elastic{on} 2017

Machine Learning in the Elastic Stack

Steve Dodson, Tech Lead, Machine Learning
Sophie Chang, Team Lead, Machine Learning

Elastic

8th March 2017

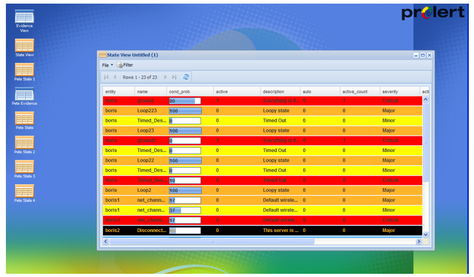
@prelertsteve

Agenda

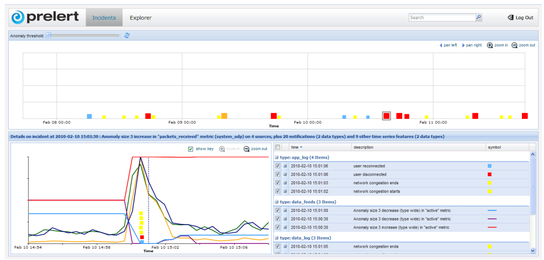
1. Background
2. Use Cases
3. Demos
4. Product Architecture and Status

Background

Prelert (founded late 2009), acquired by Elastic September 2016

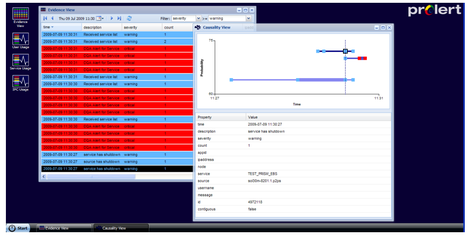


Prelert v0.9 2009-03



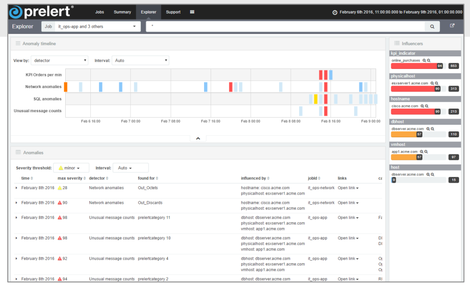
Prelert v3.0 2010-06

Prelert v5.4 2015-03

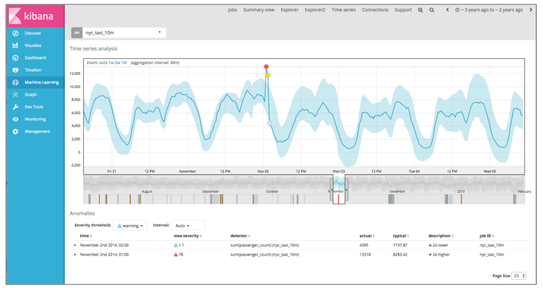


Prelert v1.0 2009-06

Prelert v3.6 2010-06



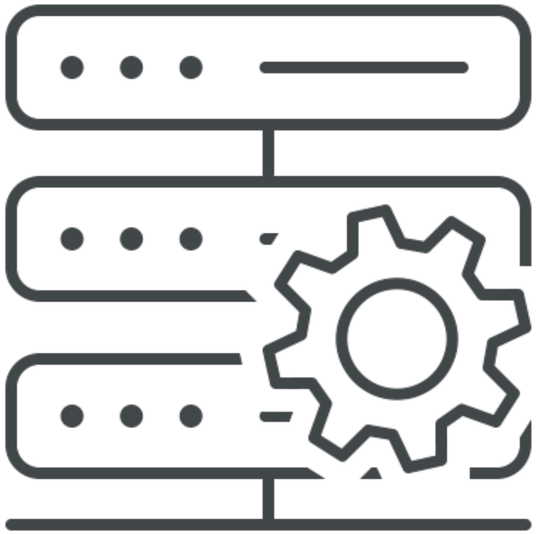
Prelert v6.1 Elastic{ON} 2016



Elastic X-Pack 5.4.0-SNAPSHOT
Elastic{ON} 2017

IT Operations

- How do I know my systems are behaving normally?
- Where to set thresholds for good alerting?
- How to find the root cause of problems?

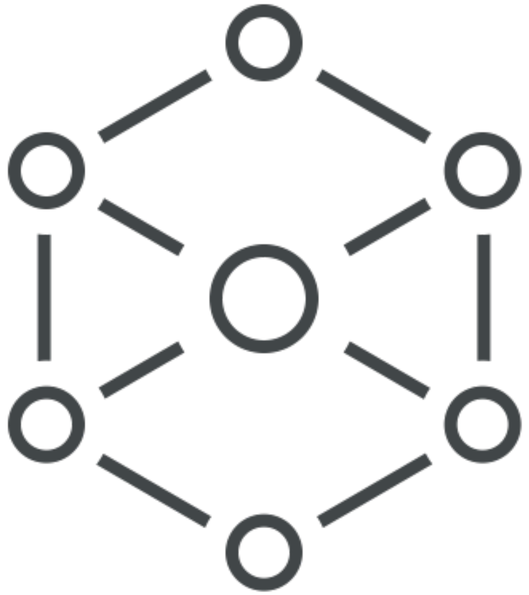


IT Security

- Do I have systems that are compromised with malware?
- Which users could be an insider threat?

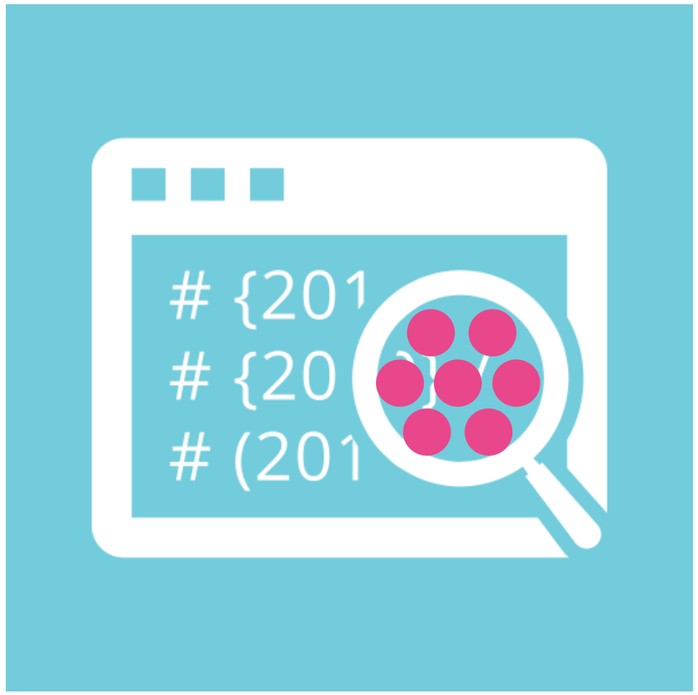
Other

- Is my factory working normally?
- What do I do with thousands of time-series data?
- Which traffic incidents are causing the most delay?



Extracting useful, valuable information is hard

Extracting useful, valuable information is hard



Search

Aggregations

Visualization

Machine Learning

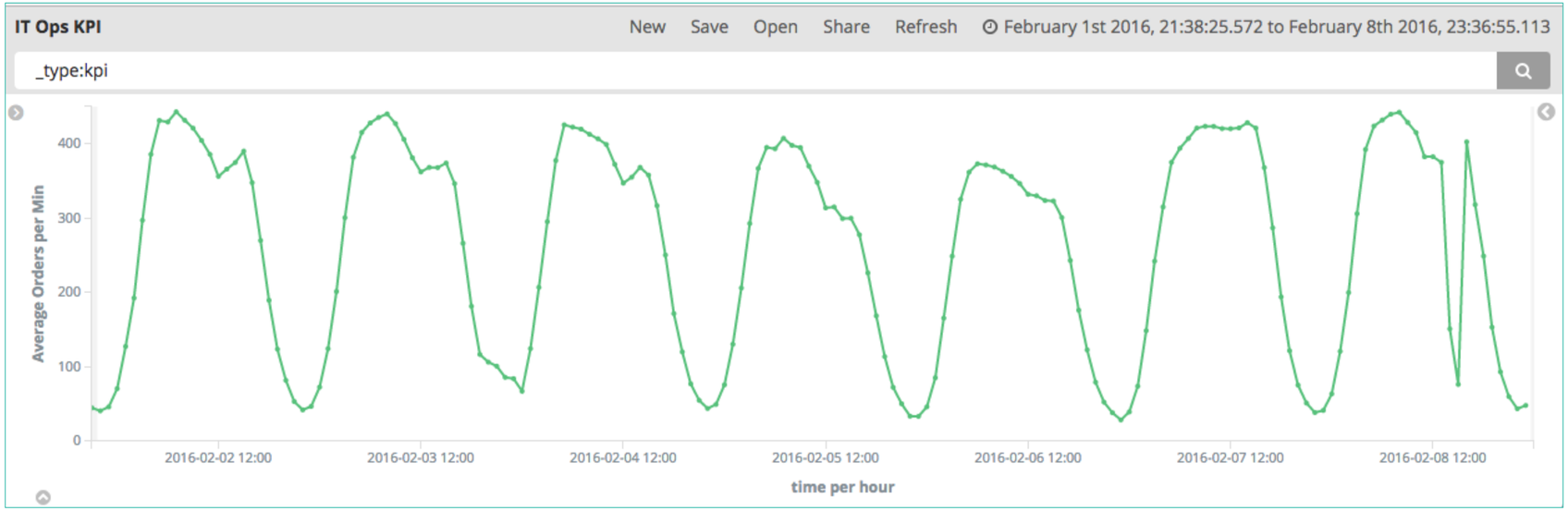
Machine Learning

- Algorithms and methods for data driven prediction, decision making, and modelling[1]

- Examples
  - Image Recognition
  - Language Translation
  - Anomaly Detection

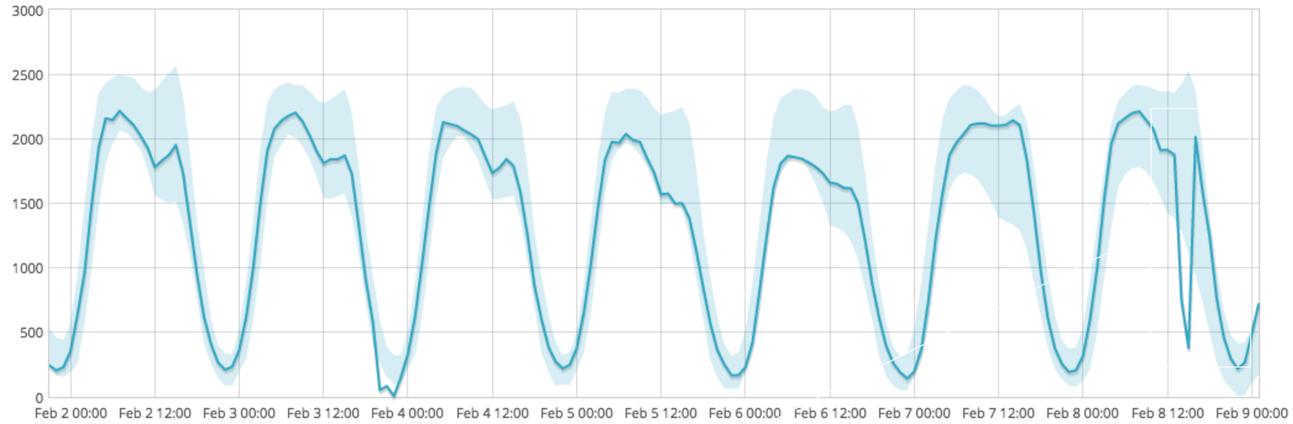
[1]Machine Learning Overview, Tommi Jaakkola, MIT

Has my order rate dropped significantly?

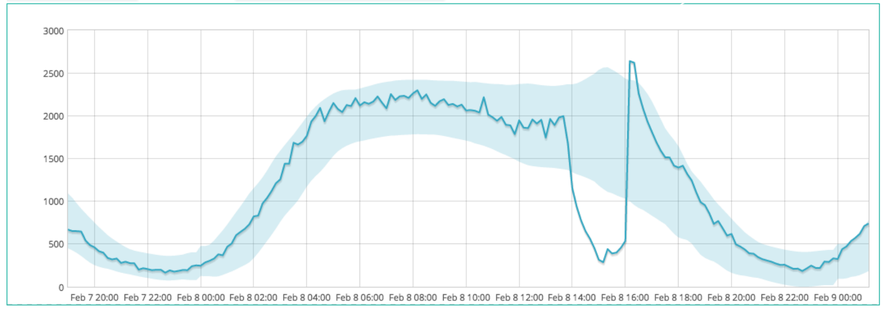


Has my order rate dropped significantly?

- Learn models from past behaviour (training, modelling)

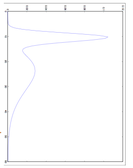


- Use models to predict future behaviour (prediction)

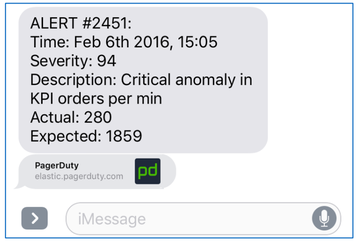


Actual value @ 15:05 = 280

Probability = 0.0000174025

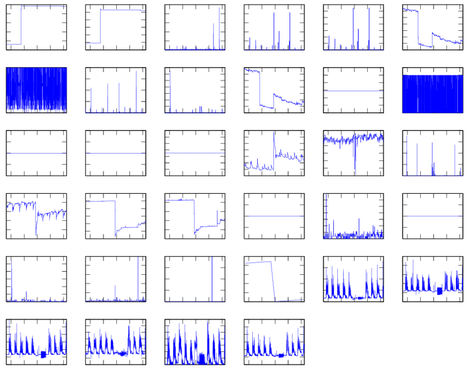


- Use predictions to make decisions

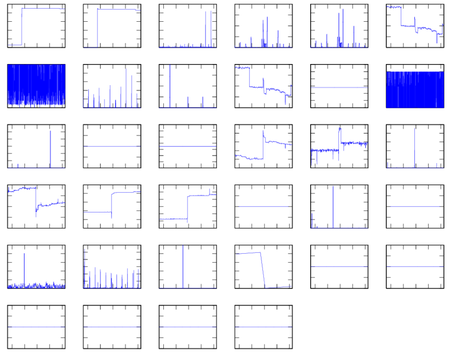


Demo: Simple Time Series

Has my system changed behaviour?

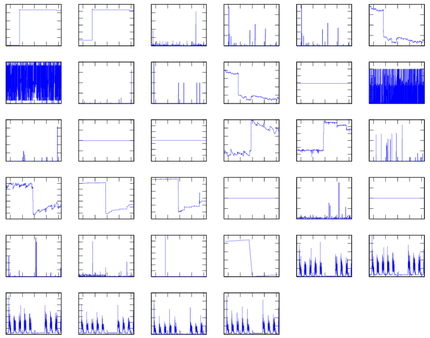


i-5cfd3dcb

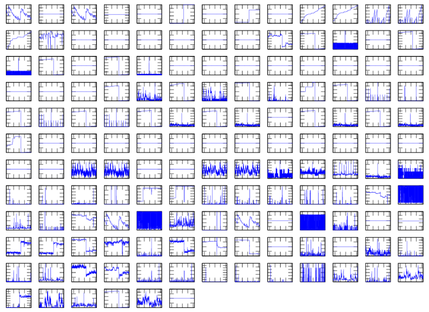


i-f1e94994

...



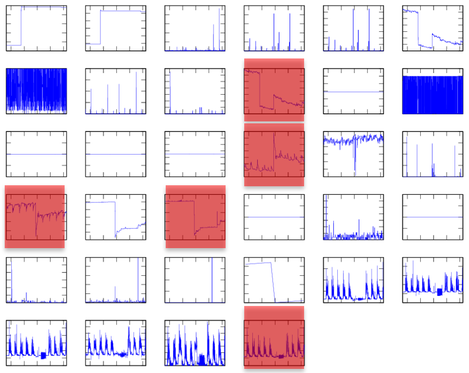
i-ece626ff



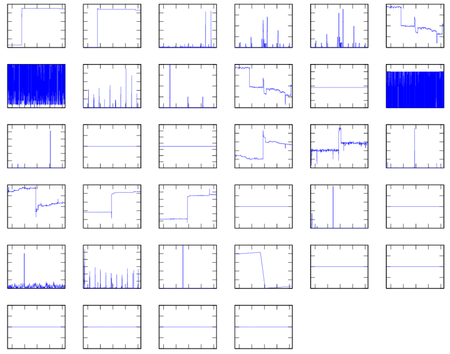
i-ebc323df

...

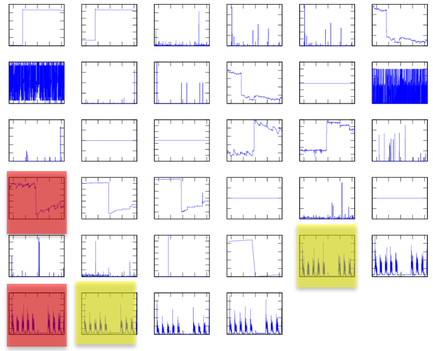
Has my system changed behaviour?



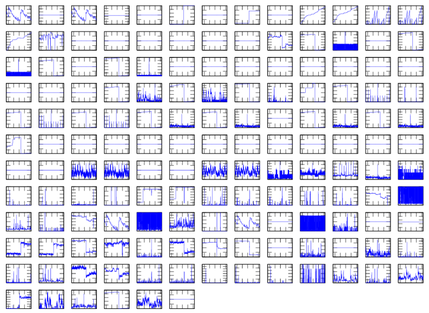
i-5cfd3dcb



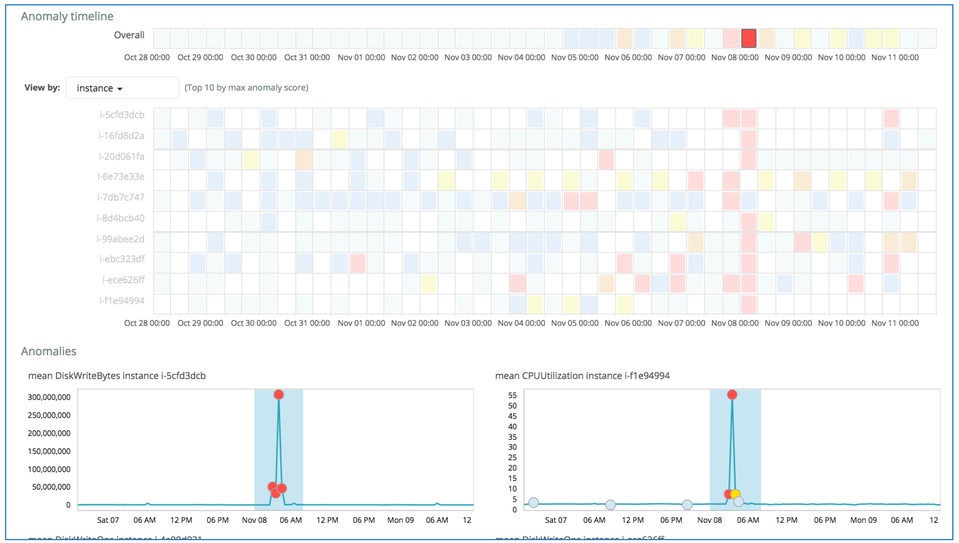
i-f1e94994



i-ece626ff

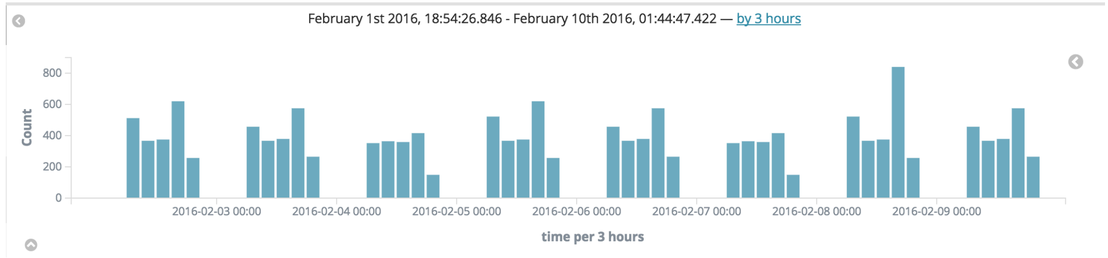


i-ebc323df



Demo: Multiple Time Series

Do my application logs contain unusual messages?

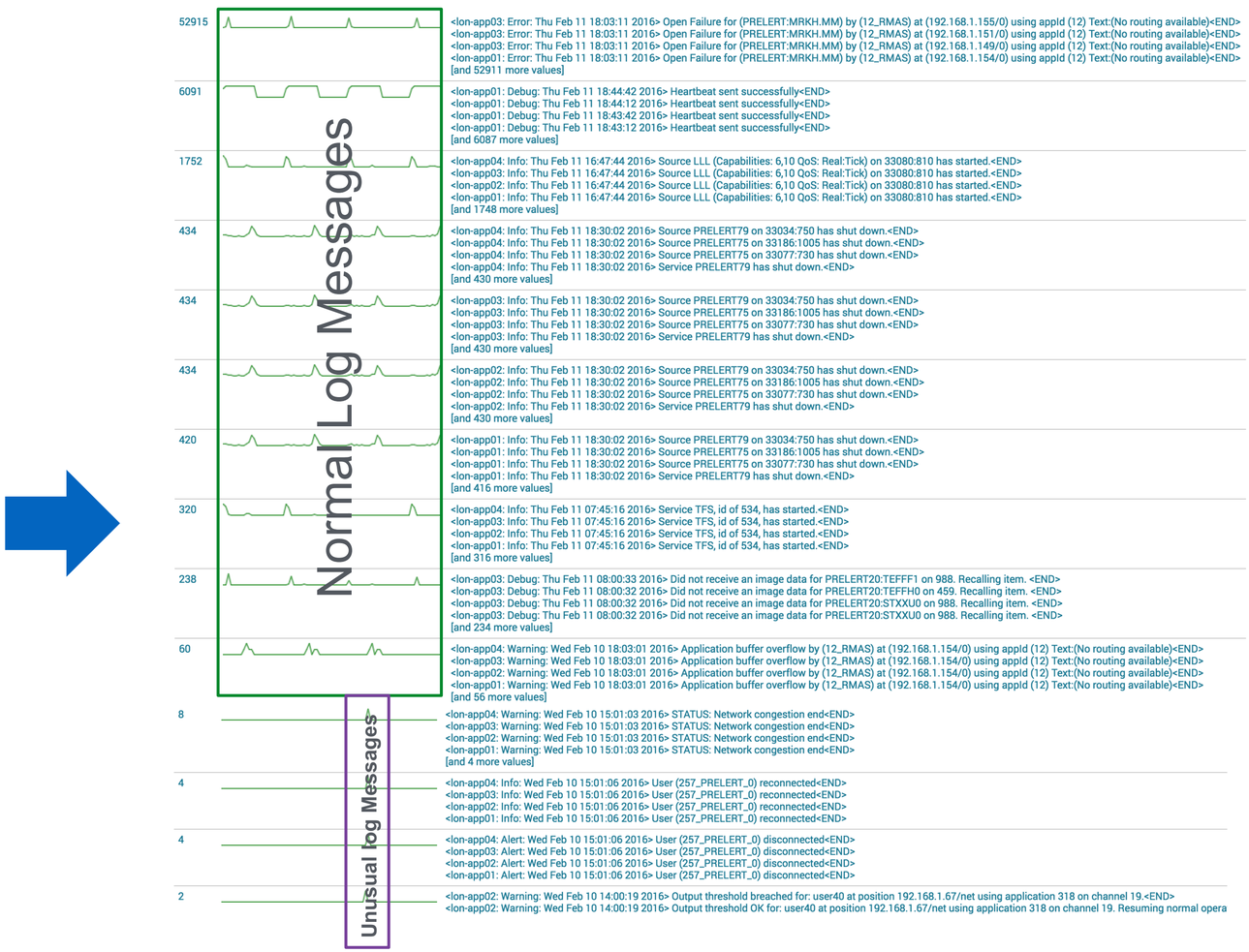


February 9th 2016, 18:03:00.000	Heartbeat sent successfully
February 9th 2016, 18:02:30.000	Heartbeat sent successfully
February 9th 2016, 18:02:00.000	Heartbeat sent successfully
February 9th 2016, 18:01:30.000	Heartbeat sent successfully
February 9th 2016, 18:01:00.000	Heartbeat sent successfully
February 9th 2016, 18:00:30.000	Heartbeat sent successfully
February 9th 2016, 18:00:02.000	Source PRELERT72 on 33034:952 has shut down.
February 9th 2016, 18:00:02.000	Source PRELERT72 on 33034:952 has shut down.
February 9th 2016, 18:00:02.000	Source PRELERT72 on 33034:952 has shut down.
February 9th 2016, 18:00:02.000	Service PRELERT72 has shut down.
February 9th 2016, 18:00:02.000	Service PRELERT72 has shut down.
February 9th 2016, 18:00:02.000	Service PRELERT72 has shut down.
February 9th 2016, 18:00:02.000	Source PRELERT72 on 33034:952 has shut down.
February 9th 2016, 18:00:02.000	Service PRELERT72 has shut down.
February 9th 2016, 18:00:01.000	Source PRELERT47 on 33077:998 has shut down.
February 9th 2016, 18:00:01.000	Source PRELERT72 on 33078:953 has shut down.

Do my application logs contain unusual messages?

Classify unstructured log messages by clustering similar messages

February 9th 2016, 18:03:00.000	Heartbeat sent successfully
February 9th 2016, 18:02:30.000	Heartbeat sent successfully
February 9th 2016, 18:02:00.000	Heartbeat sent successfully
February 9th 2016, 18:01:30.000	Heartbeat sent successfully
February 9th 2016, 18:01:00.000	Heartbeat sent successfully
February 9th 2016, 18:00:30.000	Heartbeat sent successfully
February 9th 2016, 18:00:02.000	Source PRELERT72 on 33034:952 has shut down.
February 9th 2016, 18:00:02.000	Source PRELERT72 on 33034:952 has shut down.
February 9th 2016, 18:00:02.000	Source PRELERT72 on 33034:952 has shut down.
February 9th 2016, 18:00:02.000	Service PRELERT72 has shut down.
February 9th 2016, 18:00:02.000	Service PRELERT72 has shut down.
February 9th 2016, 18:00:02.000	Service PRELERT72 has shut down.
February 9th 2016, 18:00:02.000	Source PRELERT72 on 33034:952 has shut down.
February 9th 2016, 18:00:02.000	Service PRELERT72 has shut down.
February 9th 2016, 18:00:01.000	Source PRELERT47 on 33077:998 has shut down.
February 9th 2016, 18:00:01.000	Source PRELERT72 on 33078:953 has shut down.

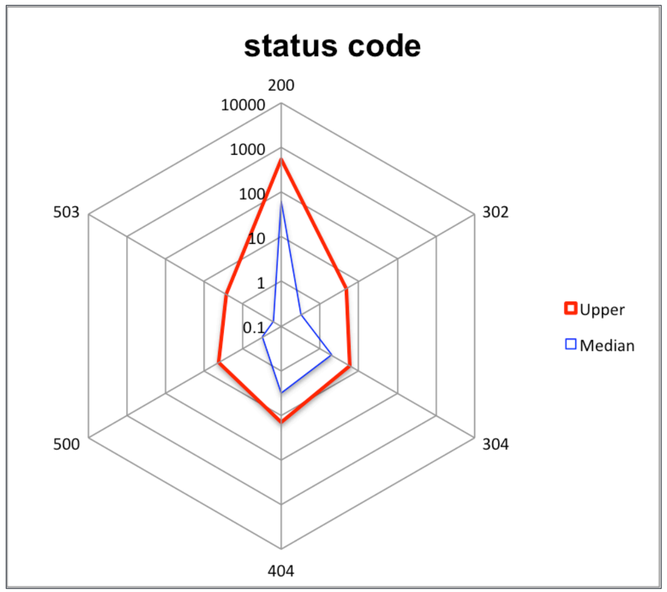


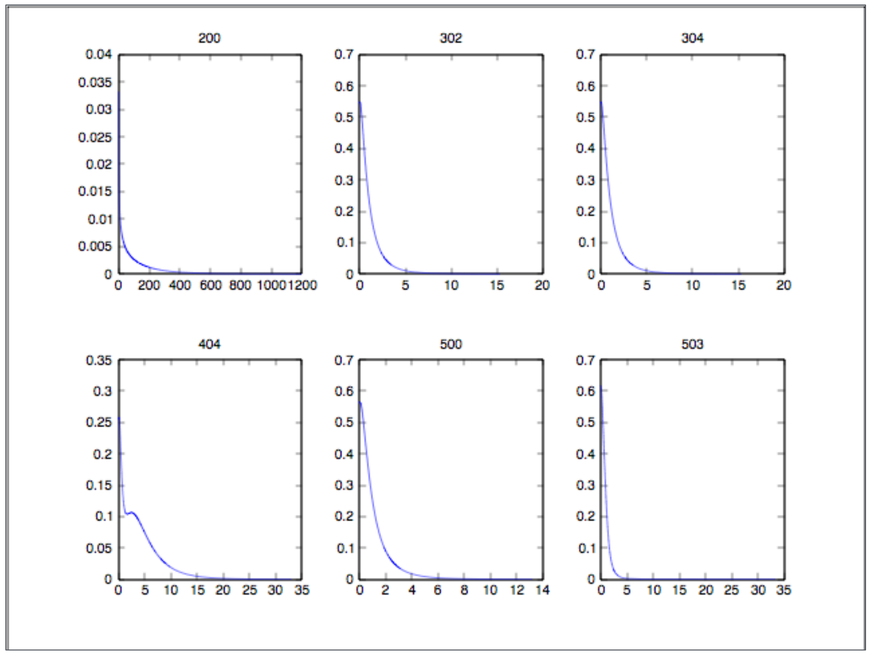
Demo: Log Messages

Entity Profiling

```
10.12.211.69 - - [01/Jan/2016:00:07:21 +0000] "GET /css/ccc_style.jsp HTTP/1.1" 200 19196 "https://www.prelertstation.com/" "Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.9.0.5) Gecko/2008120122 Firefox/3.0.5"
10.12.211.69 - - [01/Jan/2016:00:07:22 +0000] "GET /js/openWin.js HTTP/1.1" 200 2272 "https://www.prelertstation.com/" "Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.9.0.5) Gecko/2008120122 Firefox/3.0.5"
10.12.211.69 - - [01/Jan/2016:00:07:22 +0000] "GET /css/themes/ HTTP/1.1" 404 988 "https://www.prelertstation.com/" "Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.9.0.5) Gecko/2008120122 Firefox/3.0.5"
```

- Create ‘profile’ of status code responses for a typical client:

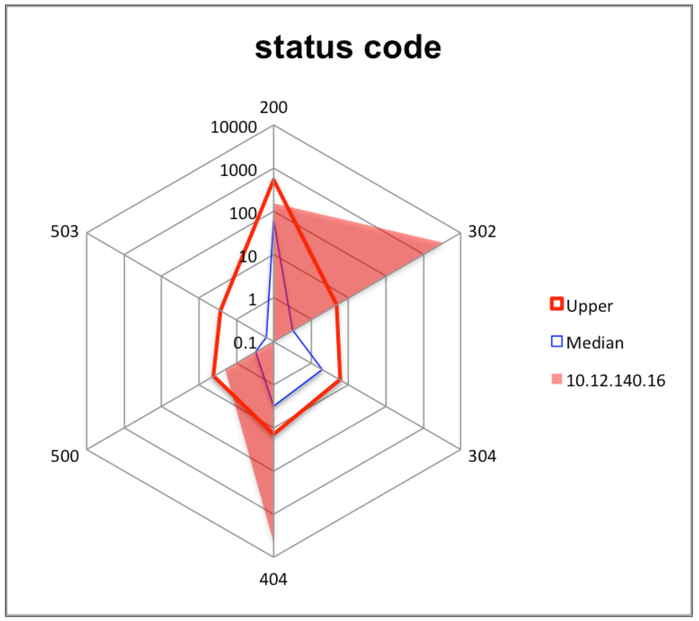




Entity Profiling

```
10.12.211.69 - - [01/Jan/2016:00:07:21 +0000] "GET /css/ccc_style.jsp HTTP/1.1" 200 19196 "https://www.prelertstation.com/" "Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.9.0.5) Gecko/2008120122 Firefox/3.0.5"
10.12.211.69 - - [01/Jan/2016:00:07:22 +0000] "GET /js/openWin.js HTTP/1.1" 200 2272 "https://www.prelertstation.com/" "Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.9.0.5) Gecko/2008120122 Firefox/3.0.5"
10.12.211.69 - - [01/Jan/2016:00:07:22 +0000] "GET /css/themes/ HTTP/1.1" 404 988 "https://www.prelertstation.com/" "Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.9.0.5) Gecko/2008120122 Firefox/3.0.5"
```

- Create ‘profile’ of status code responses for a typical client:



time	max severity	detector	found for
January 23rd 2016, 16:00	99	count	10.12.140.16

Description:
unknown anomaly in count found for clientip 10.12.140.16

Details on highest severity anomaly:

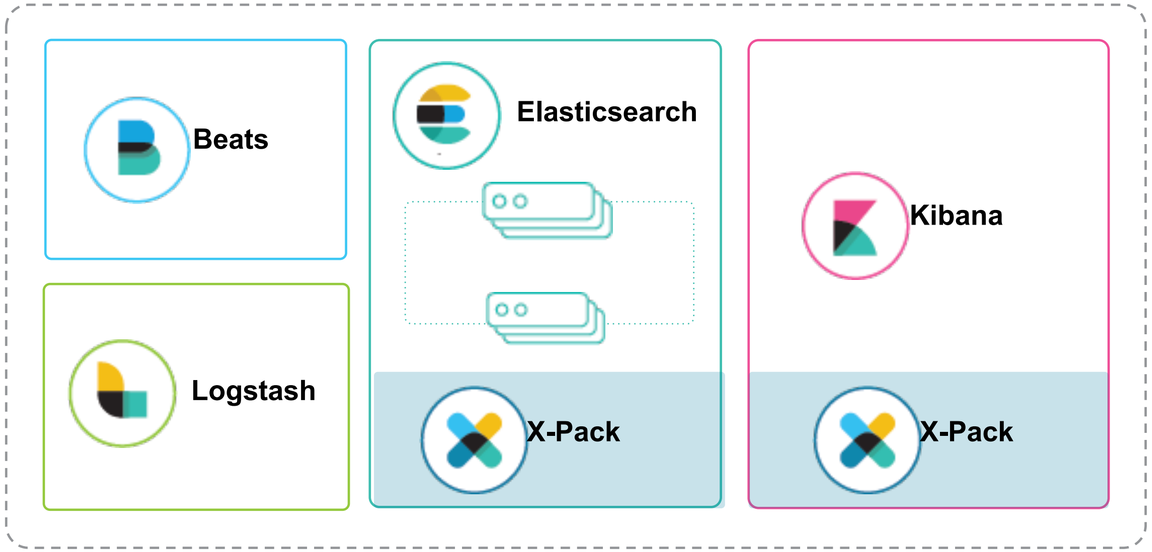
clientip:	10.12.140.16
time:	January 23rd 2016, 16:00:00 to January 23rd 2016, 17:00:00
function:	high_count
job ID:	access_logs
probability:	1.16529e-43
status values:	404 (actual 4635, typical 4.17792, probability 2.79981e-29)
	302 (actual 3502, typical 1.3176, probability 9.45046e-22)

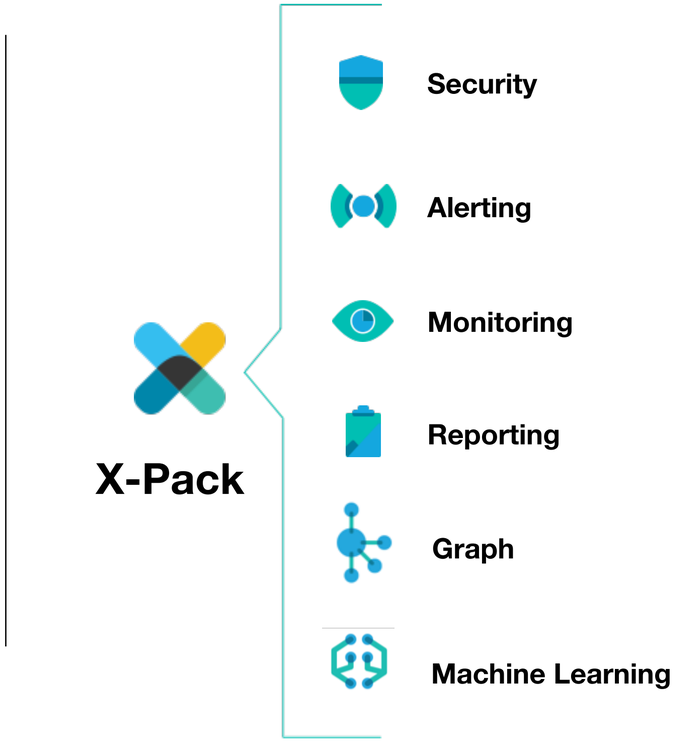
Influenced by:

clientip	10.12.140.16

Demo: Entity Profiling

Elastic Stack





- Single install - deployed with X-Pack
- Data gravity - analyzes data from the same cluster
- Contextual - anomalies and data stored together
- Scalable - jobs distributed across nodes
- Resilient - handles node failure

elastic

X-Pack

Platinum
Single install

Security

Alerting

Monitoring

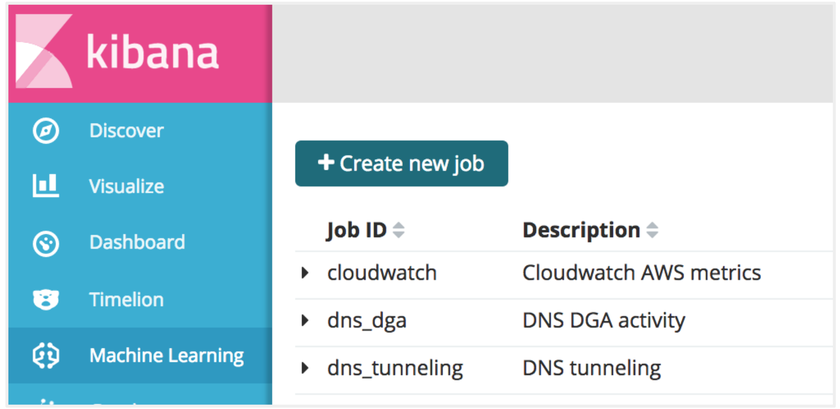
Reporting

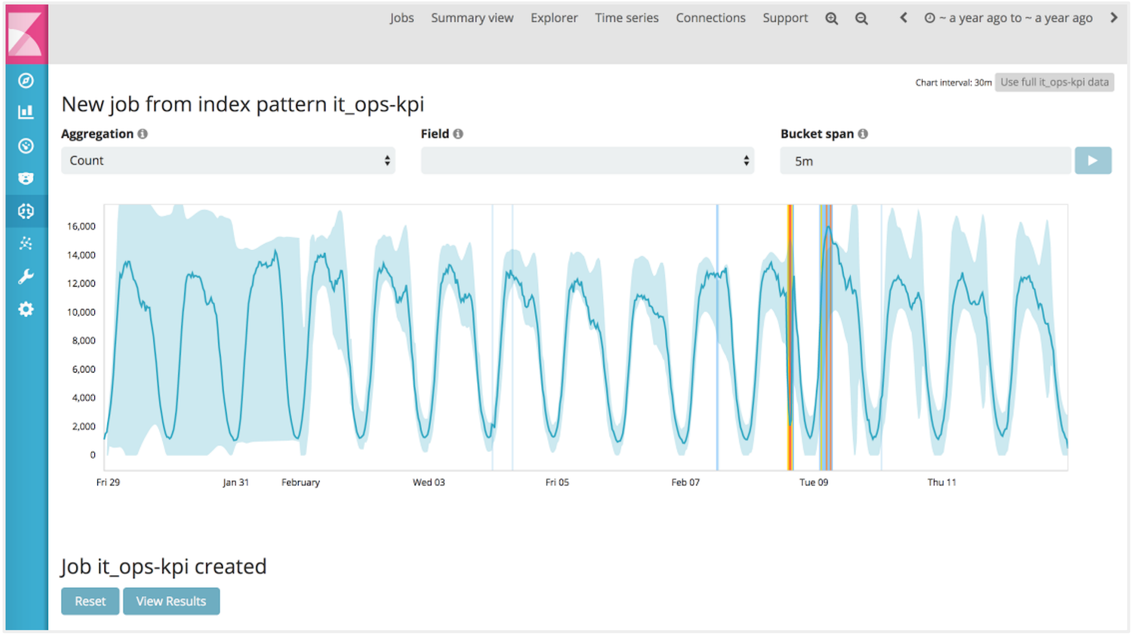
Graph

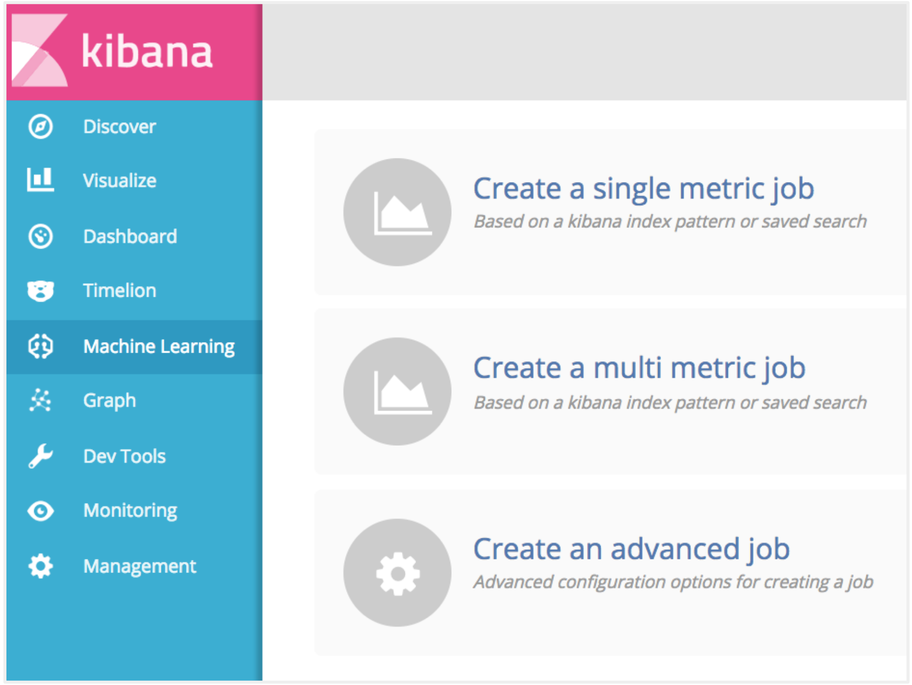
Machine Learning

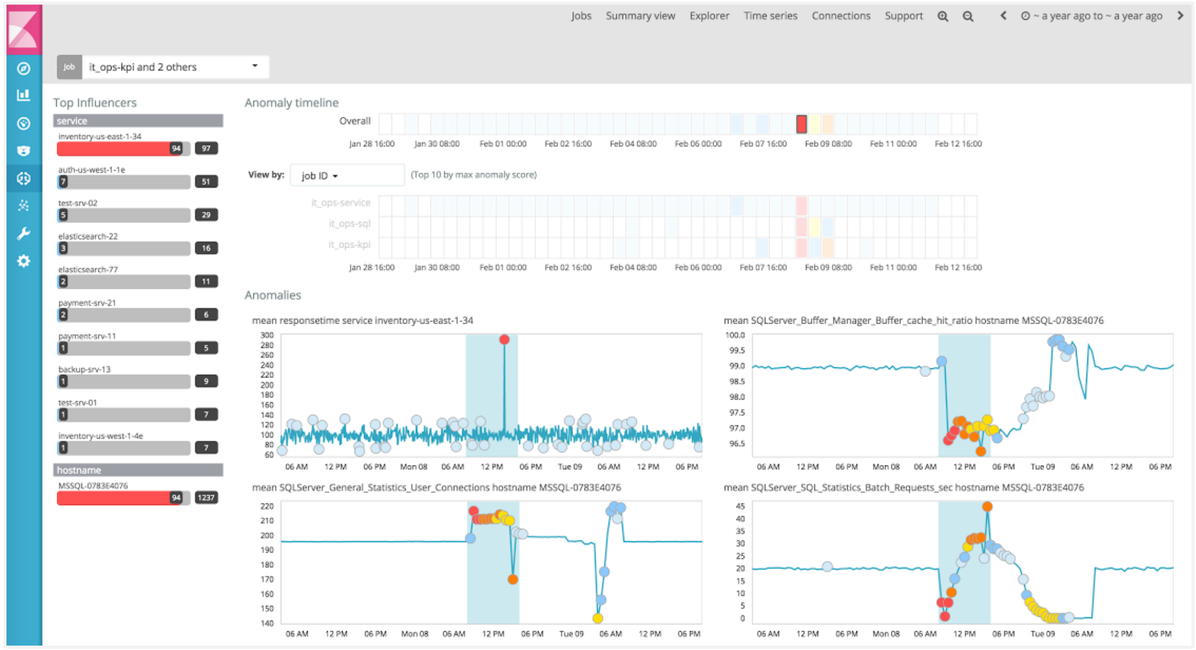
elastic

Creating a job

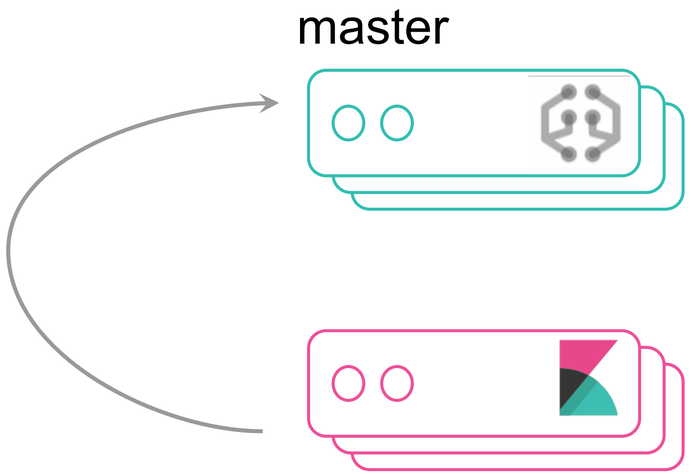


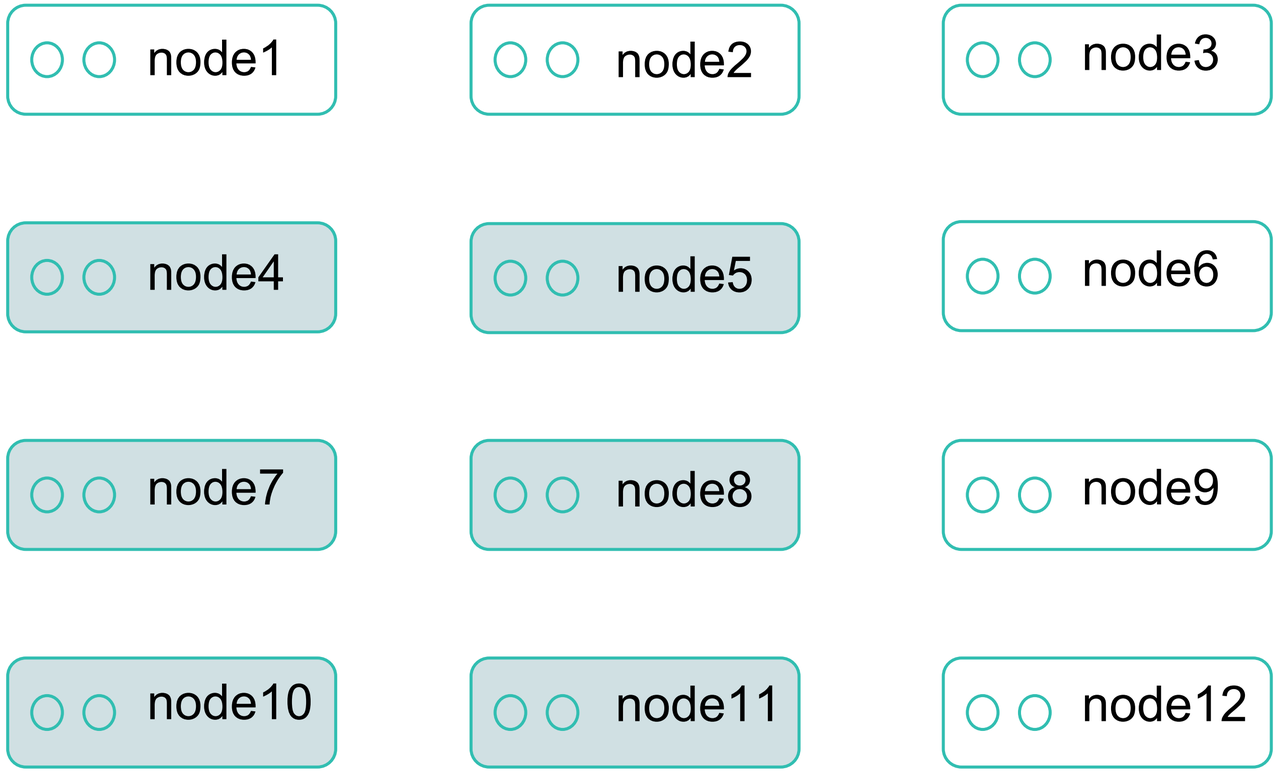






Cluster state configuration

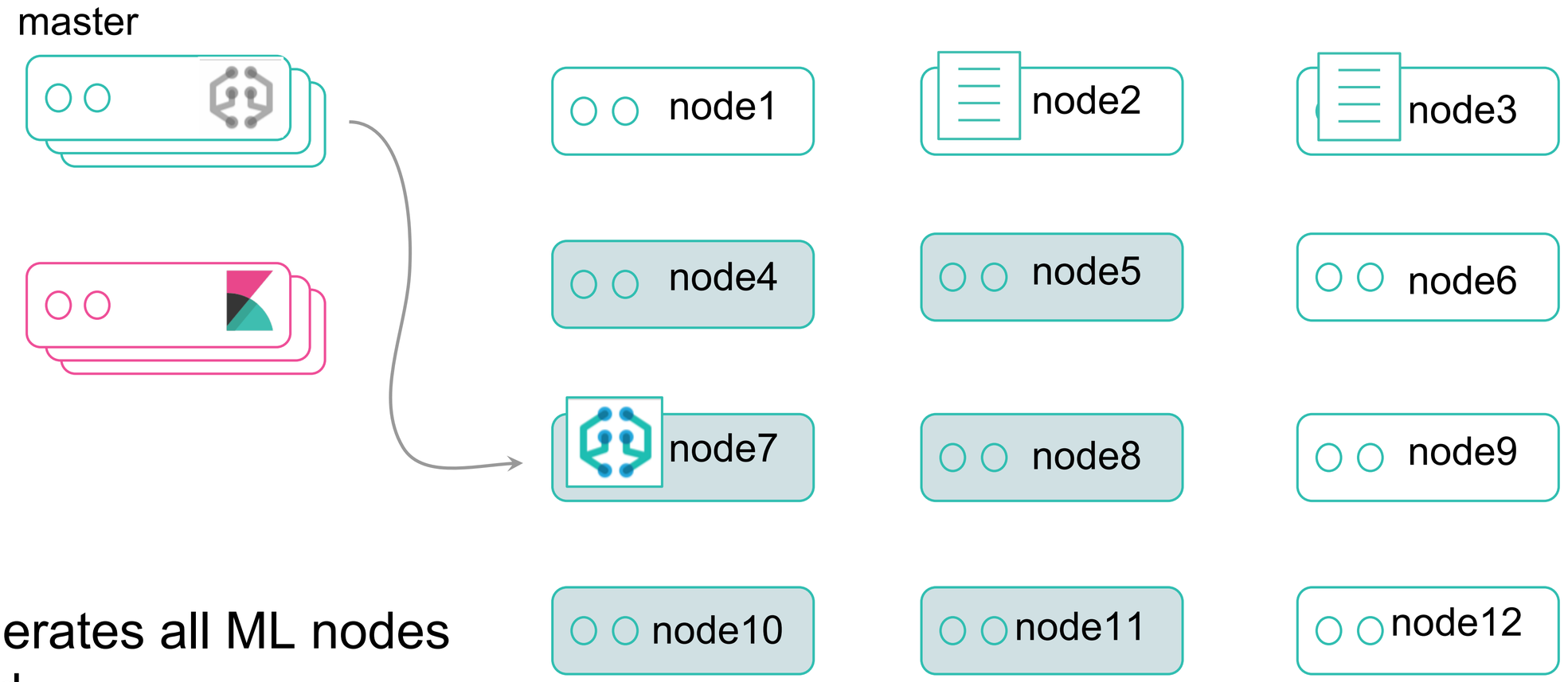




1. Create a job
2. Create a datafeed

PUT _xpack/ml/anomaly_detectors/{job_id}
PUT _xpack/ml/datafeeds/{datafeed_id}

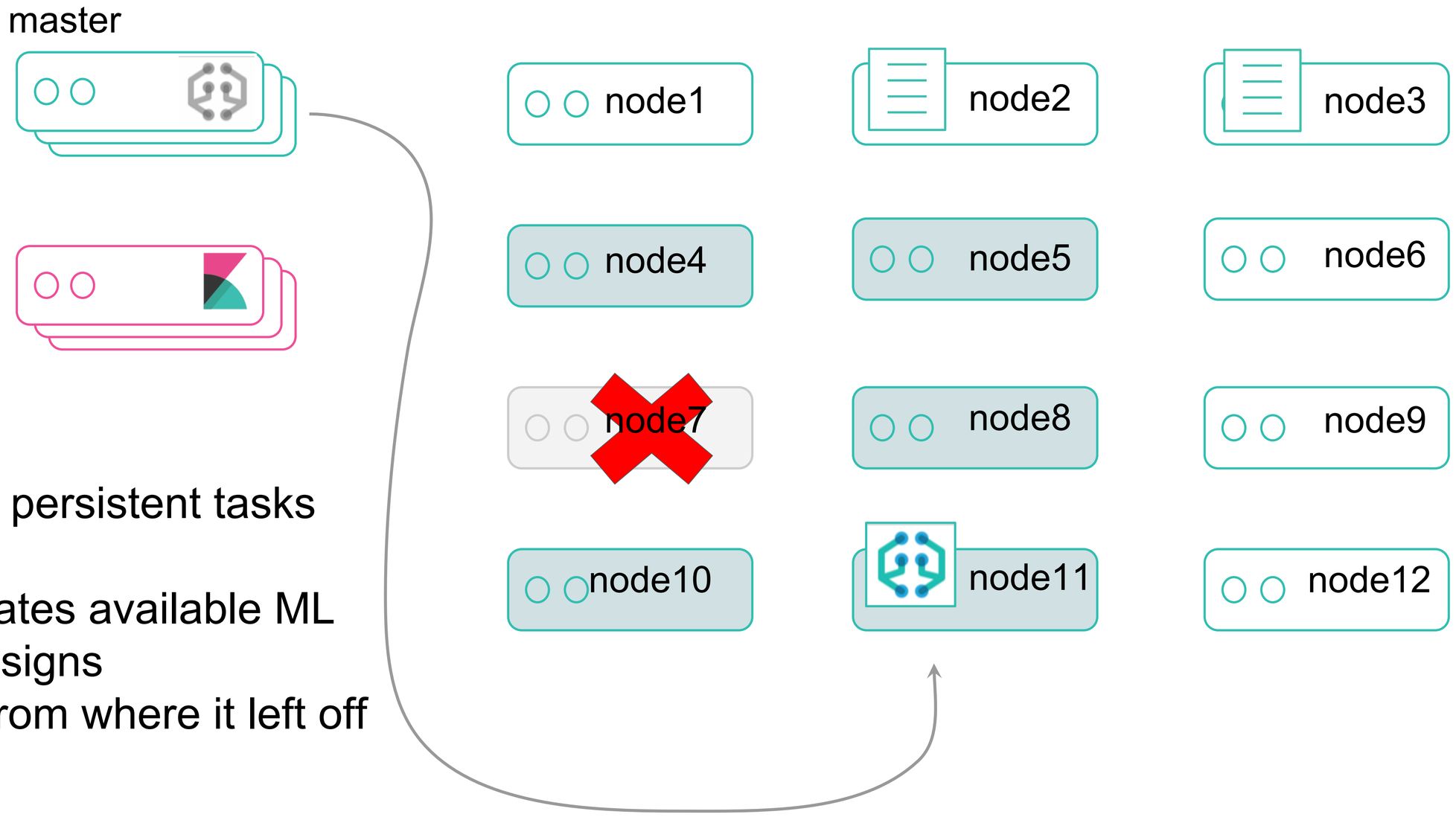
Load balanced analysis using persistent tasks



1. Master enumerates all ML nodes
2. Job is opened
3. Datafeed is started
4. Results written to index

PUT _xpack/ml/anomaly_detectors/{job_id}/_open
PUT _xpack/ml/datafeeds/{datafeed_id}/_start

Job resilience



1. If a node stops, persistent tasks continue
2. Master enumerates available ML nodes and reassigns
3. Job continues from where it left off

What's next

- Machine Learning and Statistical Methods for Time Series Analysis

Today, Stage A, 4:15pm

- Security Analytics Demo (Demo Station #2)
- AMA Booth

- Initial release planned with 5.4

More Questions?

Visit us at the AMA

www.elastic.co

Please attribute Elastic with a link to elastic.co