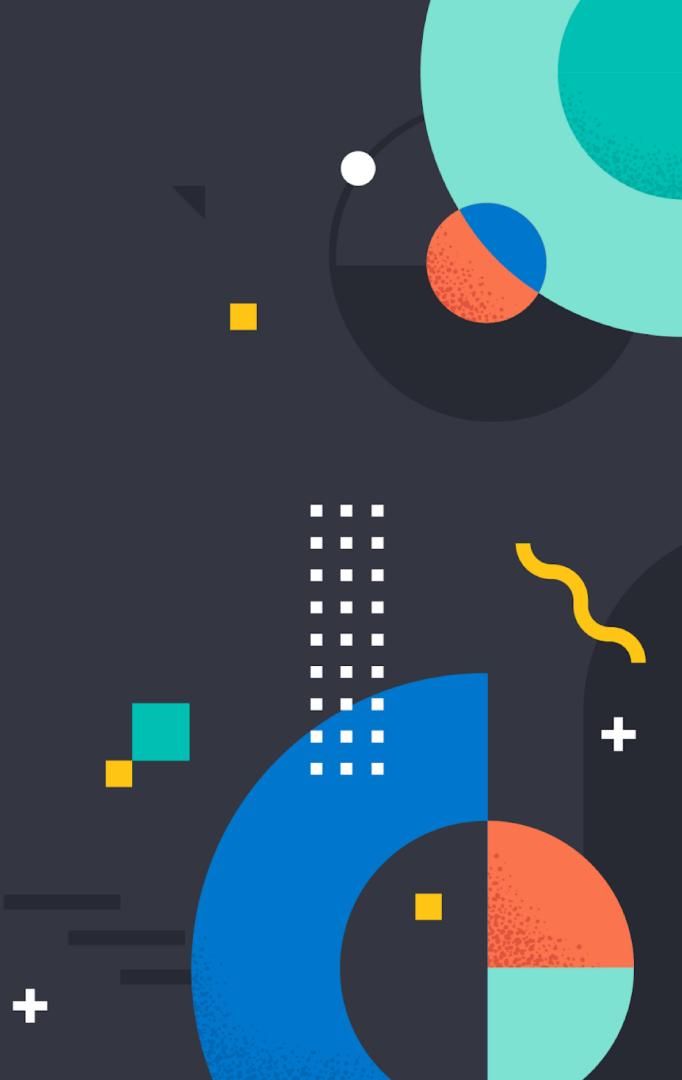




# 日志指标与可视化分析

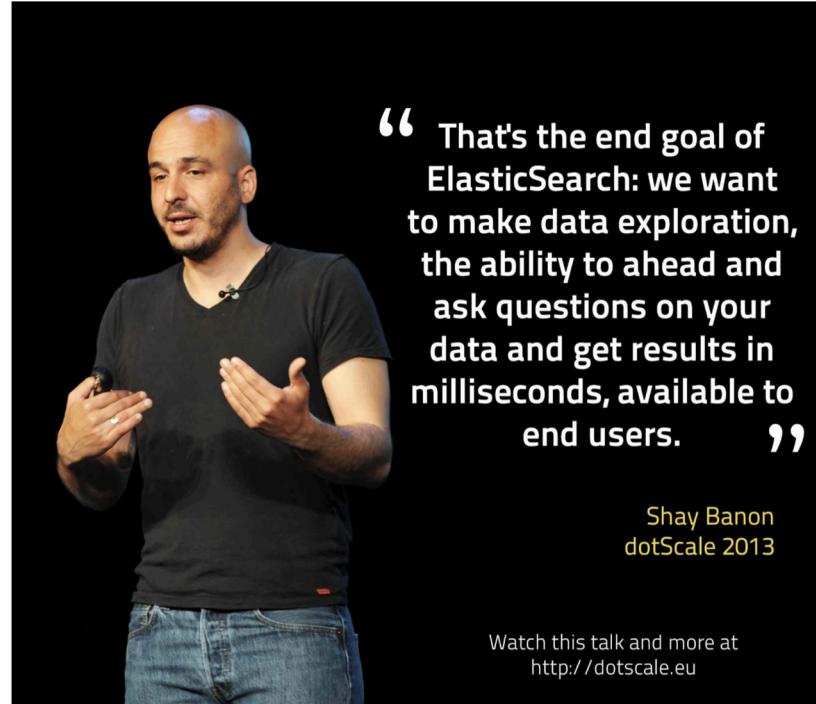
---

Jerry Zhu  
Elastic 架构师



# Elastic是 ELK背后的公司

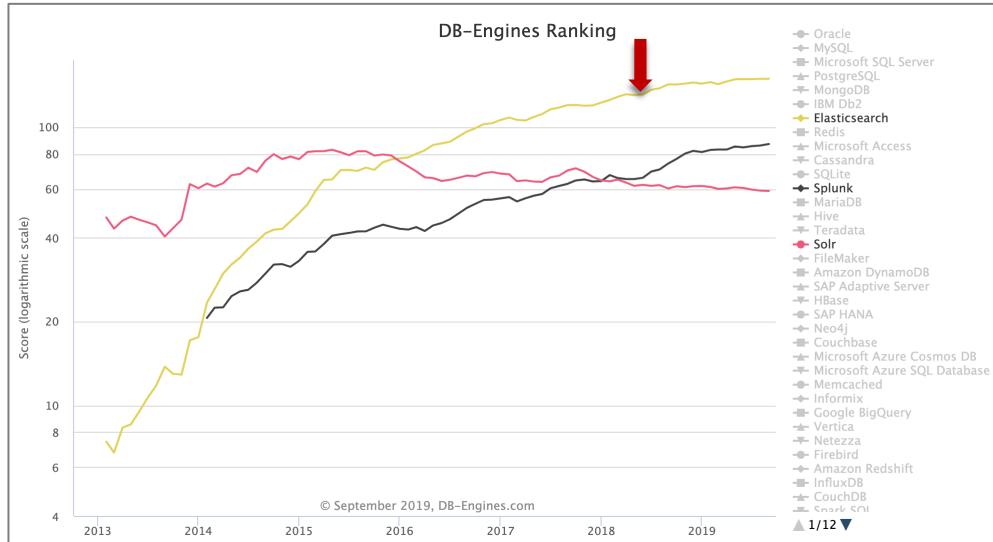
- **10万+** 全球社区参与者
- **1万+** 中国社区参与者
- **3.5亿+** 下载 (持续增长中)
- **1680+** 员工
- **8800+** 付费订阅客户
- **NYSE: ESTC**



# Elasticsearch 非常受欢迎

DB-Engines Ranking 2019 Sep

	<b>Rank</b>	Sep 2019	Aug 2019	Sep 2018	<b>DBMS</b>	<b>Database Model</b>	<b>Score</b>	Sep 2019	Aug 2019	Sep 2018
1.	1.	1.	1.	Oracle +		Relational, Multi-model <a href="#">?</a>	1346.66	+7.18	+37.54	
2.	2.	2.	2.	MySQL +		Relational, Multi-model <a href="#">?</a>	1279.07	+25.39	+98.60	
3.	3.	3.	3.	Microsoft SQL Server +		Relational, Multi-model <a href="#">?</a>	1085.06	-8.12	+33.78	
4.	4.	4.	4.	PostgreSQL +		Relational, Multi-model <a href="#">?</a>	482.25	+0.91	+75.82	
5.	5.	5.	5.	MongoDB +		Document	410.06	+5.50	+51.27	
6.	6.	6.	6.	IBM Db2 +		Relational, Multi-model <a href="#">?</a>	171.56	-1.39	-9.50	
7.	7.	7.	7.	Elasticsearch +	←	Search engine, Multi-model <a href="#">?</a>	149.27	+0.19	+6.67	
8.	8.	8.	8.	Redis +		Key-value, Multi-model <a href="#">?</a>	141.90	-2.18	+0.96	
9.	9.	9.	9.	Microsoft Access		Relational	132.71	-2.63	-0.69	
10.	10.	10.	10.	Cassandra +		Wide column	123.40	-1.81	+3.85	
11.	11.	11.	11.	SQLite +		Relational	123.36	+0.65	+7.91	
12.	12.	↑ 13.	13.	Splunk		Search engine	87.01	+1.12	+12.98	
13.	13.	↑ 14.	14.	MariaDB +		Relational, Multi-model <a href="#">?</a>	86.07	+1.11	+15.43	
14.	14.	↑ 16.	16.	Hive +		Relational	83.10	+1.30	+23.46	
15.	15.	↓ 12.	12.	Teradata +		Relational, Multi-model <a href="#">?</a>	76.97	+0.32	-0.42	
16.	16.	↓ 15.	15.	Solr		Search engine	58.97	-0.16	-1.24	
17.	17.	↑ 19.	19.	FileMaker		Relational	58.15	+0.13	+2.84	
18.	18.	↑ 20.	20.	Amazon DynamoDB +		Multi-model <a href="#">?</a>	57.82	+1.25	+4.47	
19.	↑ 20.	↓ 18.	18.	SAP Adaptive Server		Relational	56.10	+0.24	-1.94	
20.	↓ 19.	↓ 17.	17.	HBase		Wide column	55.72	-0.82	-2.75	
21.	21.	21.	21.	SAP HANA +		Relational, Multi-model <a href="#">?</a>	55.39	-0.04	+2.66	



# Elastic 产品生态

## 解决方案

搜索

日志

指标

安全

APM

企业搜索

Elastic  
云服务

### Elastic大数据平台

数据  
展示



Kibana

存储索引  
计算分析



Elasticsearch

数据  
摄取



Logstash



Beats

+



机器学习

数据关联分析

规则告警

多集群监控

报表

高级安全



Elastic  
企业云

# 日志 指标 APM IT全方位可观察性

491 total, 3 running, 488 sleeping, 3188 threads  
 2.87, 2.54, 2.28 CPU usage: 8.85% user, 6.89% sys, 84.25% idle  
 : 284M resident, 47M data, 17M linkedit.  
 : 280767 total, 4854M resident, 118M private, 1997M shared.  
 5G used (3187M wired), 1288M unused.

vsizes, 1111M framework vsizes, 341912008(16284) swapins, 348809712(0) swapouts.  
 packets: 36759414/22G in, 19823346/12G out. Disks: 87127310/2211G read, 32948289/24

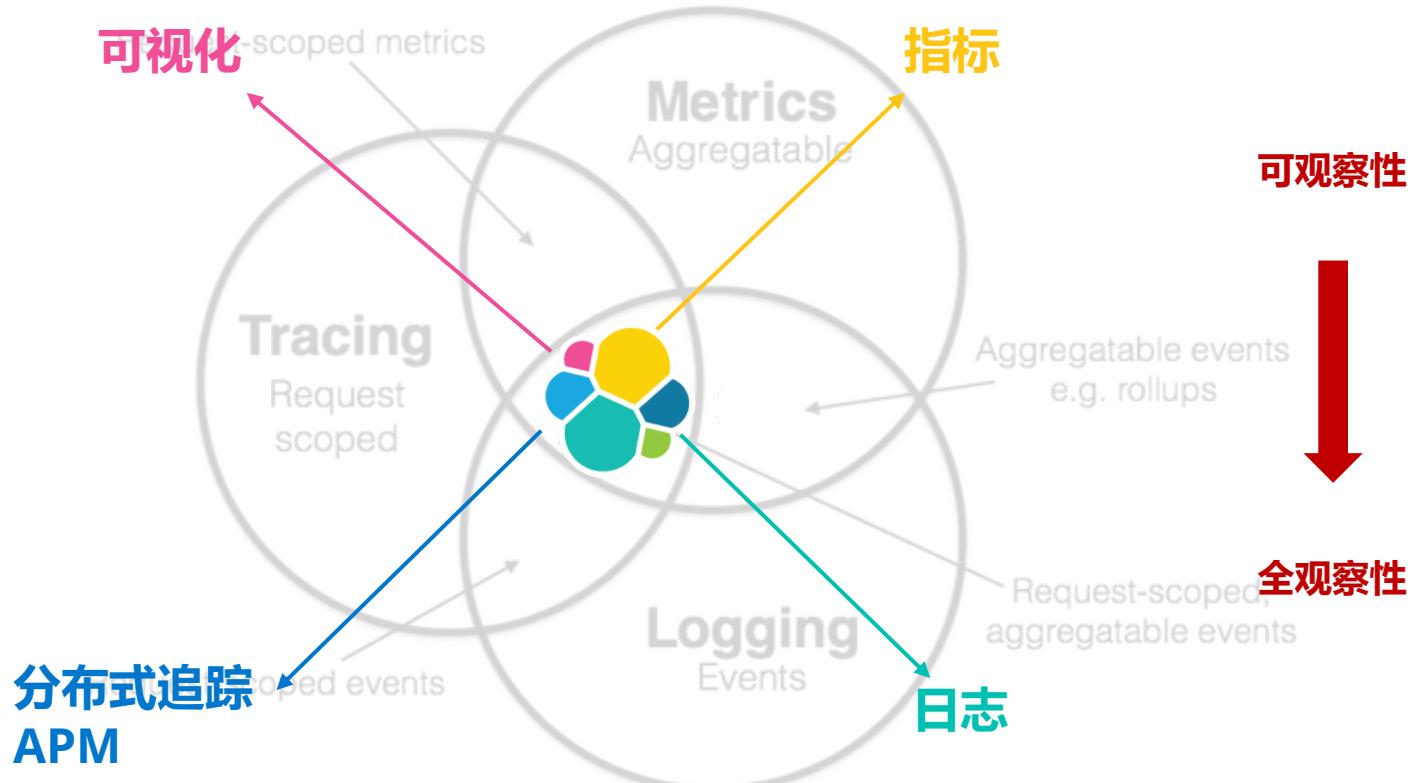
MAND	%CPU	TIME	#TH	#WQ	#PORTS	MEM	PURG	CMPRS	PGRP	PPID	STATE	DETAILS
	7.7	00:05.46	1/1	0	23	8488K	0B	0B	73698	73428	running	02:15:09 Morgans-MacBook-Pro kernel[0]: AppleThunderboltNHiType2::waitForDK2G2Sx - retries = 9
h	0.0	00:00.32	1	0	19	6764K	0B	0B	73428	73427	sleeping	02:15:09 Morgans-MacBook-Pro kernel[0]: Wake reason: RTC (Alarm)
in	0.0	00:00.04	2	1	29	2152K	0B	0B	73427	1291	sleeping	02:15:09 Morgans-MacBook-Pro kernel[0]: RTC: Maintenance 2016/5/6 07:15:09, sleep 2016/5/6 05:15:14
worker	0.0	00:00.09	3	1	54	3564K	0B	0B	73426	1	sleeping	02:15:09 Morgans-MacBook-Pro kernel[0]: AppleCamIn::systemWakeCall - messageType = 0xE8000340
worker	0.0	00:00.11	3	1	54	3548K	0B	0B	73425	1	sleeping	02:15:09 Morgans-MacBook-Pro kernel[0]: AppleCamIn::wakeEventHandlerThread
clocklookd	0.0	00:00.12	4	1	85	4760K	32K	0B	73413	1	sleeping	02:15:09 Morgans-MacBook-Pro kernel[0]: Previous sleep cause: -113
webEngineP	0.0	00:00.27	13	1	95	14M	0B	0B	73408	73408	sleeping	02:15:09 Morgans-MacBook-Pro kernel[0]: IPv6 link-local address <IPv6 address #xeaf69419aec2bd89> has no prefix
worker	0.0	00:00.18	3	1	65	5576K	0B	0B	73411	1	sleeping	02:15:09 Morgans-MacBook-Pro kernel[0]: AppleThunderboltNHiType2::prePCIWake - power up complete - took 1 us
worker	0.0	00:00.16	3	1	57	5500K	0B	0B	73410	1	sleeping	02:15:09 Morgans-MacBook-Pro kernel[0]: SystemDCharger[45]: ASL Sender Statistics
team Deck	0.0	00:02.15	35	1	649-	46M-	1476K	5292K	73408	1	sleeping	02:15:09 Morgans-MacBook-Pro kernel[0]: NetworkAnalyticsEngine::writeNetworkRecordFromSelfFingerprint[1:key@1010&fqn=1qm:isFaulty:] Hashing of the primary key failed. Dropping the journal
pd	0.0	00:00.02	2	1	40	128K	0B	1104K	73403	1	sleeping	02:15:09 Morgans-MacBook-Pro kernel[0]: listenerCallback - Thunderbolt HPD packet for route = 0x0 port = 11 unplug = 0
worker	0.0	00:00.06	3	1	54	64K	0B	3536K	73376	1	sleeping	02:15:09 Morgans-MacBook-Pro kernel[0]: AppleThunderboltNHiType2::earlyWake - complete - took 137 milliseconds
worker	0.0	00:00.07	3	1	51	60K	0B	3264K	73358	1	sleeping	02:15:09 Morgans-MacBook-Pro kernel[0]: en0: channel changed to 1
.apple.iC	0.0	00:00.23	3	2	59+	2916K+	0B	1460K	73353	1	sleeping	02:15:09 Morgans-MacBook-Pro kernel[0]: Airport: Link Up
worker	0.0	00:00.06	3	1	51	60K	0B	3256K	73303	1	sleeping	02:15:09 Morgans-MacBook-Pro kernel[0]: AppNSMClient[46]: network_reachability_changed: no response from server, reachability, 2, queryRetries, 0
worker	0.0	00:00.07	3	1	51	60K	0B	3188K	73302	1	sleeping	02:15:09 Morgans-MacBook-Pro kernel[0]: iOBDock[320]: iUAWebAppSuggestionManager::notifyBestAppChanged:type=options:bundleIdentifier:activityType:dynamicIdentifier:when:confidence:deviceName:deviceIdentifier:de
worker	0.0	00:00.08	3	1	51	60K	0B	3304K	73301	1	sleeping	02:15:09 Morgans-MacBook-Pro kernel[0]: netbiosd[2063]: network_reachability_changed : network is not reachable, netbiosd is shutting down

# 日志 + 指标 + APM融入同一个技术栈

Why?



# 日志仅仅是起点 – 终点是全方位的可观察性



# Talk is cheap. Show me the code.

Linus Torvalds

# Lucene & Elasticsearch 付出的努力

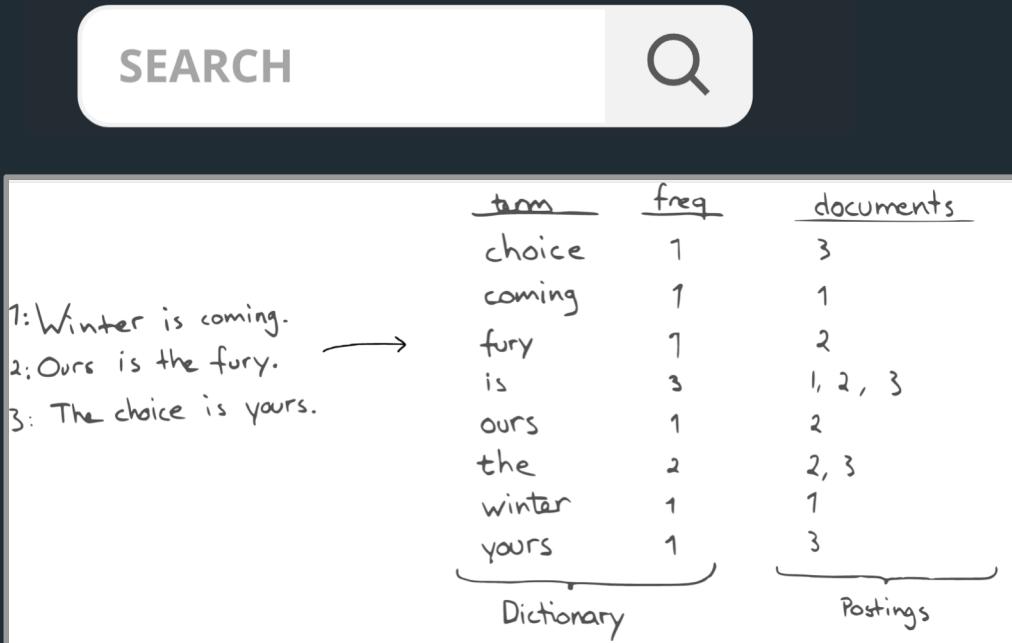
# 在 Elasticsearch 初期

它主要为应用提供搜索功能

2010 —

# 搜索引擎

Inverted index primary data structure, and  
is great for search



# Elasticsearch 进化到分析支持

基于Lucene 的 "doc values"构建的列存储

userid	first	middle	last	city	state
john123	John	James	Smith	Alamo	California
jrice	Jill	Amy	Rice		
mt123	Jeff		Twain	Toledo	Ohio
sadams	Sue		Adams		
adoe	Amy		Doe	Miami	Florida

2010

搜索引擎

Inverted index primary data structure, and is great for search

2012

列存储

Structured data storage, resulting in compact storage and faster analytics

# 聚合框架

## 不仅仅是聚合功能

{

```
"_index": "helioweb",
"_type": "voyager1",
"_id": "1977-257",
"_score": 1,
"_source": {
    "seLon": 353.8,
    "objectName": "voyager1",
    "seLat": 0.2,
    "year": 1977,
    "date": "1977-09-14T10:28:25.178Z",
    "radAU": 1.02,
    "dayOfYear": 257
}
```

}

2010

### 搜索引擎

Inverted index primary data structure, and is great for search

2012

### 列存储

Structured data storage, resulting in compact storage and faster analytics

2014

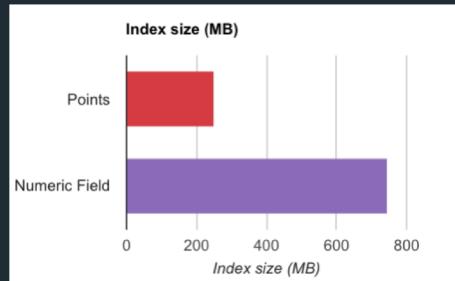
### 聚合框架

Analytics features to slice and dice data along various dimensions

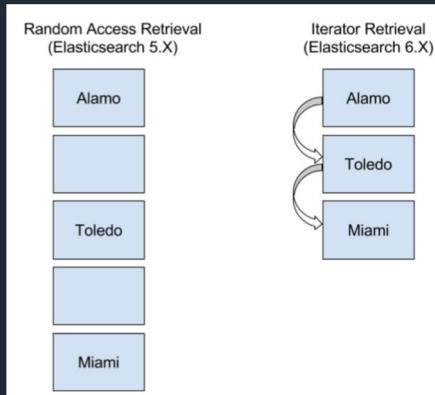
# Elasticsearch 存储效率增强

## BKD Trees & Sparse Fields

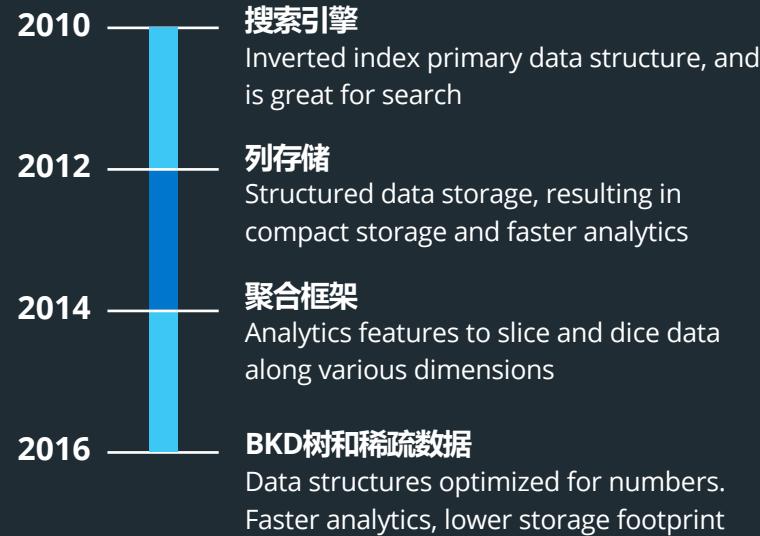
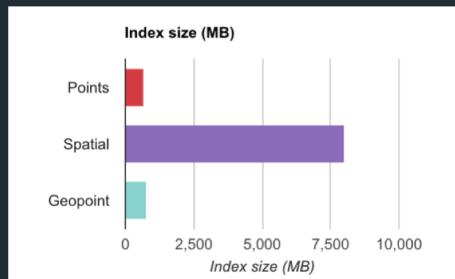
### 1-维数据



### 稀疏数据



### 2-维数据



# Rollup支持长期保留

在Elasticsearch 6.3中加入

Management

Elasticsearch  
License management  
Users  
Roles  
Watches  
Index management  
[Index rollups](#)

Kibana  
Index patterns  
Reporting  
Advanced settings  
Saved objects

Logstash  
Pipelines

Create a new rollup job

Optional: Collect metrics on important fields

You can collect metrics on as many fields as you want.

Field	Min	Max	Avg	Sum	Value count	Cardinality
system.network.out.bytes	•	•		•	•	
system.network.out.errors		•	•	•		
system.network.usage	•	•			•	

Add a metric Continue

Metric to aggregate

Field: system.network.high\_fives

Metrics to capture:  
Min, Max, Avg (unchecked)  
Sum, Value count, Cardinality (checked)

Add

Kibana Visualize / New Visualization (unsaved) Save Share Inspect Refresh Auto-refresh < May 17th 2015, 10:53:54.770 to May 21st 2015, 07:38:55.770 Options

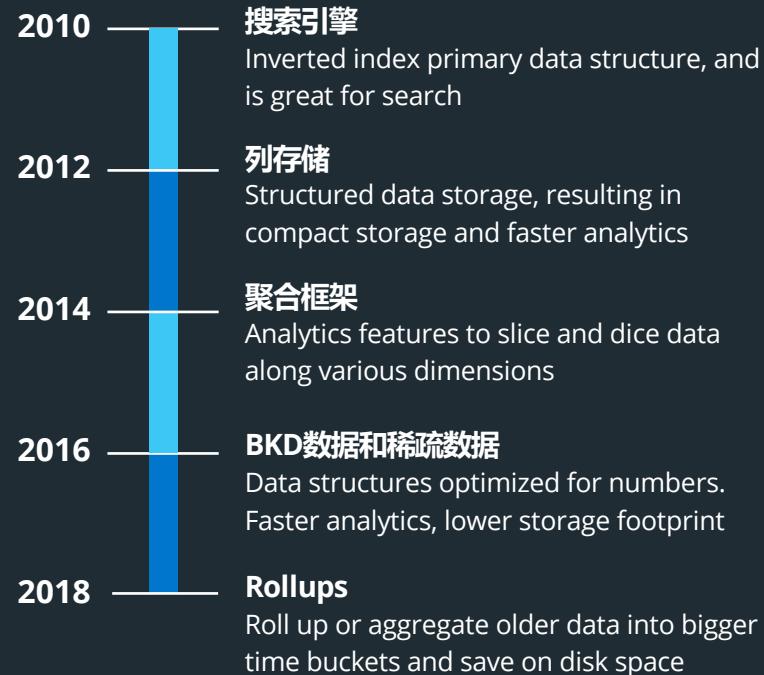
Add a filter: logstash -> system

Metrics: Sum Help Field: system.network.out.bytes

Buckets: A-Axis Aggregation: Date Histogram Data Histogram Help

Field: system.network.out.bytes

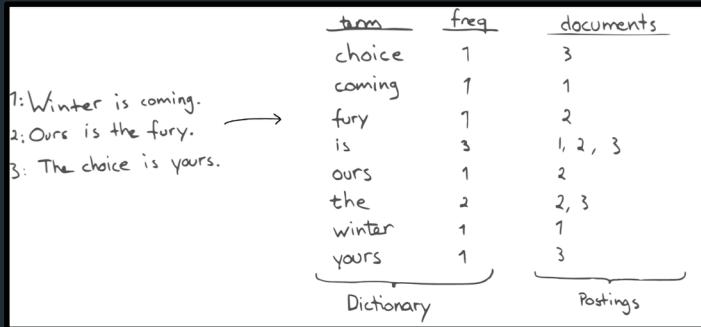
A histogram visualization showing peaks of data over time.



<https://www.elastic.co/blog/data-rollups-in-elasticsearch-you-know-for-saving-space>

# Elasticsearch 用于搜索和数值分析

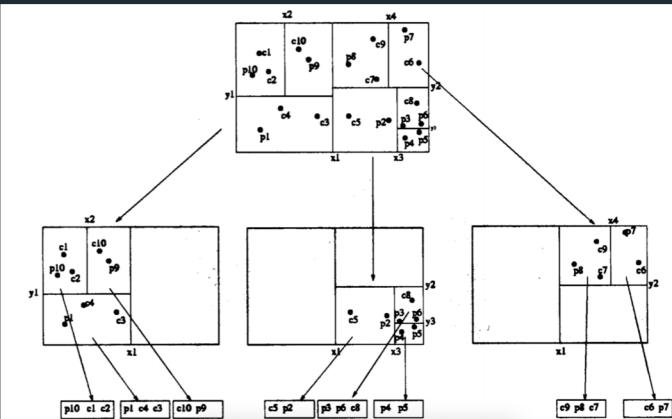
## 倒排序支持的全文搜索



## 列存储支持的结构化数据

userid	first	middle	last	city	state
john123	John	James	Smith	Alamo	California
jrice	Jill	Amy	Rice		
mt123	Jeff		Twain	Toledo	Ohio
sadams	Sue		Adams		
adoe	Amy		Doe	Miami	Florida

## BKD数支持的数值型运算



## Rollup 节省存储空间

The screenshot shows the Elasticsearch Management interface with the following sections:

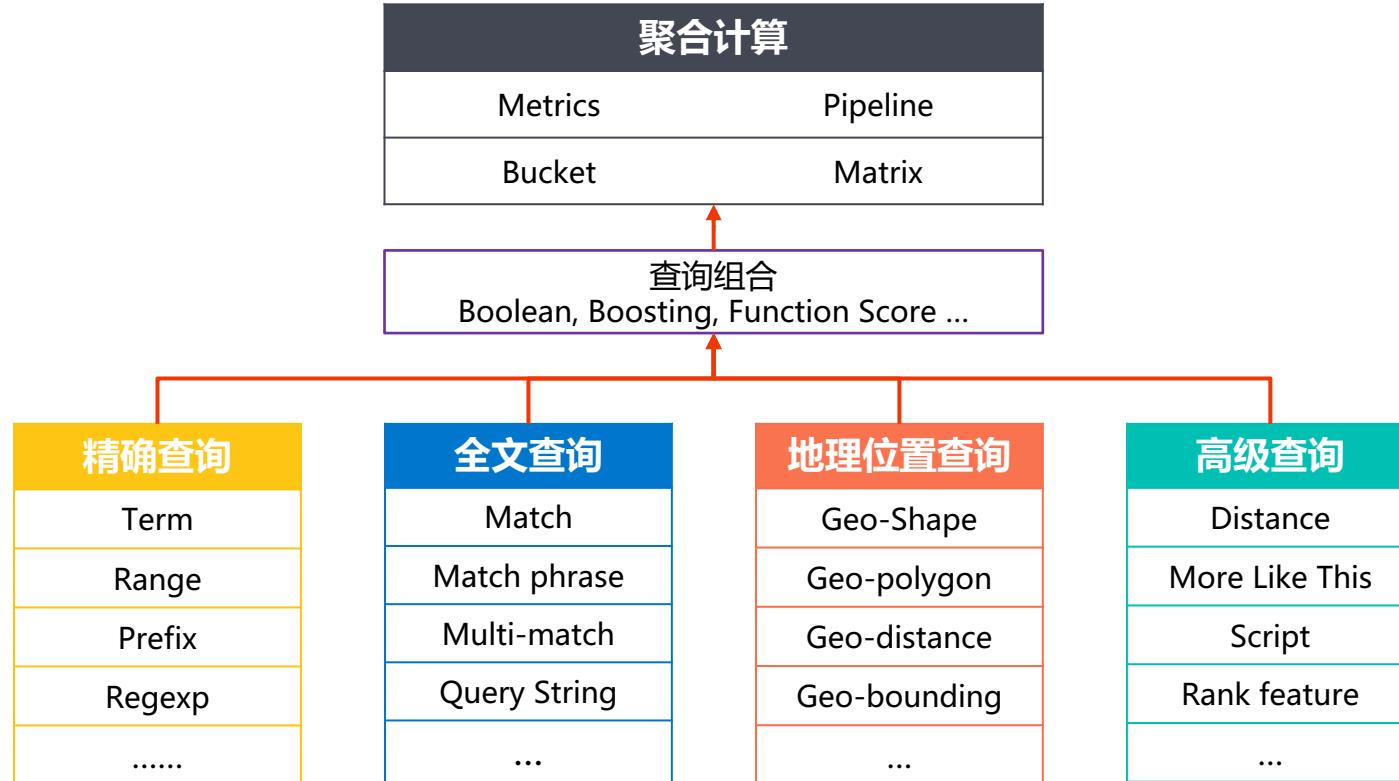
- Management** (Header)
- Elasticsearch** (Section)
  - License management
  - Users
  - Roles
  - Watches
  - Index management
  - Index refresh
- Kibana** (Section)
  - Index patterns
  - Reporting
  - Advanced settings
  - Saved objects
- Logstash** (Section)
  - Pipelines
- Create a new rollup job** (Main Action)
- Optional: Collect metrics on important fields** (Text)

You can collect metrics on as many fields as you want.

Field	Min	Max	Avg	Sum	Value count	Cardinality
system.network.out.bytes	●	●		●	●	
system.network.out.errors			●	●	●	
system.network.usage	●	●	●			●
- Add a metric** (Action Button)
- Continue** (Action Button)
- Metric to aggregate** (Section)

Field  
system.network.high\_files
- Metrics to capture** (Section)
  - Min
  - Sum
  - Max
  - Value count

# 查询能力变得空前强大



2019

# Data Frames

未来...

# Beats Logstash 付出的努力

# 日志、指标采集 Beats生态

libbeat



**FILEBEAT**  
日志文件



**METRICBEAT**  
指标数据



**PACKETBEAT**  
网络数据



**FUNCTIONBEAT**  
云服务器监控



**WINLOGBEAT**  
Window日志



**HEARTBEAT**  
服务可用性监控



**AUDITBEAT**  
Linux审核框架事件

httpbeat

apachebeat

nginxbeat

pingbeat

execbeat

dockerbeat

elasticbeat

70+  
社区制造

# 日志、指标采集 Logstash 生态



Logstash

200 + 插件

azure event hub	cloudwatch	couchdb	elasticsearch	exec
file	websocket	tcp	upd	graphite
http	http poller	imap	irc	java generator
java stdin	jdbc	jms	jmx	kafka
log4j	xmpp	redis	rss	s3
snmp	sqlite	stdin	syslog	github
更多 社区制造				

# 开放生态 第三方重磅工具



Spark  
Streaming



Prometheus



Kibana  
付出的努力

# 产品快速迭代



## 以前的 Elastic Stack

一流的引擎！  
三流的界面！

## 现在的 Elastic Stack

一流的引擎！  
一流的界面！

# 快速改进

## Kibana 5.5

Kibana 5.5 Discover interface showing a search bar and a list of fields:

- t \_id
- t \_index
- # \_score
- t \_type
- t beat.hostname
- t beat.name
- t beat.version
- t container\_id
- t container\_name
- t host
- t input\_type
- t json\_error
- t message
- t namespace
- # offset
- t pod\_name
- t tags
- t type

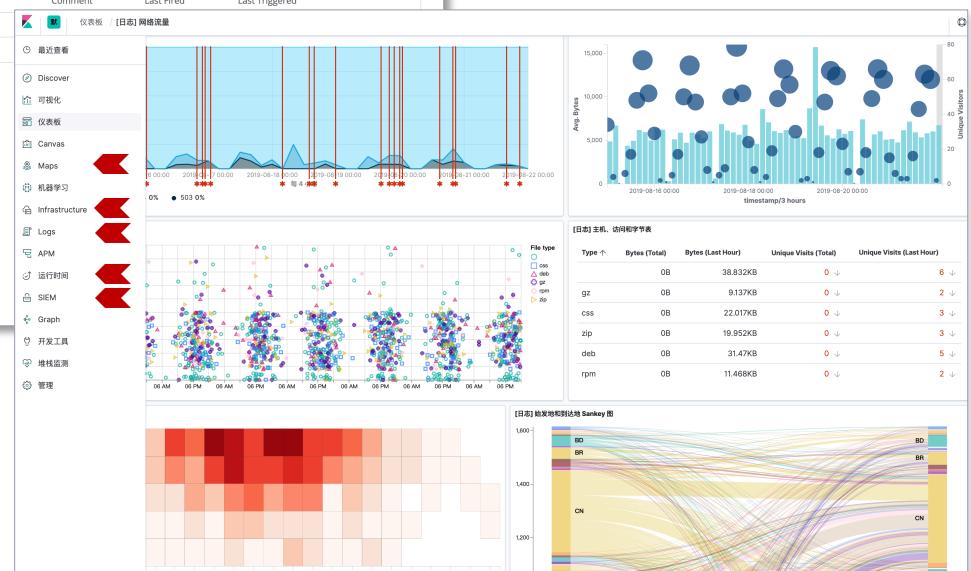
## Kibana 6.5

Kibana 6.5 Watches interface showing a table of threshold alerts and a sidebar menu:

ID	Name	State
363bx675-5638-4c...	Threshold watch	✓ OK

\_sidebar menu items include: Discover, Visualize, Dashboard, Timeline, Machine Learning, APM, Graph, Dev Tools, Monitoring, Management, elastic, Logout, and Collapse.

## Kibana 7.3



# 数据展现 – 日志分析

**新建 保存 打开 共享 检查**

Filters 搜索

+ 添加筛选

编辑筛选 编辑为查询 DSL

选定:

可用:

字段 运算符

agent.keyword 是

值

选择值

Mozilla/5.0 (X11; Linux x86\_64; rv:6.0a1) Gecko/201...

Mozilla/5.0 (X11; Linux i686) AppleWebKit/534.24 (...)

Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1;...)

# bytes

clientip

extension > Aug 21, 2019 @ 14:32:33

geo.coordinates

geo.dest

geo.src

geo.srcdest > Aug 21, 2019 @ 14:00:16

host

hour\_of\_day

Filters geo.dest:| KQL

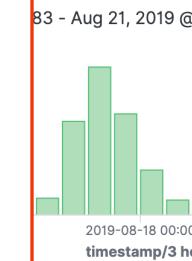
- "CN"
- "IN"
- "US"
- "ID"
- "BR"
- "PK"
- "BD"
- "NG"
- "RU"
- "MX"

83 - Aug 21, 2019 @

The histogram displays the frequency of geo.dest values over a specific timestamp range. The x-axis represents the timestamp, and the y-axis represents the count of occurrences. The distribution shows several peaks, with the highest peak occurring around 2019-08-18 00:00.

Timestamp	Count
2019-08-18 00:00	~100
2019-08-18 01:00	~80
2019-08-18 02:00	~90
2019-08-18 03:00	~70
2019-08-18 04:00	~10
2019-08-18 05:00	~10

2019-08-18 00:00  
timestamp/3 h



**Quick select**

Last
15
minutes
**Apply**

---

**Commonly used**

今日	本周
过去 15 分钟	过去 30 分钟
过去 1 小时	过去 24 小时
过去 7 天	过去 30 天
过去 90 天	过去 1 年

---

**Recently used date ranges**

- [过去 7 天](#)
- [Last 9 days](#)
- [Last 7 weeks](#)
- [Oct 11, 2013 @ 10:00:00.000 to Apr 2, 2019 @ 10:00:00.000](#)
- [Last 15 days](#)

---

**Refresh every**

15
minutes
 Stop

# 数据展现 - 时序数据分析器

默

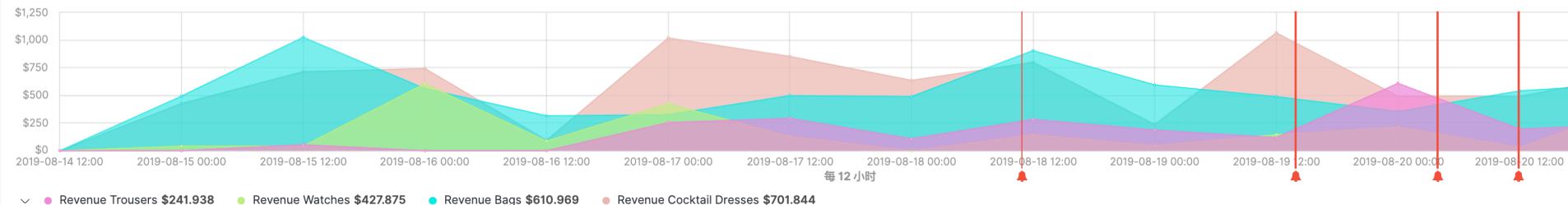
可视化 / [电子商务] 促销追踪

保存 共享 检查 刷新

过去 7 天

Show dates

时间序列 指标 前 N 个 仪表盘图 Markdown 表



自动应用

数据 面板选项 注释

## 数据源

索引模式 (必需)

kibana\_sample\_data\_ecommerce

时间字段 (必需)

order\_date

×

查询字符串

taxful\_total\_price:>250

是否忽略全局筛选? 是 否    是否忽略面板筛选? 是 否

Lucene

是 否    是 否

图标 (必需)

钟铃

▼

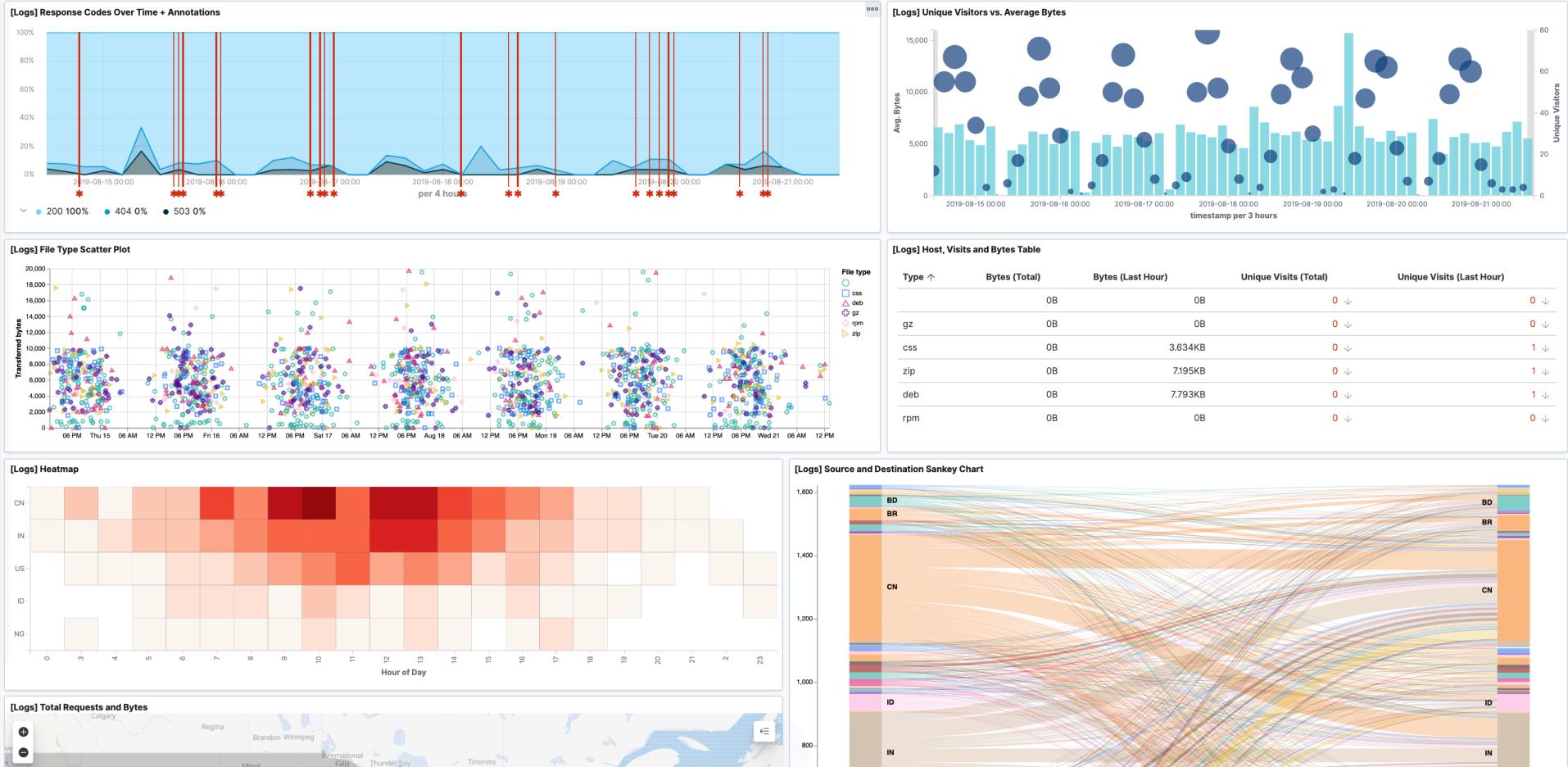
字段 (必需 - 路径以逗号分隔)

taxful\_total\_price

行模板 (必需)

Ring the bell! \${taxful\_total\_price}

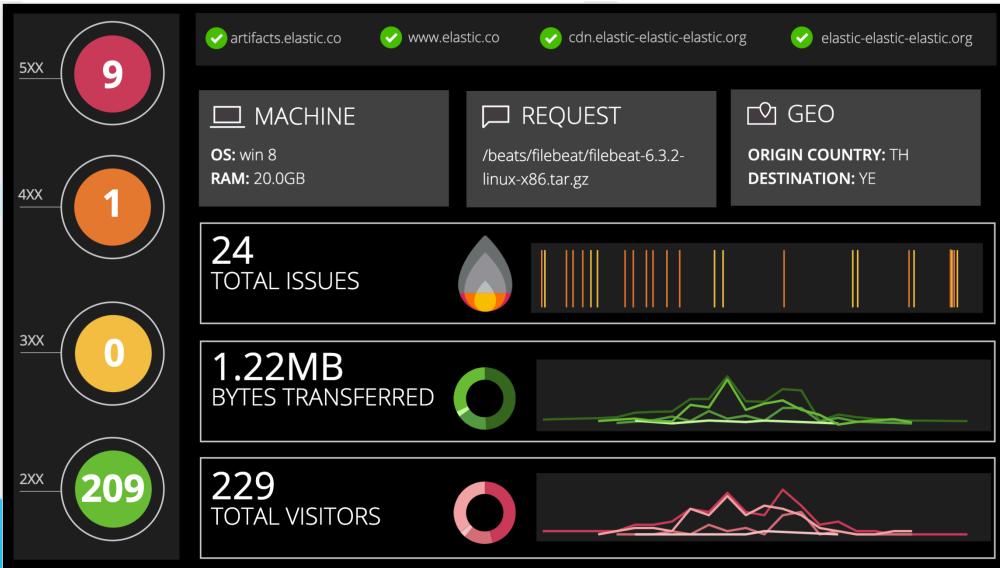
# 数据展现 – 丰富的图表



# 数据展现 - 丰富的开放生态，支持开源Vega图表组件



# 数据展现 – 像PPT一样的全动态大屏报表



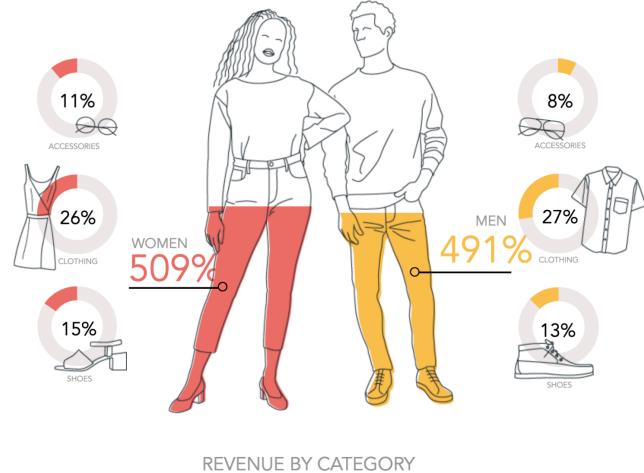
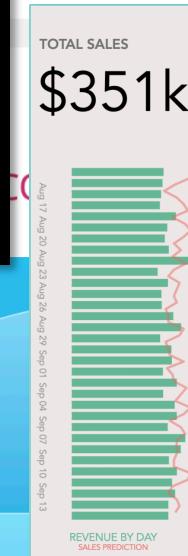
## FLIGHTS

12.5 MILES

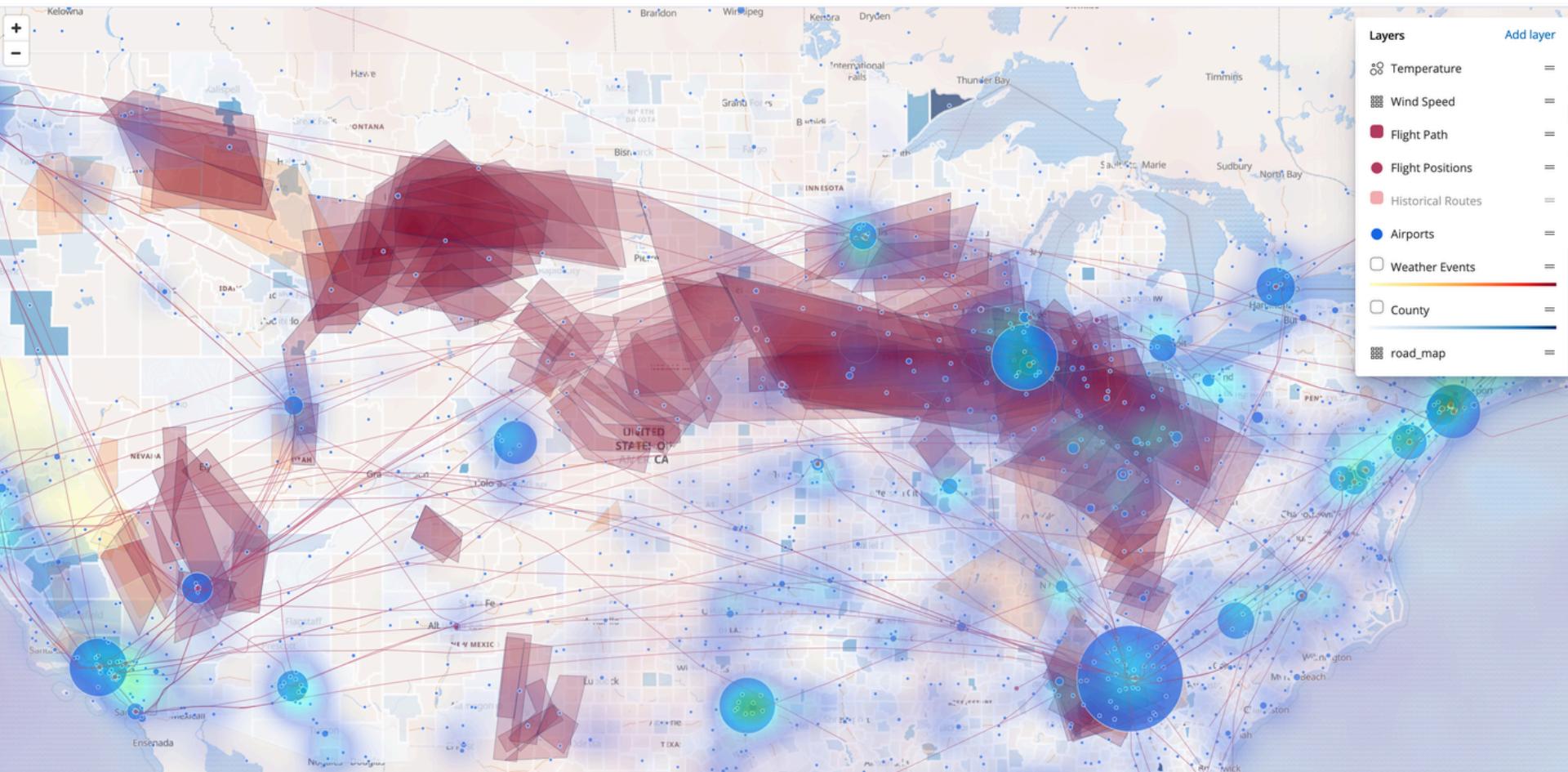
LONGEST FLIGHT

## AIRPORTS

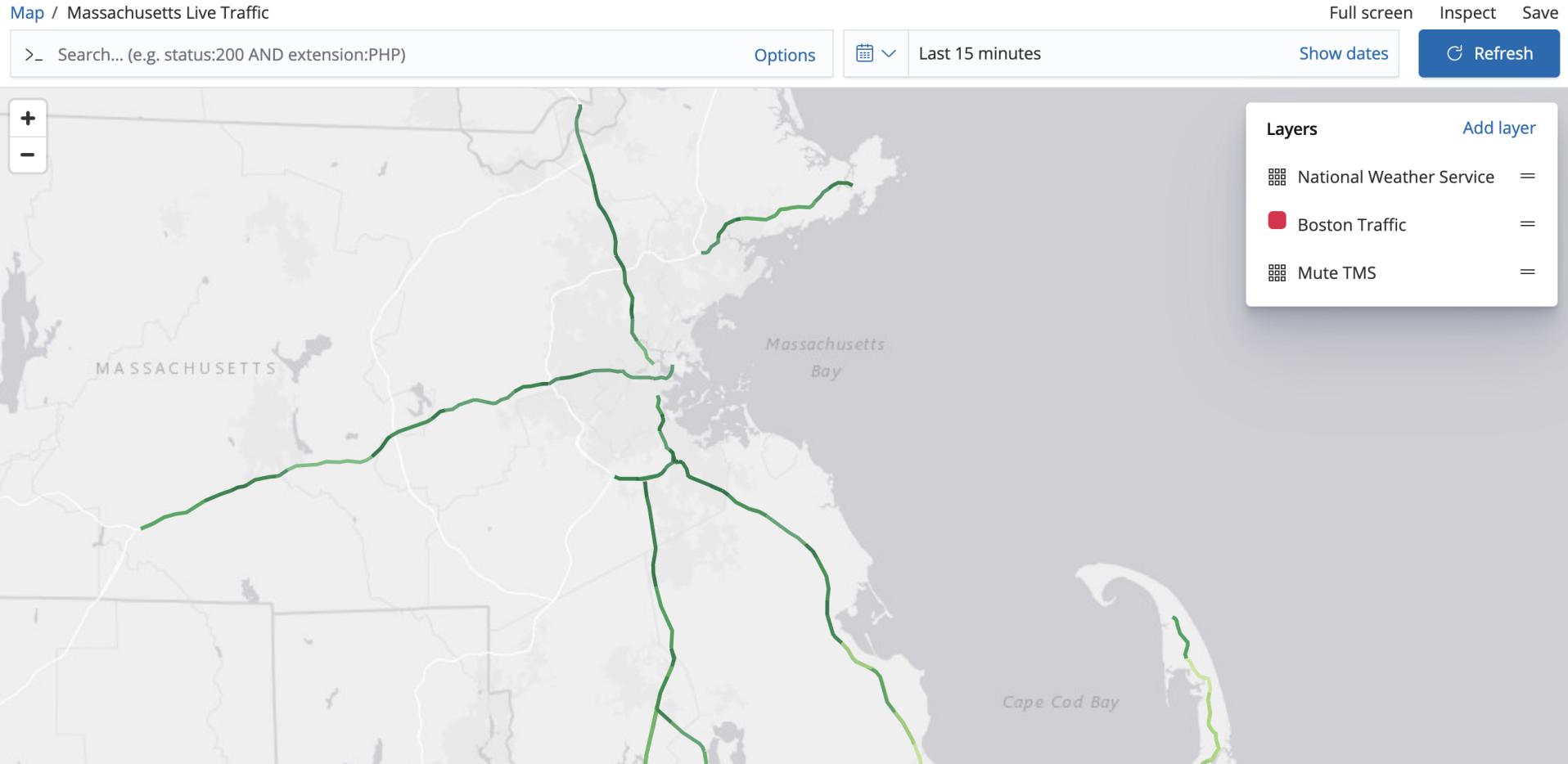
10.6k MILES



# 数据展现 – 在地图上分析数据



# 数据展现 – 地图上画路径也能做到

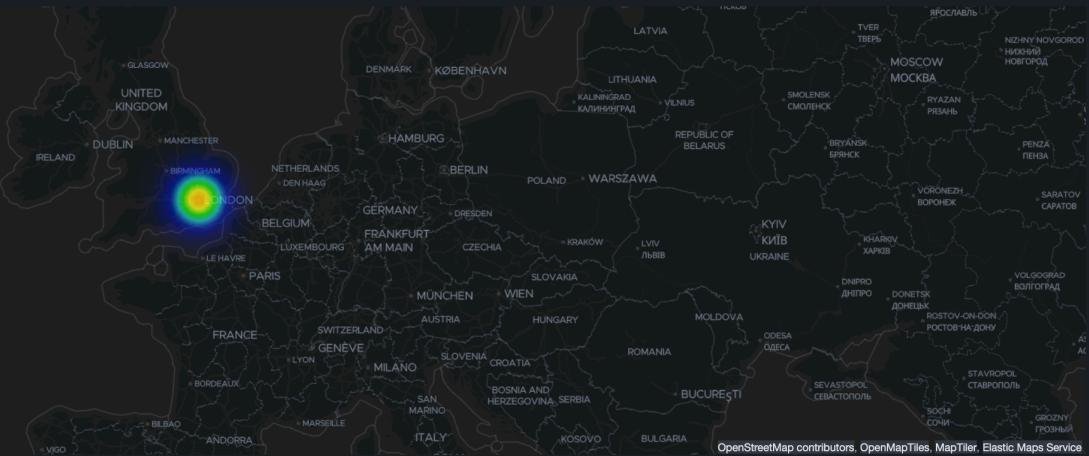


**1,021**  
Total Attacks    **4**  
Unique Rules Hit    **5**  
Unique Sources    **3**  
Source Attack Countries    **2**  
Reporting Hosts

WAF - Chart of Top 10 Rules



WAF - Attack Heat Map



WAF - Chart Of Severities



WAF - OS Pie Chart



WAF - Events Over Time, By Severity

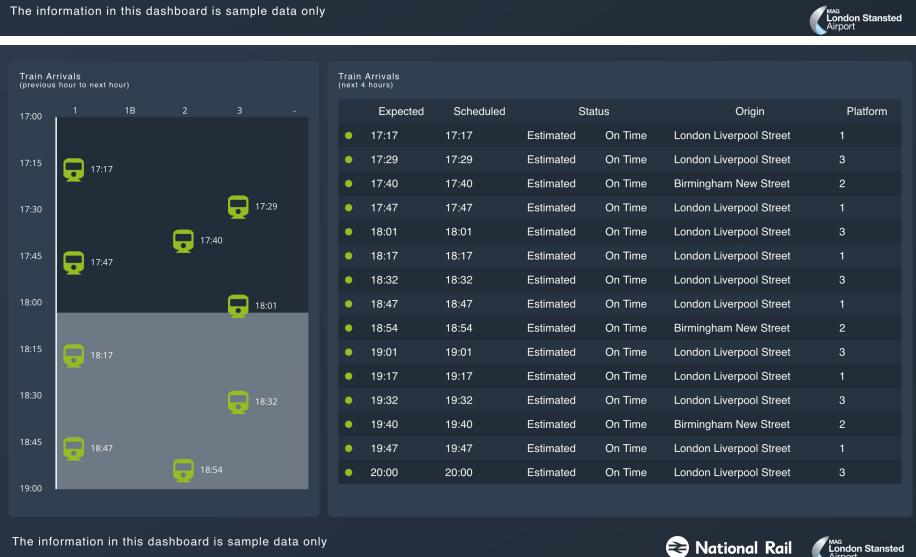
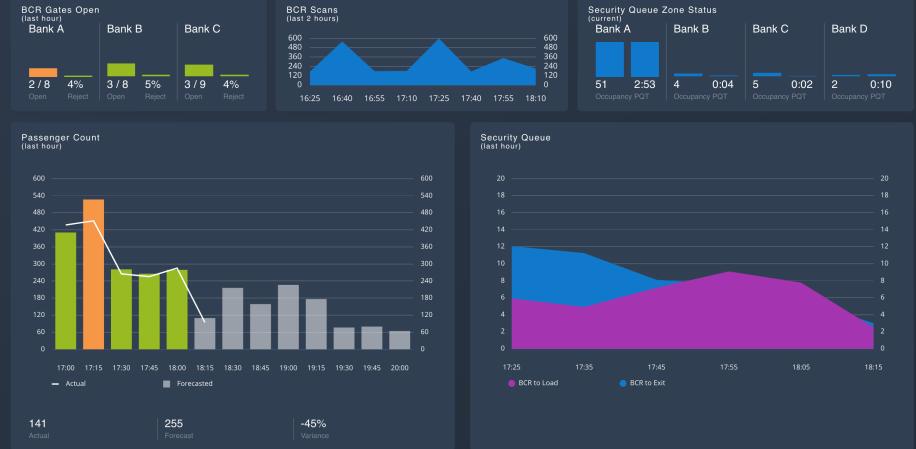
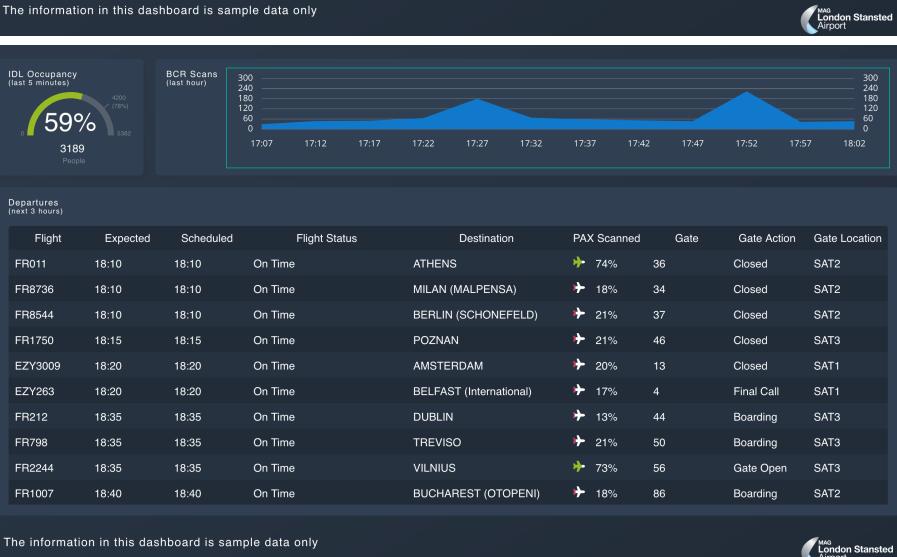


WAF - Top 5 Targeted URI's



WAF - Messages Over Times





# APM 从无到有

# APM

统一 日志 , 指标 , APM

## Open Source

### 各种语言客户端

Java, .Net, Go, RUM, Node, Python, Ruby

### Kibana UI

APM 工作流程耗时图  
分布式的追踪

### 只是多了一种索引

和其它数据关联  
借助于ES的功能





## Transaction sample

Actions ▾

[View full trace](#)

Timestamp

9 months ago (November 27th 2018, 08:00:32.947)

URL

http://infra-petclinic-client-7b9644b6c8-hjjsq/api/owners

Duration

492 ms

% of trace

N/A

Result

HTTP 2xx

Errors

None

User ID

lwoollaston52

[Timeline](#)[Metadata](#)

Services

petclinic-node petclinic-spring

0 ms 50 ms 100 ms 150 ms 200 ms 250 ms 300 ms 350 ms 400 ms 450 ms 492 ms

HTTP 2xx /api/owners 492 ms

GET infra-petclinic-client-7b9644b6c8-hjjsq/petclinic/api/owners 489 ms

HTTP 2xx OwnerRestController#getOwners 418 ms

SELECT 2,893 µs

SELECT 3,204 µs

SELECT 2,124 µs

# 其它Kibana APP

# Logs 实时流式日志查看器

搜索日志条目..... (例如 host.name:host-1)			配置	定制	Highlights	08/21/2019 2:39:09 PM	实时流式传输	
Timestamp	event.dataset	Message						
Aug 21, 2019 @ 13:48:08.489		6195 "-" "Mozilla/4.0 (compatible; MSIE 6.0; Windo 123.24.232.117 - - [2018-08-01T05:48:08.489Z] "GET /apm-server/apm-server-6.3.2-windows-x64.zip HTTP/1.1" 200 8503 "-" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1; .NET CLR 1.1.4322)" 151.60.185.99 - - [2018-08-01T05:52:18.588Z] "GET /people/type:astronauts/name:eric-a-boe/profile HTTP/1. 1" 200 8661 "-" "Mozilla/5.0 (X11; Linux i686) AppleWebKit/534.24 (KHTML, like Gecko) Chrome/11.0.696.50 S afari/534.24" 197.207.29.94 - - [2018-08-01T05:57:34.742Z] "GET /beats/filebeat/filebeat-6.3.2-linux-x86_64.tar.gz HTTP/ 1.1" 200 5906 "-" "Mozilla/5.0 (X11; Linux i686) AppleWebKit/534.24 (KHTML, like Gecko) Chrome/11.0.696.50 Safari/534.24" 208.144.38.239 - - [2018-08-01T05:58:42.764Z] "GET /beats/metricbeat HTTP/1.1" 200 5883 "-" "Mozilla/5.0 (X11; Linux x86_64; rv:6.0a1) Gecko/20110421 Firefox/6.0a1" 153.95.253.45 - - [2018-08-01T06:00:16.260Z] "GET /beats/filebeat/filebeat-6.3.2-linux-x86.tar.gz HTTP/1. 1" 200 8202 "-" "Mozilla/5.0 (X11; Linux x86_64; rv:6.0a1) Gecko/20110421 Firefox/6.0a1" 186.85.119.253 - - [2018-08-01T06:32:33.368Z] "GET /beats/filebeat/filebeat-6.3.2-linux-x86.tar.gz HTTP/1. 1" 200 2690 "-" "Mozilla/5.0 (X11; Linux x86_64; rv:6.0a1) Gecko/20110421 Firefox/6.0a1" 138.234.210.146 - - [2018-08-01T06:33:59.193Z] "GET /beats HTTP/1.1" 200 6320 "-" "Mozilla/4.0 (compatib e; MSIE 6.0; Windows NT 5.1; SV1; .NET CLR 1.1.4322)" 231.54.75.65 - - [2018-08-01T06:39:09.880Z] "GET /styles/ads.css HTTP/1.1" 200 7789 "-" "Mozilla/5.0 (X11; Linux i686) AppleWebKit/534.24 (KHTML, like Gecko) Chrome/11.0.696.50 Safari/534.24" 55.89.141.209 - - [2018-08-01T06:54:29.360Z] "GET /apm-server/apm-server-6.3.2-windows-x86.zip HTTP/1.1" 2 00 4763 "-" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1; .NET CLR 1.1.4322)" 123.217.24.241 - - [2018-08-01T07:02:46.200Z] "GET /enterprise HTTP/1.1" 200 4460 "-" "Mozilla/5.0 (X11; L inux i686) AppleWebKit/534.24 (KHTML, like Gecko) Chrome/11.0.696.50 Safari/534.24" 243.236.31.15 - - [2018-08-01T07:24:41.231Z] "GET /styles/semantic-ui.css HTTP/1.1" 200 6381 "-" "Mozilla/ 4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1; .NET CLR 1.1.4322)" 70.133.115.149 - - [2018-08-01T07:29:39.275Z] "GET /elasticsearch/elasticsearch-6.3.2.zip HTTP/1.1" 200 29 62 "-" "Mozilla/5.0 (X11; Linux i686) AppleWebKit/534.24 (KHTML, like Gecko) Chrome/11.0.696.50 Safari/53 4.24" 19.112.190.54 - - [2018-08-01T07:30:13.527Z] "GET /people/type:astronauts/name:guy-gardner/profile HTTP/1. 1" 200 5615 "-" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1; .NET CLR 1.1.4322)" 110.248.232.101 - - [2018-08-01T07:43:19.649Z] "GET /beats/filebeat HTTP/1.1" 200 9268 "-" "Mozilla/4.0 (c o	03 AM					
Aug 21, 2019 @ 13:52:18.588								
Aug 21, 2019 @ 13:57:34.742								
Aug 21, 2019 @ 13:58:42.764								
Aug 21, 2019 @ 14:00:16.260								
Aug 21, 2019 @ 14:32:33.368								
Aug 21, 2019 @ 14:33:59.193								
Aug 21, 2019 @ 14:39:09.880								
Aug 21, 2019 @ 14:54:29.360								
Aug 21, 2019 @ 15:02:46.200								
Aug 21, 2019 @ 15:24:41.231								
Aug 21, 2019 @ 15:29:39.275								
Aug 21, 2019 @ 15:30:13.527								
Aug 21, 2019 @ 15:43:19.649								

# Infrastructure

Inventory Metrics Explorer

Search for infrastructure data... (e.g. host.name:host-1)

09/05/2019 8:15:32 PM

Auto-refresh

Hosts

Kubernetes

Docker

Metric: CPU Usage

Group By: All

Configuration

Map View

Table View

Showing the last 1 minute of data from the time period

- [View logs](#)
- [View metrics](#)
- [View container APM traces](#)
- [View container in Uptime](#)

keen\_mon...

0%

sync

0.5%

ltp-0

0%

tiweb-0

0%

esafe-core

0.3%

esafe-no...

0.1%

transmissi...

0.2%

fastdfs-0

0.1%

mongo-0

0.4%

esafe-py...

0.1%

autotopic-0

0%

mobweb-0

0%

heritrix-0

0.4%

syncserve...

0.1%

es-0

0.4%

# Uptime

D Uptime

Last 15 minutes Show dates Refresh

## Overview

Up Down Location ID Name URL Port Scheme

### Current status

Up 3 Down 0 Total 3

### Pings over time

Time	Pings
20:23	11
20:24	11
20:25	9
20:26	11
20:27	11
20:28	9
20:29	11
20:30	11
20:31	9
20:32	11
20:33	11
20:34	9
20:35	11
20:36	11
20:37	11

### Monitor status

Status	Name	URL	Downtime history	Integrations
Up a few seconds ago	Unnamed - auto-http-0XC5358829EE16F210-4428eac569fc9423	http://10.0.0.4:5000/api/membership_profile ↗		...
Up a few seconds ago	Unnamed - auto-http-0XC5358829EE16F210-64c0df773eb00927	http://10.0.0.4:8088/topiclist/topics ↗		...
Up a few seconds ago	Unnamed - auto-http-0XC5358829EE16F210-ed18b544cc9262fb	http://10.0.0.4:5000/api/hello ↗		...

Elastic Stack  
需要 融入开放生态



fluentd



CloudBees



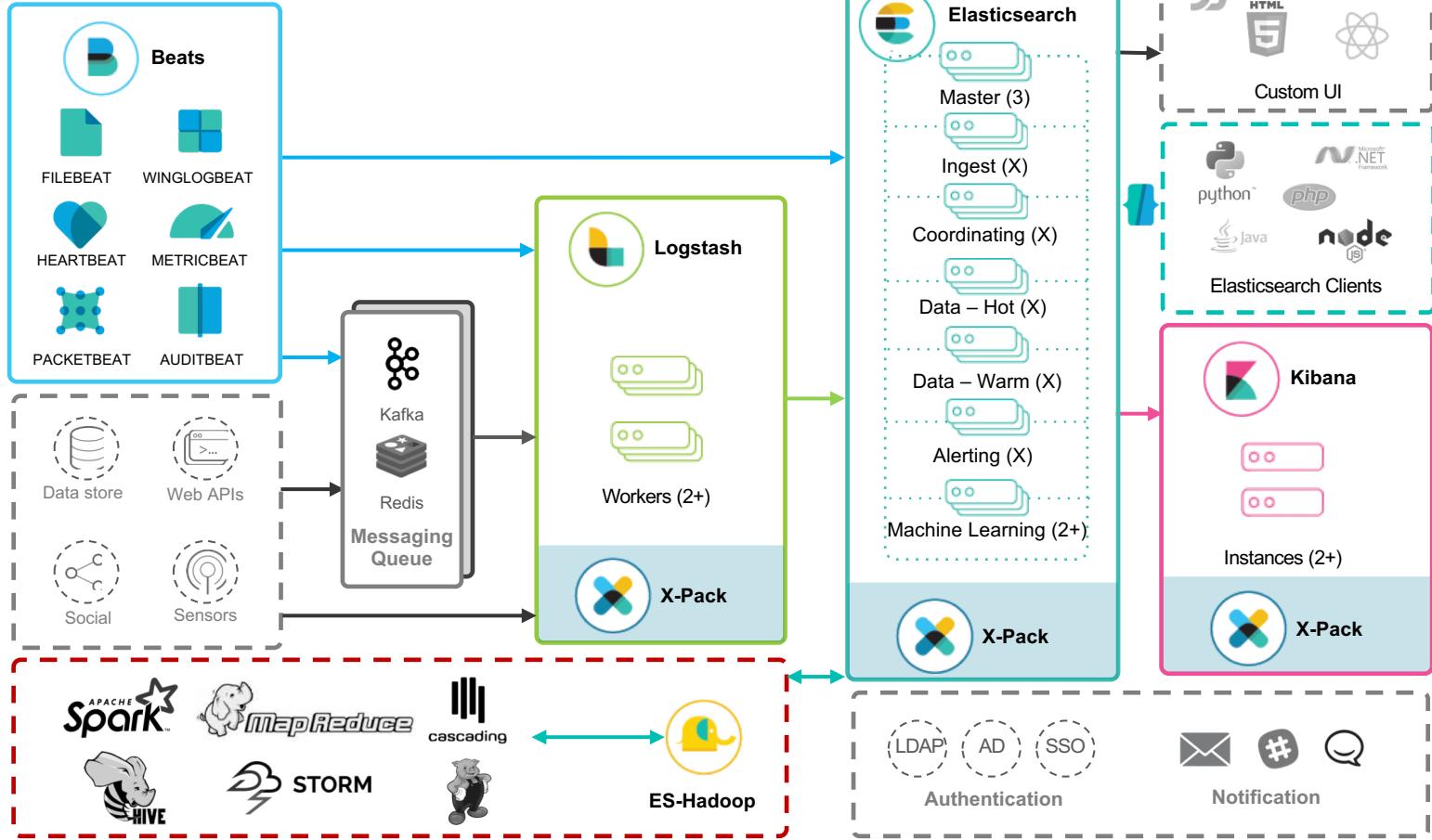
Apache Flink



Elastic

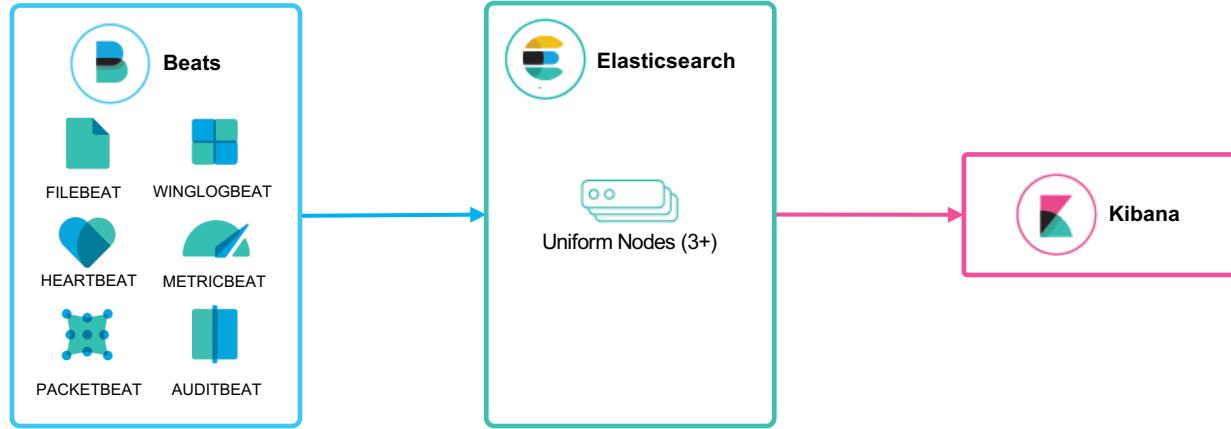


# ES和Hadoop配合很流行

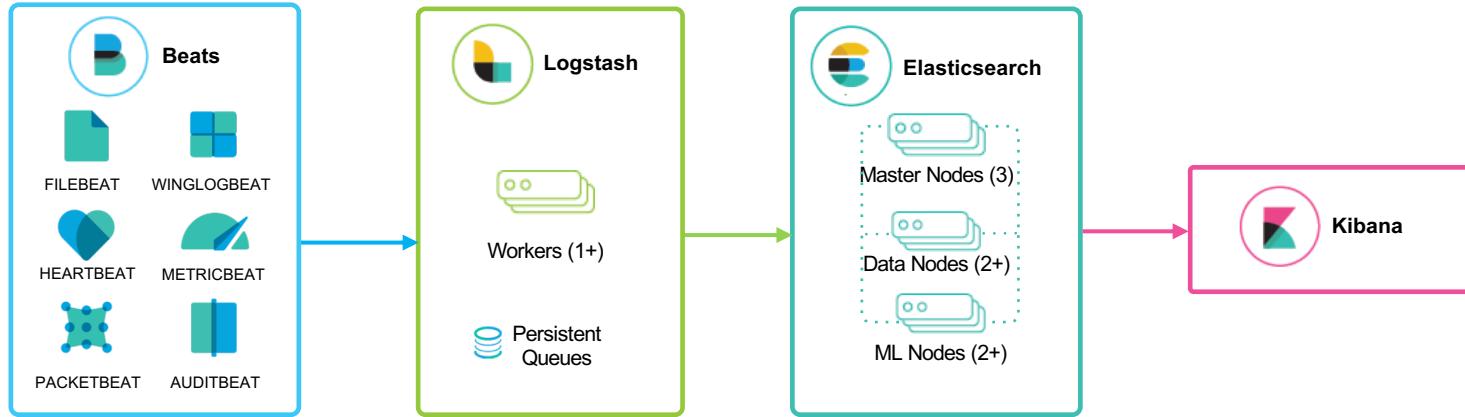


# 日志指标APM 整体架构

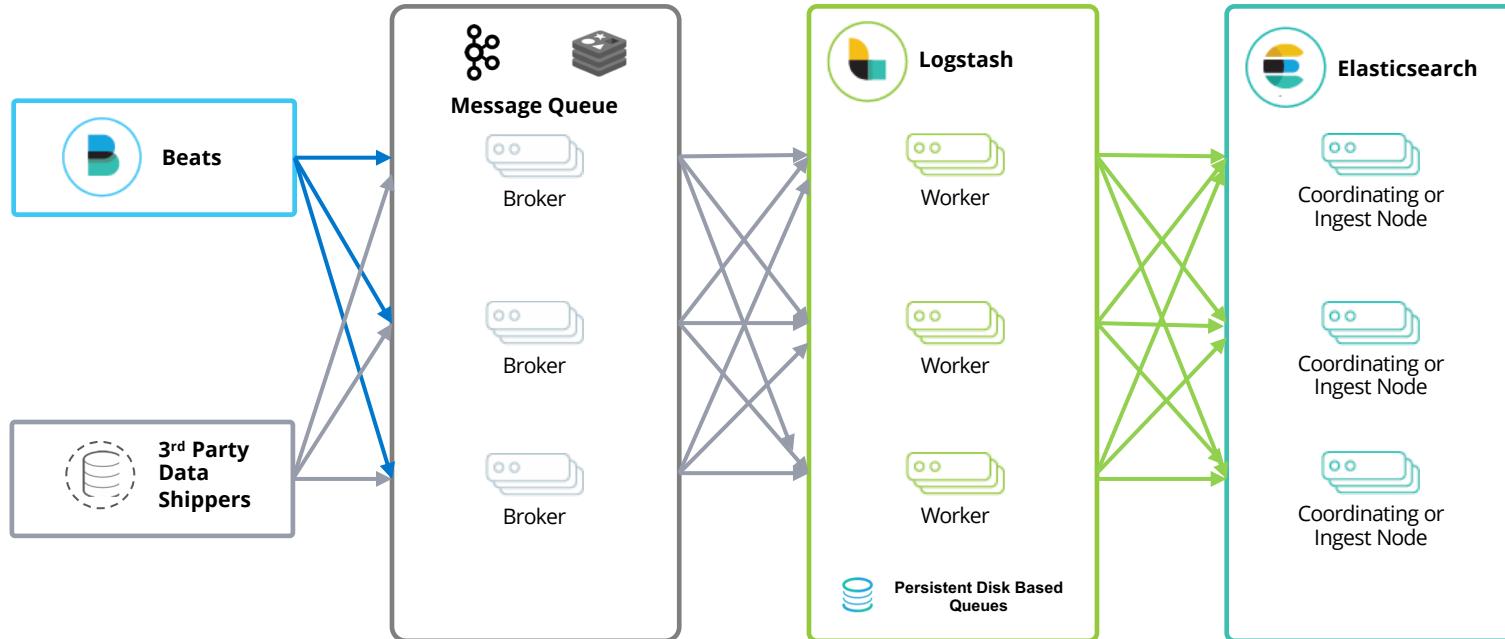
# 整体架构图



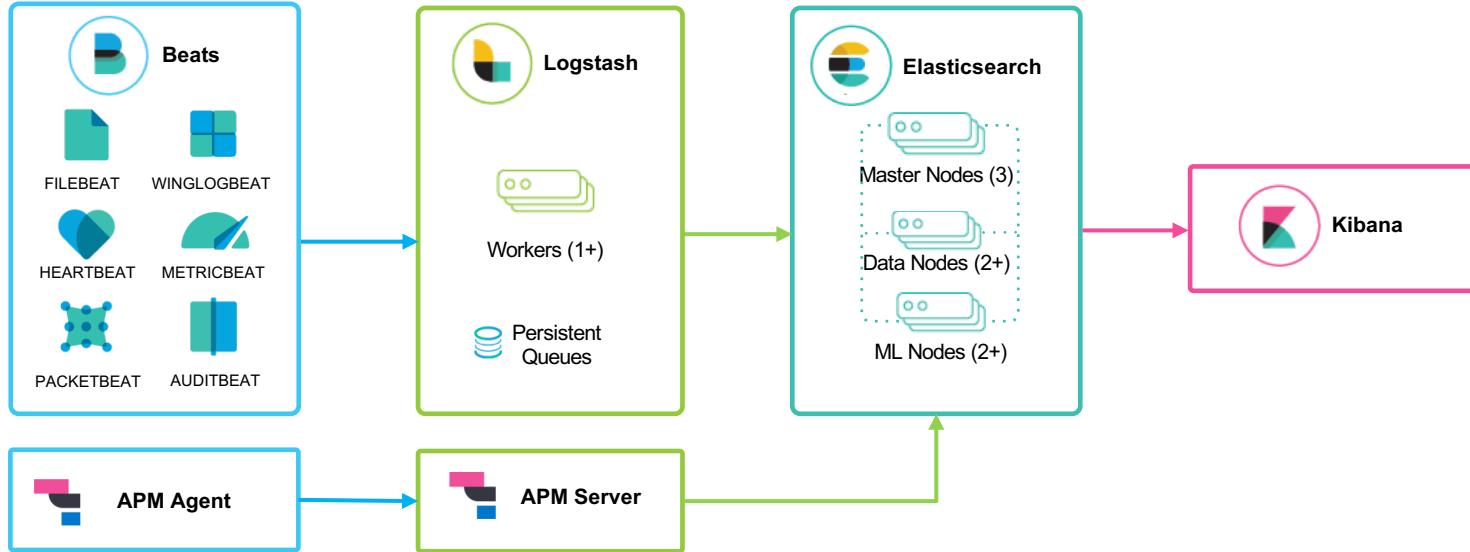
# 整体架构图



# 更复杂一些的架构图



# 整体架构图



# Filebeat使用详解

# Filebeat详解

 data	▶
 logs	▶
 README.md	
 filebeat	
 modules.d	▶
 module	▶
 kibana	▶
 filebeat.yml	
 filebeat.reference.yml	
 fields.yml	
 NOTICE.txt	
 LICENSE.txt	

模块开关，配置

模块文件

Kibana Dashboard模板

Filebeat 配置文件

# Filebeat module作用

对一些常用日志格式做到开箱即用

- 缺省的配置
- Elasticsearch pipeline 定义
- Kibana Dashboard 配置

# Filebeat命令行

```
./filebeat --help
```

```
Usage:  
  filebeat [flags]  
  filebeat [command]  
  
Available Commands:  
  enroll      Enroll in Kibana for Central Management  
  export       Export current config or index template  
  generate    Generate Filebeat modules, filesets and fields.yml  
  help        Help about any command  
  keystore    Manage secrets keystore  
  modules     Manage configured modules  
  run         Run filebeat  
  setup       Setup index template, dashboards and ML jobs  
  test        Test config  
  version    Show current version info
```

# Filebeat命令行举例

```
./filebeat modules list
```

```
./filebeat modules enable system
```

```
./filebeat modules disable system
```

# Filebeat重要配置

```
filebeat.inputs:
- type: log

  # Change to true to enable this input configuration.
  enabled: false

  # Paths that should be crawled and fetched. Glob based paths.
  paths:
    - /var/log/*.log

  # Optional additional fields. These fields can be freely picked
  # to add additional information to the crawled log files for filtering
  #fields:
  #  level: debug
  #  review: 1

setup.kibana:

  # Kibana Host
  # Scheme and port can be left out and will be set to the default (http and 5601)
  host: "localhost:5601"

  # Kibana Space ID
  # ID of the Kibana Space into which the dashboards should be loaded. By default,
  # the Default Space will be used.
  #space.id:
```

# Filebeat重要配置

```
output.elasticsearch:  
  # Array of hosts to connect to.  
  hosts: ["localhost:9200"]  
  
  # Optional protocol and basic auth credentials.  
  #protocol: "https"  
  #username: "elastic"  
  #password: "changeme"
```

```
output.logstash:  
  # The Logstash hosts  
  #hosts: ["localhost:5044"]
```

# Lab1 Filebeat Metricbeat使用

# APM 使用详解

# APM

统一 日志 , 指标 , APM

## Open Source

### 各种语言客户端

Java, .Net, Go, RUM, Node, Python, Ruby, 等等.

### 有专门UI

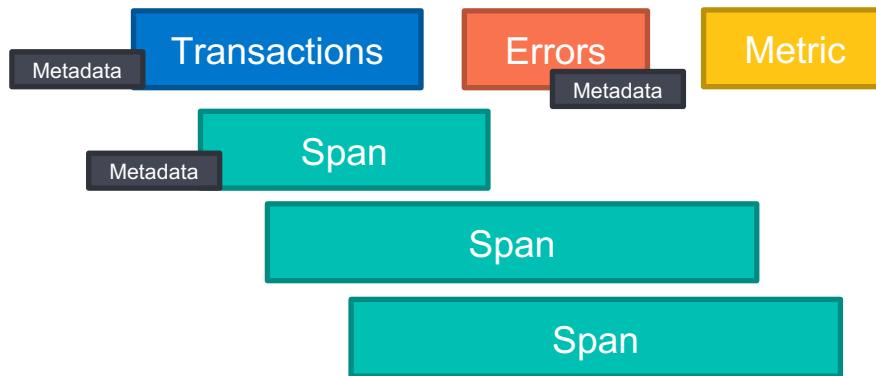
APM 工作流程耗时图  
分布式的追踪

### 仅仅多了一些索引

和其它数据关联  
借助于ES的功能



# APM Date model



# APM Agent 举例 java agent

加载 java apm agent的三种方式

方式一

```
java -javaagent:/path/to/elastic-apm-agent-<version>.jar
```

方式二

```
java -jar apm-agent-attach-standalone.jar --pid 42
```

方式三

```
import co.elastic.apm.attach.ElasticApmAttacher;
import org.springframework.boot.SpringApplication;

@SpringBootApplication
public class MyApplication {
    public static void main(String[] args) {
        ElasticApmAttacher.attach();
        SpringApplication.run(MyApplication.class, args);
    }
}
```



# Java Agent Supported Technologies

web-frameworks	Servlet API, Spring Web MVC, Spring Boot, JAX-RS, JAX-WS
Application Servers Servlet Containers	Tomcat, Jboss EAP, Jetty, WebSphere, WebLogic, WildFly, ...
Data Stores	JDBC, Elasticsearch, Hibernate Search
Networking frameworks	Apache HttpClient, Spring RestTemplate, OkHttp ...
Asynchronous frameworks	ExecutorService
Messaging frameworks	JMS
Scheduling frameworks	Scheduling Annotation, Quartz
Logging frameworks	slf4j

不支持的可以通过API来创建 Transactions 和 Spans

# Java Agent Metric

System Metric	JVM Metric
system.cpu.total.norm.pct	jvm.memory.heap.used
system.process.cpu.total.norm.pct	jvm.memory.heap.committed
system.memory.total	jvm.memory.heap.max
system.memory.actual.free	jvm.memory.non_heap.used
system.process.memory.size	jvm.memory.non_heap.committed
...	...

# Java Agent 配置举例

```
service_name=escafe-core
application_packages=org.cafe.web.controller
server_urls=http://10.0.0.2:8200
trace_methods=com.cafe.web.controller.OrderReceiver#process
```

## trace\_methods 举例

```
org.example.* [1.4.0]Added in 1.4.0. Omitting the method is possible since 1.4.0
org.example.*#* (before 1.4.0, you need to specify a method matcher)
org.example.MyClass#myMethod
org.example.MyClass#myMethod()
org.example.MyClass#myMethod(java.lang.String)
org.example.MyClass#myMe*od(java.lang.String, int)
private org.example.MyClass#myMe*od(java.lang.String, *)
* org.example.MyClas*#myMe*od(*.String, int[])
public org.example.services.*Service#*
public @java.inject.ApplicationScoped org.example.*
public @java.inject.* org.example.*
public @@javax.enterprise.context.NormalScope org.example.*
```

# APM Agent 举例 python agent

```
import elasticapm
from elasticapm.contrib.flask import ElasticAPM

app = Flask(__name__)

app.config['ELASTIC_APM'] = {
    'SERVICE_NAME': 'escafe-python',
    'SERVER_URL': config['elasticapm.url'],
    'INSTRUMENT': config['elasticapm.instrument']
}
apm = ElasticAPM(app)

@app.route("/api/hello")
def hello():
    do_calculation()
    return "Hello World!"

@elasticapm.capture_span()
def do_calculation():
    if avg_service_time != 0:
        delay = abs(gauss(avg_service_time, stdev_service_time))
        time.sleep(delay)
    return
```

# Logstash 使用详解

# 日志、指标采集 Logstash 生态

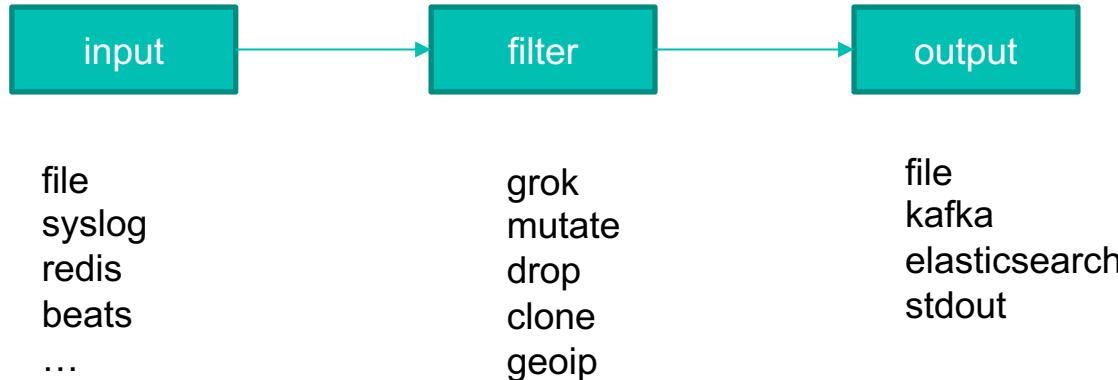


Logstash

200 + 插件

azure event hub	cloudwatch	couchdb	elasticsearch	exec
file	websocket	tcp	upd	graphite
http	http poller	imap	irc	java generator
java stdin	jdbc	jms	jmx	kafka
log4j	xmpp	redis	rss	s3
snmp	sqlite	stdin	syslog	github
更多 社区制造				

# Logstash 工作原理



# Logstash配置文件基本结构

```
input {
  file {
    path => "/tmp/access_log"
    start_position => "beginning"
  }
}

filter {
  if [path] =~ "access" {
    mutate { replace => { "type" => "apache_access" } }
    grok {
      match => { "message" => "%{COMBINEDAPACHELOG}" }
    }
  }
  date {
    match => [ "timestamp" , "dd/MMM/yyyy:HH:mm:ss Z" ]
  }
}

output {
  elasticsearch {
    hosts => ["localhost:9200"]
  }
  stdout { codec => rubydebug }
}
```

# Multiple Pipelines

pipelines.yml

```
- pipeline.id: my-pipeline_1
  path.config: "/etc/path/to/p1.config"
  pipeline.workers: 3

- pipeline.id: my-other-pipeline
  path.config: "/etc/different/path/p2.cfg"
  queue.type: persisted
```

# grok基础

```
55.3.244.1 GET /index.html 15824 0.043
```

```
%{IP:client} %{WORD:method} %{URIPATHPARAM:request} %{NUMBER:bytes} %{NUMBER:duration}
```

内置pattern定义 <https://github.com/elastic/logstash/tree/v1.4.2/patterns>

```
IP  (?::%{IPV6} | %{IPV4})
WORD \b\w+\b
URIPATHPARAM %{URIPATH} (?::%{URIPARAM})?
NUMBER (?::%{BASE10NUM})
BASE10NUM (?<! [0-9.+-]) (?>[+-]?(?:(?:(?:[0-9]+(?:\.[0-9]+)?)|(?:\.\.[0-9]+))))
```

```
grok {
  match => { "message" => ["%{SYSLOGTIMESTAMP:[system][syslog][timestamp]} %{{SYSLOGHOST:[system][syslog][hostname]}}
%{DATA:[system][syslog][program]} (?::\[%{POSINT:[system][syslog][pid]}\]\])?:
%{GREEDYMULTILINE:[system][syslog][message]}"]
    pattern_definitions => { "GREEDYMULTILINE" => "(.|\\n)*" }
    patterns_dir => ["./patterns"]
}

# contents of ./patterns/postfix:
POSTFIX_QUEUEID [0-9A-F]{10,11}
```

可以是数组

定义正则

正则文件

# grok测试器

The screenshot shows the Grok Debugger interface, which includes the following sections:

- Dev Tools**: A sidebar on the left containing icons for various developer tools.
- Console**: Shows the command line interface.
- Search Profiler**: Shows search profiler metrics.
- Grok Debugger**: The active tab.
- Sample Data**: Displays the input log line: 1 55.3.244.1 GET /index.html 15824 0.043.
- Grok Pattern**: Shows the grok pattern: 1 %{IP:client} %{WORD:method} %{URIPATHPARAM:request} %{NUMBER:bytes} %{NUMBER:duration}.
- Custom Patterns**: A link to view custom patterns.
- Simulate**: A large blue button.
- Structured Data**: Displays the simulated structured data as JSON:

```
1+ {  
2   "duration": "0.043",  
3   "request": "/index.html",  
4   "method": "GET",  
5   "bytes": "15824",  
6   "client": "55.3.244.1"  
7 }
```

# Lab2 Logstash使用

# Kibana可视化分析

# END