ELK在实时流量分析中的应用

Clearclouds-global.com

杨润达 Tech.lead



Tech.lead

杨润达

runda.yang@clearclouds-global.com

ELK在实时流量分析中的应用

Overview

- 1 概 述
- 2 流量分析平台的比较
- 3 为什么使用ES
- 4 提高ES的导入速度
- 5 网络数据分析

- 6 网络流量的可视化
- **7** Kibana 插件开发
- ❷ 功能演示
- 9 Questions & Answers

Thank you for being here today

流量分析的方式

overview



采样、精度不够 增加设备压力



粒度最细、分析全面 数据量太大



根据需求选择数据全流量

为什么使用Elasticsearch

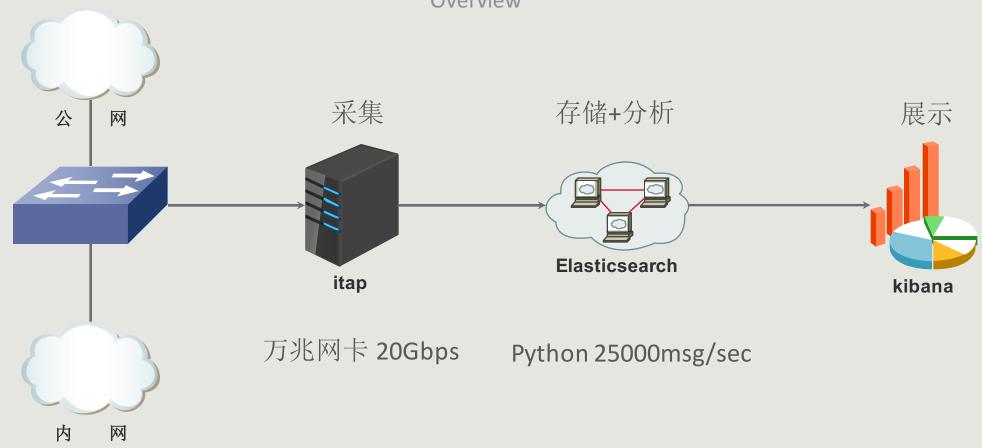
overview







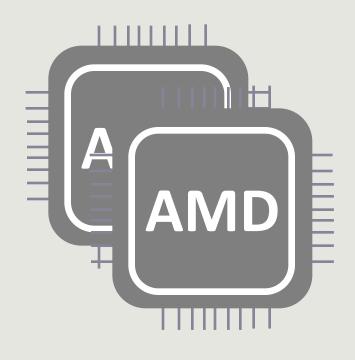
概 述 Overview



提高数据导入的速度



机器配置



2 x AMD Opteron(tm) Processor 6172

12 cores per CPU



32GB



SATA 1TB

多进程优化方法

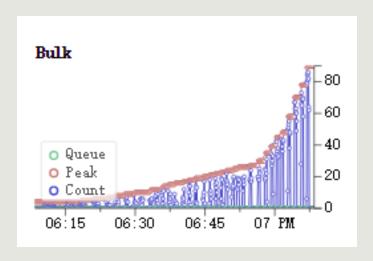
5个字段



100万条真实TCP数据



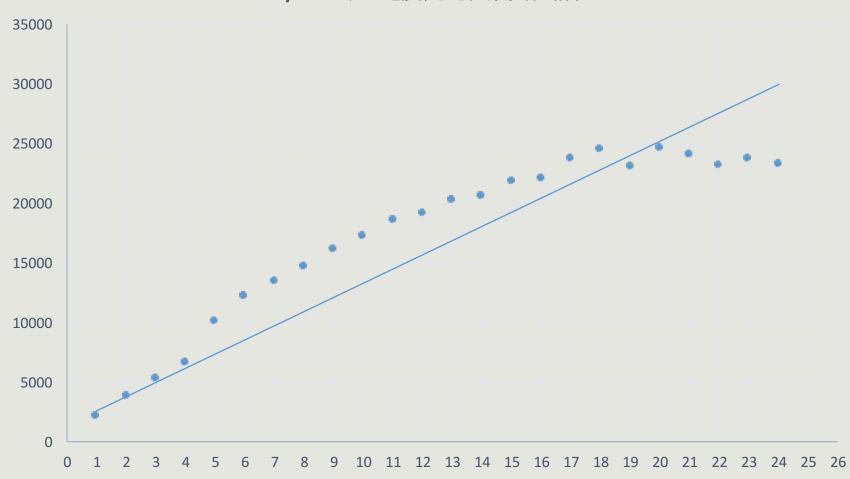
多进程并行导入



开启bulk线程池

优化结果

Python导入速度随进程数变化情况



json优化前

```
共计: 1155s
       平均每次:1016us
               占总时间:61.2%
 1155422158
           1016.2
                   61.3
```

使用uJson

```
共计: 26s
    平均每次:23us
            占总时间:10.9%
                10.9
26331870
         23.2
```

IP段聚合分析

内网

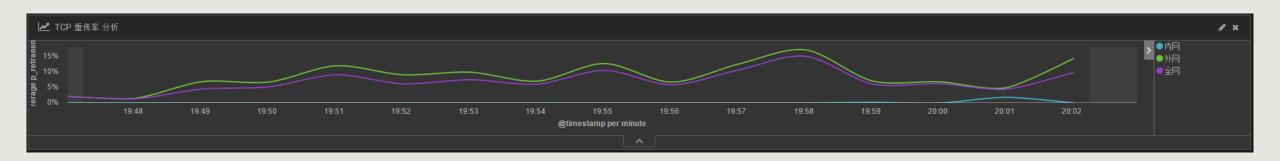
sip: [192.168.0.0 TO 192.168.1.255] AND

dip: [192.168.0.0 TO 192.168.1.255]

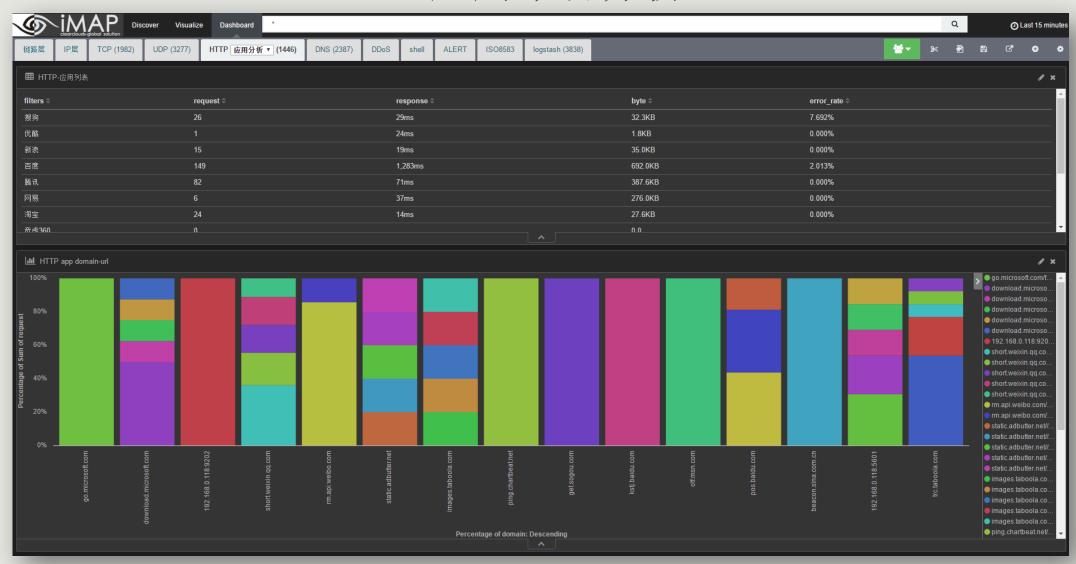
外网

NOT (sip:[192.168.0.0 TO 192.168.0.255] AND

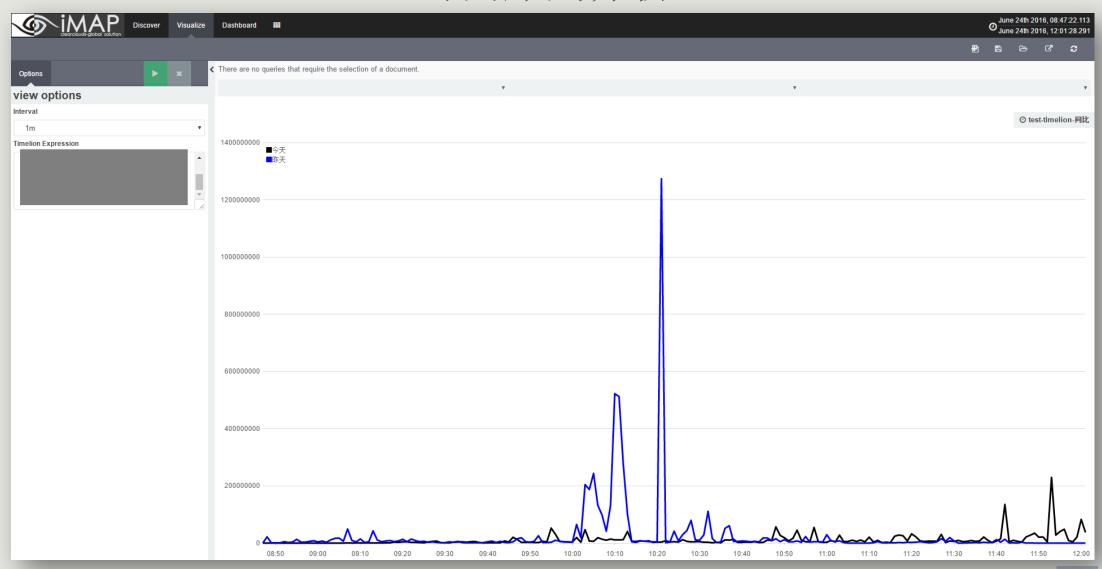
dip:[192.168.0.0 TO 192.168.1.255])



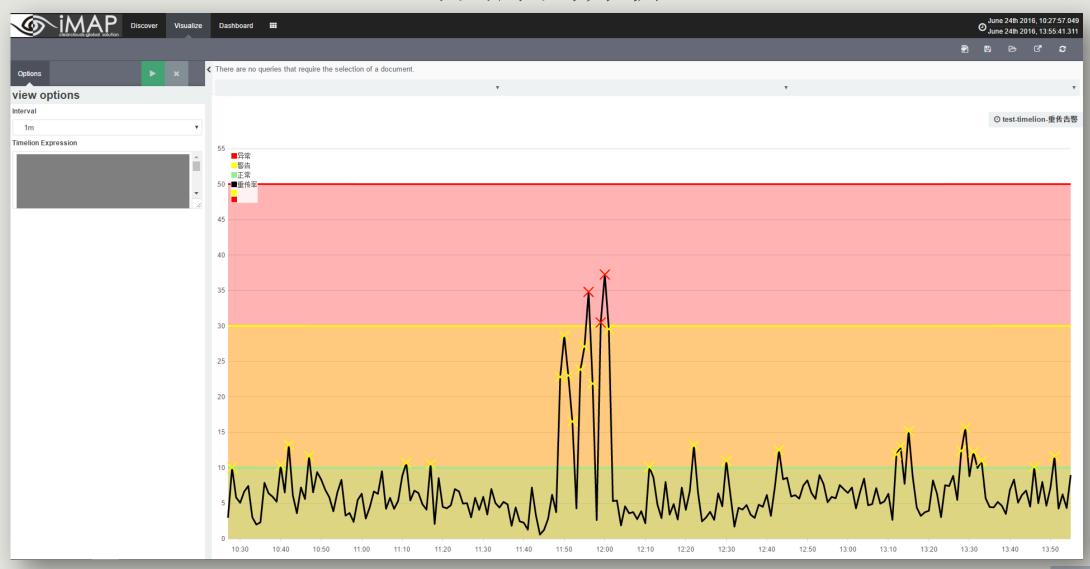
应用聚合分析



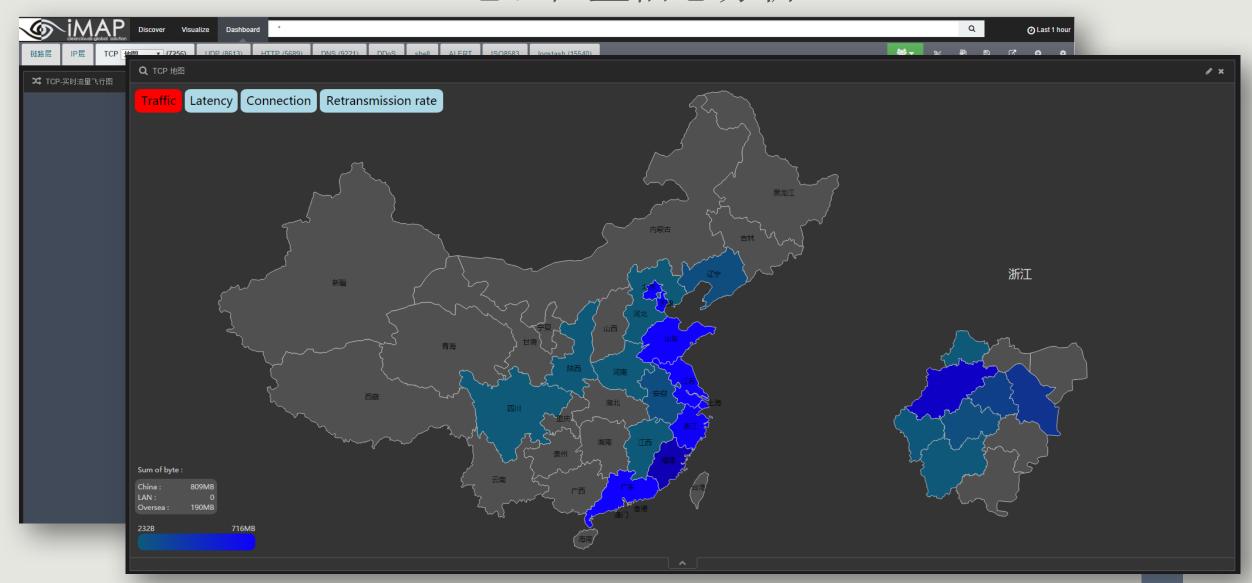
时间序列分析



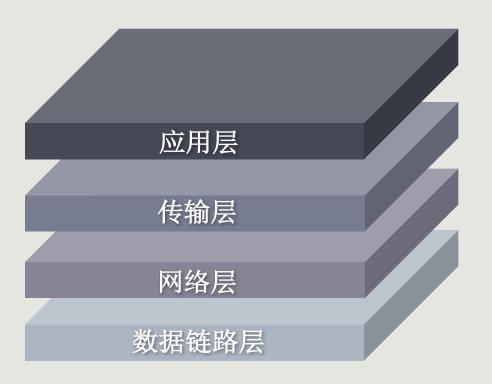
时间序列分析

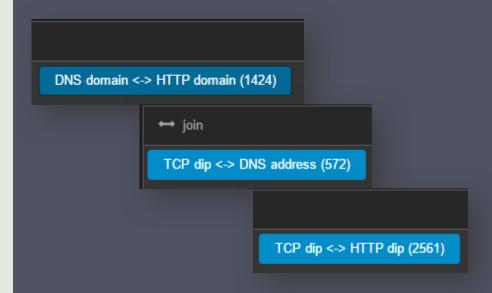


地理位置信息分析



关联查询 siren-join





将不同层次 联系在一起

