elastic | 中国开发者大会

Elasticsearch在ebay的落地和平台化管理

阮一鸣
eBay



战略级赞助商 HUAWEI | 钻石级赞助商 普翔 | 白金级赞助商 华夏博格 神州数码 Digital China | 金牌级赞助商 iDataAPI

合作伙伴 开源中国 oschina.net 掘金 谱时 IT大咖说 otpub Broadview 百格活动 bagevent.com MAXHUB



Pronto
Elasticsearch as Service in eBay

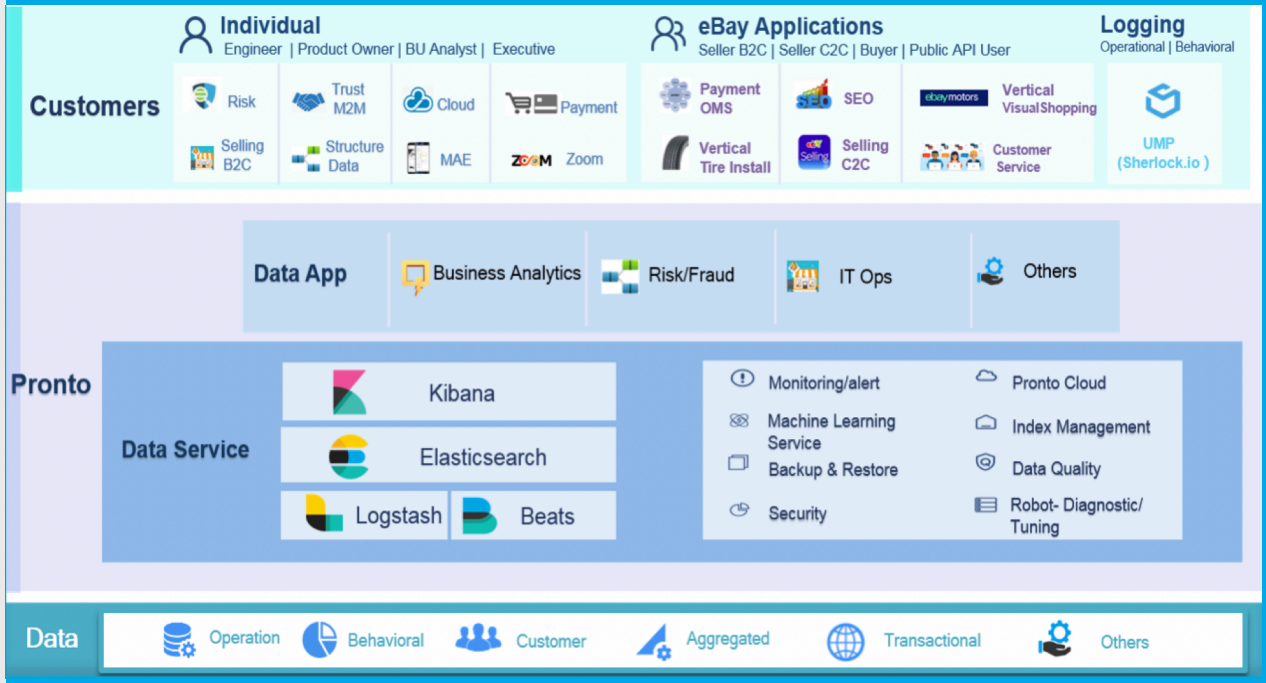
Ruan, Yiming
11/10/2018, Pronto Team, eBay

Agenda

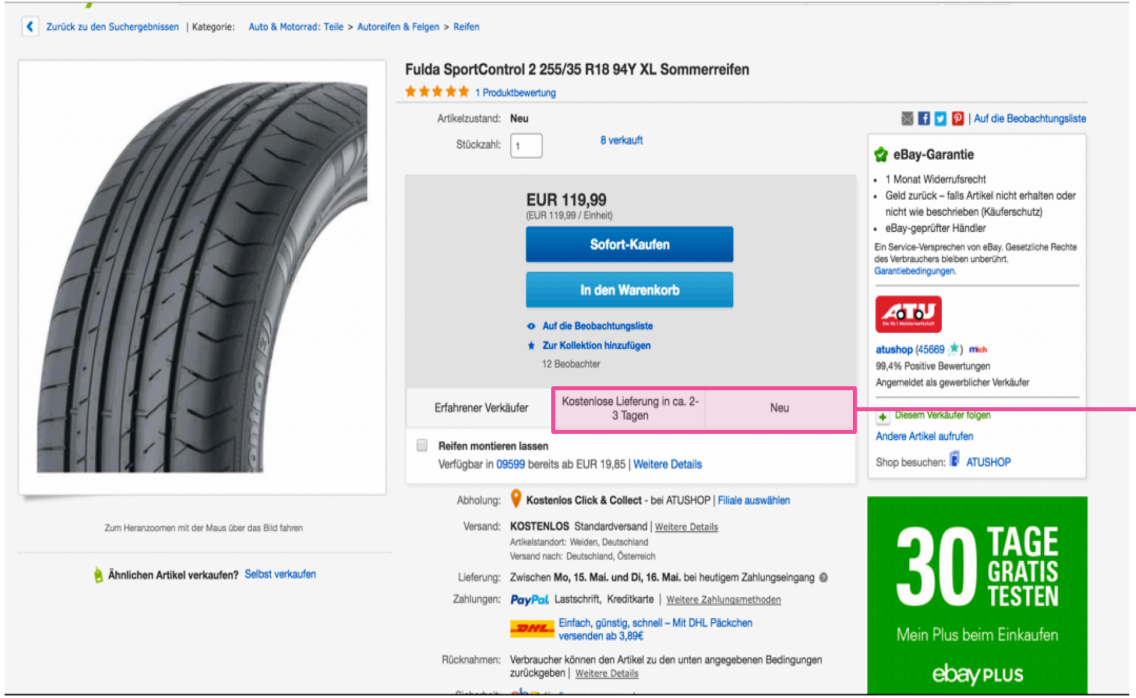
Pronto - Elasticsearch as Service in eBay

1. Use Cases in eBay
2. Pronto Overview
3. Provision & Deployment & Monitoring & Remediation
4. Shared & OOTB Solution
5. Pronto Services & Tools

elastic


Customers
Individual
Engineer | Product Owner | BU Analyst | Executive
Risk
Trust M2M
Cloud
Payment
Selling B2C
Structure Data
MAE
Zoom
eBay Applications
Seller B2C | Seller C2C | Buyer | Public API User
Payment OMS
SEO
Vertical Visual Shopping
Vertical Tire Install
Selling C2C
Customer Service
Logging
Operational | Behavioral
UMP (Sherlock.io)
Data App
Business Analytics
Risk/Fraud
IT Ops
Others
Pronto
Data Service
Kibana
Elasticsearch
Logstash
Beats
Monitoring/alert
Machine Learning Service
Backup & Restore
Security
Pronto Cloud
Index Management
Data Quality
Robot- Diagnostic/ Tuning
Data
Operation
Behavioral
Customer
Aggregated
Transactional
Others

Vertical Tire Installation



Business Value

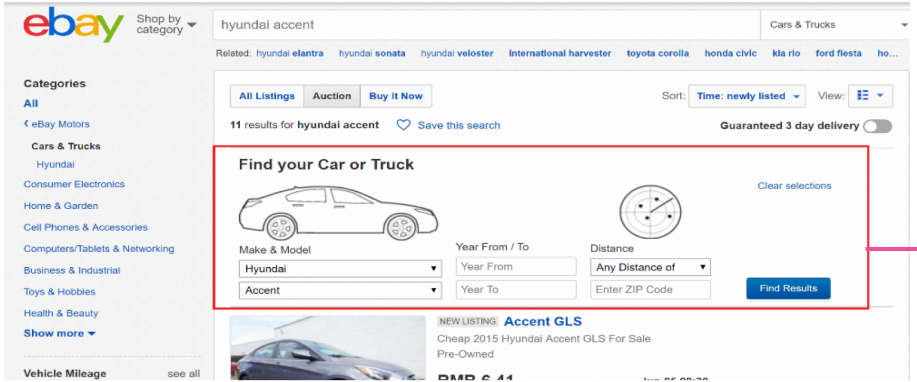
- Fast search of local tire store with geo search
- Improved purchase experience

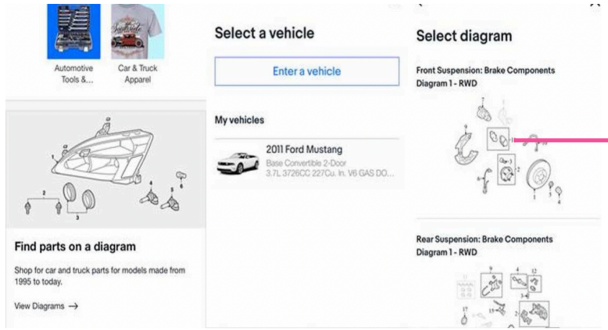
1 Geo Location

elastic

Vertical Shopping

For purchasing Car & Truck parts





1 Search filter

2 Shopping by searching parts on a diagram

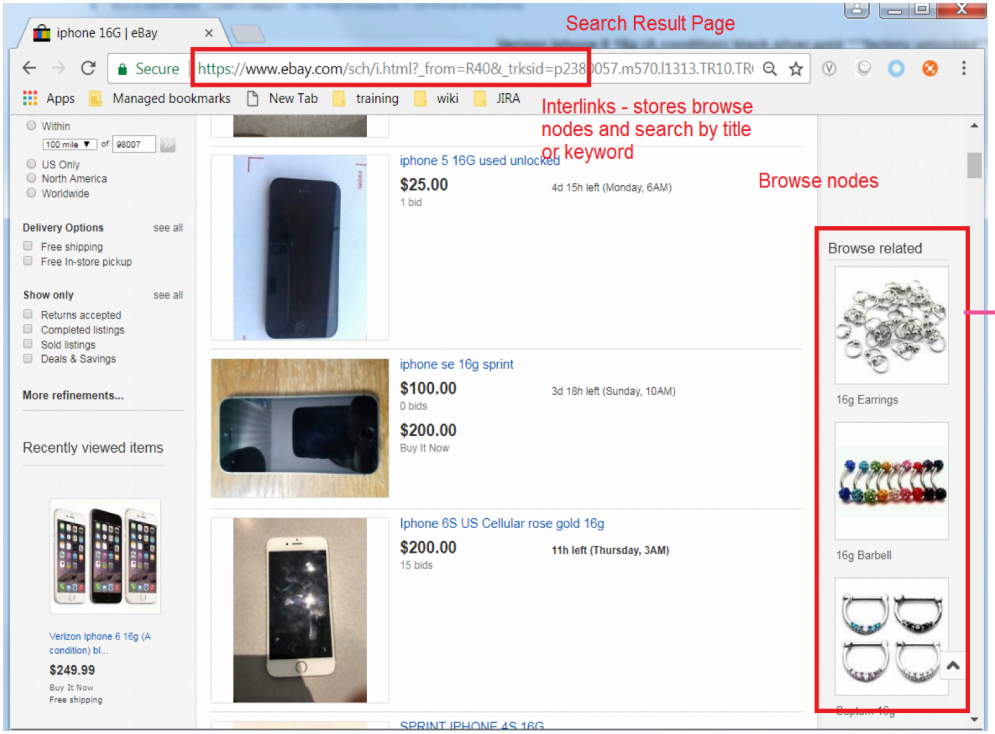
Business Value

- eFast application search delivered to auto buyers

	Availability Guarantee	DBA/Team Support	Flexibility	Efficient Partial String Search	Efficient Nested Search	Scalability	Consistency	Less Maintenance	No Data Duplication
ORACLE	✓	✓	✗	✗	✗	✓	✓	✗	✗
MONGO	✗	✓	✓	✗	✗	✓	✗	✗	✗
ELASTIC SEARCH	✗	✓	✓	✓	✓	✓	✓	✓	✗
NEO4J	✗	✗	✓	✓	✓	✓	✓	✗	✓
COUCH BASE	✗	✓	✓	✗	✗	✓	✓	✗	✗
IN-MEMORY	✓	✓	✓	✓	✓	✗	✗	✓	✓

elastic

Search Engine Optimization



Business Value

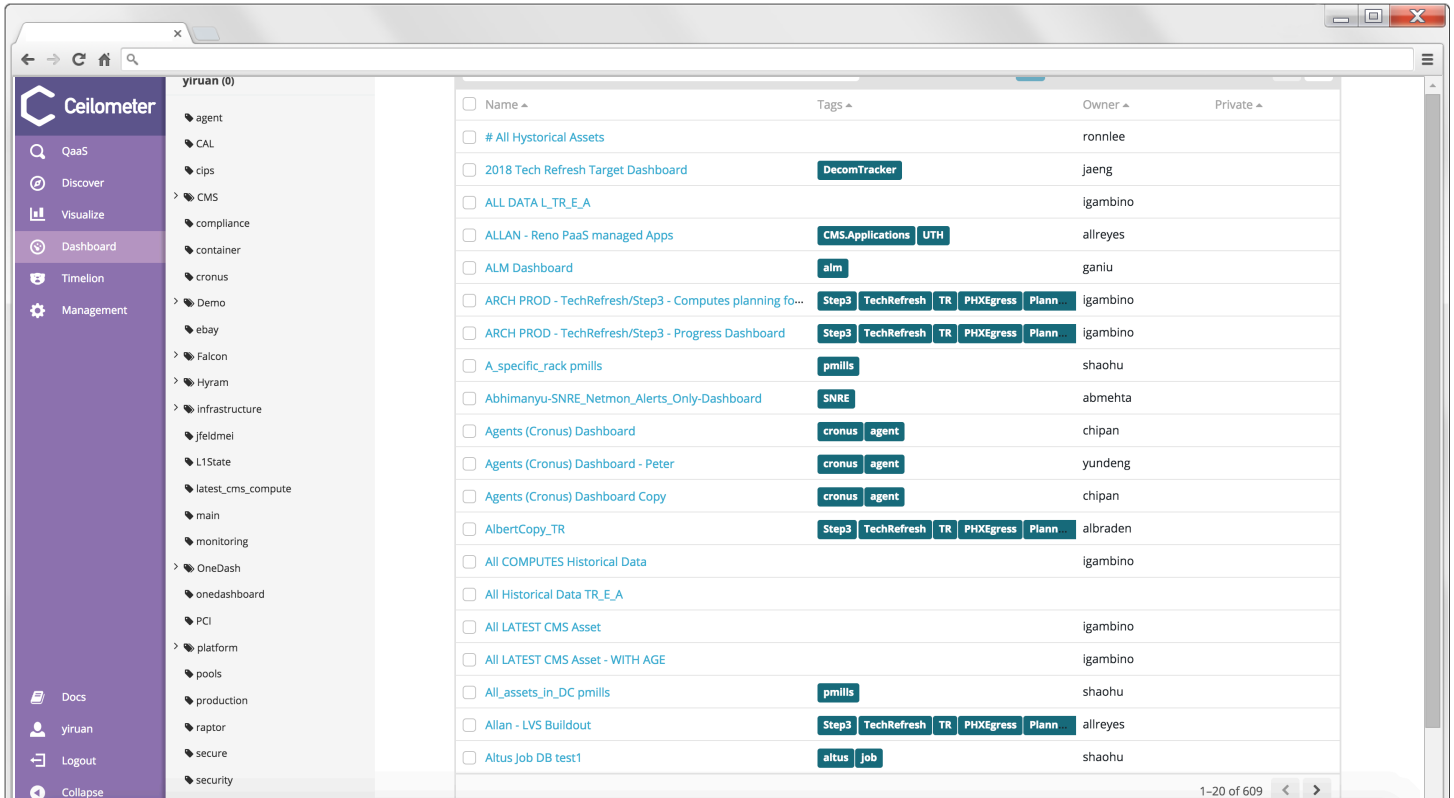
- Secured 600M daily traffic with less than 50ms response time
- Increase BN coverage
- Improved precision and relevance

1 Browser related

elastic

Ceilometer - IT Operation Analytics

Query as Service



elastic

UMP

Unify Monitoring Platform

- Applications are built with internal frameworks.
- Internal PaaS provides lifecycle management for applications built on supported frameworks.
- Internal Logging and Monitoring APIs provide ability to ship operational data.








100K +
Computes

100B +
URL Hits

1.2 PB
Application Logs

5M/s
Metric Data points

https://www.elastic.co/blog/monitoring-petabytes-of-logs-at-ebay-with-beats

elastic

Agenda

Pronto - Elasticsearch as Service in eBay

1 Use Cases in eBay

2 Pronto Overview

3 Provision & Deployment & Monitoring & Remediation

4 Shared & OOTB Solution

5 Pronto Services & Tools

elastic

Pronto



80+ Use Cases on 2 Providers

4000+ VMs in 4 Regions

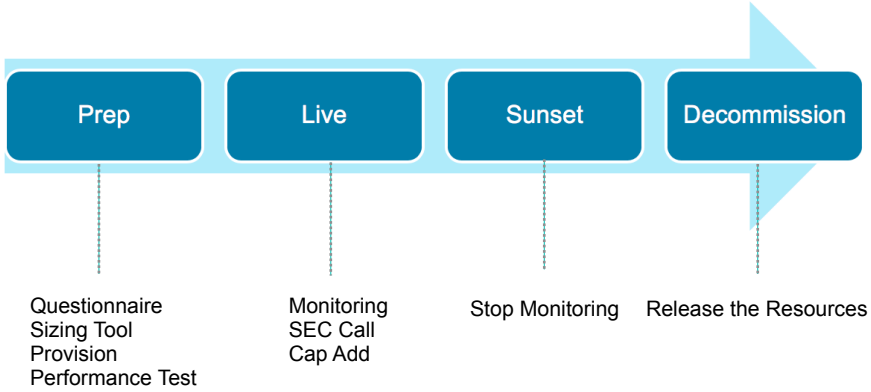
Deployments

- *Hot-warm architecture*
- *X-Pack*

elastic

Manage at Scale Requirements

Manage Elasticsearch fleets at scale

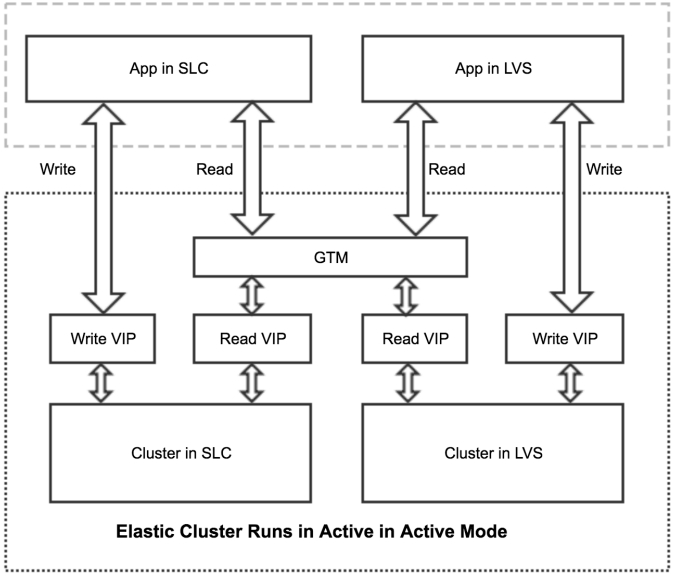


- **Provision** and **Deployment** for high availability and high performance in production
- **Monitoring** for getting visibility across all resources
- Ensure consistent **Security & Compliance**
- **Backup** and restore to avoid the data loss
- **Multi-tenant** to maximize resource utilization
- LOM for cluster auto **Remediation**
- OOTB **Solutions** to reduce the time to market
- eBay **Integrations** & other requirements from domain teams

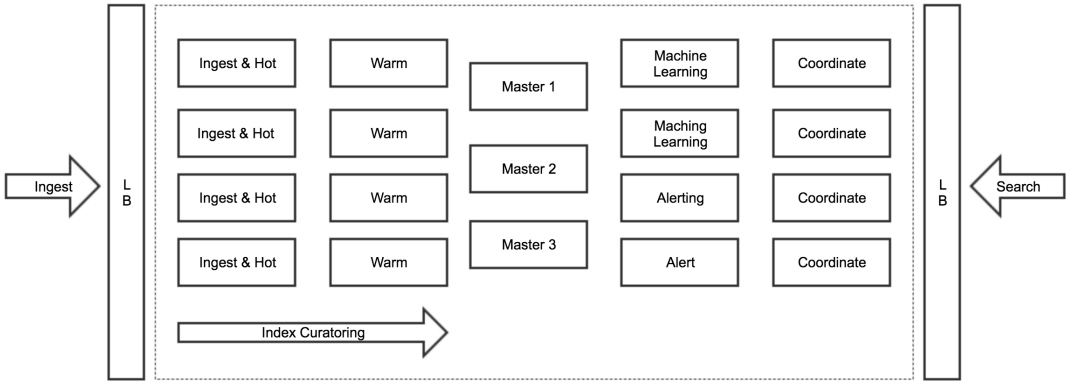
elastic

Deployment in Production

Deployment for Search & Logging



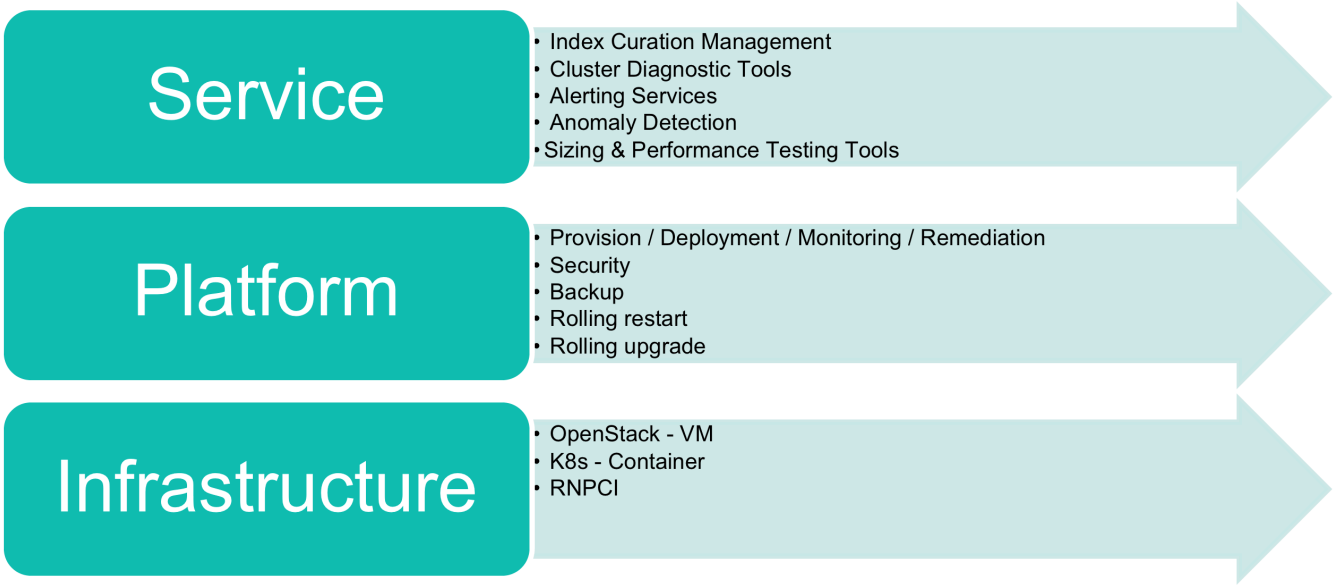
Deployment for Search



Deployment for Logging & Analytic

elastic

Overview of Pronto Services



elastic

Use Case On-boarding

Questionnaires & Sizing

- Onboard Questionnaire
  - **Use Case Description**
  - **Sizing Information** - # of docs, average doc size, YoY growth estimation
  - **Data Read/Write Information -** throughput, # of users, specific queries, pagination or sorting operations
  - **Data Source & Retention -** create/purge strategy
  - **Data Classification -** Internal / Public / Restricted / Confidential / PII / PCI
- Performance and Sizing
  - Web based sizing tool
  - JMeter Scripts & Web based perf testing tool

elastic

Agenda

Pronto - Elasticsearch as Service in eBay

1 Use Cases in eBay

2 Pronto Overview

3 Provision & Deployment & Monitoring & Remediation

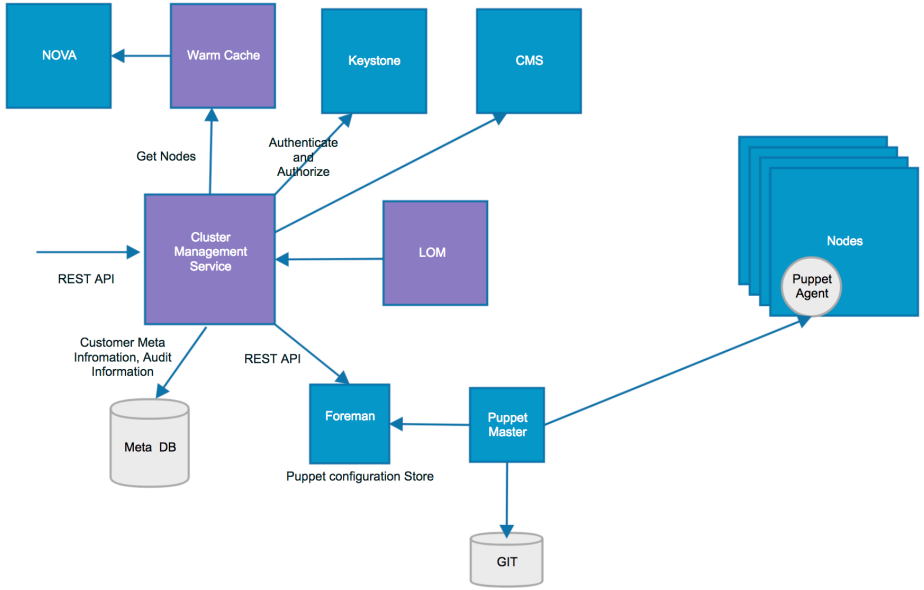
4 Shared & OOTB Solution

5 Pronto Services & Tools

elastic

Provision and Deployment

Pronto on OpenStack



- Provision
  - Based on VMs
  - Warm-Cache
  - Anti Affinity
- Deployment
  - Puppet Module
  - Puppet / Foreman
  - Support Hot-Warm Architecture
  - Support X-PACK
- Orchestration
  - WISB & WIRI

1. https://www.ebayinc.com/stories/blogs/tech/elasticsearch-cluster-lifecycle-at-ebay/
2. https://www.ebayinc.com/stories/blogs/tech/ready-to-use-virtual-machine-pool-store-via-warm-cache/

elastic

Hardware Profiles

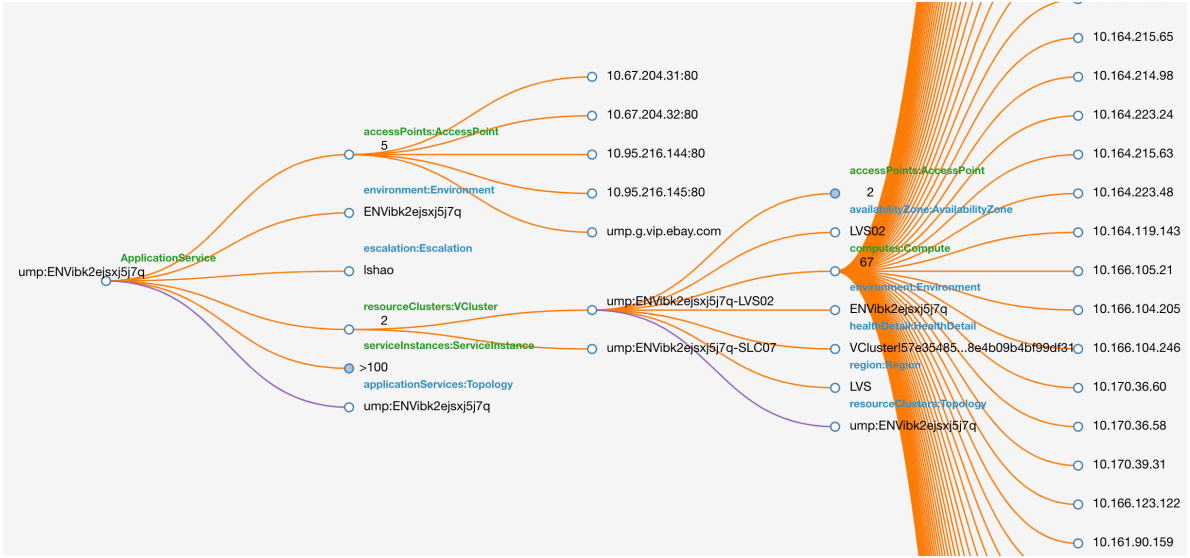
Best Practices for high Performance

- **Hardware profiles**
  - SSD - Fast data access. Expensive
  - HDD - Slow data access
  - Lot of CPU - Processing at ingestion, many queries
  - Lot of RAM - Aggregations, Kibana, master nodes
- **Characteristic of different node roles**
  - Data Node
    - High I/O
    - High storage
  - ML Node
    - Ample CPU & RAM
  - Kibana
    - Amply RAM

- **Hardware Best Practices**
  - In general, choose medium machines over large machines
  - Avoid running multiple node on one server
  - Large machine can be helpful as cold nodes
    - Disable concurrent merges
- **Storage Best Practices**
  - Prefer SSD
    - Use nook or deadline scheduler in the OS
    - Trim your SSD
  - Local disk is King (avoid NFS or SMB)

elastic

CMS - Infrastructure & Configuration



ApplicationService[@architecture="elasticsearch"].resourceClusters.computes[@resource="ipaddress"].nodeServer.nodeServer!AssetServer.managementServer.nodeServer!AssetServer.asset.assets!Rack{@resourceId}

Topology
- ApplicationService
- VCluster
- AccessPoint

Query
- Graph database

Integration
- Fireware
- LOM

elastic

Monitoring

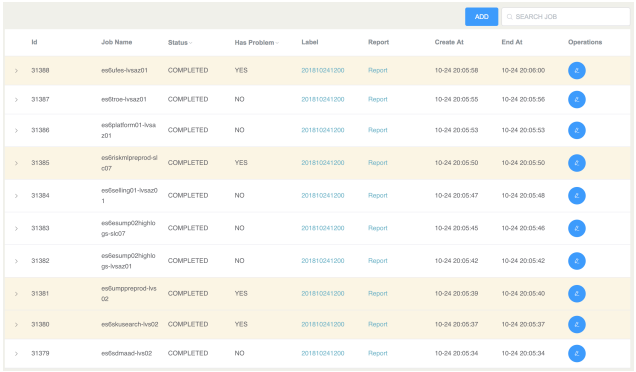


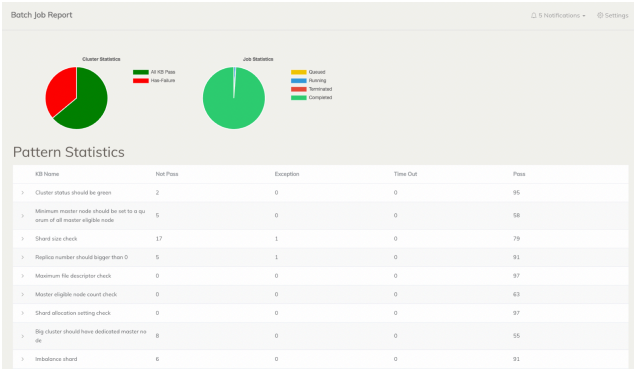
- Features
  - Monitoring all ES Clusters 2.x, 5.x, 6.x
    - Nodes, Indices, Shards, and general cluster metrics
    - Load Balancer
  - Alerting
    - Sherlock
    - Integrate with SLACK and PageDuty
    - SEC 7*24 on call

- Architecture
  - V1 - **Plugin** to push metrics to sherlock.io
  - V2 - **Beat** to push metrics to Elasticsearch
    - Node Stats / Cluster Stats / Index Stats API
    - Cluster Health API / Pending Task API
  - Grafana Dashboard over Elasticsearch

elastic

Diagnostic Tool for Elasticsearch Clusters

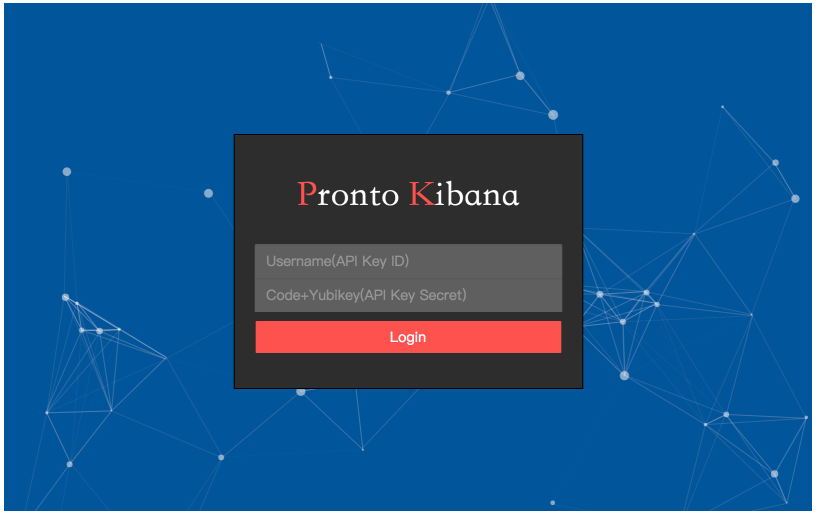




- Shard size check (20GB-40GB should be good)
- Too many shards
- Cluster status should be green
- Minimum master node should be set to a quorum of all master eligible node
- Replica number should bigger than 0
- Maximum file descriptor check
- Master eligible node count check
- Shard allocation setting check
- Big cluster should have dedicated master node
- Imbalance shards
- Index have too many fields
- Rack id attribute missed

elastic

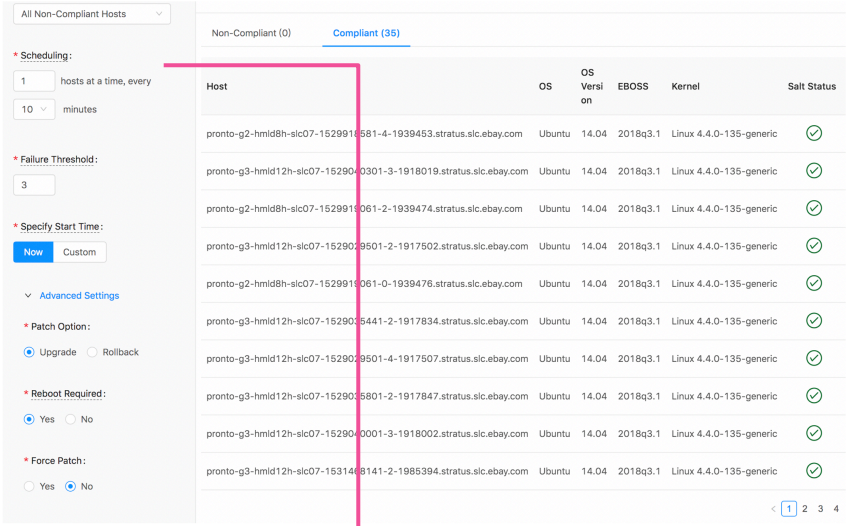
Pronto Security Plugin



- X-Pack Security is not free
- PII data / RNPCI
- Authentication & authorization
  - Keystone API & 2FA
- TLS encryption for data in transit
- Whitelist for application IP list
  - Firewall policies based on CMS model
- Audit logs
- REST management API
- Kibana multi tenancy

elastic

Data Platform Security Control Matrix



- Disk Encryption
- Change Management
- Firewall Rules
- Vulnerability Management

elastic

Agenda

Pronto - Elasticsearch as Service in eBay

1 Use Cases in eBay

2 Pronto Overview

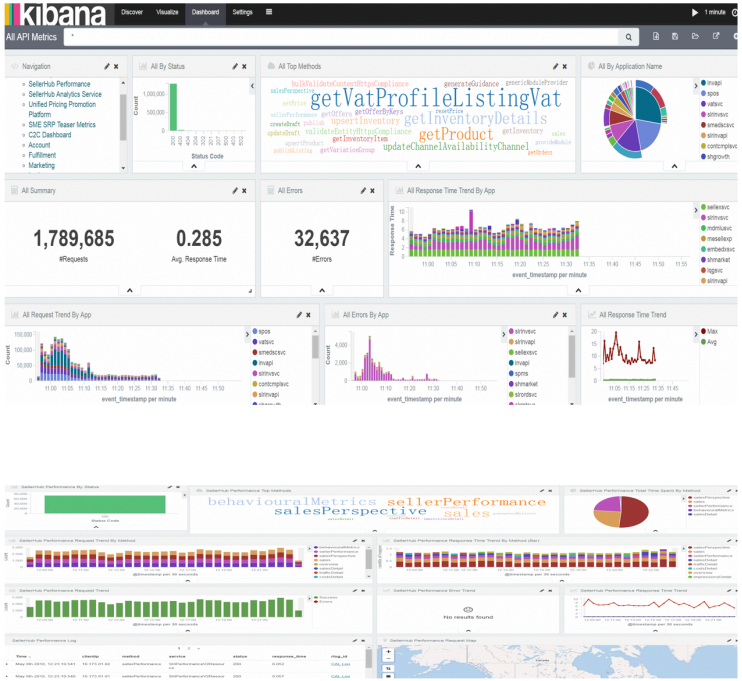
3 Provision & Deployment & Monitoring & Remediation

4 Shared & OOTB Solution

5 Pronto Services & Tools

elastic

Shared Cluster & OOTB Solution for Selling Team

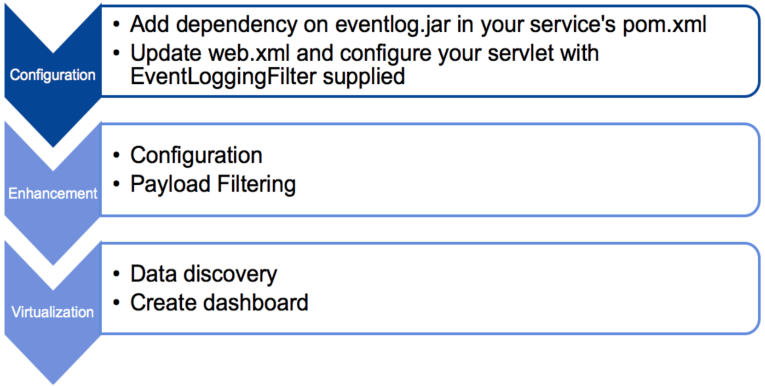


- Dedicated v.s Shared
  - Save the on boarding time and hardware utilization
  - Multi-tangency
    - Security for Index Level Control
    - Rate Limiting for read/write and disk usage
- Out-of-the-box
  - Reduce the time for on boarding
  - Less code
- Tracking team
  - Want to derive meaningful insights from API usage
  - Include use behavioral metrics
  - Want to gather operational metrics quickly and efficiently
  - Including usage volume, performance trends, keyword

elastic

Simple Steps to onboard a new Use Case

Connect data source and create dashboard



- pom.xml

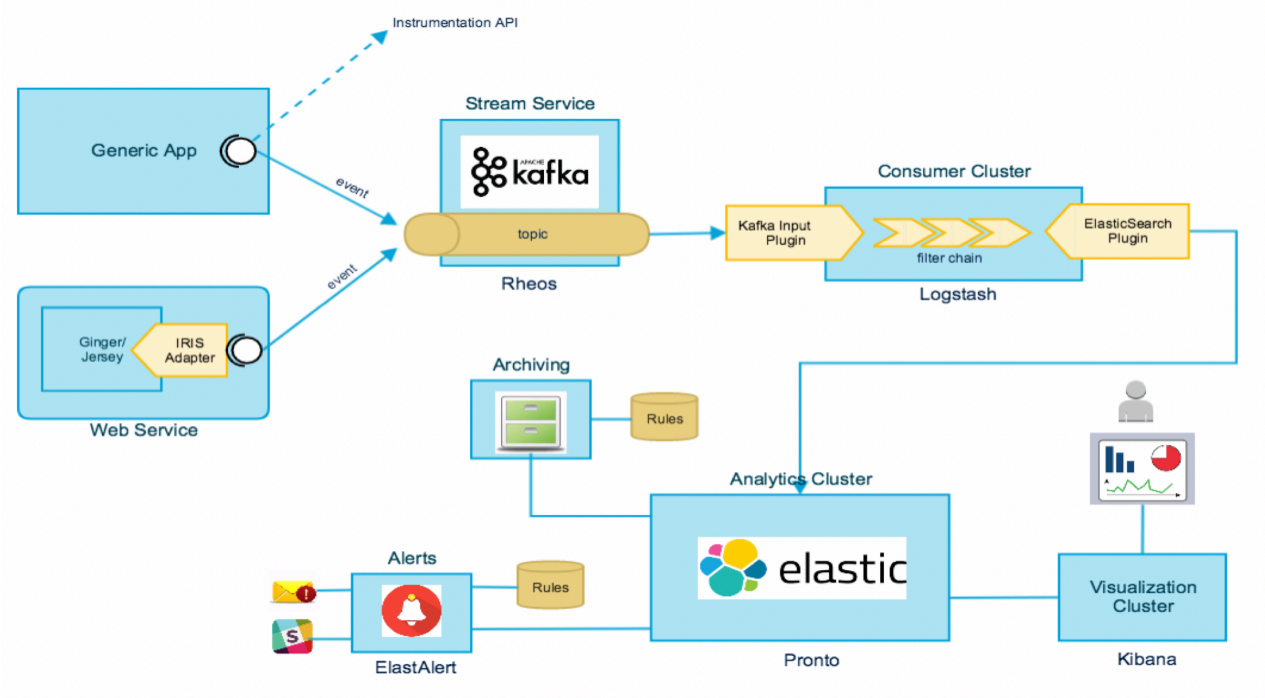
```
<dependency>
        <groupId>com.ebay.app.raptor.selling</groupId>
        <artifactId>eventlog</artifactId>
        <version>1.0.2-SNAPSHOT</version>
</dependency>
```

- web.xml

```
<init-param>
        <param-name>com.sun.jersey.spi.container.ContainerResponseFilters</param-name>
        <param-value>com.ebay.raptor.eventlog.web.EventLoggingFilter</param-value>
</init-param>
<init-param>
        <param-name>com.sun.jersey.spi.container.ContainerRequestFilters</param-name>
        <param-value>com.ebay.raptor.eventlog.web.EventLoggingFilter</param-value>
</init-param>
```


elastic

Architecture for Selling



elastic

Agenda

Pronto - Elasticsearch as Service in eBay

1 Use Cases in eBay

2 Pronto Overview

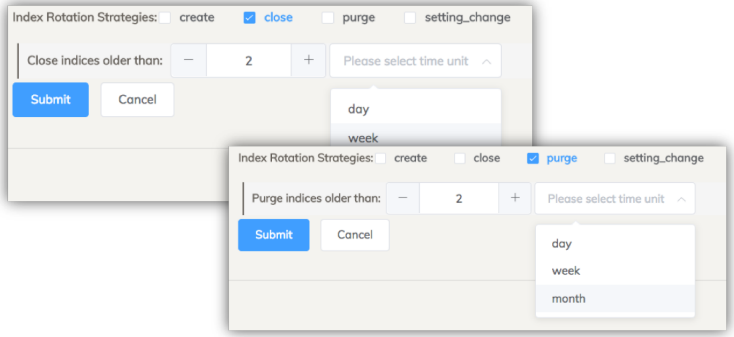
3 Provision & Deployment & Monitoring & Remediation

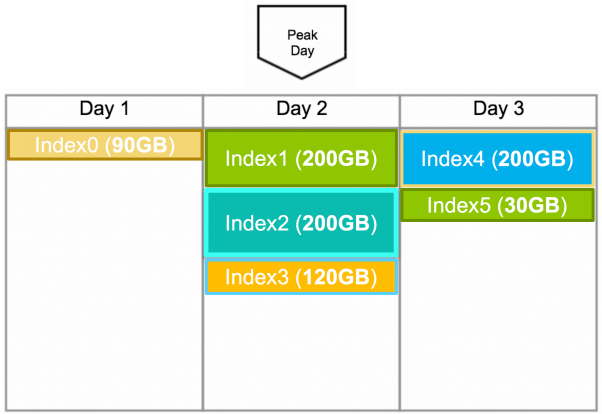
4 Shared & OOTB Solution

5 Pronto Services & Tools

elastic

Index Curation





- Features
  - High availability index lifecycle management
  - Self Service - One stop management portal for multiple clusters and multiple tenants
  - Support backup and restore (plugin for SWIFT)
- Value
  - Help customers to manage ES data through its full lifecycle
  - Let important business applications survive from peak days
    - No worries to configure primary shard number for daily indices

elastic

Anomaly Detection Use Cases



- How do I know my system are behaving normally?
- Where to set thresholds for good alerting?
- How to find the root cause of problems when I don't know what to look for



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

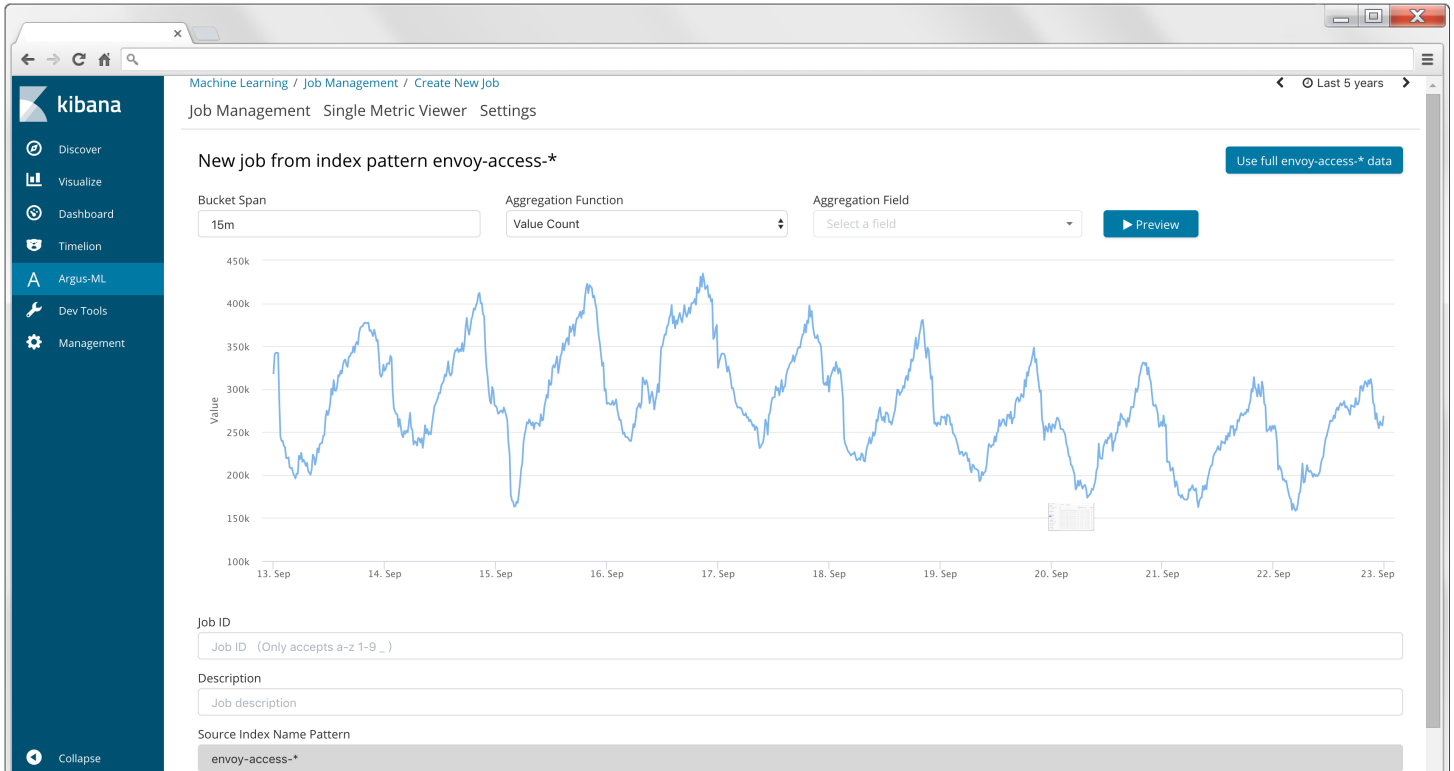


- Is my factory working normally?
- What do I do with thousands of time-series data?
- Which traffic incidents are causing the most delay?

elastic

Argus - Anomaly Detection

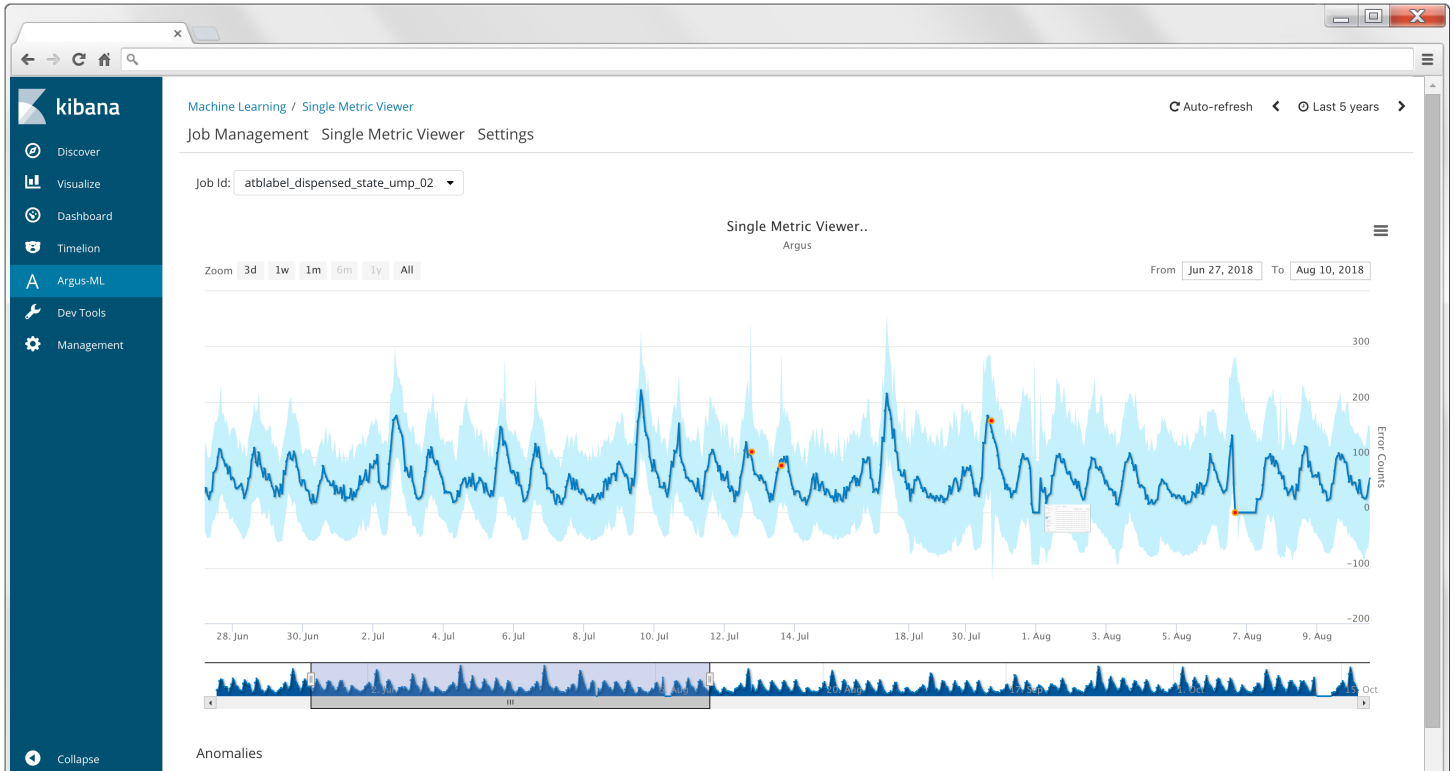
Job Management



elastic

ML for Anomaly Detection

Single Metric Viewer

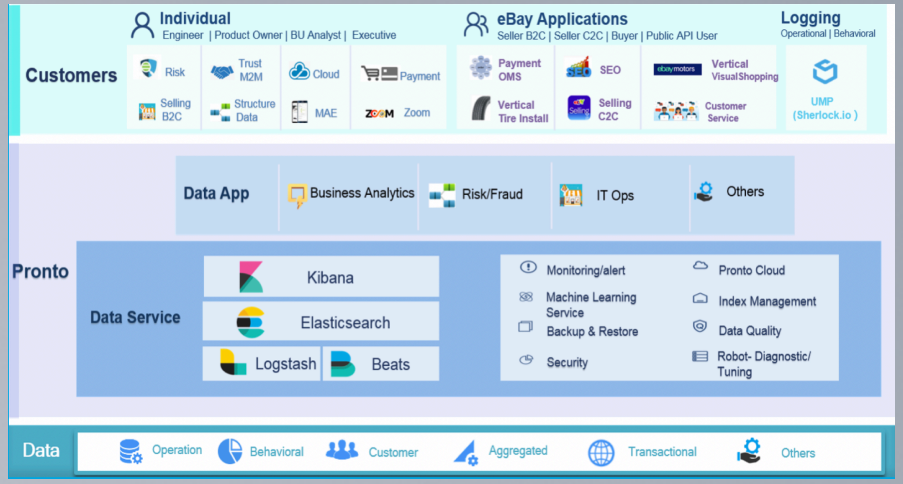


elastic

Review

Pronto - Elasticsearch as Service in eBay

- Use Cases
  - Near Realtime Search
  - Logging
  - Analytics
- Pronto Services
  - Infrastructure
  - Platform
  - Service
  - eBay Integration & Tools we developed



Thank You!

yiruan@ebay.com



elastic



elastic
中文社区

专业、垂直、纯粹的 Elastic 开源技术交流社区
https://elasticsearch.cn/

elastic