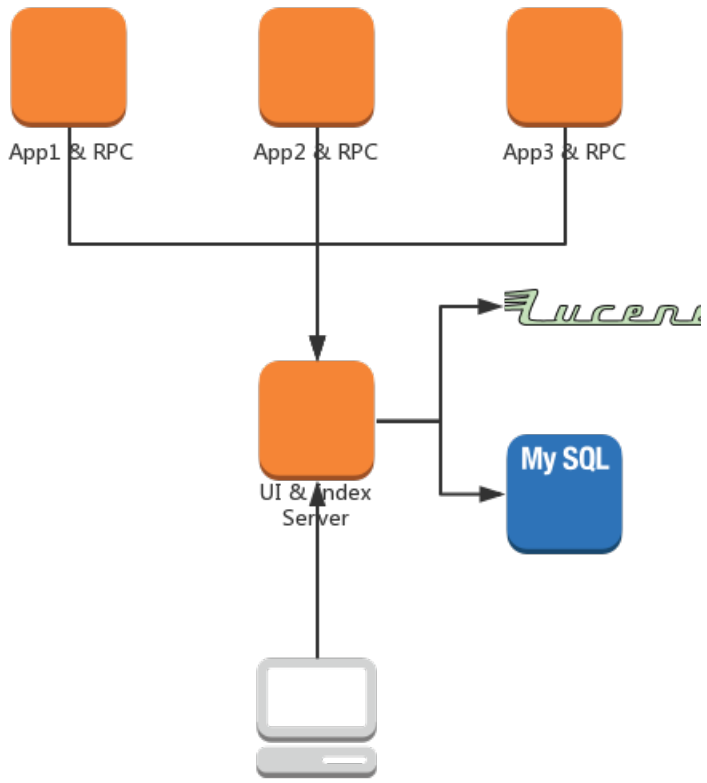
为什么要日志分析平台

- 查找日志
- 一个服务跑在多台服务器上
- 性能监控
- 错误分析
- On Call
- …

为什么要日志分析平台

- 采集
  - 收集日志
- 处理
  - 对日志进行过滤, 预处理, 格式化
- 计算
  - 检索, 分类, 统计
- 应用
  - 报表, 监控, 决策

DerbySoft日志平台演进 – V1



- 2008 ~ 2012
- 基于Web Service实现的Log4j Appender & Lucene
- 通用RPC接口实现日志实时查看
- ViewLog/SSH命令行 & 手工下载

DerbySoft日志平台演进 – V1

存在的问题

1. 只有关键日志进入系统
2. 功能上仅限于日志查找以及追踪调用链
3. 即使按天分表, 单表数据也越来越大, 索引和查询越来越慢
4. 日志类型单一, 数据提取困难, 更谈不上分析日志

DerbySoft日志平台演进 - Splunk

- 优点: 好用
- 缺点: 太贵

影响

1. 第一个标准化的日志 – Perf Log
2. 初步形成的监控机制

DerbySoft日志平台演进 – V2

elastic

logstash

DerbySoft日志平台演进 – V2

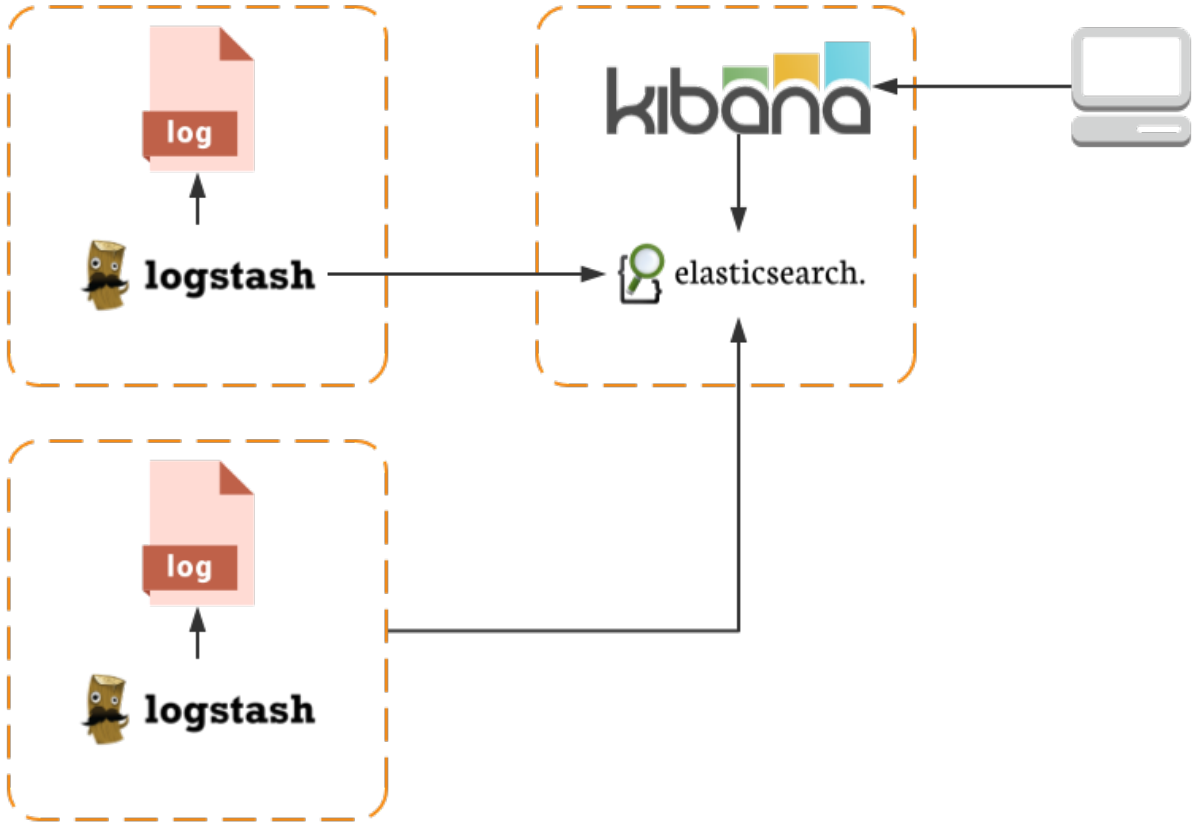
ELK

Elasticsearch: 基于lucene的存储, 索引, 搜索引擎

Logstash: 用来搜集和过滤日志的工具

Kibana: 基于Web用于查询和可视化存储在Elasticsearch中的数据图形界面

DerbySoft日志平台演进 – V2



- 2013
- 开源, 易于搭建
- 不侵入系统, 对工程师透明

DerbySoft日志平台演进 – V2

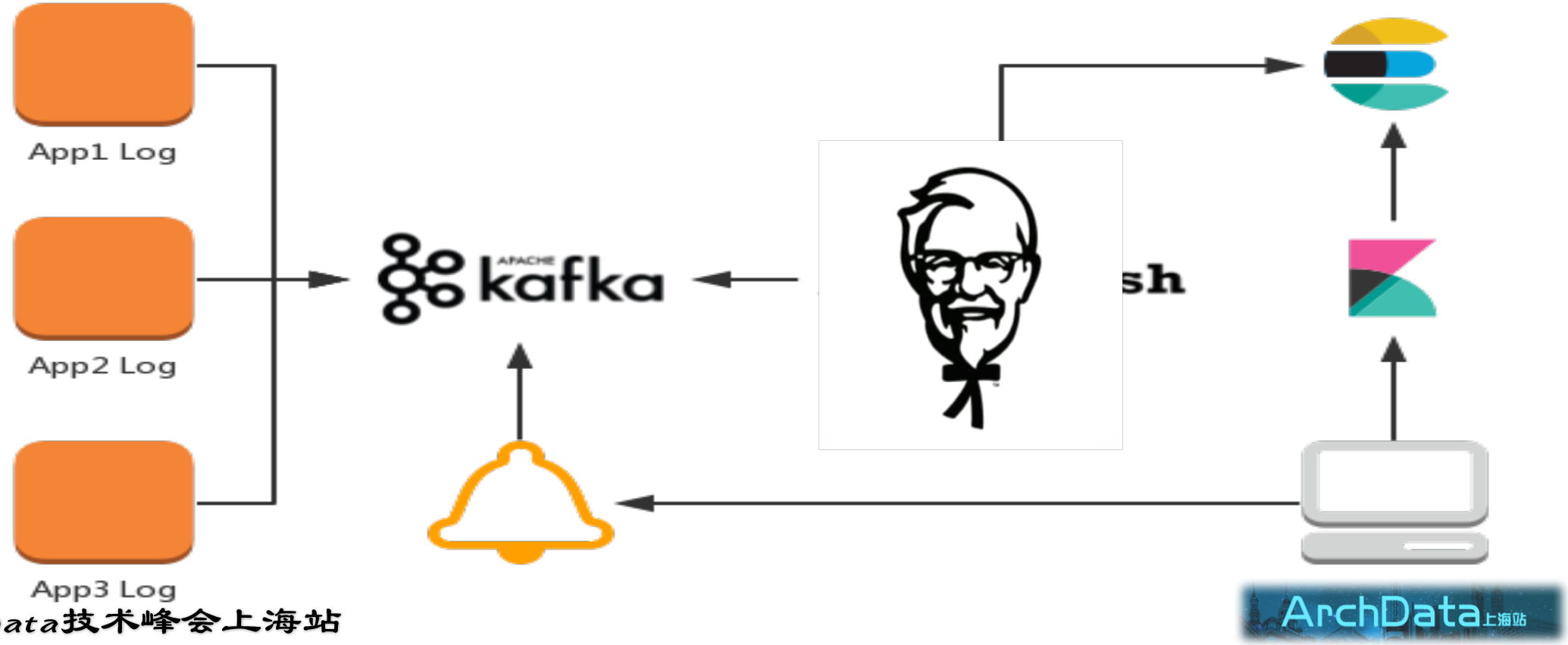
存在的问题

1. Logstash太重, 占用CPU和内存资源较大, 影响应用稳定性
2. 日志格式没有统一标准
3. 新增日志类型难以实施
4. 数据丢失

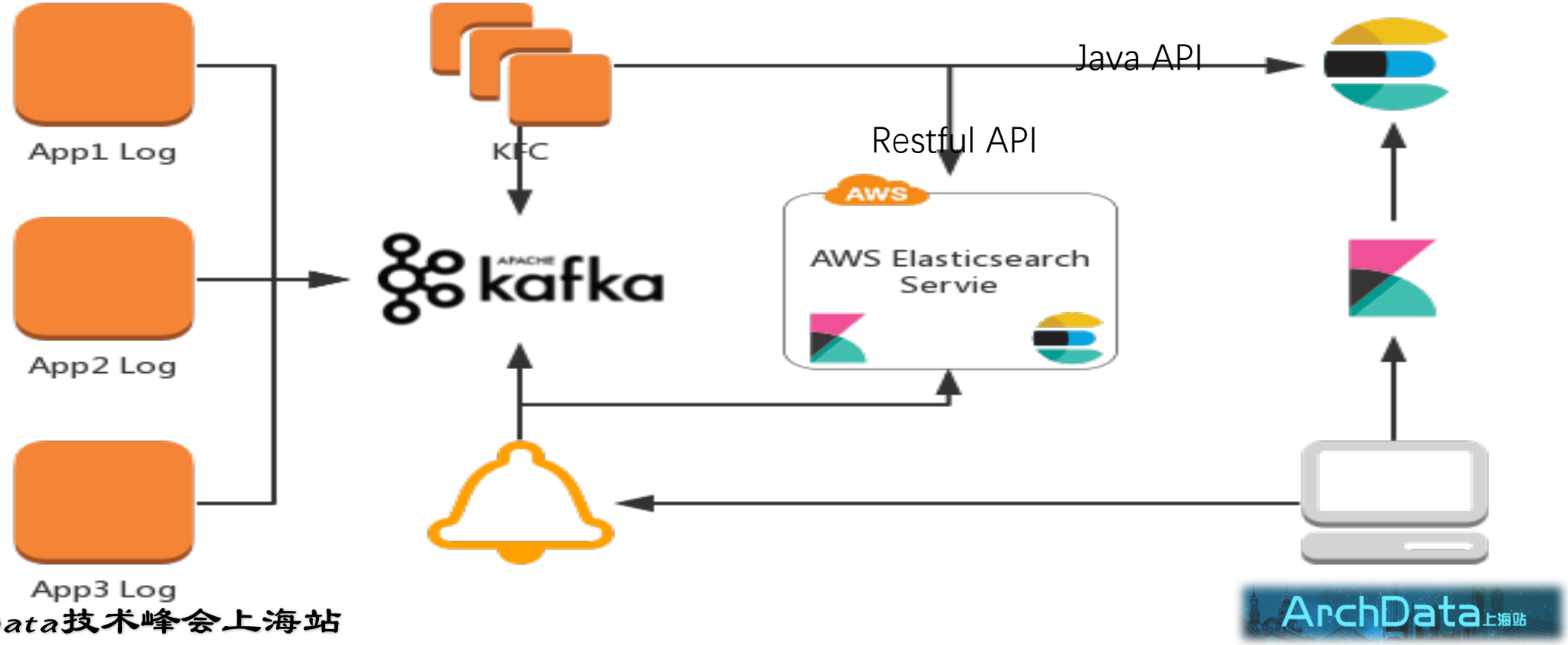
DerbySoft日志平台演进 – V3

- 分布式的流式数据平台
- 高吞吐
- 容错性

DerbySoft日志平台演进 – V3



DerbySoft日志平台演进 – V3



DerbySoft日志平台演进 – V3

新的问题

1. 日志数据需要长期存储, 客户有时要求我们提供半年前甚至是一年前的相关日志
2. Elasticsearch集群越来越大, 硬件资源投入陡增
3. Elasticsearch集群维护开销越来越大

DerbySoft日志平台演进 – V3

实际用例

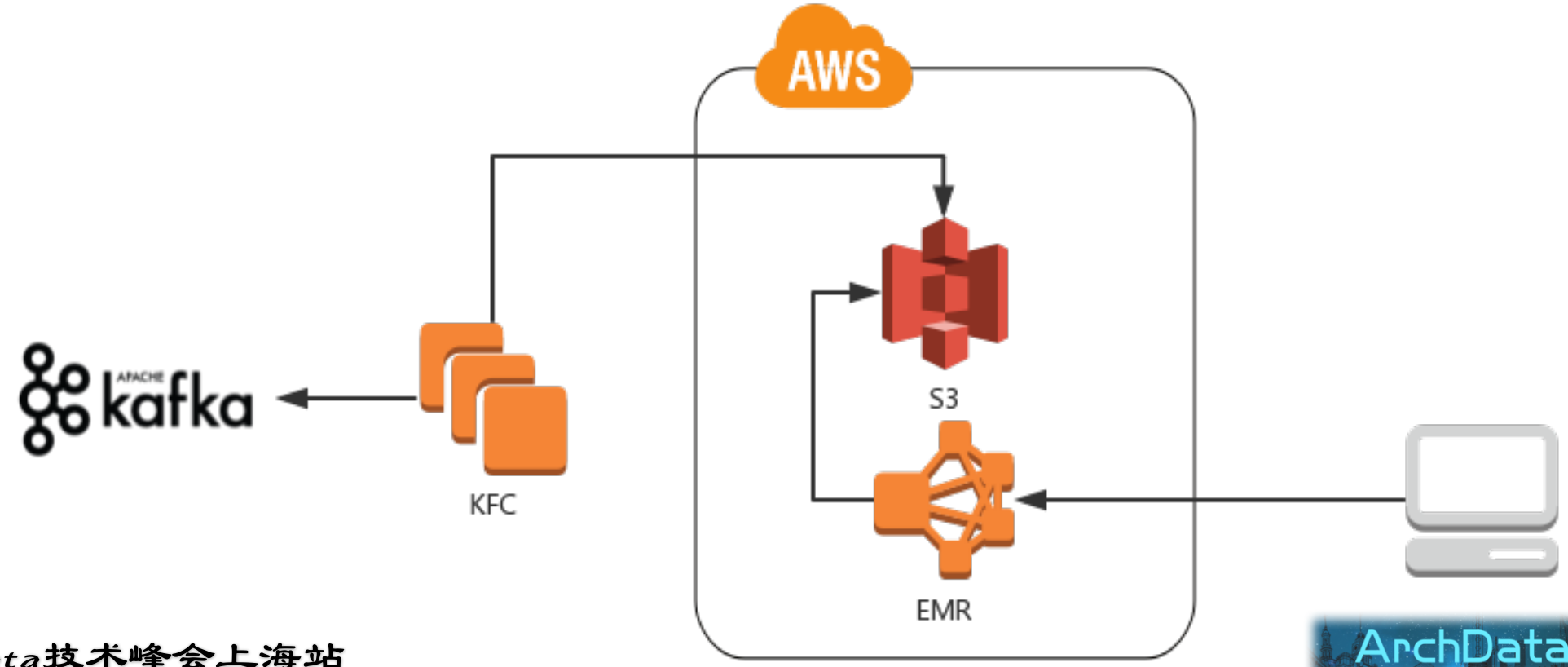
1. 30天以内的日志查询最多
2. 90天以内的日志查询时有发生
3. 90天以外的数据很少用到

基于成本和性能的妥协

- 常开
- 关闭, 按需使用
- 删除

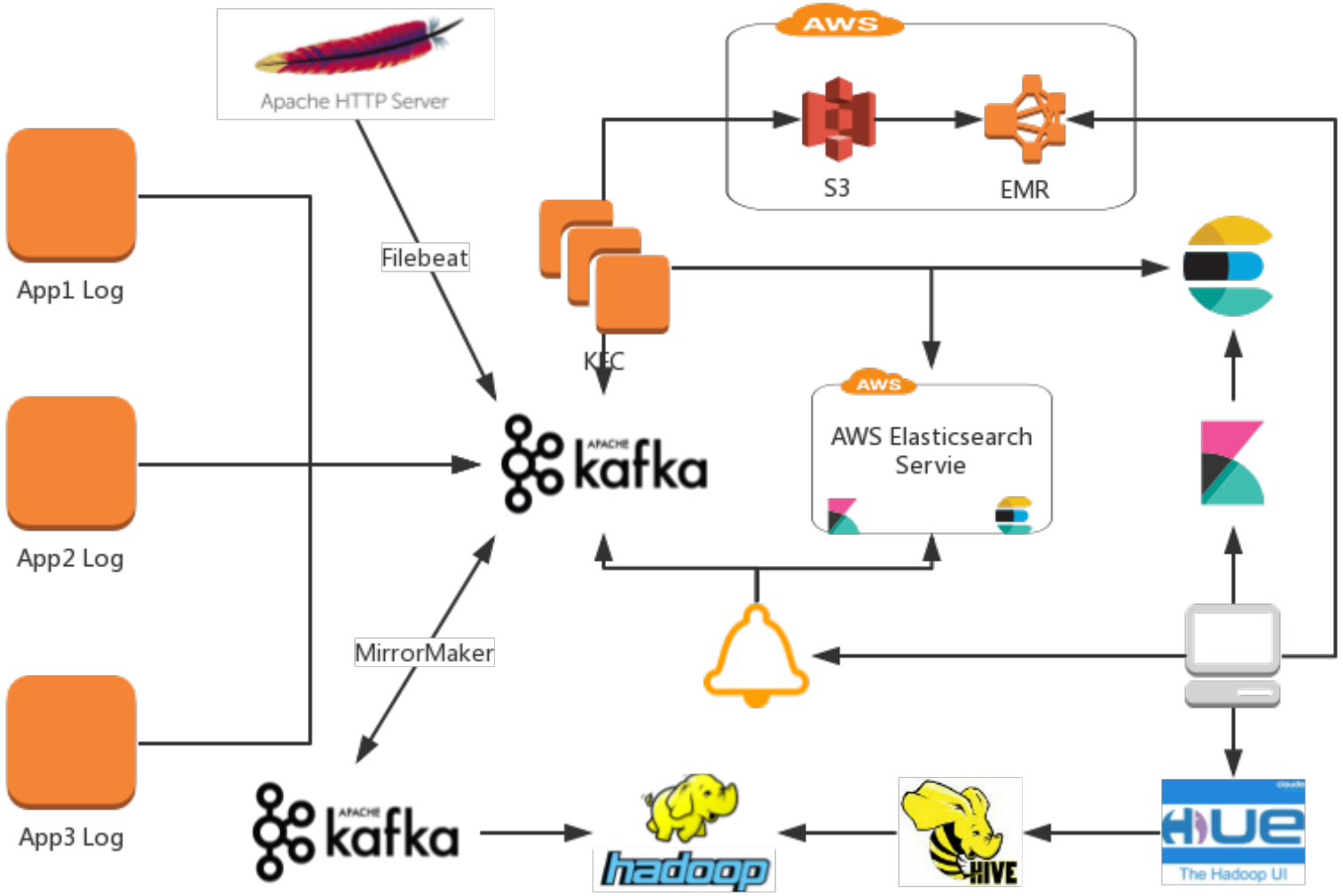
DerbySoft日志平台演进 – V4

引入S3



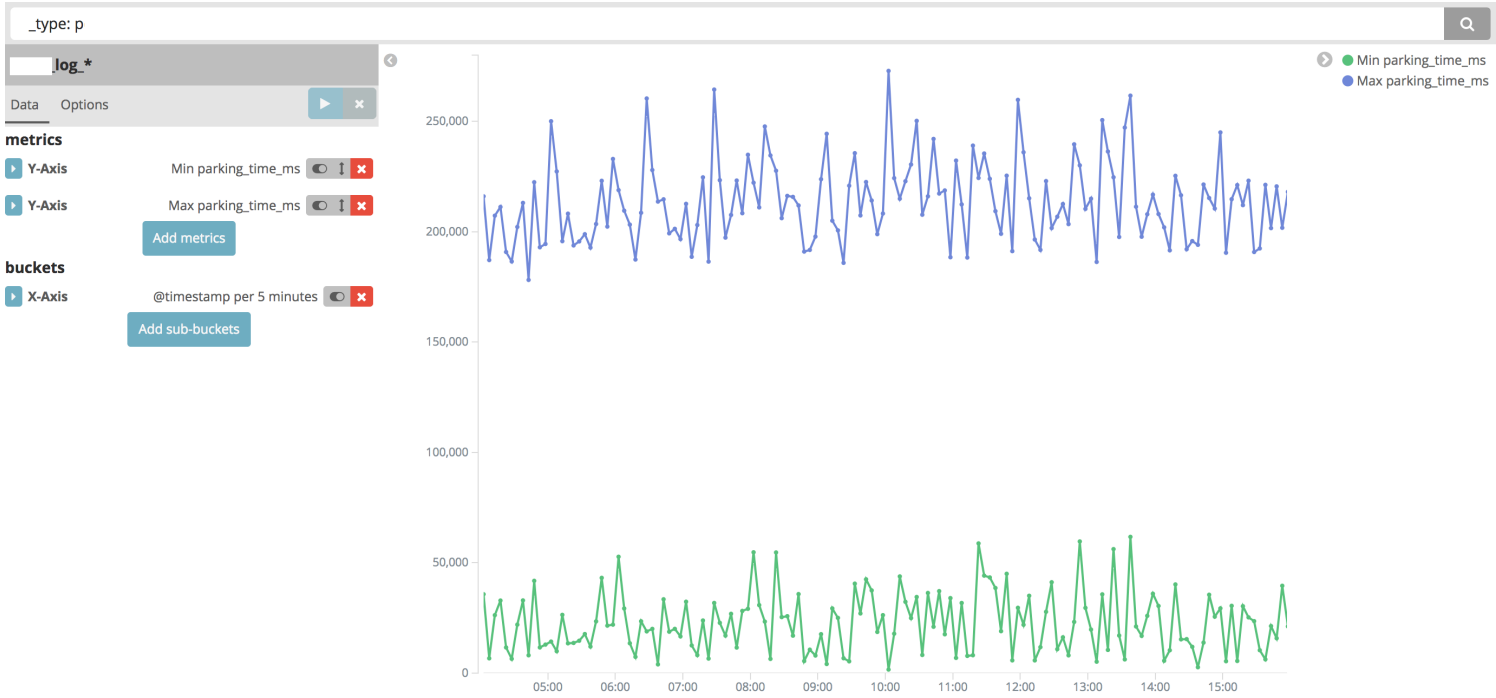
DerbySoft日志平台演进 – V4

- XML & StreamLog的处理
  - 提取关键信息
  - 结构扁平化, 用空间降低复杂度
- 基于列的数据存储 & Spark SQL
- 多种方案并存
  - S3 + DLog + AWS Elasticsearch Service
  - S3 + EMR
  - Hadoop + Hive

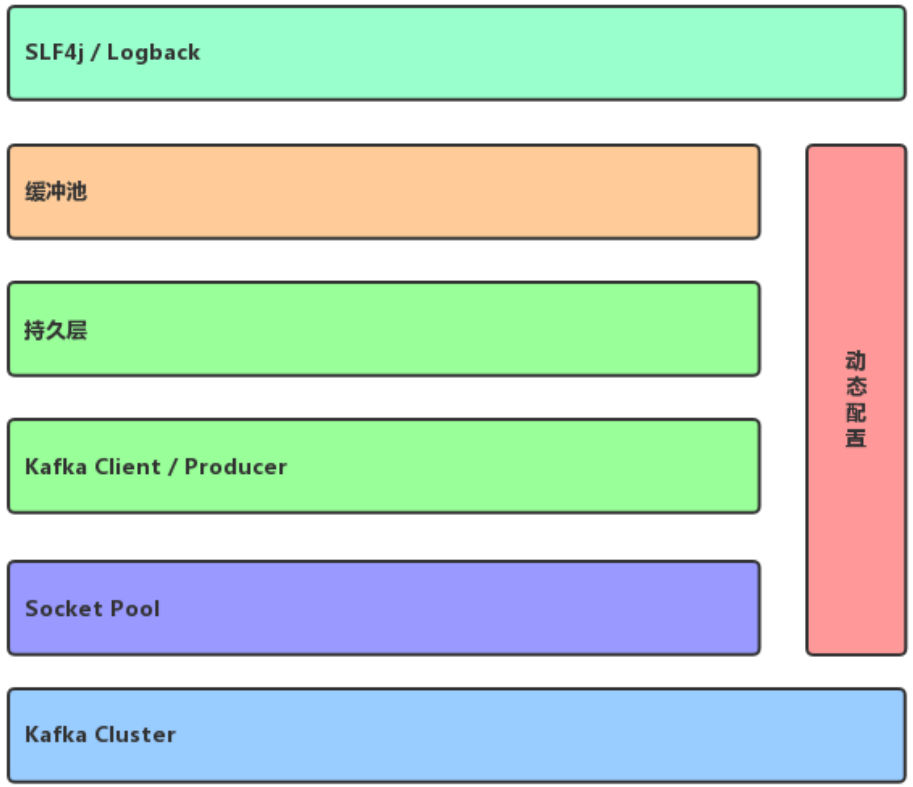

Apache HTTP Server
AWS
S3
EMR
App1 Log
App2 Log
App3 Log
Filebeat
KEC
kafka
AWS
AWS Elasticsearch Servie
MirrorMaker
kafka
hadoop
HIVE
HUE
The Hadoop UI

DerbySoft日志平台现状

- 1000 + topics
- 5分钟
- 50 Billion
- 1 TB



SLF4J Kafka Appender



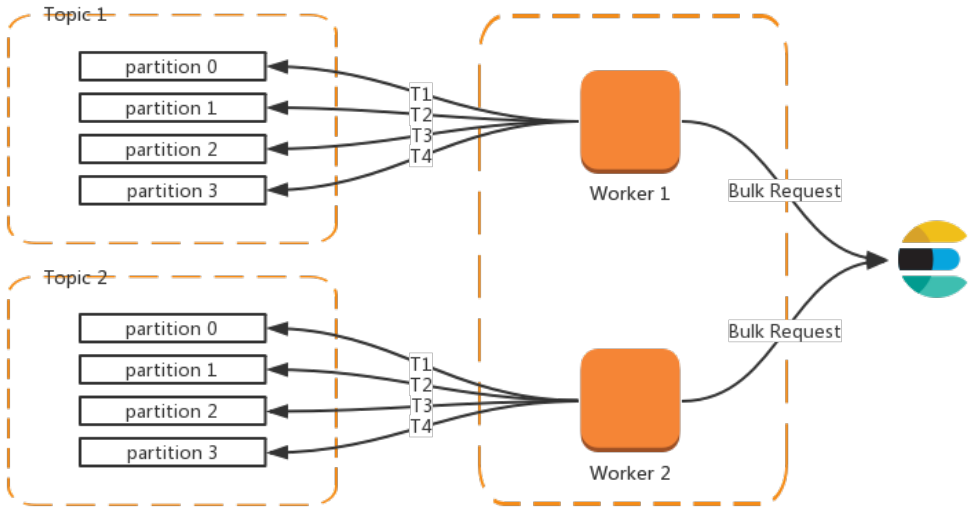
- 缓冲
- 持久化 & 批量处理
- 重试
- 监控
- 动态配置

KFC

一个可高度扩展的Kafka消息处理程序.

1. 基于时间间隔的批次处理
2. 擅长处理无序消息
3. 高度灵活(基于Groovy的通用处理框架)
4. 优先级调度, 分组处理…

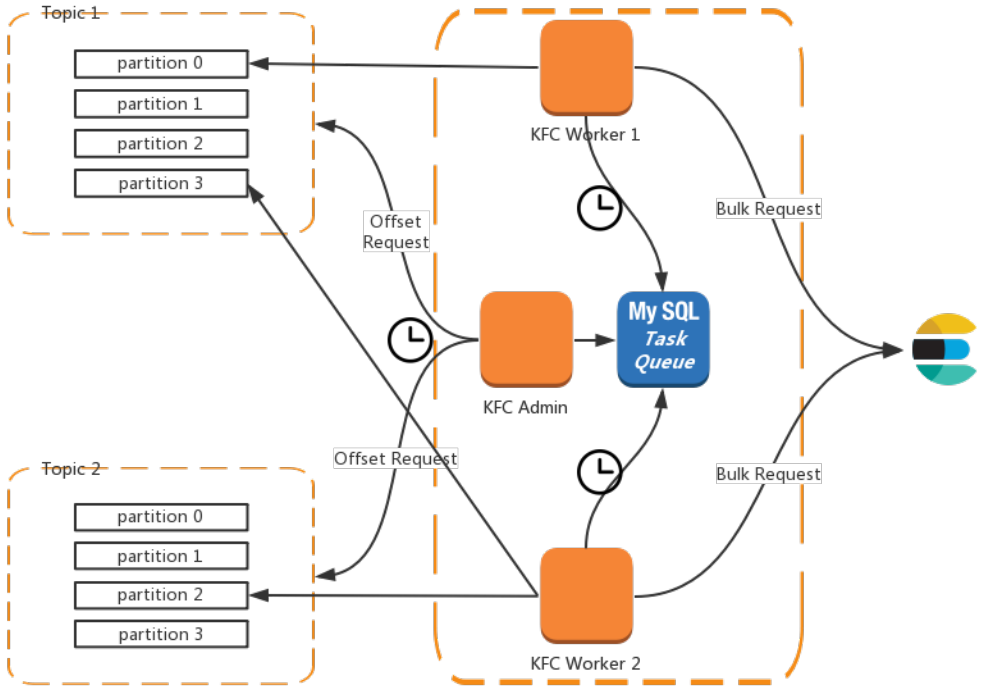
KFC – V1



- 有序实时处理
- 性能最佳

- 线程爆发
- 资源浪费
- 可控性差

KFC – V2



- 近实时处理
- 扩展性好
- 资源可控

- 资源竞争
- 无序

Kafka Proxy Server

- Python & Twisted
- 端口映射
- 集群

```
47      def _send_metadata_response(self, correlation_id, metadata_rq):
48          with self.metadata_lock:
49              metadata_response = self.factory.kafka_client.send_metadata_request(
50                  metadata_rq.topics)
51              proxy_brokers = []
52              for broker in metadata_response.brokers:
53                  proxy_broker = self.factory.get_proxy_broker(
54                      self.transport.getPeer().host, broker[0])
55                  proxy_brokers.append(proxy_broker + broker[3:])
```

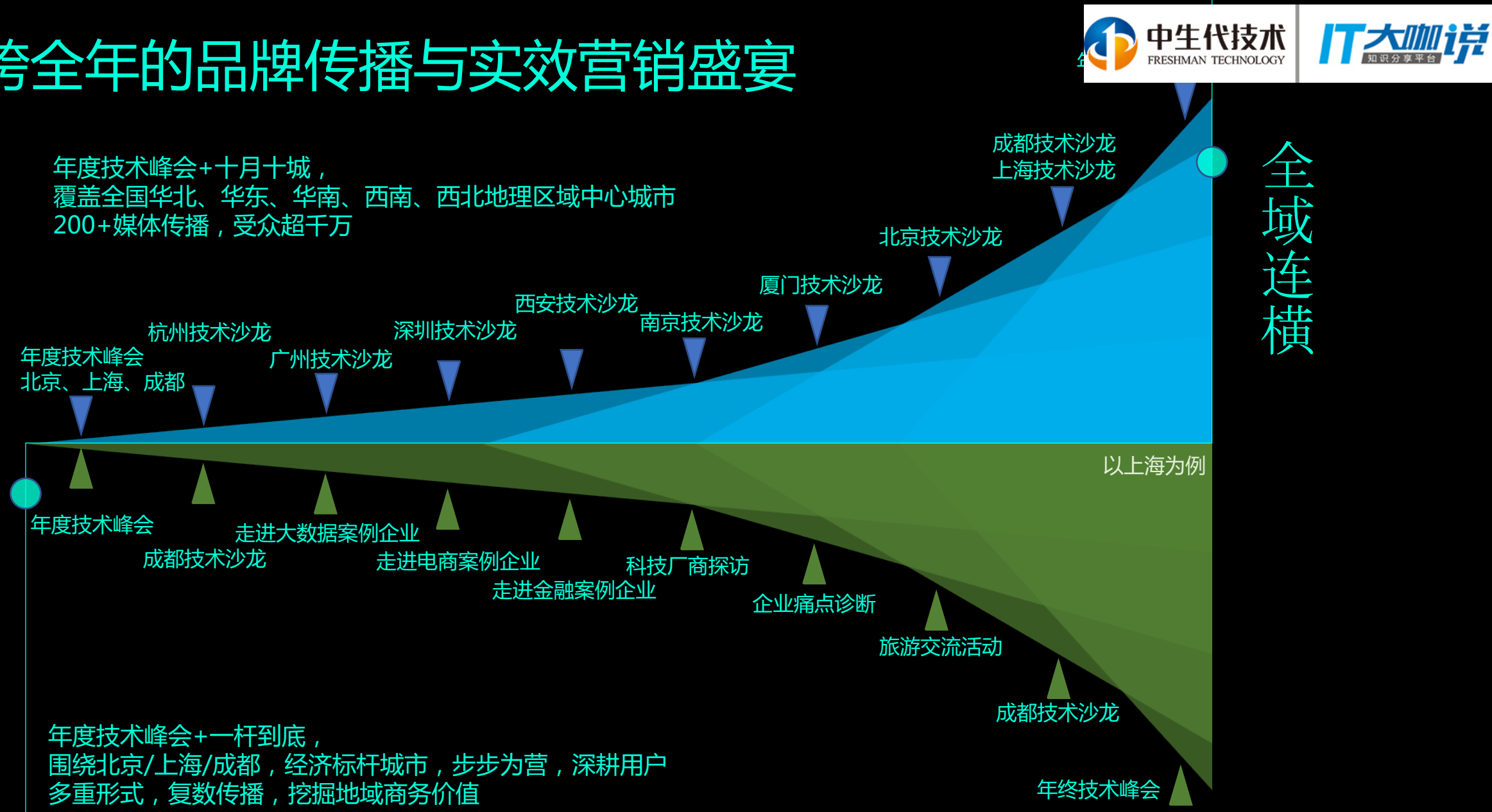
```
28
29      def stringReceived(self, string):
30          api_key, api_version, correlation_id, client_id_size, message = \
31              struct.unpack("!hhih%ds" % (len(string) - 10,), string)
32          self.log.debug(
33              "Received client[%s]'s request with Api(%s, %s), CorrelationId: %s"
34              % (self.transport.getPeer().host, api_key, api_version,
35                 correlation_id))
36          if api_key == MetadataRequest[0].API_KEY:
37              _, message = struct.unpack(
38                  "!%ds%ds" % (client_id_size,
39                               len(message) - client_id_size), message)
40              d = threads.deferToThread(
41                  self._send_metadata_response, correlation_id,
42                  MetadataRequest[api_version].decode(message))
43              d.addCallback(self.sendString)
44              return
45          self.peer.sendString(string)
```

Thank You!

德比软件(DerbySoft) 付学良

2017-09-09

横跨全年的品牌传播与实效营销盛宴



合纵连横，在中国开发者群体中缔造品牌营销奇迹

预告：ArchData峰会（北京）精彩内容抢先看！9月24日 关键词：AI、区块链、大数据、

09:00–09:50	史凯	AI驱动的企业创新架构
09:50–10:40	李艳鹏	区块链原理解析
10:40–10:50	短歇	
10:50–11:40	王东	微服务下的APM全链路监控
13:30–14:20	严静	漏斗转换运算的优化过程
13:30–14:20	吴金龙	深度学习与智能对话机器人
14:20–15:10	裴丹	智能运维中的AI问题
15:10–15:20	短歇	
15:20–16:10	涂威威	Towards AI for Everyone
16:10–16:50	余军	分布式数据库在金融行业的创新实践
16:50–18:00	何文斌	主题待定

ArchData上海站