

# 使用Elastic Stack快速搭建SIEM系统



Elastic首席架构师 吴斌

# 什么是SIEM?

- 由Gartner的Mark Nicolett和Amrit Williams在2005年提出了Security Information & Event Management (SIEM) 的概念
- Describes the product capabilities of gathering, analyzing and presenting information from network and security devices; identity and access management applications; vulnerability management and policy compliance tools; operating system, database and application logs; and external threat data.

# 核心目标

- 识别
  - 潜在的威胁
  - 已经造成的破坏
- 审计
  - 收集安全以及合规日志
- 分析
  - 调查
  - 关联性



# 安全事件 workflow



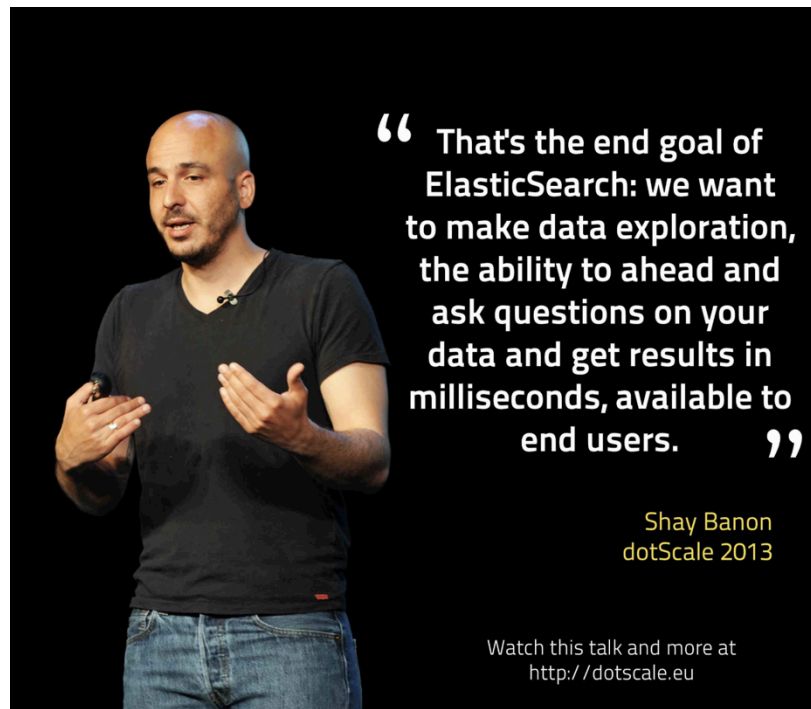
NIST Cybersecurity Framework

<https://www.nist.gov/cyberframework>



# 我们致力于帮助客户解决所有实时数据处理问题

- **100,000+** 社区参与者
- **350M+** 下载 (持续增长中)
- 政府, 金融保险, 电信, 高科技
- **1000+** 员工
- **5,000+** 订阅客户



“ That's the end goal of ElasticSearch: we want to make data exploration, the ability to ahead and ask questions on your data and get results in milliseconds, available to end users. ”

Shay Banon  
dotScale 2013

Watch this talk and more at  
<http://dotscale.eu>



# Elastic 产品选择



# Elastic Stack

100% 开源  
无商业/企业版本  
全新6.x版本



Kibana



Elasticsearch

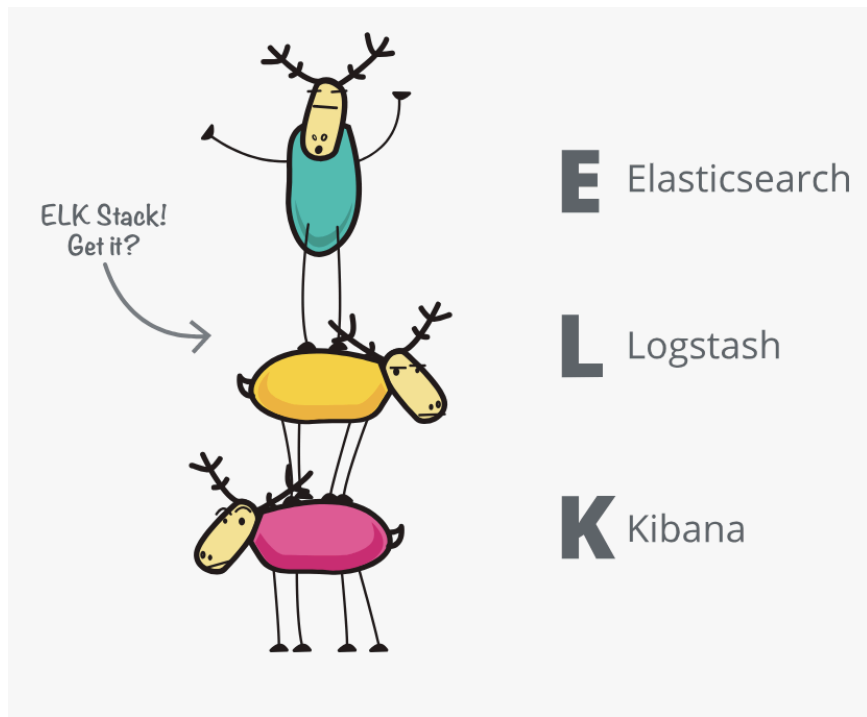


Beats



Logstash

## ELK to ELKB (Elastic Stack)





# 订阅服务

技术支持服务  
一次安装  
Elastic Stack的商业扩展  
订阅式软件许可证



数据安全



告警



监控



报表

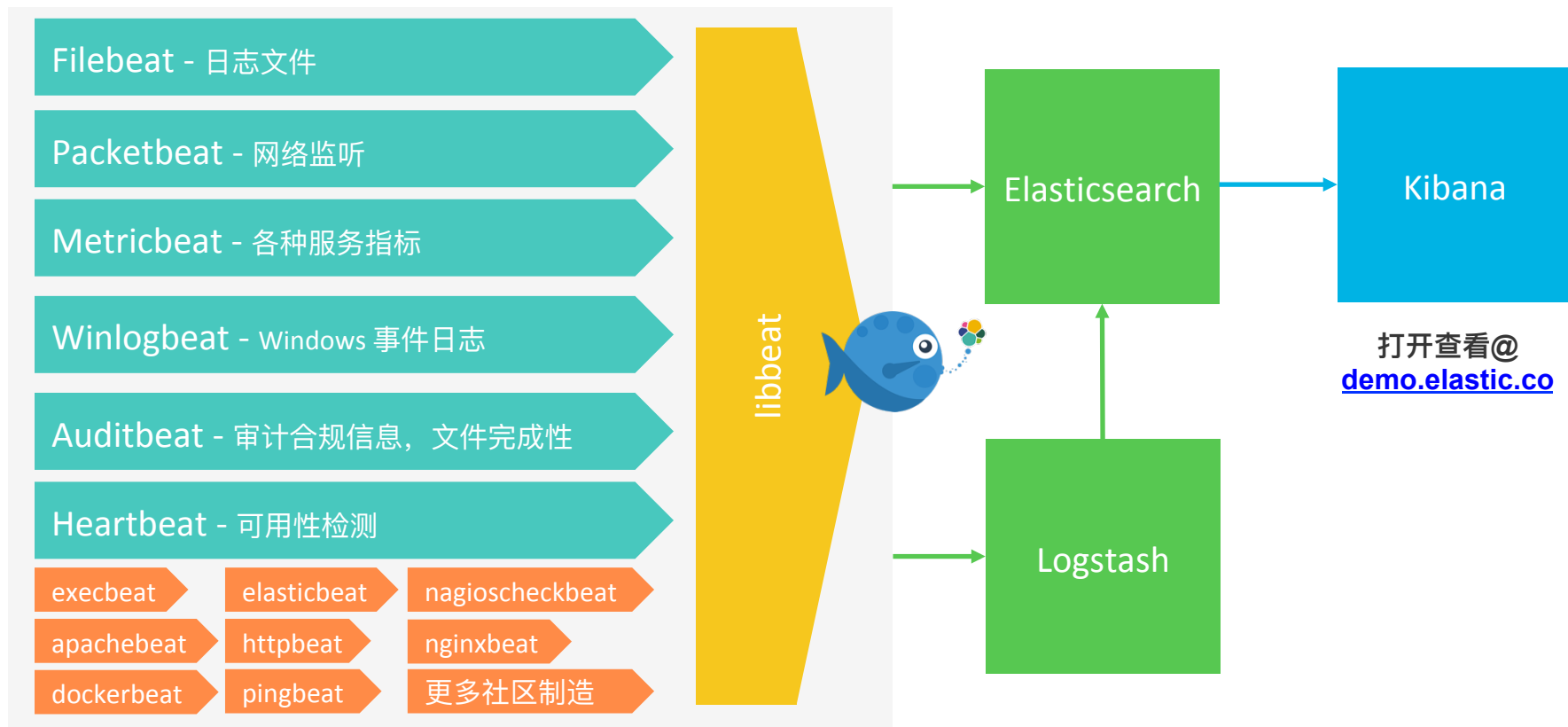


图查询



机器学习

# Beats: 轻量级数据搬运工





# Beats 预置模块

简化收集、解析和可视化常见日志格式

System

Apache

Kafka

Couchbase

Redis

HAProxy

Elasticsearch

Windows

NGINX

Zoo  
keeper

CEPH

Docker

Golang

Postgres

Dropwizard

Aerospike

Prometheus

MySQL

Memcache

Jolokia

PHP-FPM

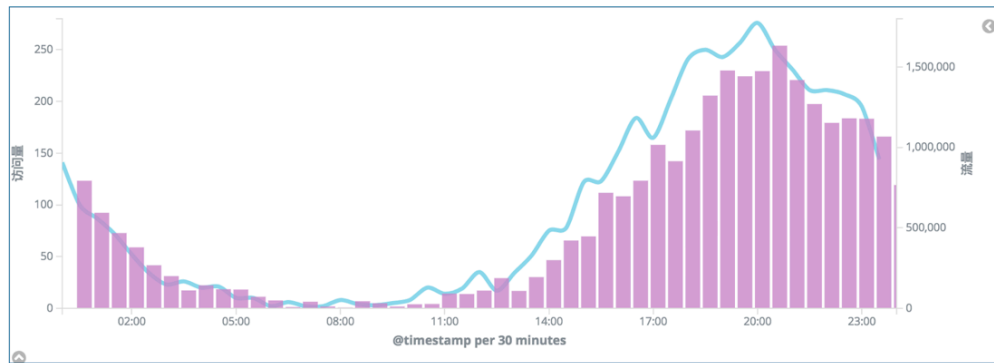
Kibana

HTTP

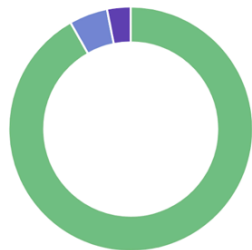
Auditd

# Beats预置面板

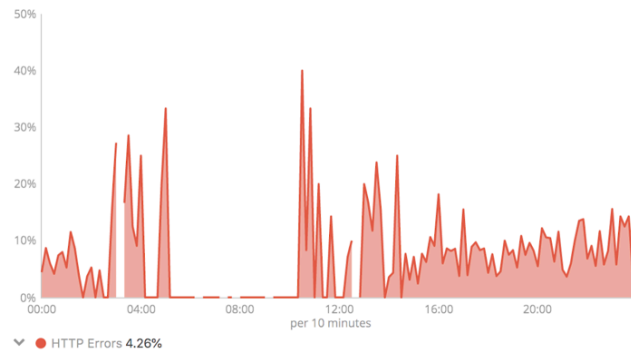
bw-logs-流量-访问量



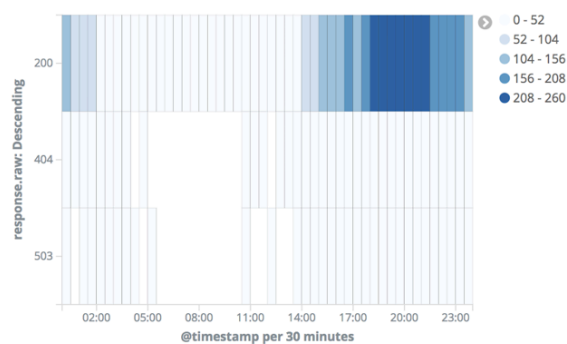
bw-logs-返回值-饼图



bw-logs-http-errors



bw-logs-返回值



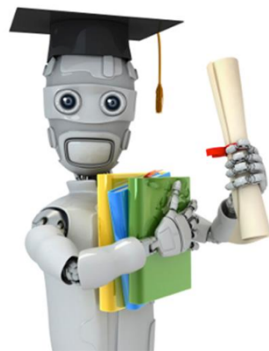


# X-Pack

机器学习

## 时序型数据的非监督、增量学习

- 异常检测自动化
- 加速归因分析



# Elastic

## *X-Pack*

# Elastic Stack满足SIEM场景一些基本诉求

安全分析

日志分析

指标分析

商业分析

搜索



保护你的数据



数据告警



异常检测



监测管理  
Elastic Stack

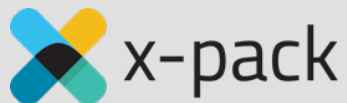


找到数据关联



共享你的报表

# 保护你的 Elastic Stack



安全



告警



机器学习



监控



图查询



报表



## 限制数据访问

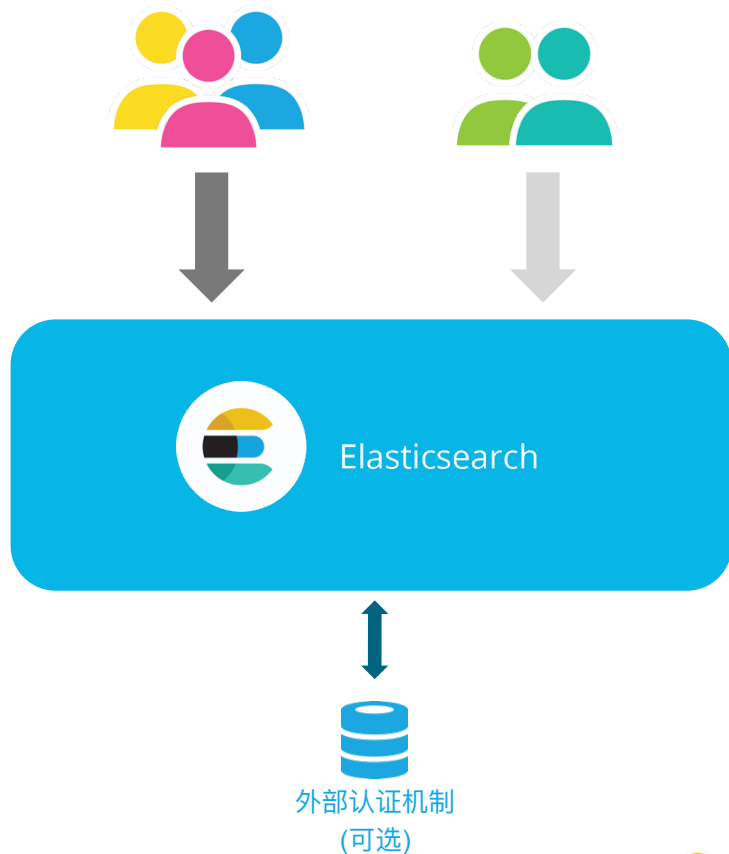
### Elastic 认证:

- 原生认证 (用户名/密码)
- LDAP, Active Directory
- 定制 (like Kerberos, SSO)
- SAML 6.2

### Elastic 授权:

- 索引, 文档, 字段
- APIs

### IP 过滤规则



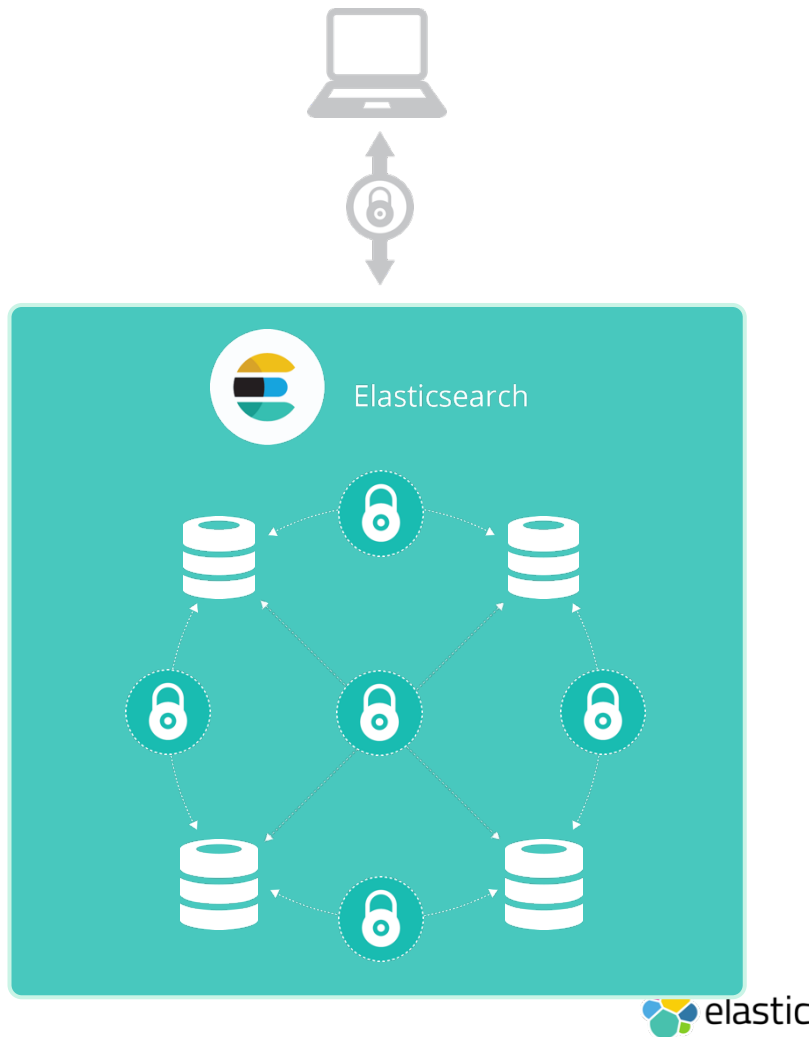
## 数据加密

### 通讯:

- 和Elasticsearch集群的交互
- Elasticsearch节点间通讯

### 磁盘数据:

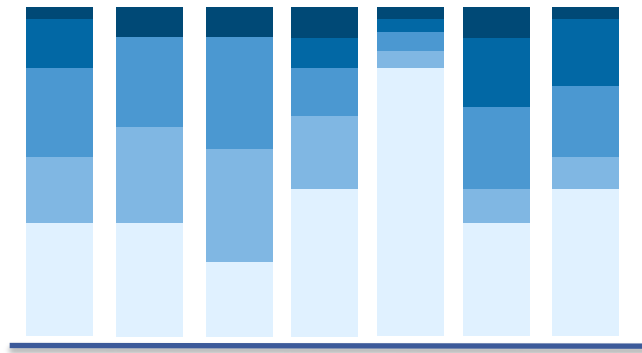
- 支持 *dm-crypt* 加密



## 审计/合规

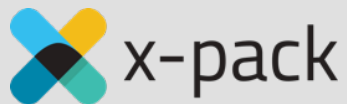
### 使用安全审计功能:

- 洞悉正常和非常规行为
- 对非常规行为进行告警
- 构建真对访问控制的面板



Access vs. time by users

# 对异常数据/行为进行告警



安全



告警



机器学习



监控



图查询



报表

## IT 运营分析

### 使用规则

在下列情况发生时通知:

- 磁盘剩余空间不足 5%
- Elasticsearch 集群健康状态红色预警

### 使用异常检测

在下列情况发生时通知:

- 突发的高错误率
- 未知的高CPU使用率



# 安全分析

## 使用规则

在下列情况发生时通知:

- 过去5分钟内单个机器上超过5次登录失败
- 进程 X 在任意服务器上启动

## 使用异常检测

在下列情况发生时通知:

- 在非常规的时间/地点登陆
- 向系统外可疑的数据传输

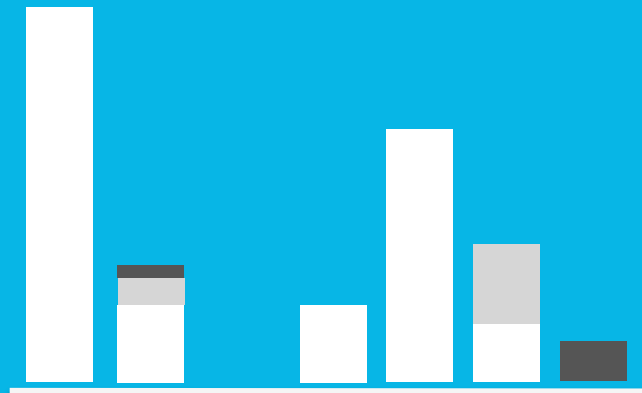




## 追踪记录所有告警

所有告警记录可追溯:

- SLAs是否达标?
- 安全事件的趋势?
- 哪台服务器故障问题最多?
- 哪些事件与异常数据存在相关性?



内存空间告警

-- 服务器 1 -- 服务器 2 -- 服务器 3

# 机器学习

Image Classification **Recommendations**

Autonomous cars Voice Recognition **Predictive Medicine**

*Fraud detection*

**异常检测**

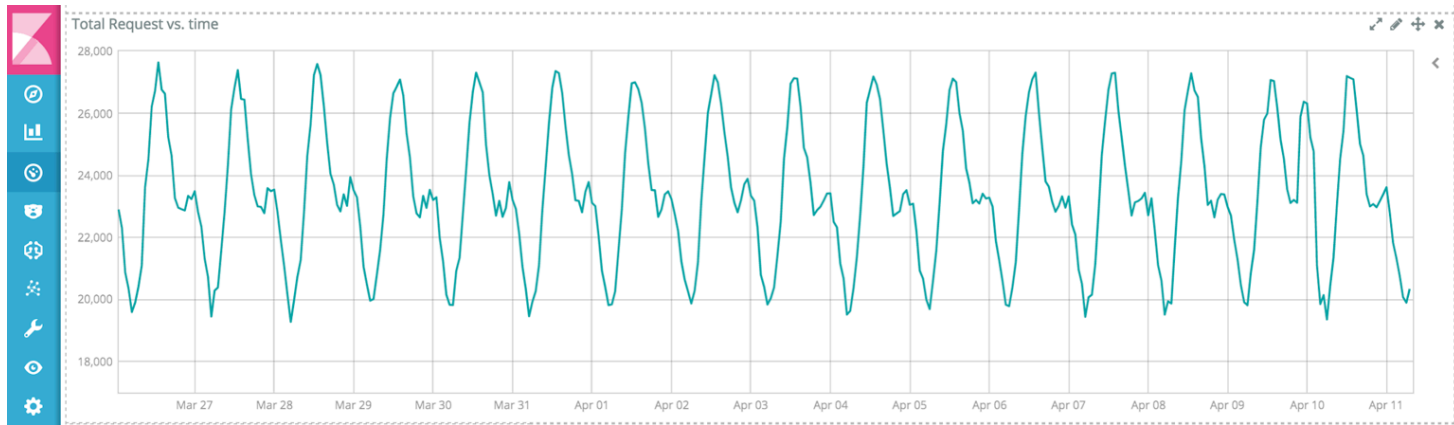
*Learn to Rank* Speech Recognition

*Language Translation* **Entity Resolution**

# 异常检测是规则引擎很好的补充

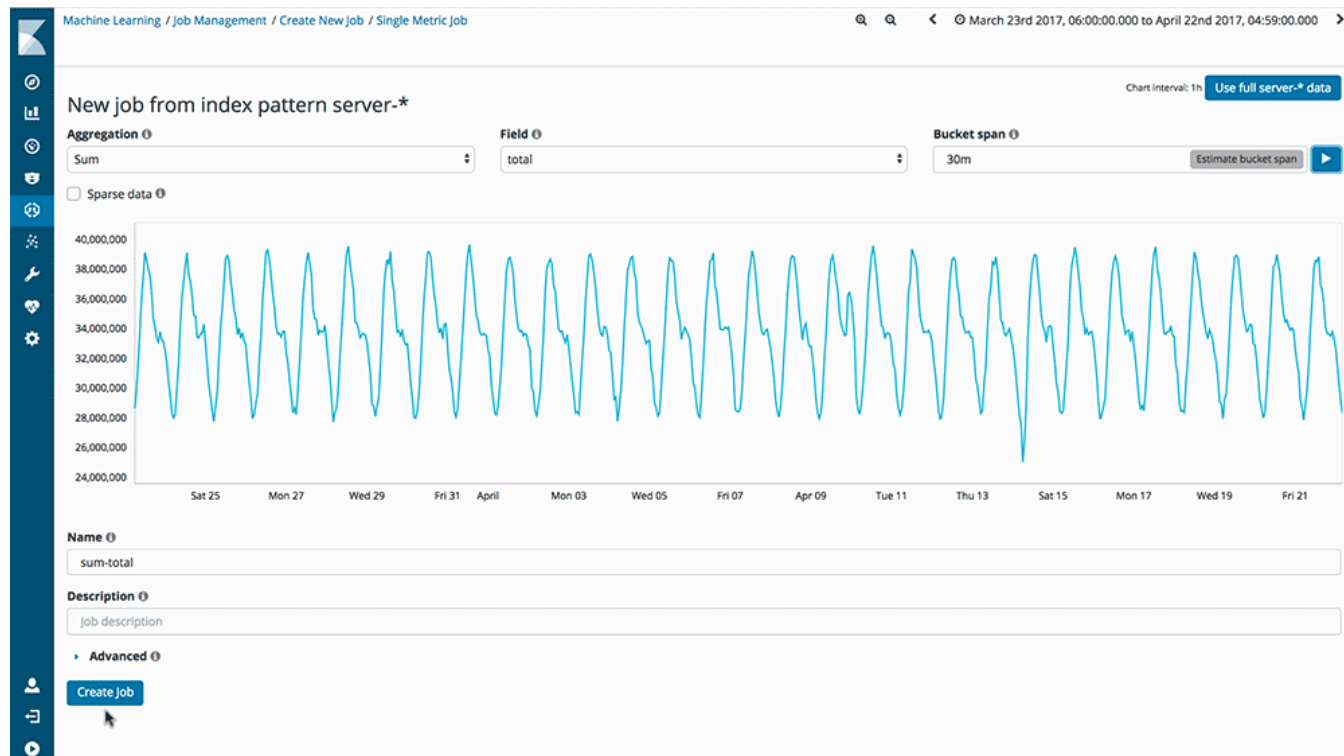
- 规则并不是定义正常/异常最好的方法
- 定义再好的规则，也不会随着数据改变而持续演化、改进

如何设置阈值？

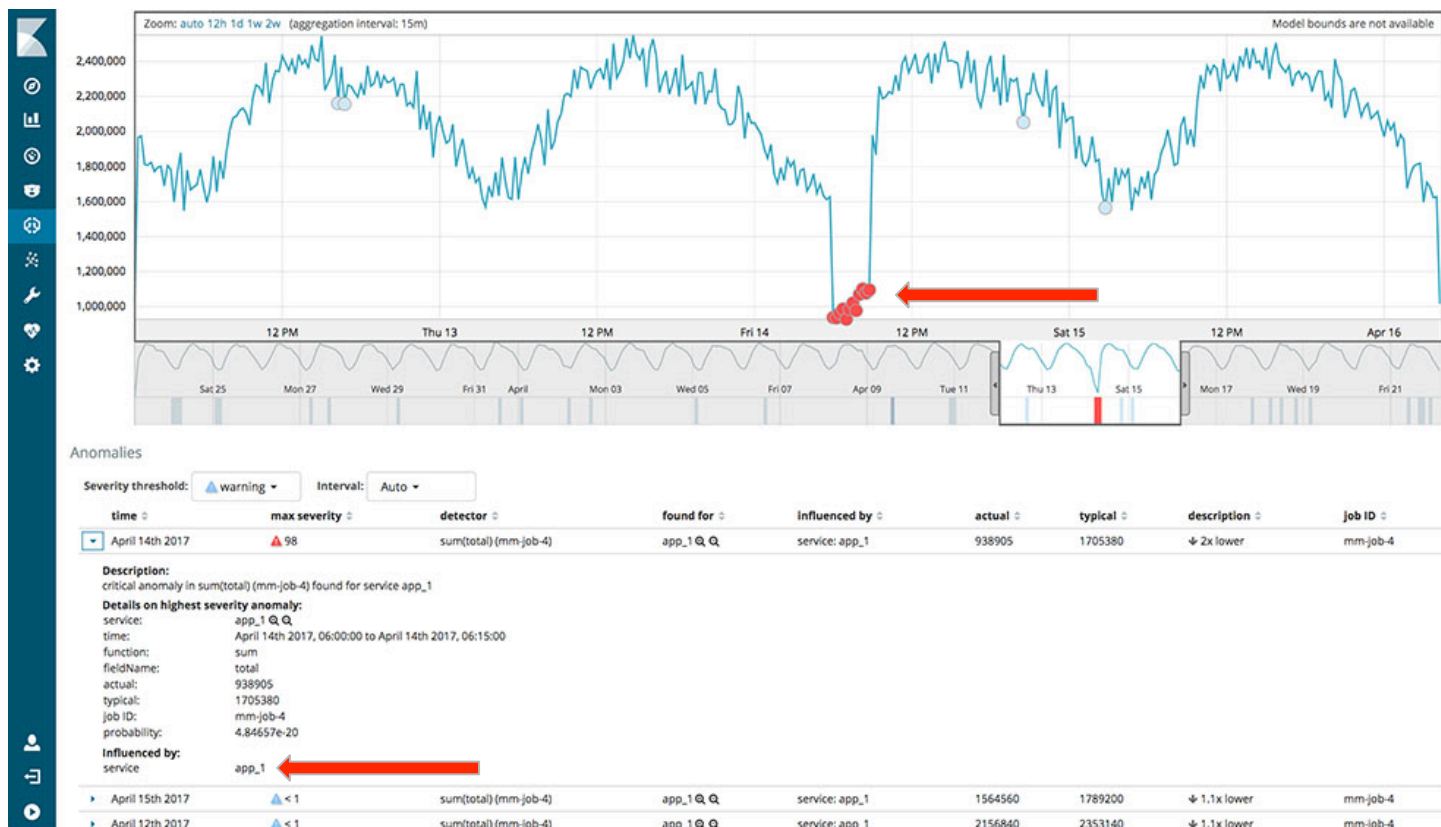


# 非监督机器学习

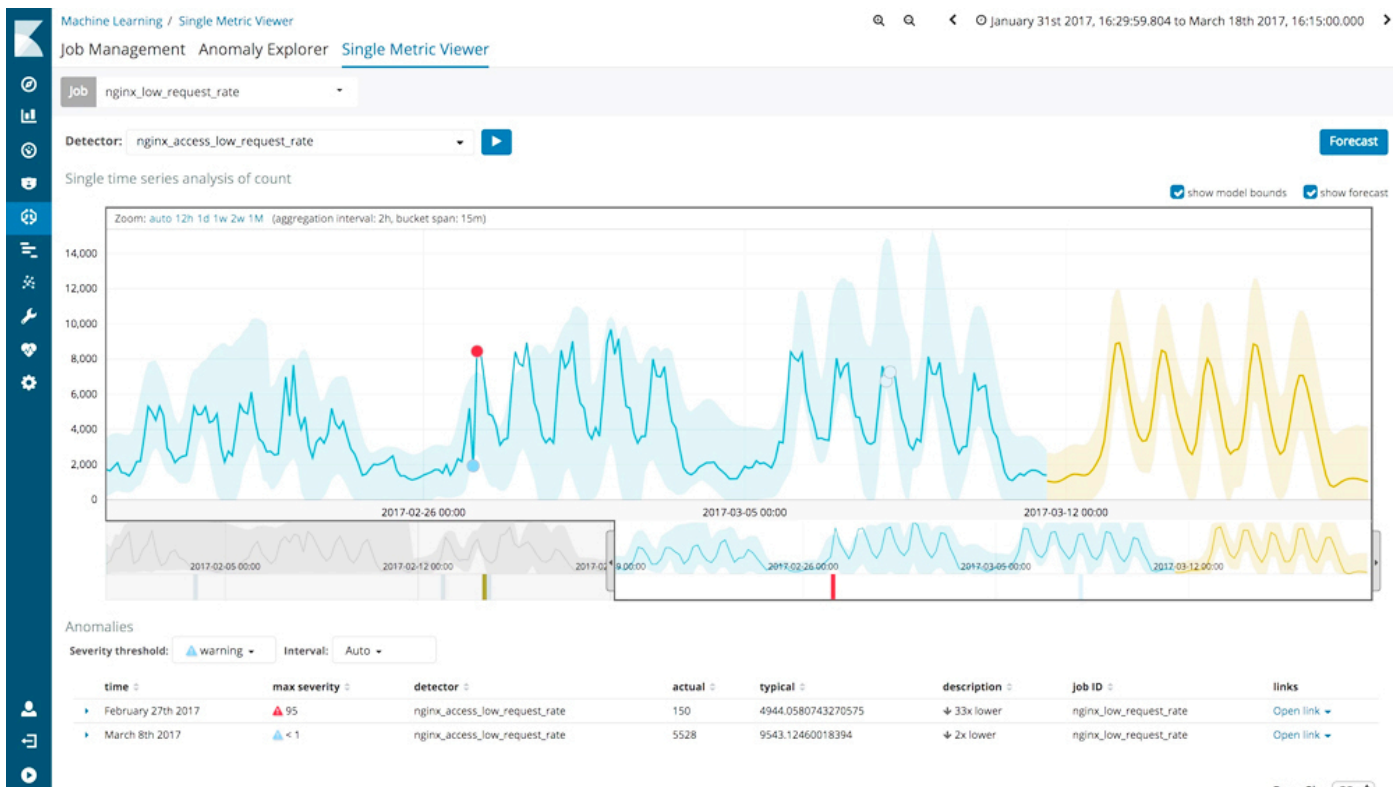
全自动察觉数据正常范围，随数据变化而持续改进、更新



# 加速归因分析，锁定影响因子

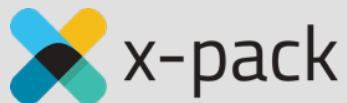


# 预测、洞悉未来





# 监控 Elastic Stack



安全



告警



机器学习



监控



图查询



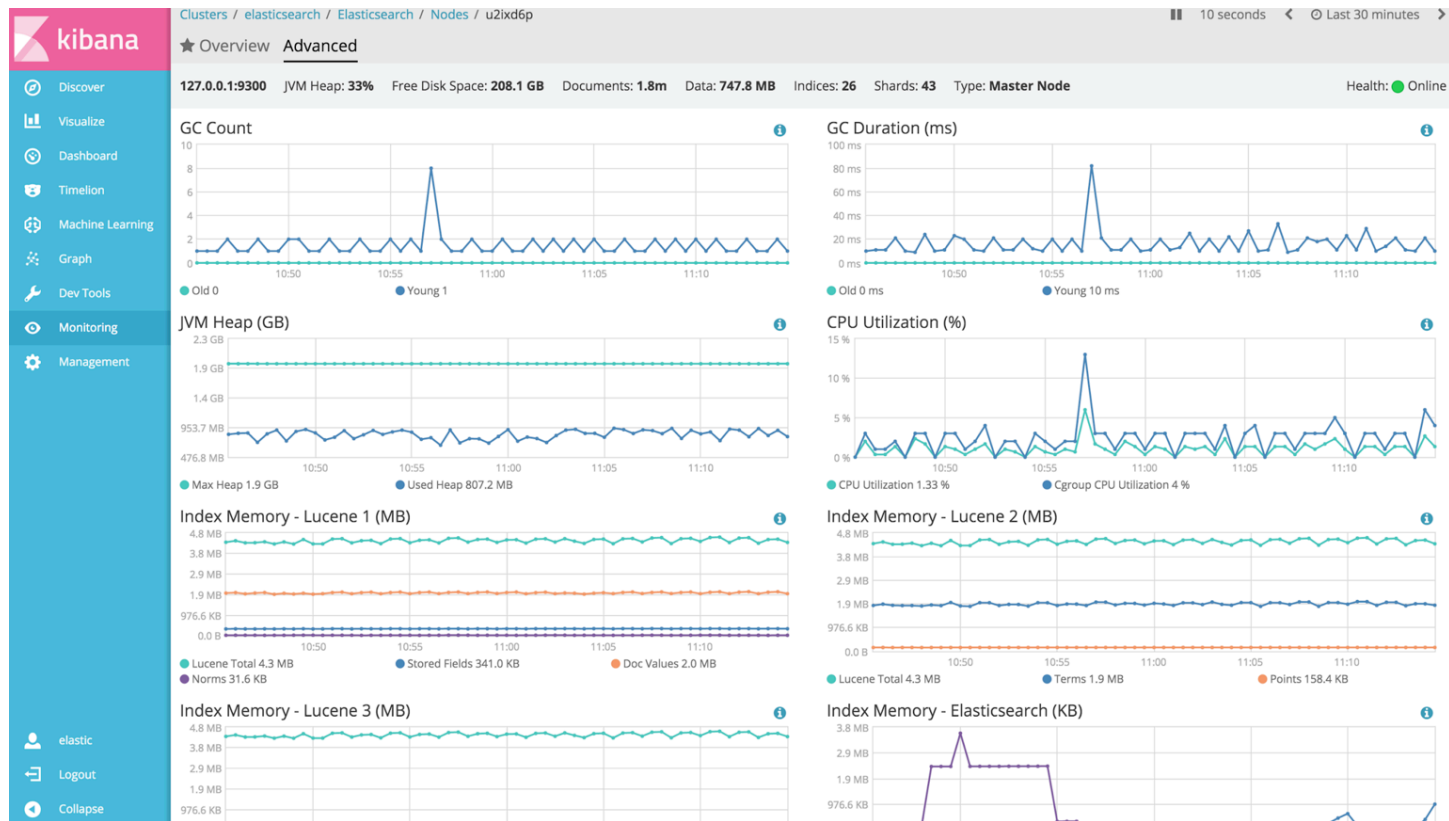
报表

# 实时监控

- 集群整体的工作状态、压力
- 是否有压力暴增的单个节点？
- 在灾难发生前洞察问题并着手解决



# 统一的监控入口，掌控集群整体状态



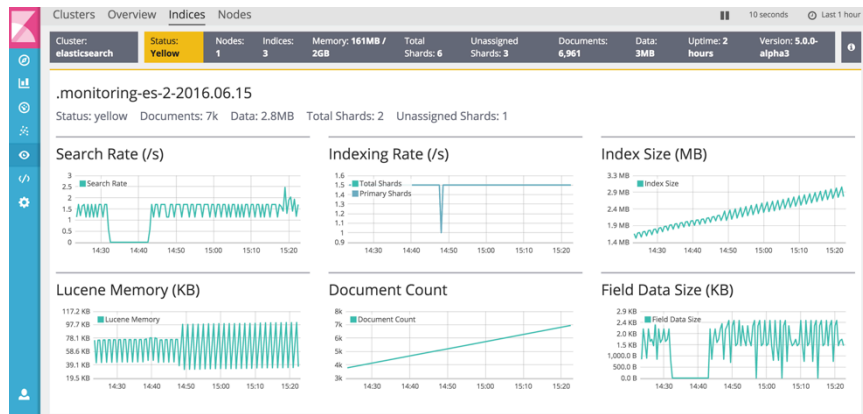
# 监控数据也会被索引到 Elasticsearch

数据可以留作日后对集群状态的分析 and 调优结果评估

Table

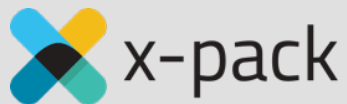
JSON

```
1 {
2   "_index": ".marvel-es-1-2016.06.06",
3   "_type": "node_stats",
4   "_id": "AVUnC_T4qKi2Snrly9N",
5   "_score": null,
6   "_source": {
7     "cluster_uuid": "ReVr5CXFTi2fucBt2LbLNA",
8     "timestamp": "2016-06-06T18:49:02.968Z",
9     "source_node": {
10      "uuid": "0xWM08tpSn-ijbx-kY17VA",
11      "host": "10.137.126.111",
12      "transport_address": "10.137.126.111:19168",
13      "ip": "10.137.126.111",
14      "name": "tiebreaker-0000000060",
15      "attributes": {
16        "logical_availability_zone": "tiebreaker",
17        "availability_zone": "us-east-1b",
18        "data": "false",
19        "max_local_storage_nodes": "1",
20        "region": "us-east-1",
21        "master": "true"
22      }
23    },
24    "node_stats": {
25      "node_id": "0xWM08tpSn-ijbx-kY17VA",
26      "node_master": false,
27      "mlockall": false,
28      "disk_threshold_enabled": false,
29      "disk_threshold_watermark_high": 90,
30      "indices": {
```



JVM 堆内存使用过高!

# 找到数据之间的关联



安全



告警



机器学习



监控



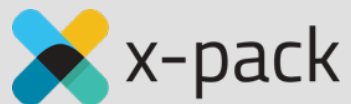
图查询



报表



# 洞悉数据的共享



安全



告警



机器学习



监控



图查询



报表

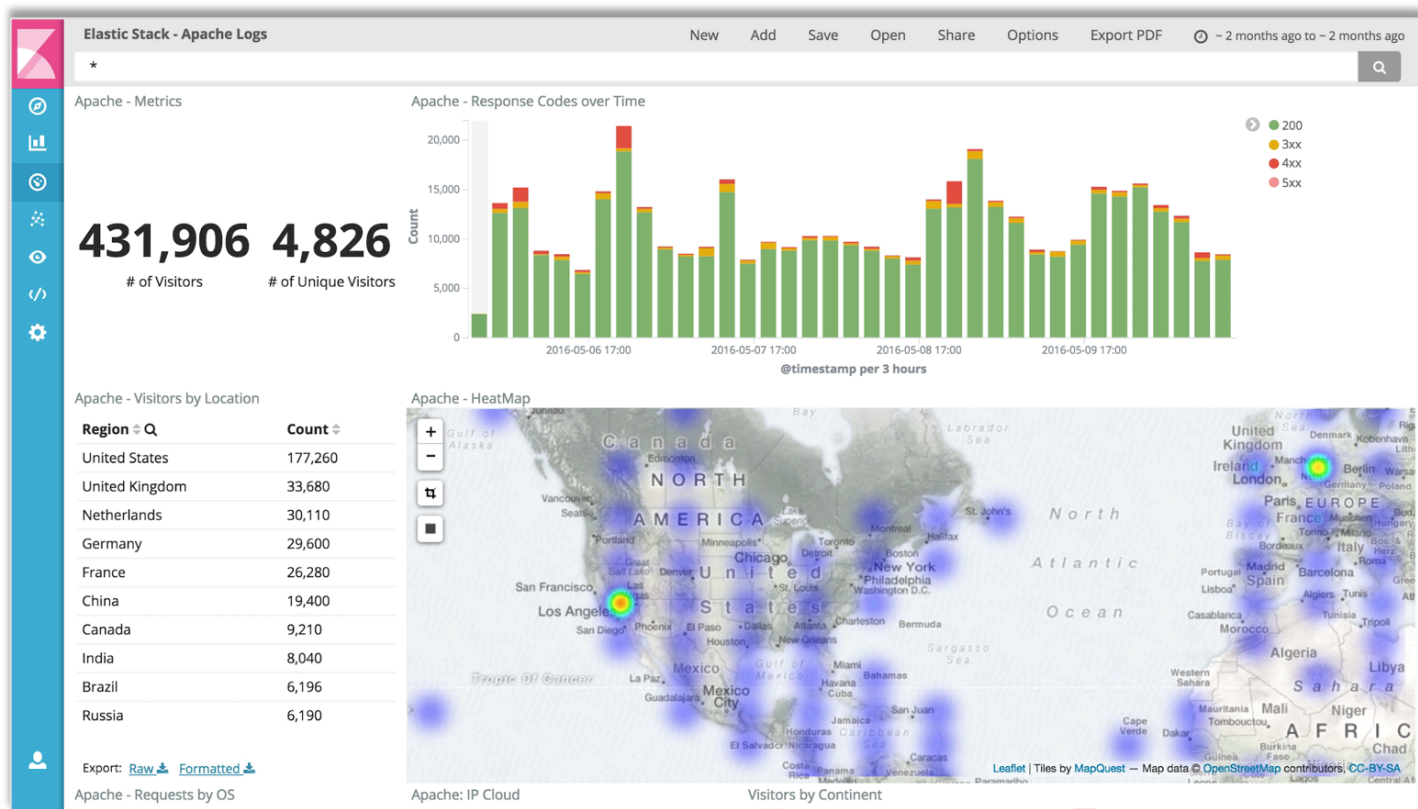
## 按需或者按计划生成报表

- 把面板导出为 PDF 文档
- 给 Elastic Stack 用户共享
- 提供离线数据访问
- 定时发送报表到指定邮箱
  - 节省重复工作时间

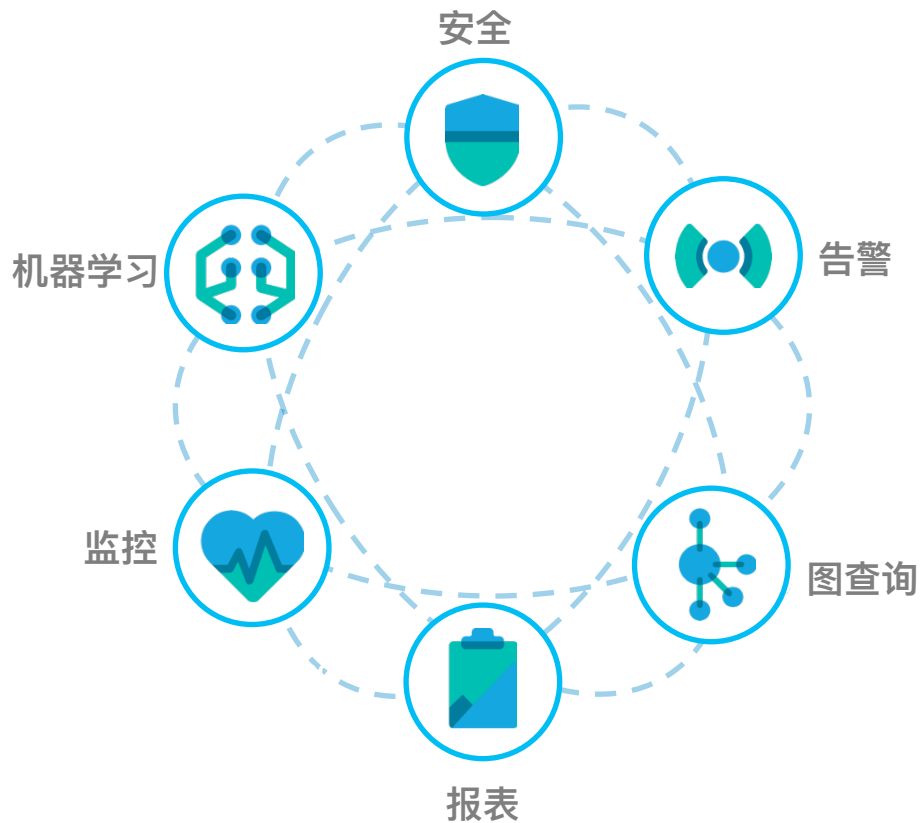




# 生成报表来共享面板数据



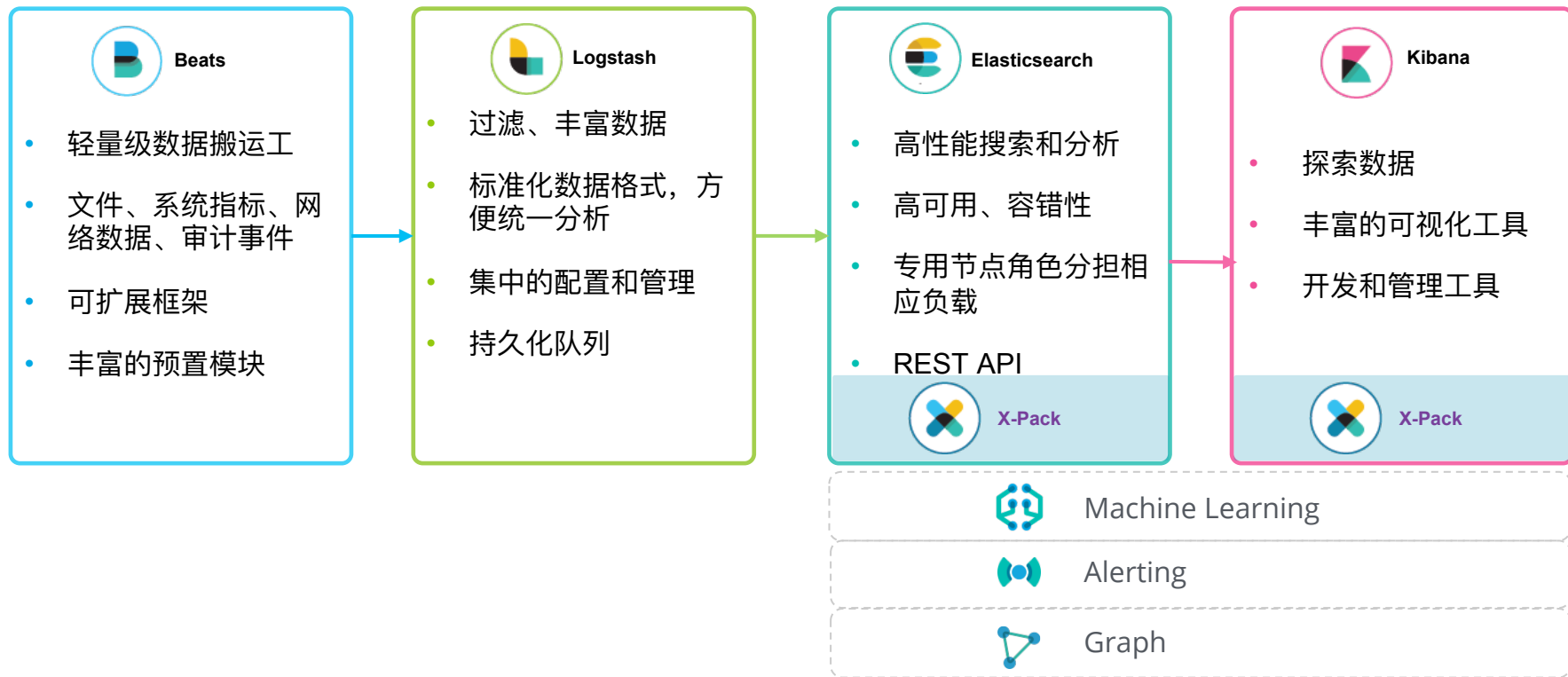
## 在SIEM应用中的价值



# 基于Elastic Stack的SIEM

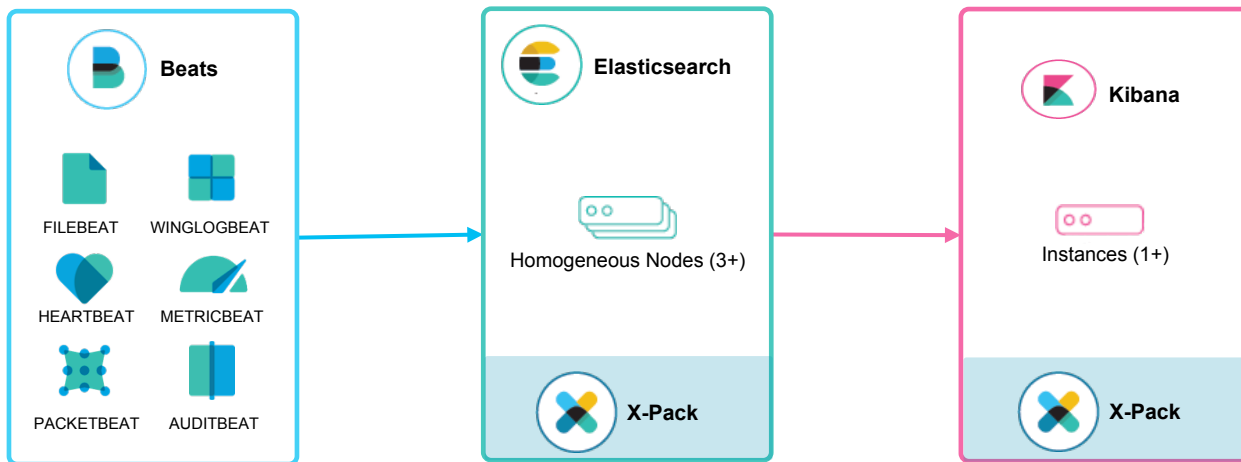
部署快速上手

# 数据处理管道逻辑图



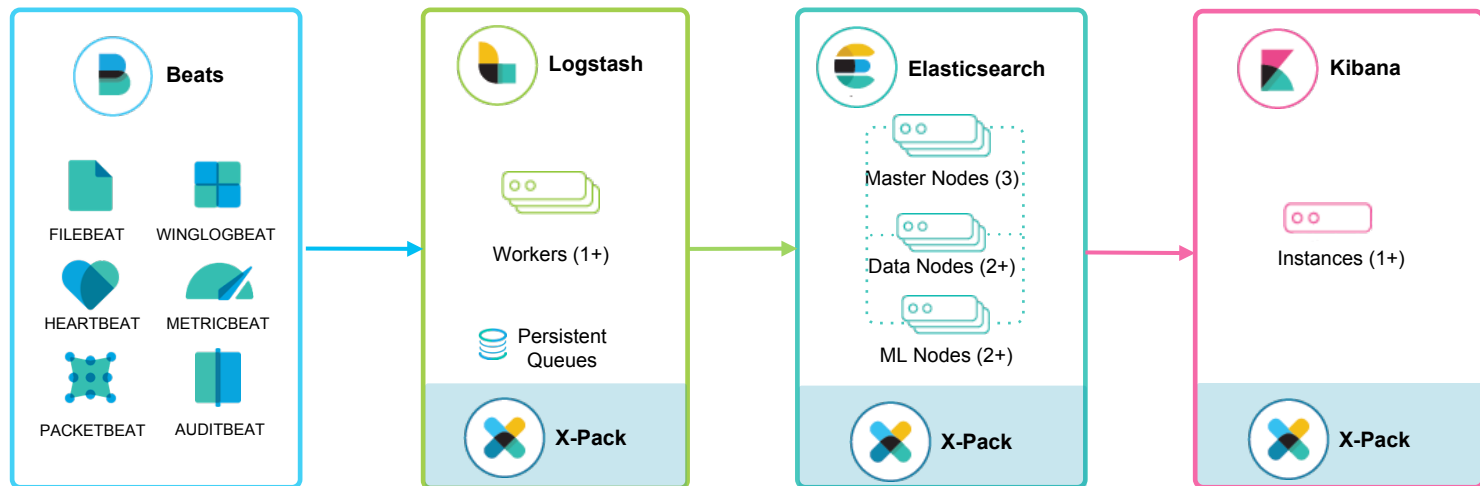
# 快速开始

## Beats, Elasticsearch and Kibana

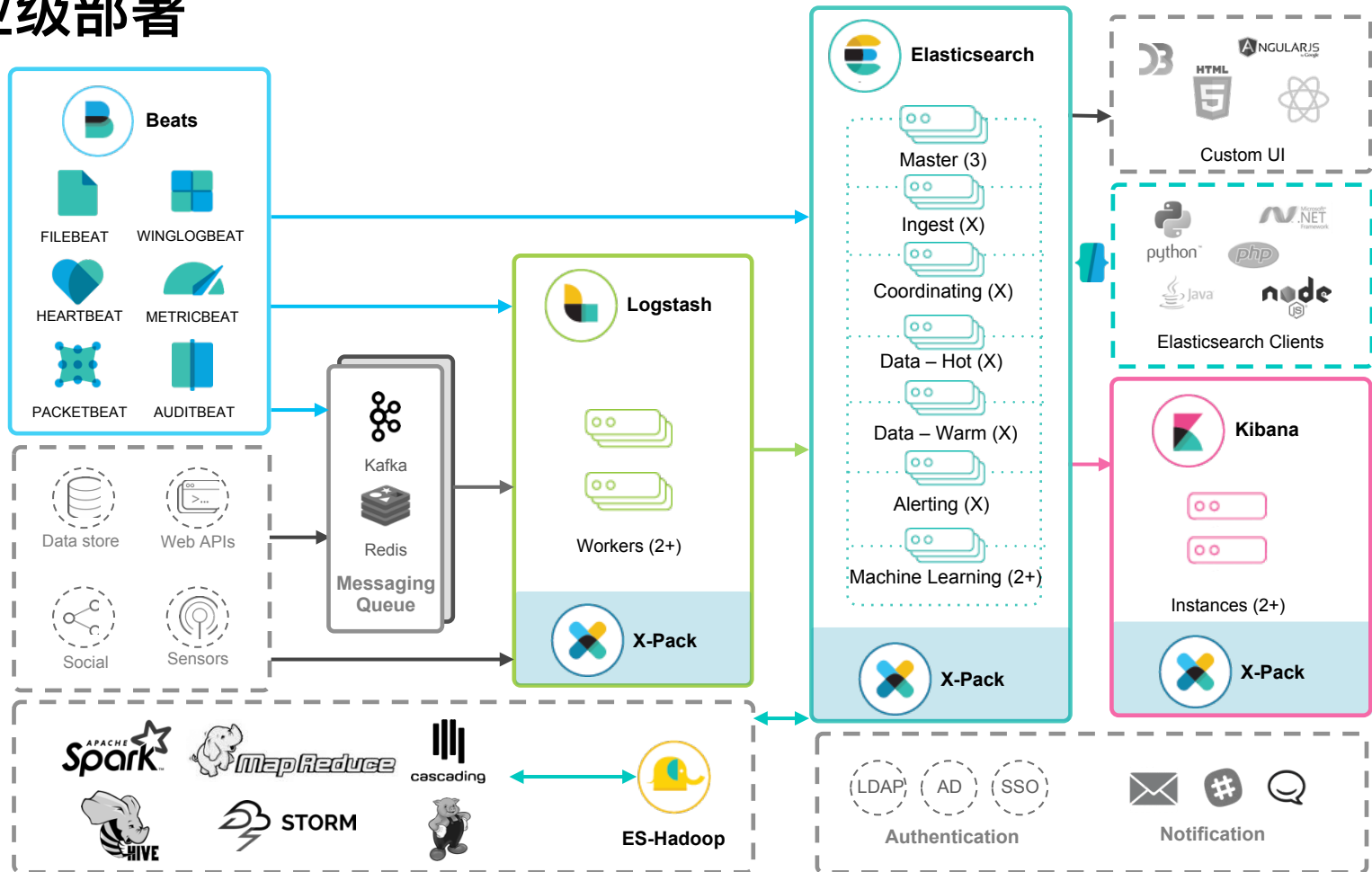


# 更加完善和容错更好的数据管道

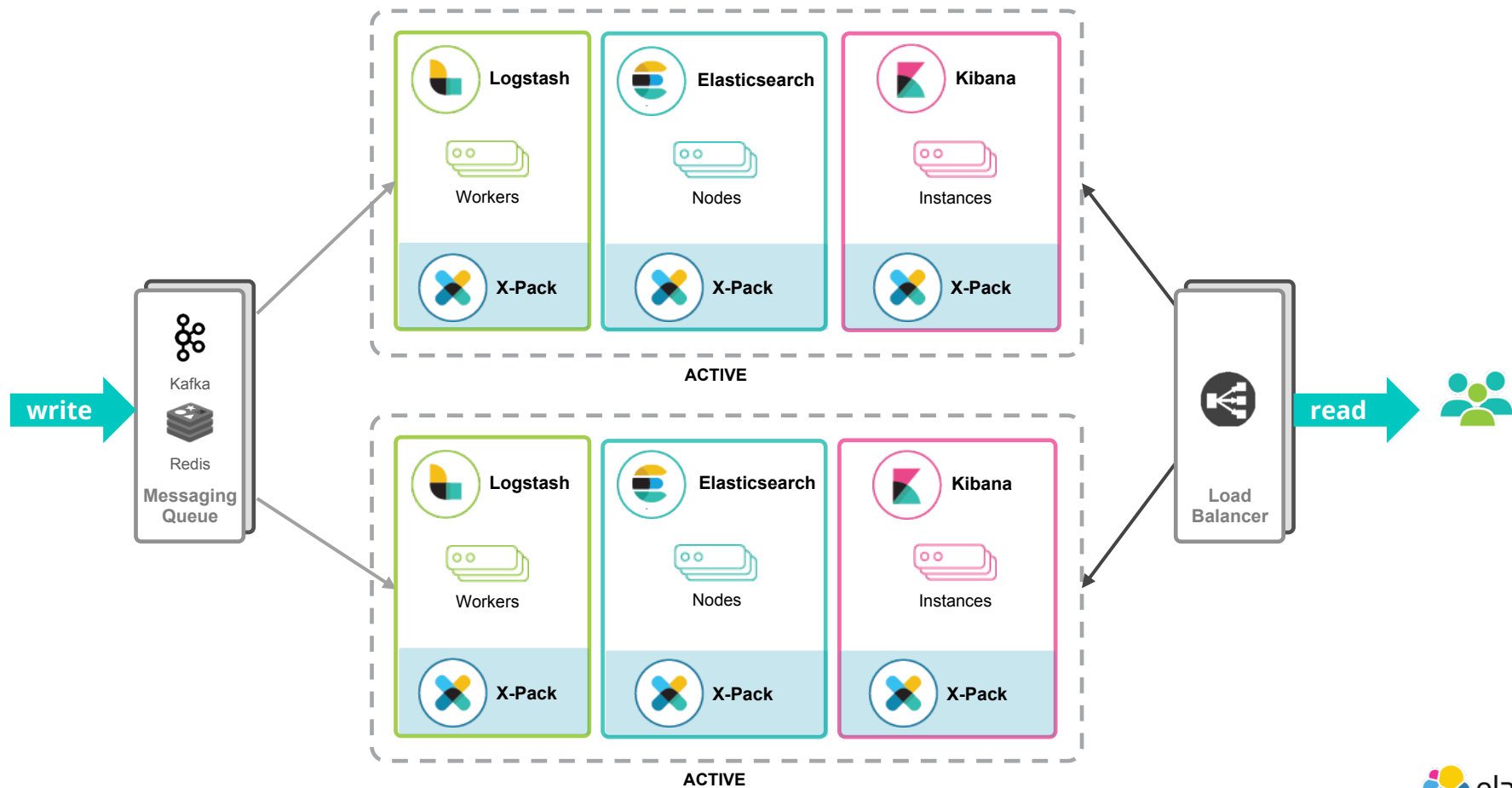
添加了Logstash和按照角色分配Elasticsearch的任务



# 企业级部署

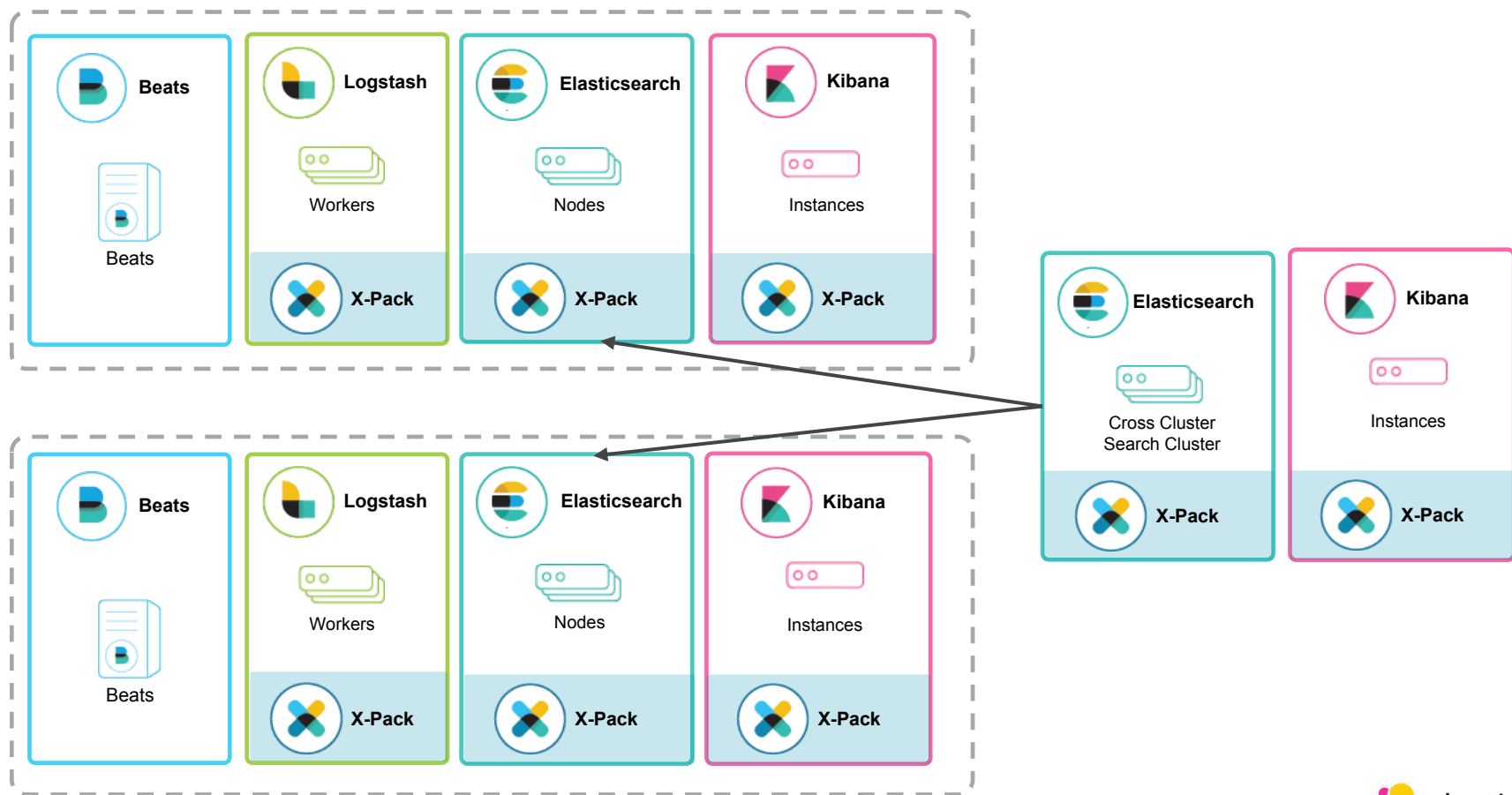


# 跨数据中心冗余数据





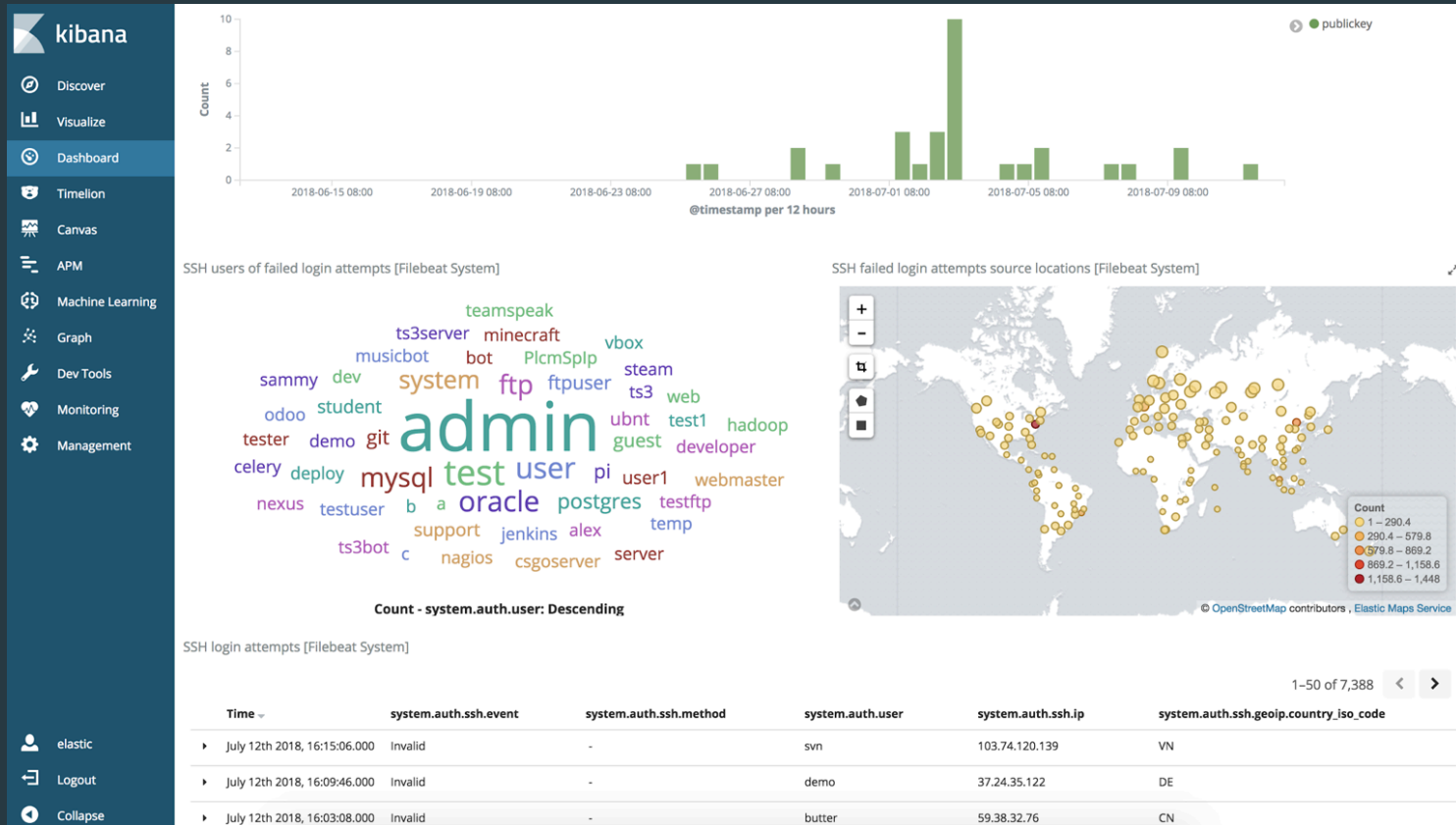
# 多数据中心非冗余数据跨区查询



# 基于Elastic Stack的SIEM

业务场景构建

# SSH 登陆情况面板



# SSH 登陆场景

## 业务实体

- 用户名（谁）
- 时间
- IP（地点）

## 登陆成功

- 异常时间
- 异常地点

如何确定异常？

## 登陆失败

- 暴力登陆
  - 多少次算暴力登陆呢？

# SSH 登陆信息探索

```
62-   "filter": {
63-     "range": {
64-       "@timestamp": {
65-         "gte": "2018-07-01T00:00:00"
66-       }
67-     }
68-   }
69- }
70- }
71- }
```

GET filebeat-\*/\_search

```
74- {
75-   "size": 0,
76-   "aggs": {
77-     "NAME": {
78-       "cardinality": {
79-         "field": "system.auth.ssh.event"
80-       }
81-     }
82-   }
83- }
```

GET filebeat-\*/\_search

```
86- {
87-   "size": 0,
88-   "aggs": {
89-     "NAME": {
90-       "terms": {
91-         "field": "system.auth.ssh.event",
92-         "size": 10
93-       }
94-     }
95-   }
96- }
```

```
1- {
2-   "took": 3,
3-   "timed_out": false,
4-   "_shards": {
5-     "total": 1,
6-     "successful": 1,
7-     "skipped": 0,
8-     "failed": 0
9-   },
10-   "hits": {
11-     "total": 93549,
12-     "max_score": 0,
13-     "hits": []
14-   },
15-   "aggregations": {
16-     "NAME": {
17-       "value": 3
18-     }
19-   }
20- }
```

```

70- }
71- }
72-
73- GET filebeat-*/_search
74- {
75-   "size": 0,
76-   "aggs": {
77-     "NAME": {
78-       "cardinality": {
79-         "field": "system.auth.ssh.event"
80-       }
81-     }
82-   }
83- }

```

```

84-
85- GET filebeat-*/_search
86- {
87-   "size": 0,
88-   "aggs": {
89-     "NAME": {
90-       "terms": {
91-         "field": "system.auth.ssh.event",
92-         "size": 10
93-       }
94-     }
95-   }
96- }

```

```

97-
98- GET /_settings/index.version.created
99-
100- GET _xpack/license
101-
102- GET server-metrics/_search
103-
104- PUT _ingest/pipeline/plus_one_year
105- {
106-   "processors": [
107-     {
108-       "script": {
109-         "lang": "painless",
110-         "source": ""

```

```

1- {
2-   "took": 8,
3-   "timed_out": false,
4-   "_shards": {
5-     "total": 1,
6-     "successful": 1,
7-     "skipped": 0,
8-     "failed": 0
9-   },
10-   "hits": {
11-     "total": 93549,
12-     "max_score": 0,
13-     "hits": []
14-   },
15-   "aggregations": {
16-     "NAME": {
17-       "doc_count_error_upper_bound": 0,
18-       "sum_other_doc_count": 0,
19-       "buckets": [
20-         {
21-           "key": "Invalid",
22-           "doc_count": 10625
23-         },
24-         {
25-           "key": "error:",
26-           "doc_count": 45
27-         },
28-         {
29-           "key": "Accepted",
30-           "doc_count": 38
31-         }
32-       ]
33-     }
34-   }
35- }

```

Search... (e.g. status:200 AND extension:PHP)

system.auth.ssh.event: "exists"

Add a filter +

filebeat-\*

Selected Fields

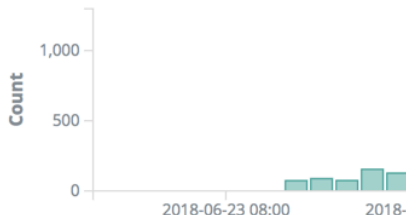
? \_source

Available Fields

@timestamp

\_id

\_index



```
"@timestamp": "2018-07-20T08:39:07.000Z",
"system": {
  "auth": {
    "hostname": "ip-172-31-13-50",
    "ssh": {
      "geoip": {
        "continent_name": "Asia",
        "city_name": "Shenzhen",
        "country_iso_code": "CN",
        "region_name": "Guangdong",
        "location": {
          "lon": 114.1333,
          "lat": 22.5333
        }
      },
      "ip": "36.36.201.21",
      "event": "Invalid"
    },
    "pid": "25062",
    "user": "nexus",
    "timestamp": "Jul 20 08:39:07"
  }
},
"beat": {
  "hostname": "ip-172-31-13-50",
  "name": "ip-172-31-13-50",
  "version": "6.3.0"
},
"host": {
  "name": "ip-172-31-13-50"
}
```

# SSH暴力破解异常检测



# 选择学习数据

Job Details

Analysis Configuration

Datafeed

Edit JSON



Datafeed job ⓘ

**Query** ⓘ

```
"terms":{"system.auth.ssh.event":["Invalid"],"boost":1}}
```

# 指定观测指标

## Add new detector

Description ⓘ

high\_count("system.auth.ssh.ip")

function ⓘ

high\_count × ▾

field\_name ⓘ

system.auth.ssh.ip × ▾

by\_field\_name ⓘ

Select... ▾

over\_field\_name ⓘ

Select... ▾

partition\_field\_name ⓘ

Select... ▾

exclude\_frequent ⓘ

Select... ▾

[Help for high\\_count](#)

## Influencers

- ☐ system.auth.ssh.geoip.country\_iso\_code
- ☐ system.auth.ssh.geoip.region\_name
- ☒ system.auth.ssh.ip
- ☐ system.auth.ssh.method
- ☐ system.auth.ssh.signature
- ☐ system.auth.sudo.command
- ☐ system.auth.sudo.error
- ☐ system.auth.sudo.pwd
- ☐ system.auth.sudo.tty
- ☐ system.auth.sudo.user
- ☐ system.auth.timestamp
- ☒ system.auth.user
- ☐ system.auth.useradd.home

## Watch JSON ( Syntax )

```
196     }
197   }
198   }
199   }
200   }
201   }
202   }
203   }
204 },
205 "condition": {
206   "compare": {
207     "ctx.payload.aggregations.bucket_1
208     "gt": 0
209   }
210 },
211 },
212 "actions": {
213   "log": {
214     "logging": {
215       "level": "info",
216       "text": "Alert for job [{{ctx.payload.
217     }
218   }
219 }
220 }
```

Edit

Simulate

Simulation Results

Simulation Status: ✓ OK

ActionType	Mode	State	Reason
log	logging	simulate	✓ OK

Simulation Output:

```
1 {
2   "watch_id": "_inlined_",
3   "node": "0CZaRj4AR_Gzuex-sDE5Xw",
4   "state": "execution_not_needed",
5   "status": {
6     "state": {
7       "active": true,
8       "timestamp": "2018-07-20T15:57:50.132Z"
9     },
10    "last_checked": "2018-07-20T15:57:50.142Z",
11    "actions": {
12      "log": {
13        "ack": {
14          "timestamp": "2018-07-20T15:57:50.132Z",
15          "state": "awaits_successful_execution"
16        }
17      }
18    },
19    "execution_state": "execution_not_needed",
20    "version": -1
21  },
22  "trigger_event": {
23    "type": "manual",
24    "triggered_time": "2018-07-20T15:57:50.139Z",
25    "manual": {
26      "text": "Alert for job [{{ctx.payload.
27    }
28  }
29 }
```

rt\_ssh\_brute\_force

10,662okopenedstarted2018-07-20 23:28:14

Job settings

Job config

Datafeed

Counts

JSON

Job messages

Datafeed preview

Forecasts

General

job\_id

rt\_ssh\_brute\_force

job\_type

anomaly\_detector

job\_version

6.3.0

groups

realtime

description

create\_time

2018-07-04 10:49:15

established\_model\_memory

636.9 KB

model\_snapshot\_retention\_days

1

model\_snapshot\_id

1532073064

results\_index\_name

shared

state

opened

open\_time

448534s

Actions

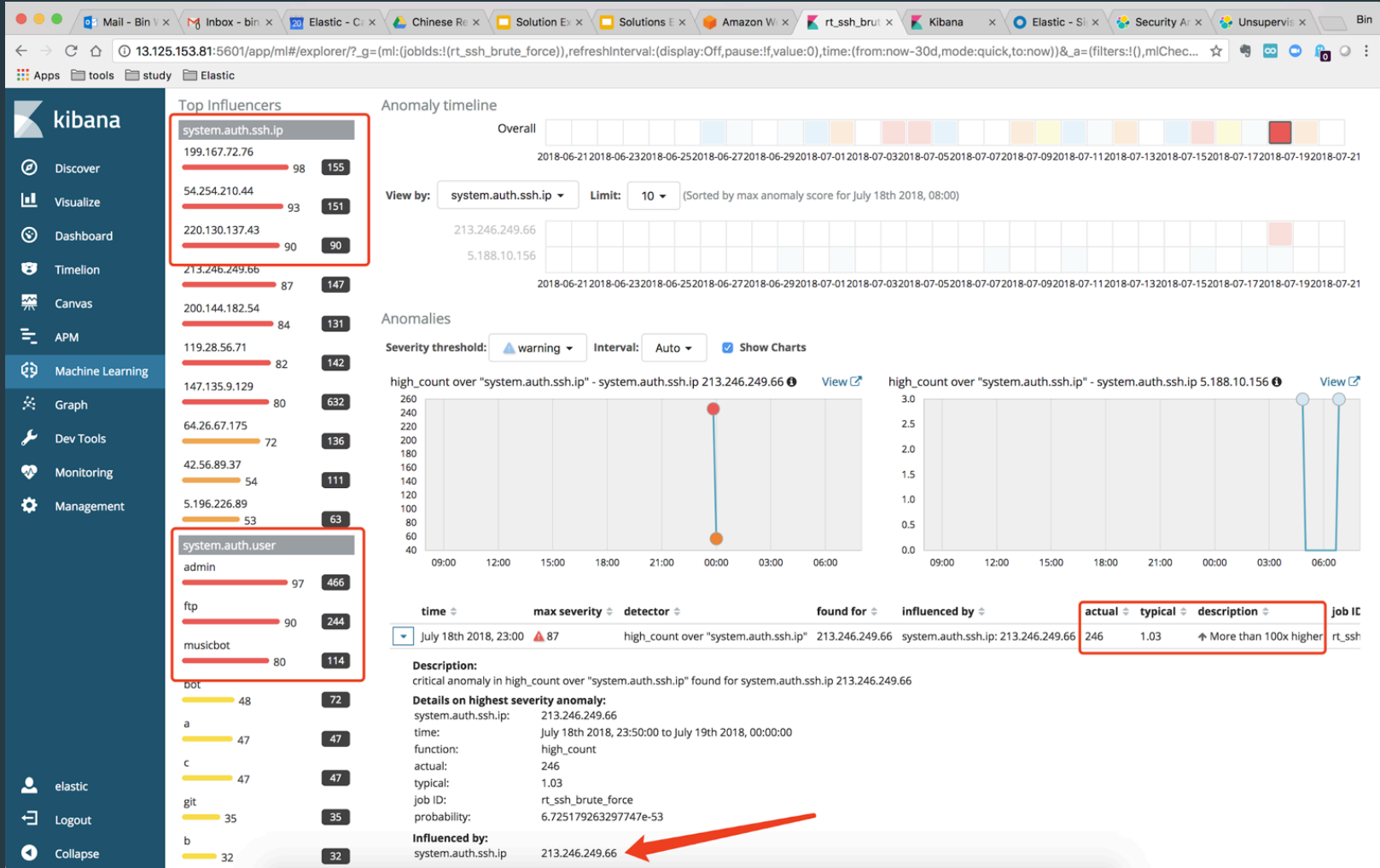
Close Job

Validate Job ?

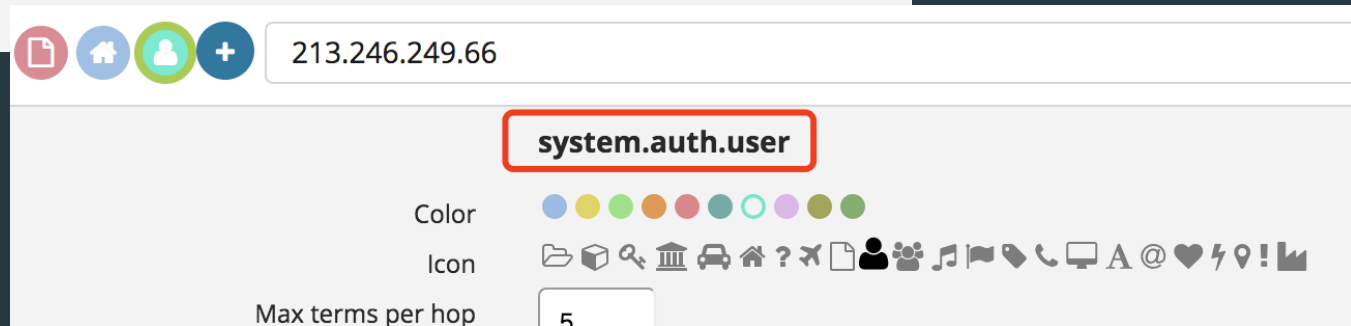
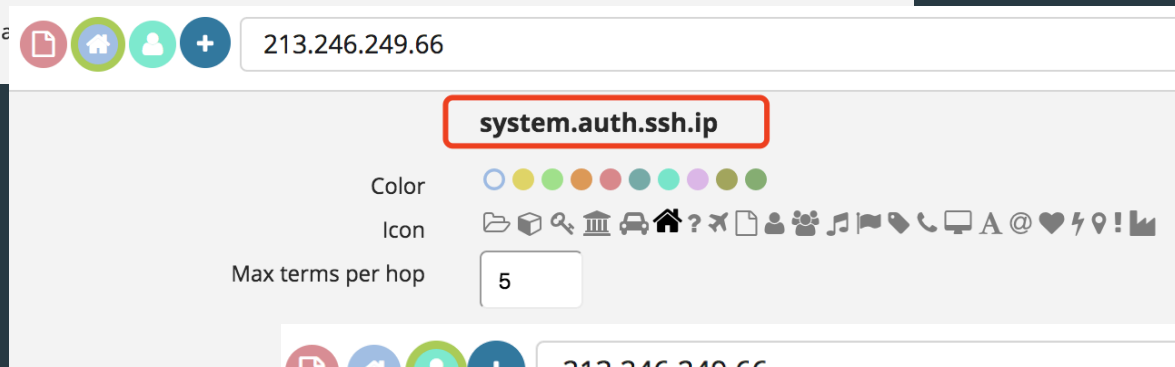
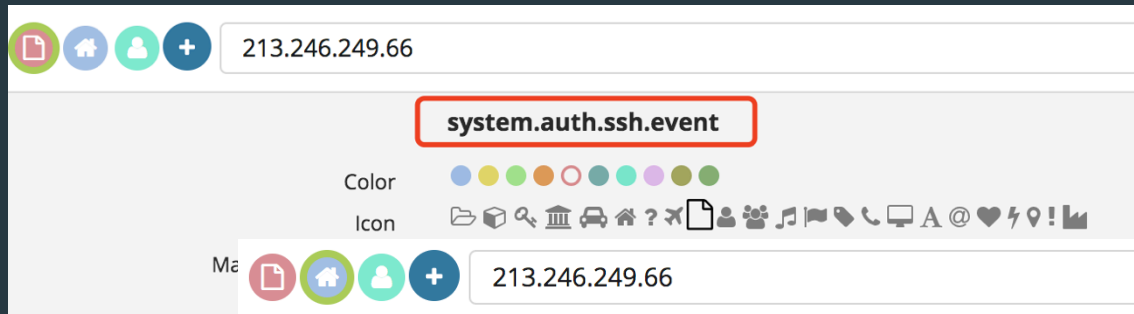
Node

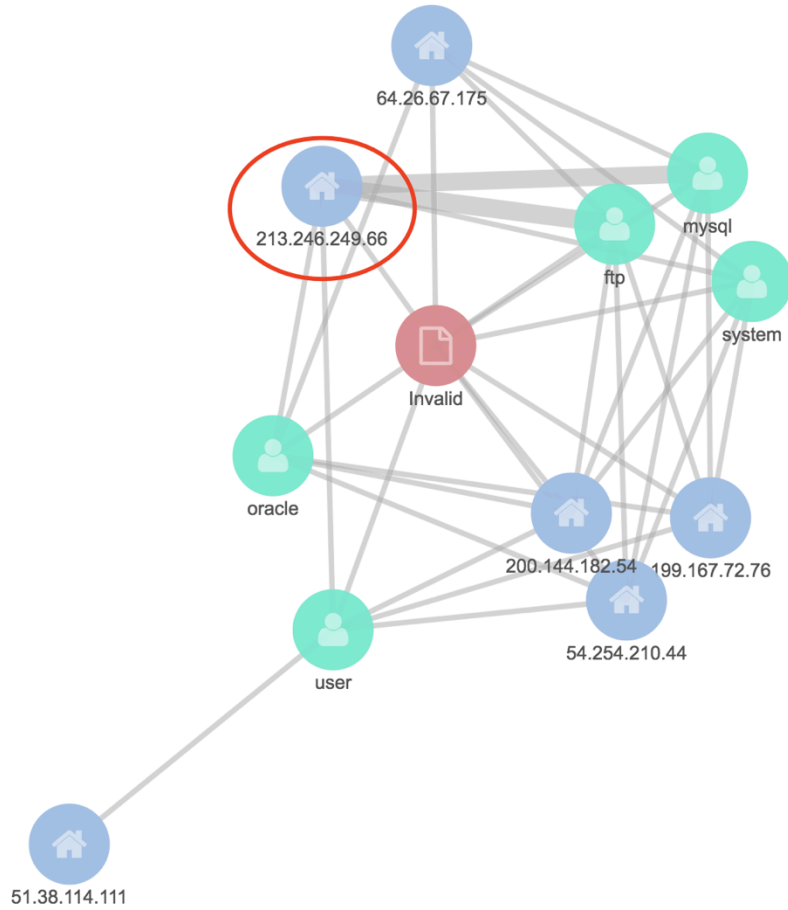
name

tiger

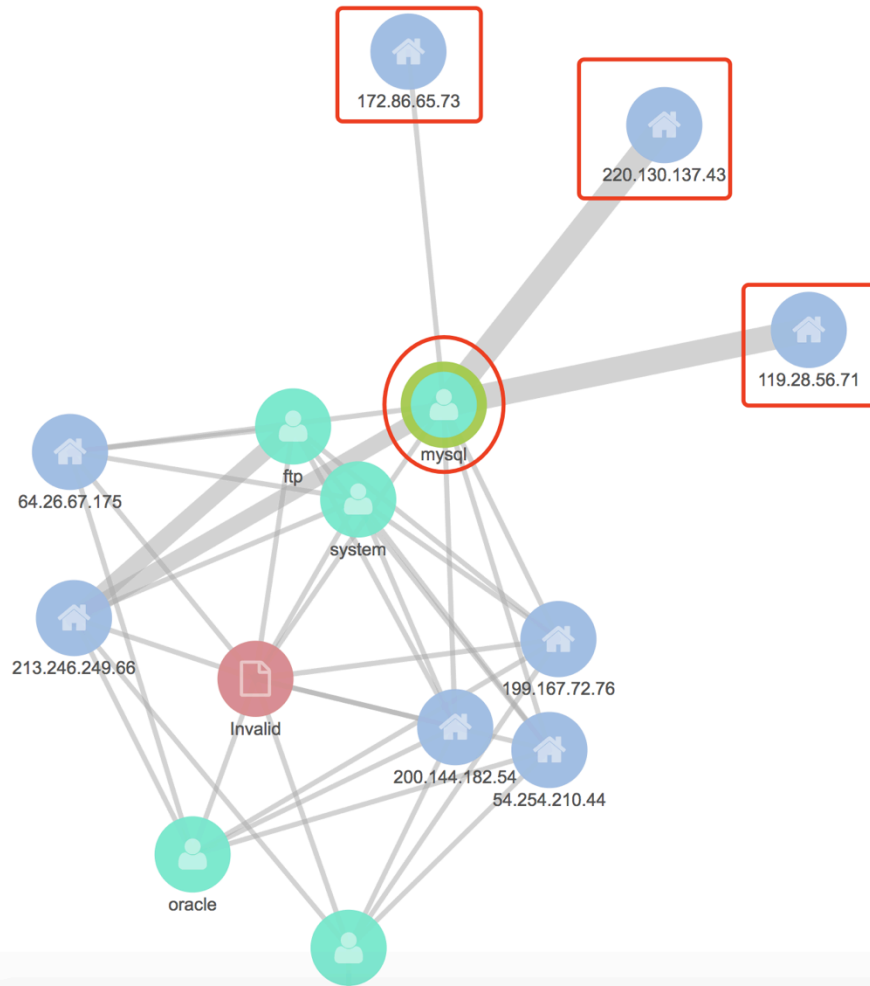


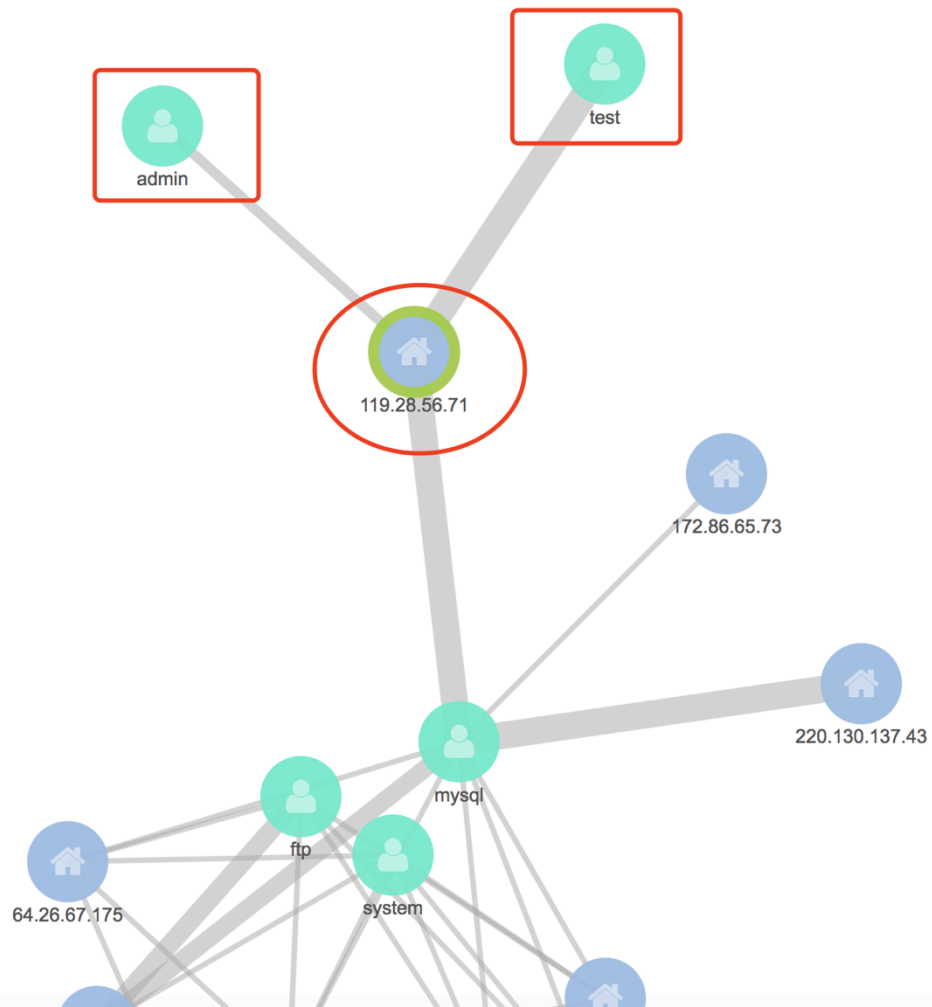
## 图关联查询


























# 异常检测任务组

▶ rt_port_scan	3,642,361	ok	opened	started	2018-07-21 00:09:50	  
▶ rt_ssh_brute_force	10,665	ok	opened	started	2018-07-21 00:00:10	  
▶ rt_unusual_login_location	36	ok	opened	started	2018-07-19 08:37:18	  
▶ rt_unusual_login_time	36	ok	opened	started	2018-07-19 08:37:18	  
▶ rt_unusual_process	7,495,861	ok	opened	started	2018-07-21 00:04:59	  

## Detectors

high\_distinct\_count("dest.port")

## Influencers

dest.ip

## Detectors

rare by "system.auth.ssh.geoip.city\_name"

## Influencers

system.auth.ssh.ip, system.auth.user, system.auth.ssh.geoip.country\_iso\_code,  
system.auth.ssh.geoip.continent\_name, system.auth.ssh.geoip.city\_name

## Detectors

time\_of\_day by "system.auth.user"

## Influencers

system.auth.ssh.ip, system.auth.user

## Detectors

rare by "system.process.name"

## Influencers

system.process.cmdline, system.process.cwd, system.process.username,  
system.process.state

Auth Logs

Audit Events

ArcSight

DNS Traffic

NetFlow

1

## Download & Unpack packages

Close ^

Elasticsearch

Kibana

Packetbeat

2

## Get started

Details v

3

## Open Kibana. Play.

Details v

<https://www.elastic.co/products/x-pack/machine-learning/recipes>

IT Operations

### Service Response Change (Response Code)

Analyze response code metrics to detect service issues

[Learn More](#)

IT Operations

### System Metric Change (CPU Utilization)

Analyze CPU metrics to detect system problems

[Learn More](#)

Security Analytics

### DNS Data Exfiltration (Tunneling)

Analyze DNS logs to detect DNS Tunneling

[Learn More](#)

Security Analytics

### Suspicious Process Activity (Host)

Analyze endpoint proxy logs to detect rare processes

[Learn More](#)

Security Analytics

### HTTP Data Exfiltration (Proxy)

Analyze web proxy logs to detect HTTP exfiltration

[Learn More](#)

Security Analytics


### Suspicious Login Activity (Volume)

Analyze server logs to detect brute force login attacks

[Learn More](#)



2018 杭州·云栖大会  
THE COMPUTING CONFERENCE

 **Alibaba** Group  
阿里巴巴集团

# 驱动数字中国

EMPOWER DIGITAL CHINA



elastic  
中文社区

专业、垂直、纯粹的 Elastic 开源技术交流社区

<https://elasticsearch.cn/>