elastic

What's new in Elastic Stack 7.0?

Jun Ohtani, Developer | Evangelist
2019/05/23 Middlewares Deep Talks

elastic

What's new in Elastic Stack ~~7.0~~ 7.1?

Jun Ohtani, Developer | Evangelist

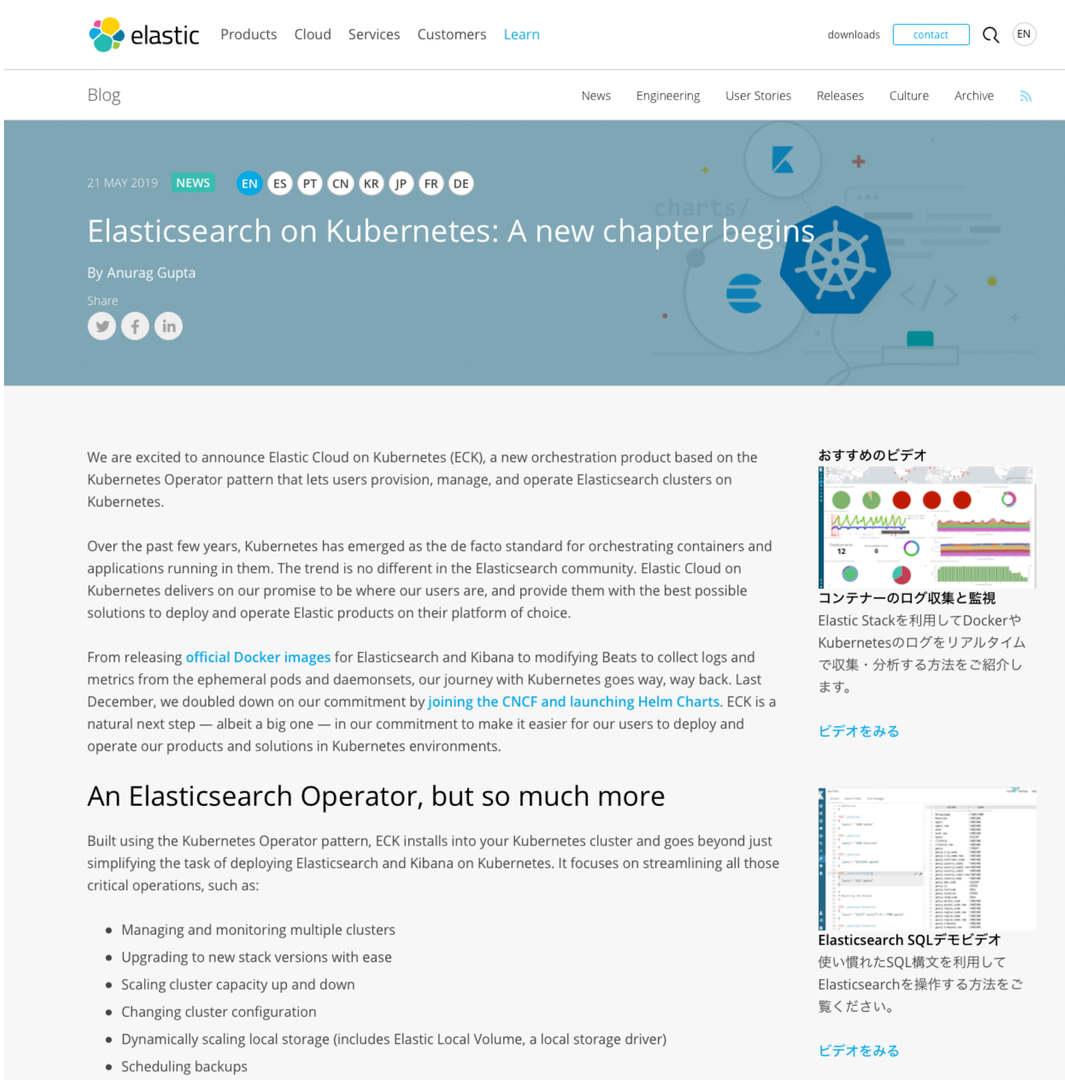
2019/05/23 Middlewares Deep Talks

about

- Me, Jun Ohtani / Community Engineer
  - Elasticsearch勉強会 / 検索技術勉強会
  - データ分析基盤構築入門 共著
  - http://blog.johtani.info
- Elastic, founded in 2012
  - Products: Elasticsearch, Logstash, Kibana, Beats
    Elastic APM,
    Elastic Cloud, Swiftype
    Professional services: Support & development subscriptions
    Trainings, Consulting, SaaS

7.1リリース

- Elastic Cloud on Kubernetes
  - Kubernetes Operator pattern
  - Dynamically scaling local storage w/ Elastic Local Volume



Elasticsearch on Kubernetes: A new chapter begins

By Anurag Gupta

We are excited to announce Elastic Cloud on Kubernetes (ECK), a new orchestration product based on the Kubernetes Operator pattern that lets users provision, manage, and operate Elasticsearch clusters on Kubernetes.

Over the past few years, Kubernetes has emerged as the de facto standard for orchestrating containers and applications running in them. The trend is no different in the Elasticsearch community. Elastic Cloud on Kubernetes delivers on our promise to be where our users are, and provide them with the best possible solutions to deploy and operate Elastic products on their platform of choice.

From releasing official Docker images for Elasticsearch and Kibana to modifying Beats to collect logs and metrics from the ephemeral pods and daemonsets, our journey with Kubernetes goes way, way back. Last December, we doubled down on our commitment by joining the CNCF and launching Helm Charts. ECK is a natural next step — albeit a big one — in our commitment to make it easier for our users to deploy and operate our products and solutions in Kubernetes environments.

An Elasticsearch Operator, but so much more

Built using the Kubernetes Operator pattern, ECK installs into your Kubernetes cluster and goes beyond just simplifying the task of deploying Elasticsearch and Kibana on Kubernetes. It focuses on streamlining all those critical operations, such as:

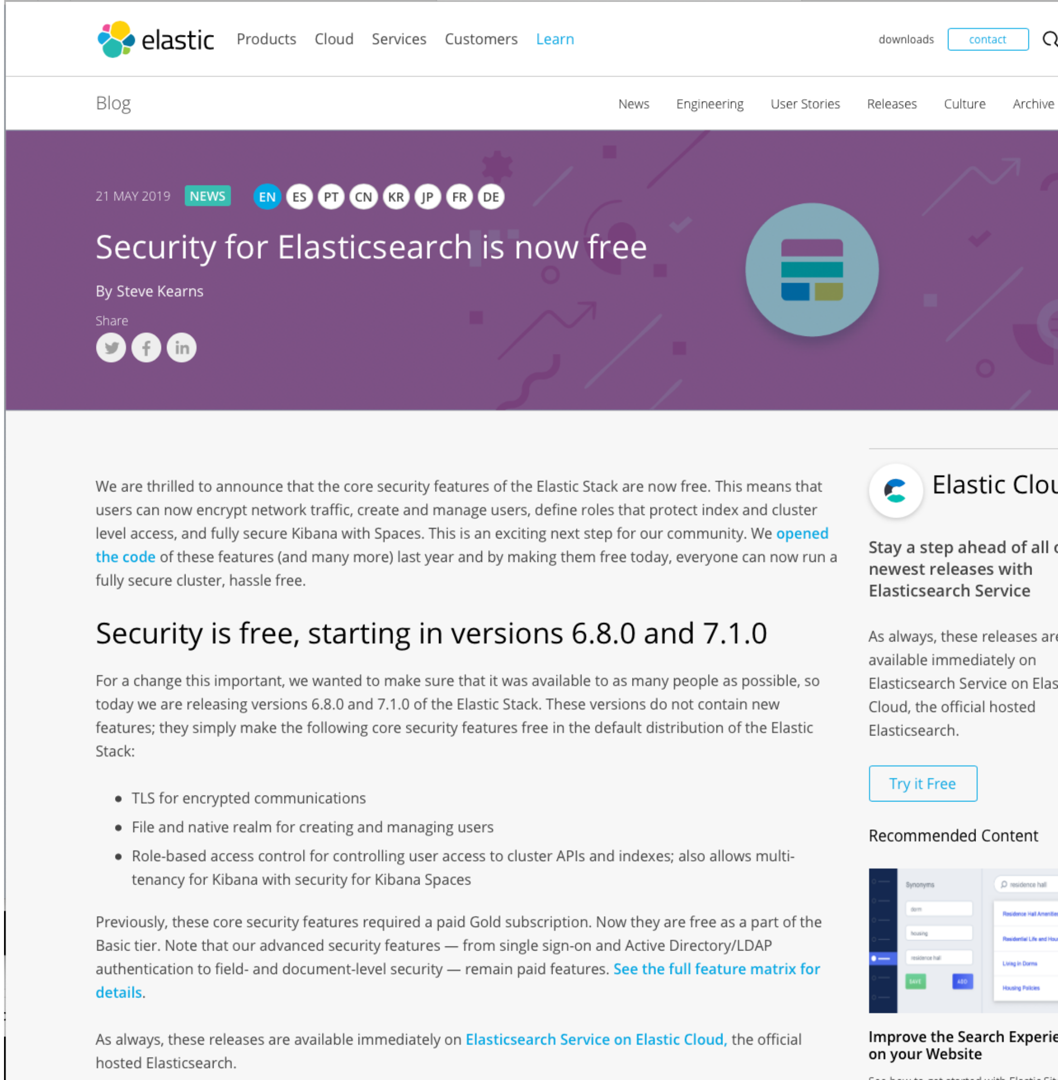
- Managing and monitoring multiple clusters
- Upgrading to new stack versions with ease
- Scaling cluster capacity up and down
- Changing cluster configuration
- Dynamically scaling local storage (includes Elastic Local Volume, a local storage driver)
- Scheduling backups

コンテナーのログ収集と監視
Elastic Stackを利用してDockerやKubernetesのログをリアルタイムで収集・分析する方法をご紹介します。

Elasticsearch SQLデモビデオ
使い慣れたSQL構文を利用してElasticsearchを操作する方法をご覧ください。

7.1リリース

- Securityの機能がBasicの対象に
  - TLS
  - File and native realm
  - RBAC



Security for Elasticsearch is now free

By Steve Kearns

We are thrilled to announce that the core security features of the Elastic Stack are now free. This means that users can now encrypt network traffic, create and manage users, define roles that protect index and cluster level access, and fully secure Kibana with Spaces. This is an exciting next step for our community. We opened the code of these features (and many more) last year and by making them free today, everyone can now run a fully secure cluster, hassle free.

Security is free, starting in versions 6.8.0 and 7.1.0

For a change this important, we wanted to make sure that it was available to as many people as possible, so today we are releasing versions 6.8.0 and 7.1.0 of the Elastic Stack. These versions do not contain new features; they simply make the following core security features free in the default distribution of the Elastic Stack:

- TLS for encrypted communications
- File and native realm for creating and managing users
- Role-based access control for controlling user access to cluster APIs and indexes; also allows multi-tenancy for Kibana with security for Kibana Spaces

Previously, these core security features required a paid Gold subscription. Now they are free as a part of the Basic tier. Note that our advanced security features — from single sign-on and Active Directory/LDAP authentication to field- and document-level security — remain paid features. See the full feature matrix for details.

As always, these releases are available immediately on Elasticsearch Service on Elastic Cloud, the official hosted Elasticsearch.

Elastic Clou

Stay a step ahead of all
newest releases with
Elasticsearch Service

As always, these releases are
available immediately on
Elasticsearch Service on Elas
Cloud, the official hosted
Elasticsearch.

Recommended Content

Improve the Search Experie
on your Website

Elasticsearch

新世代クラスター管理層

クラスター管理を将来のための基盤として再構築

強固な理論と広範囲なテスト

形式モデルで検証済み

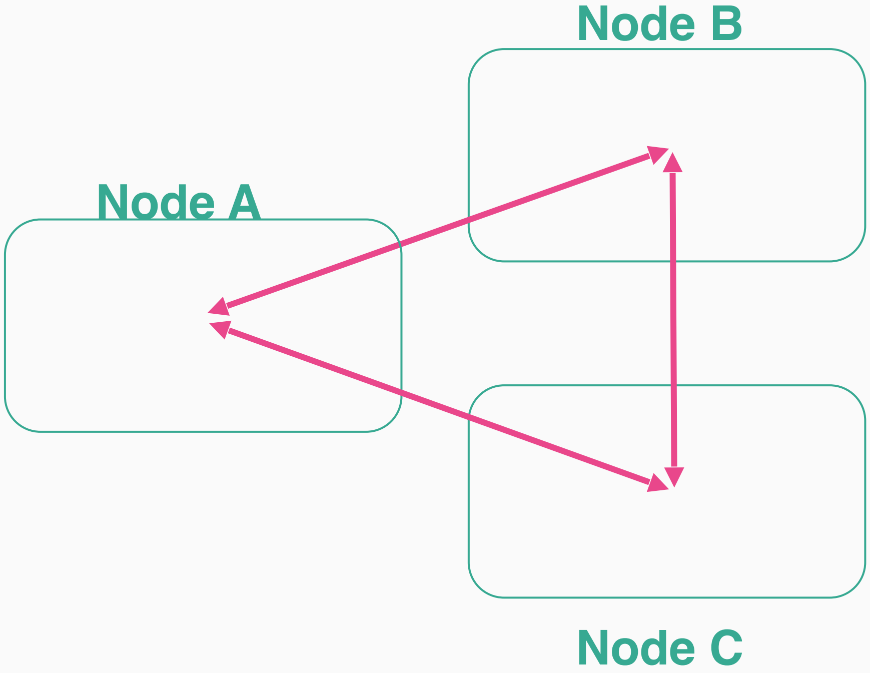
利点

`minimum_master_nodes`設定の排除

1秒以内でのマスター選出

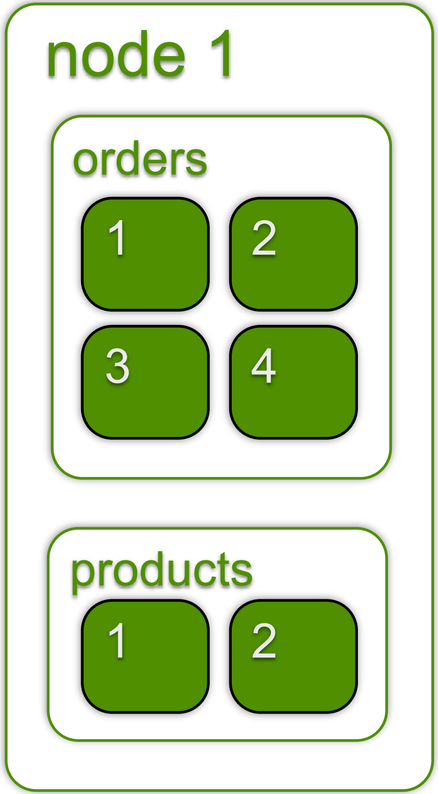
ラグやゾンビノードの迅速な除去

https://ela.st/new-cluster-coordination

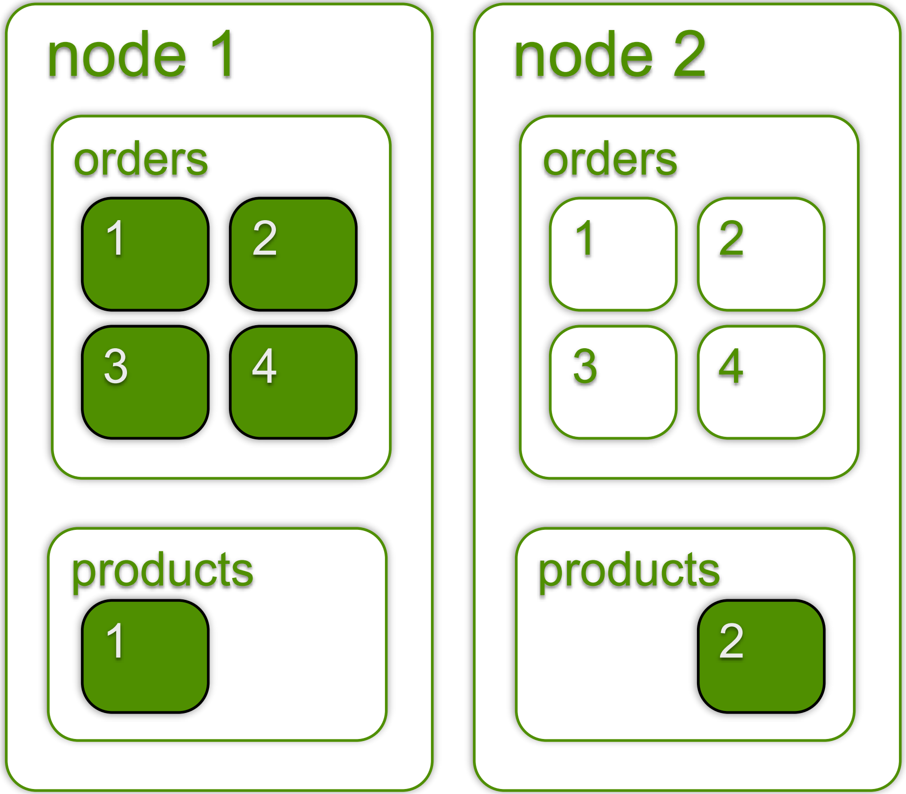


Cluster Coordination?

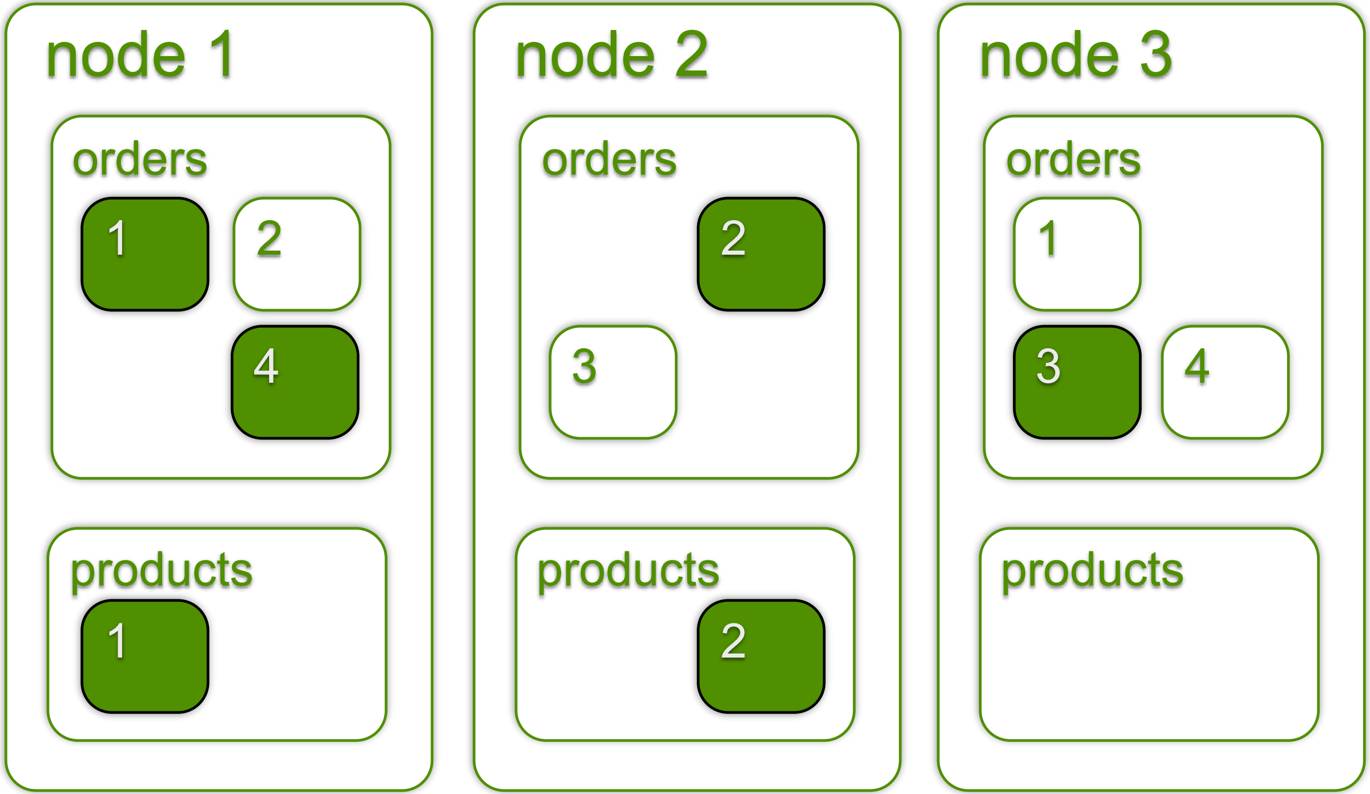
自動的な分散



自動的な分散



自動的な分散


node 1
orders
1
2
4
products
1
node 2
orders
2
3
products
2
node 3
orders
1
3
4
products

Cluster State?

Cluster State

Metadata about a cluster

- Cluster settings
- Index metadata
    - settings
    - mappings
    - where are my shards?
    - which shards are in sync?
- Lots more besides...

```
GET _cluster/state
```

Cluster Coordination

- Making sure all nodes have the latest cluster state
- Making sure cluster state updates are not lost

- (Making sure all nodes are healthy)

Scope

How does it work?

Discovery

- Where are the master-eligible nodes?
- Is there already a master?

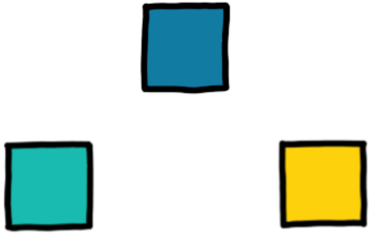
Master Election

- Agree on a node to take charge
- Form a cluster

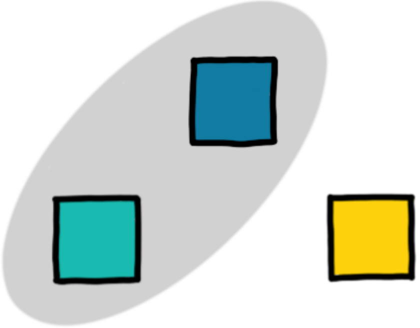
Cluster State Publication

- Agree on updates to cluster state
- Broadcast updates to all nodes

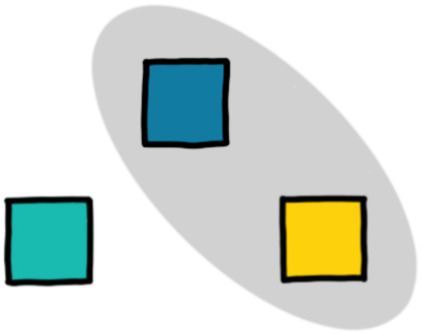
Agree?



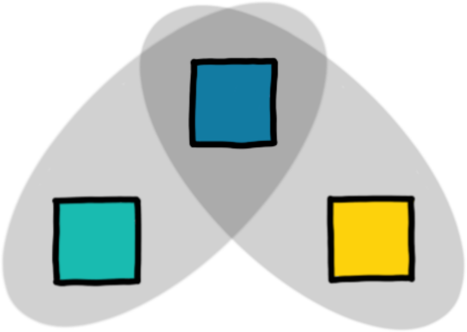
Agree?



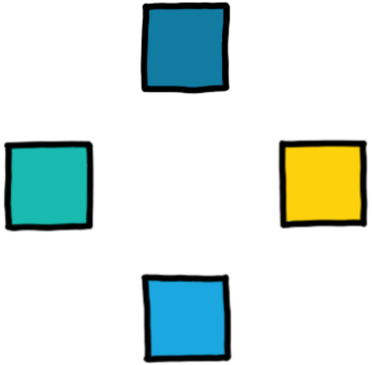
Agree?



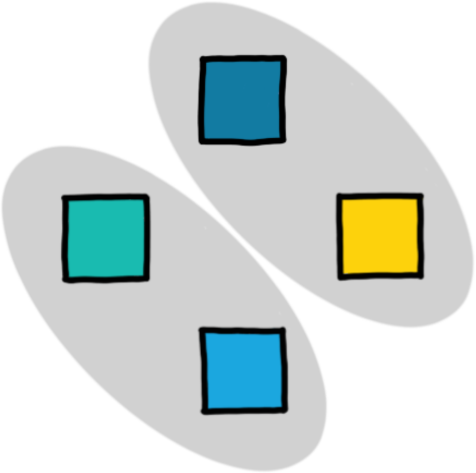
Agree?



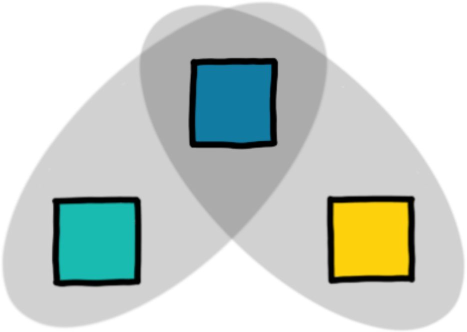
Agree?



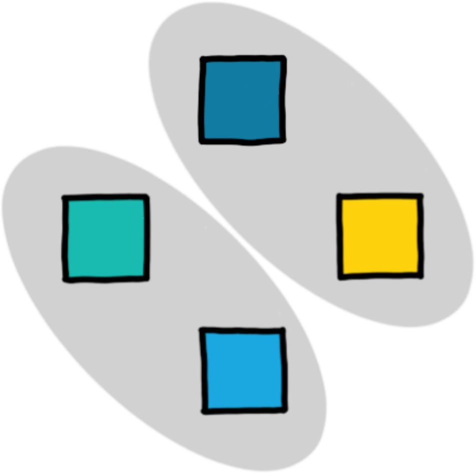
Agree?



Agree?



Agree?



Voting configurations

```
GET /_cluster/state?filter_path=nodes.*.name,metadata.cluster_coordination.last_committed_config
=>
```

```
{
  "nodes": {
    "ATdJrEZ6Rj62Sh7-UxhP5w": {"name": "node-0"},
    "vfMBDxJ-R0SdXyAu9wN-4Q": {"name": "node-1"},
    "kEyYc7lRTcqDjiWJkgG7qA": {"name": "node-2"},
    "w0GviS2DTNajk85-_dJ6nQ": {"name": "node-3"}
  },
  "metadata": {
    "cluster_coordination": {
      "last_committed_config": [
        "ATdJrEZ6Rj62Sh7-UxhP5w",
        "kEyYc7lRTcqDjiWJkgG7qA",
        "w0GviS2DTNajk85-_dJ6nQ"
      ]
    }
  }
}
```

How do we use it?

Discovery

“Where are all the master-eligible nodes?”

- Settings
  - `discovery.seed_nodes` (formerly ~~`discovery.zen.ping.unicast.hosts`~~)
    - Hostnames or IP addresses (maybe with transport ports)
  - `discovery.seed_providers` (formerly ~~`discovery.zen.hosts_provider`~~)
    - `file`, `ec2`, `gce`, …
    - some providers have extra settings too

How do we use it?

Cluster bootstrapping

“Which nodes count in the very first election?”

- Settings
  - `cluster.initial_master_nodes`
    - node names or transport addresses (IP addresses only)
    - ok to set this on multiple nodes, as long as they’re all consistent
    - ignored once node has joined a cluster, even if restarted
    - unnecessary when joining a new node to an existing cluster

How do we use it?

Other settings

“What else can we adjust?”

- Lots of other settings
  - see reference manual
- Defaults are good
- Changing them is **not recommended**

How do we use it?

Dynamic clusters

“How do we add or remove a node?”

- Master-ineligible nodes
  - no different from earlier versions
- Adding master-eligible nodes
  - just do it, no settings to adjust
- Removing master-eligible nodes
  - just do it, **as long as you remove less than half of them at once**

Troubleshooting

Logging when things go wrong

- `ClusterFormationFailureHelper`
  - `WARN` comprehensive periodic status update while cluster not formed

```
[2019-04-04T10:02:10,276][WARN ][o.e.c.c.ClusterFormationFailureHelper] [node-
5] master not discovered or elected yet, an election requires at least 2 nodes
with ids from [QsZt-170SBeRpHJbWb-W6Q, 9becUMgvTF-DbSDZ2uf_Og,
EpyPd38tQhKES91p_23WgA], have discovered [{node-0}{S68ml-4eREaGpmkVrnjorw}
{127.0.0.1:9300}] which is not a quorum; discovery will continue using
[127.0.0.1:9300, 127.0.0.1:9301, 127.0.0.1:9302, 127.0.0.1:9303,
127.0.0.1:9304] from hosts providers and [{node-4}{9becUMgvTF-DbSDZ2uf_Og}
{127.0.0.1:9304}, {node-5}{EpyPd38tQhKES91p_23WgA}{127.0.0.1:9305}] from last-
known cluster state; node term 42, last-accepted version 165 in term 42
```

Troubleshooting

Logging when things go wrong

- `ClusterFormationFailureHelper`
  - `WARN` comprehensive periodic status update while cluster not formed

```
[2019-04-04T10:02:10,276][WARN ][o.e.c.c.ClusterFormationFailureHelper] [node-
5] master not discovered or elected yet, an election requires at least 2 nodes
with ids from [QsZt-170SBeRpHJbWb-W6Q, 9becUMgvTF-DbSDZ2uf_Og,
EpyPd38tQhKES91p_23WgA], have discovered [{node-0}{S68ml-4eREaGpmkVrnjorw}
{127.0.0.1:9300}] which is not a quorum; discovery will continue using
[127.0.0.1:9300, 127.0.0.1:9301, 127.0.0.1:9302, 127.0.0.1:9303,
127.0.0.1:9304] from hosts providers and [{node-4}{9becUMgvTF-DbSDZ2uf_Og}
{127.0.0.1:9304}, {node-5}{EpyPd38tQhKES91p_23WgA}{127.0.0.1:9305}] from last-
known cluster state; node term 42, last-accepted version 165 in term 42
```

トップヒットクエリの高速化

クエリ性能の改善によるユーザー体験の向上

Block-MAX WANDアルゴリズムによる新実装

アプリケーション検索、エンタープライズ検索のようなユースケースにマッチ

以下のユースケースは対象外:

- aggregation
- Kibana (aggregationを利用)
- 正確なヒット件数が必要

Script Score Query

カスタムスコアのための関数を定義

全フィールドが対象

_scoreも対象

正規化

- Saturation
- Logarithm
- Sigmoid

Painlessで記述も可能

もちろん提供済みの関数もOK

```
"script" : {
    "source" : "decayGeoExp(params.origin,
params.scale, params.offset, params.decay,
doc['location'].value)",
    "params": {
        "origin": "40, -70.12",
        "scale": "200km",
        "offset": "0km",
        "decay" : 0.2
    }
  }
```

Rank Features Query

新しいデータタイプ:

- rank_feature
- rank_features
  (rank_featureのベクトル)

関連ランキングスコアに追加可能

追加前の正規化も可能:

- Saturation
- Logarithm
- Sigmoid

top-kクエリの高速化にも利用できる設計のため、性能も向上

```
PUT my_index
{
  "mappings": {
    "properties": {
      "pagerank": {
        "type": "rank_feature"
      },
      "url_length": {
        "type": "rank_feature",

"positive_score_impact": false
      }
    }
  }
}
```

クラスター横断検索 (CCS) の改善

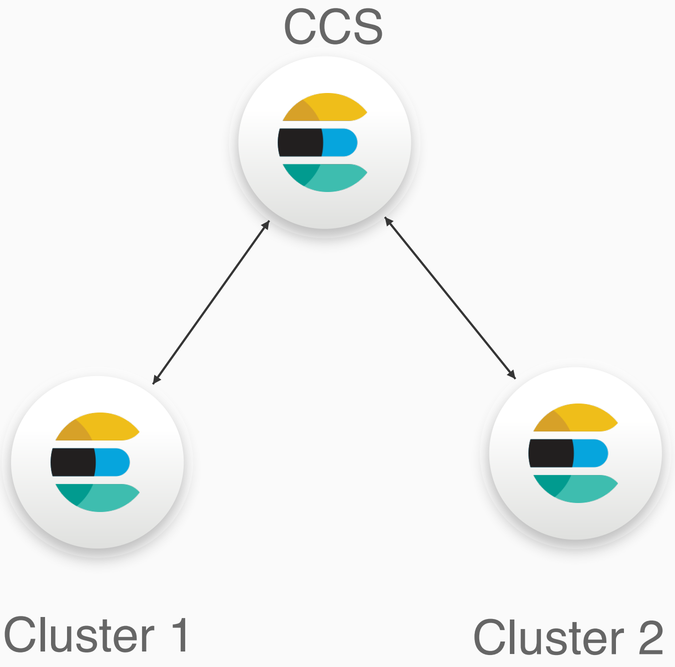
WANのために最適化された新実行モードの登場 (`ccs_minimize_roundtrips`)

6.7以前:

リモートクラスターにある各シャードからの応答により、ネットワーク上を多数の小さなリクエストが発生

7.0以降:

各リモートクラスターのcoordinatingノードから1度の応答により、ネットワーク上のリクエスト回数を削減



elastic

JVM のバンドル

ElasticsearchにJDK (OpenJDKを利用)をバンドルして配布

Javaのインストール手順をなくすことでインストールを簡易化

必要に応じてJVMを含まないバージョンもダウンロード可

7.0の大きな変更点

- number_of_indices=1がデフォルト（これまでは5）
  - _splitでもちろんこれまで通り分割も可能

- Hitsの形式の変更（後述）

- Aggregationの最大bucketsの制御が可能
  （ユーザーによる巨大なBucketsの指定）
  - もし、ユーザーが巨大な値を設定しても、real memory circuit-breakerが保護

7.0の大きな変更点

- number_of_indices=1がデフォルト（これまでは5）
  - _splitでもちろんこれまで通り分割も可能

- Hitsの形式の変更

- Aggregationの最大bucketsの制御が可能

  （ユーザーによる巨大なBucketsの指定）
  - もし、ユーザーが巨大な値を設定しても、real memory circuit-breakerが保護

そのほかの改善

- Adaptive replica selectionがデフォルトに

- 検索されないシャードのバックグラウンドRefreshをスキップ
  ⇒ 多くのユーザーのインデックススループットが向上

- High level Java REST clientが全機能に対応
  Transport clientが7.0で非推奨、8.0で廃止予定

- Nano-secondsのデータに対応
  date_nanos データタイプの登場

Kibana

新しい
ルック・アンド・フィール

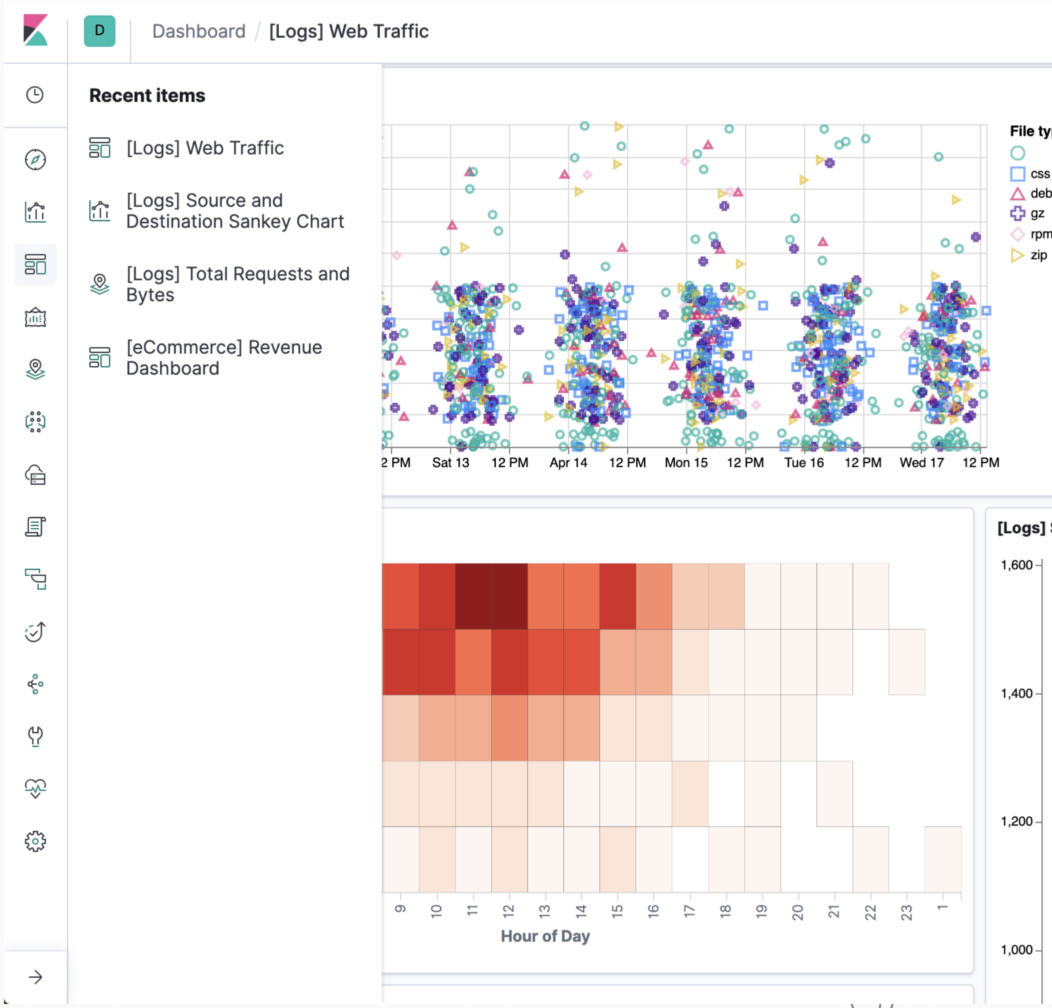
新しいグローバル・ナビゲーション

どこでもダークモード

レスポンシブなダッシュボード

タイムピッカーの改善

Kibanaクエリー言語がデフォルトに



Dashboard / [Logs] Web Traffic

Recent items

[Logs] Web Traffic

[Logs] Source and Destination Sankey Chart

[Logs] Total Requests and Bytes

[eCommerce] Revenue Dashboard

新しい
ルック·アンド·フィール

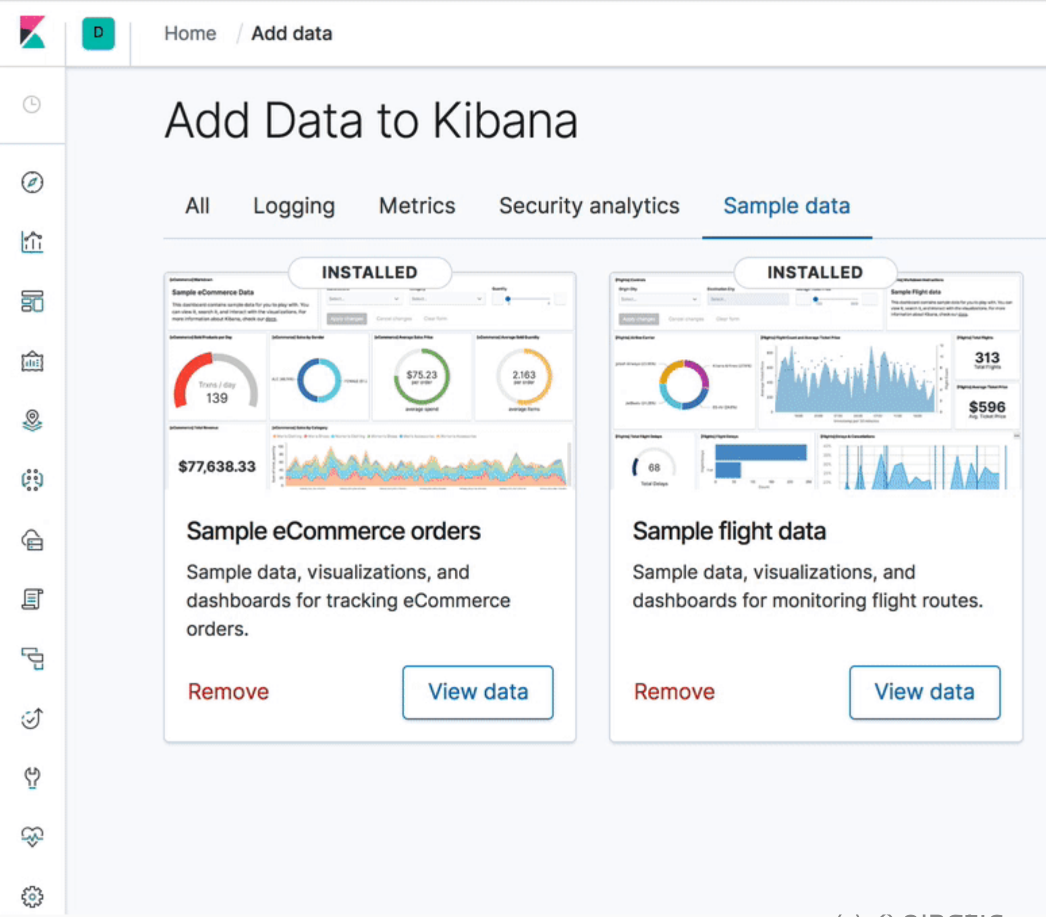
新しいグローバル・ナビゲーション

どこでもダークモード

レスポンシブなダッシュボード

タイムピッカーの改善

Kibanaクエリー言語がデフォルトに



Home / Add data

Add Data to Kibana

All | Logging | Metrics | Security analytics | Sample data

INSTALLED

Sample eCommerce orders

Sample data, visualizations, and dashboards for tracking eCommerce orders.

Remove | View data

INSTALLED

Sample flight data

Sample data, visualizations, and dashboards for monitoring flight routes.

Remove | View data

新しい
ルック・アンド・フィール

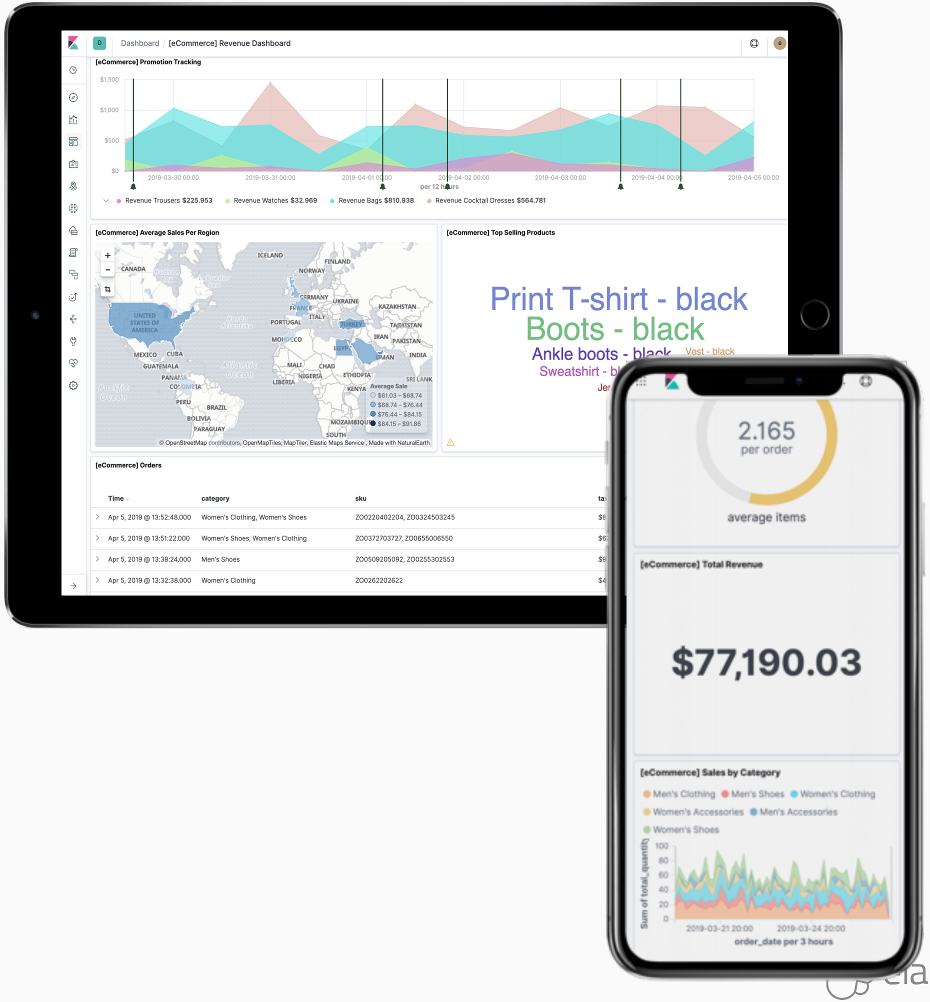
新しいグローバル・ナビゲーション

どこでもダークモード

レスポンシブなダッシュボード

タイムピッカーの改善

Kibanaクエリー言語がデフォルトに



Time	category	sku	ta
Apr 5, 2019 @ 13:52:48.000	Women's Clothing, Women's Shoes	ZO0220402204, ZO0324503245	$8
Apr 5, 2019 @ 13:51:22.000	Women's Shoes, Women's Clothing	ZO0372703727, ZO0655006550	$6
Apr 5, 2019 @ 13:38:24.000	Men's Shoes	ZO0509205092, ZO0255302553	$9
Apr 5, 2019 @ 13:32:38.000	Women's Clothing	ZO0262202622	$4

新しい
ルック・アンド・フィール

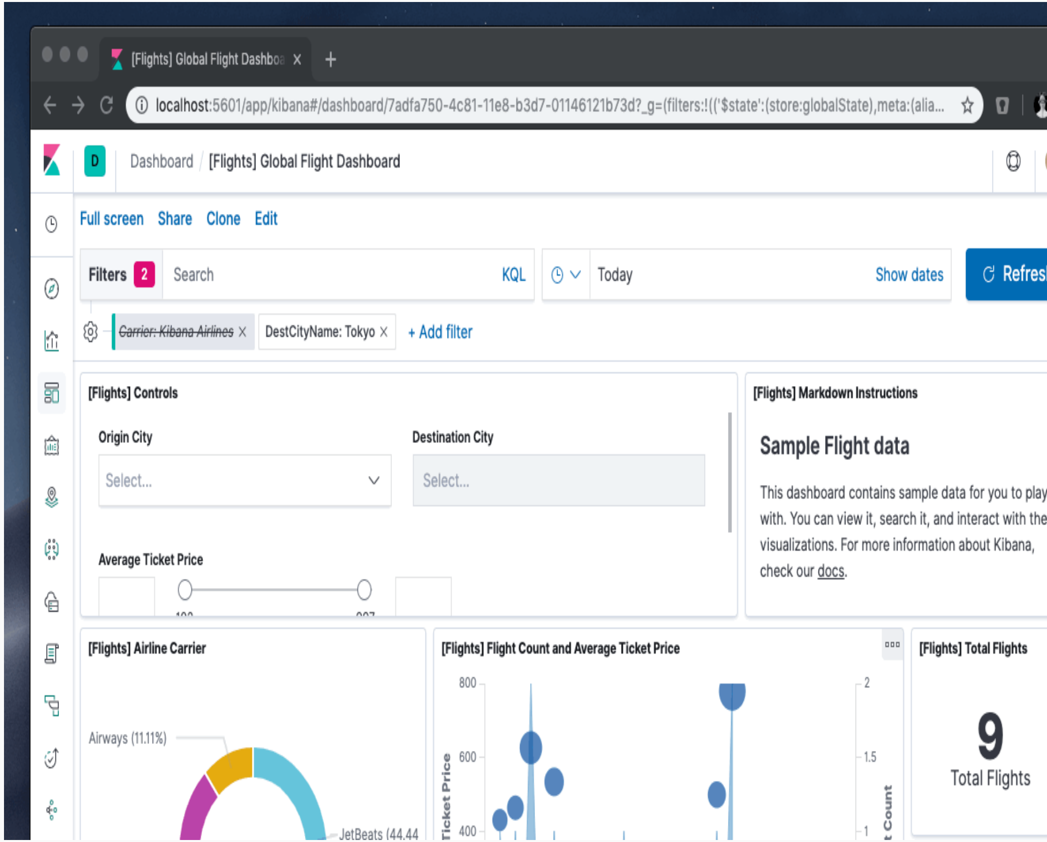
新しいグローバル・ナビゲーション

どこでもダークモード

レスポンシブなダッシュボード

タイムピッカーの改善

Kibanaクエリー言語がデフォルトに



localhost:5601/app/kibana#/dashboard/7adfa750-4c81-11e8-b3d7-01146121b73d?_g=(filters:!(('$state':(store:globalState),meta:(alia...

Dashboard / [Flights] Global Flight Dashboard

~~Carrier: Kibana Airlines~~ X

DestCityName: Tokyo X

Sample Flight data

This dashboard contains sample data for you to play with. You can view it, search it, and interact with the visualizations. For more information about Kibana, check our docs.

新しい
ルック·アンド·フィール

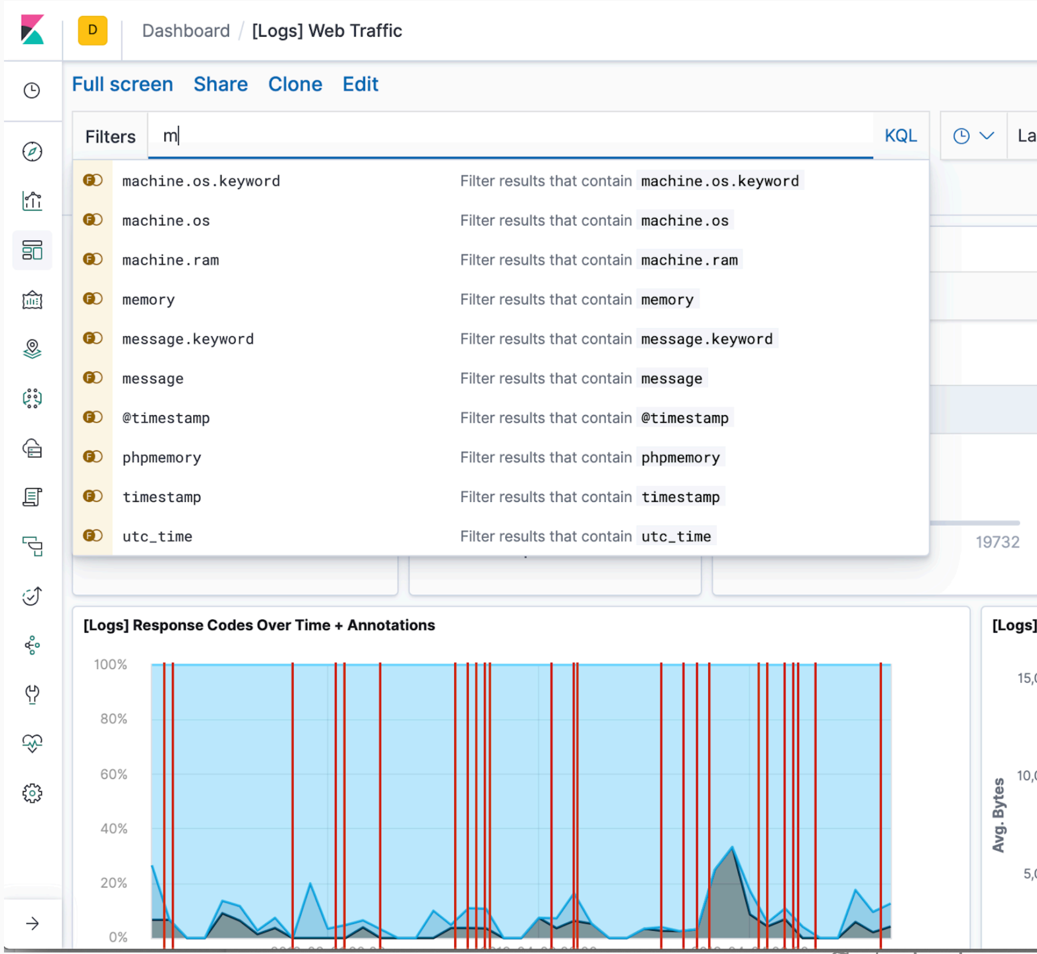
新しいグローバル・ナビゲーション

どこでもダークモード

レスポンシブなダッシュボード

タイムピッカーの改善

Kibanaクエリー言語がデフォルトに



Dashboard / [Logs] Web Traffic

Full screen Share Clone Edit

Filters m| KQL

machine.os.keyword	Filter results that contain machine.os.keyword
machine.os	Filter results that contain machine.os
machine.ram	Filter results that contain machine.ram
memory	Filter results that contain memory
message.keyword	Filter results that contain message.keyword
message	Filter results that contain message
@timestamp	Filter results that contain @timestamp
phpmemory	Filter results that contain phpmemory
timestamp	Filter results that contain timestamp
utc_time	Filter results that contain utc_time

[Logs] Response Codes Over Time + Annotations

Beats

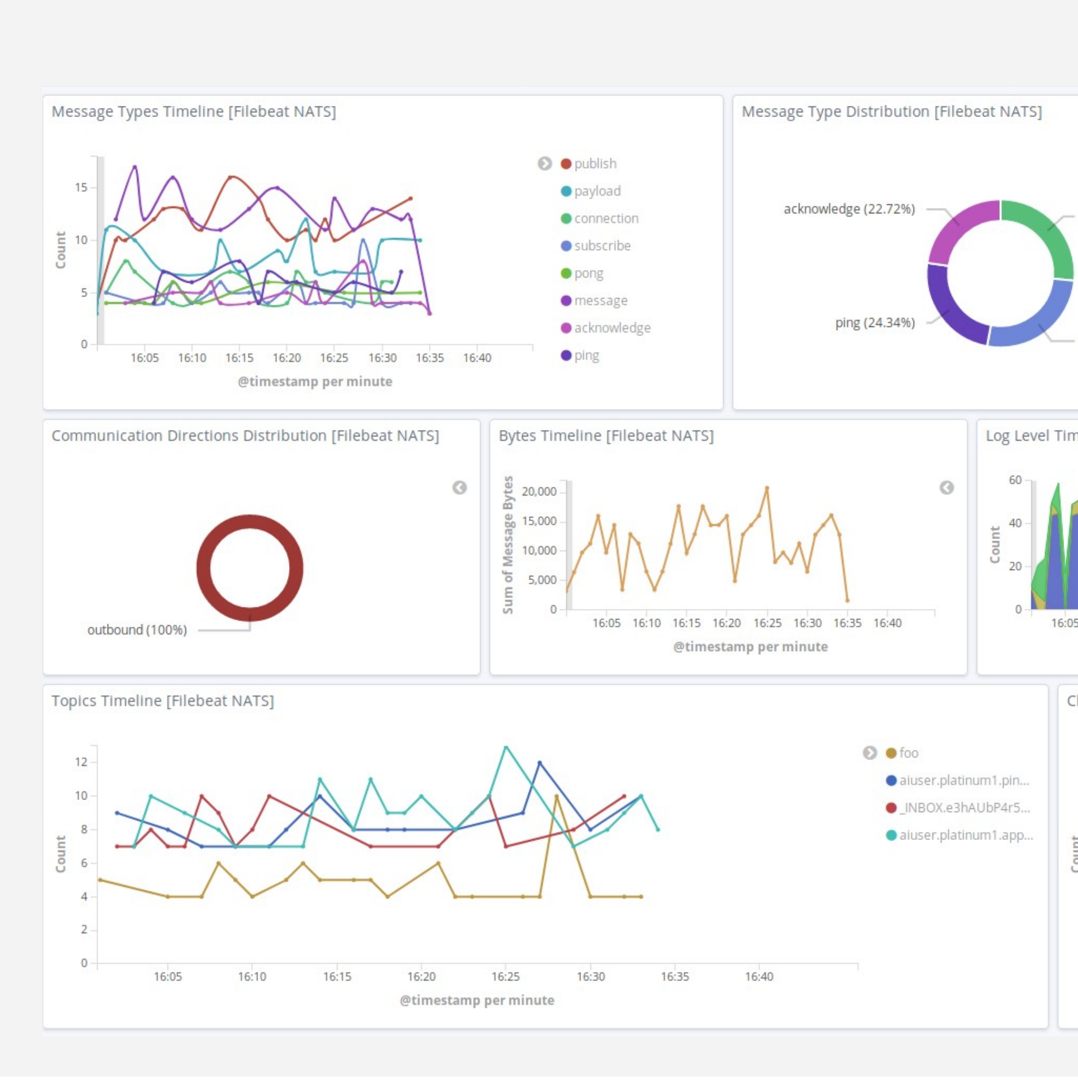
新しいBeatsモジュール

新しいFilebeatモジュール

- Zeek (fka Bro) (Basic)
- IPtables (Basic)
- Santa (OSS)

新しいMetricbeatモジュール

- AWS EC2 (Basic)
- Microsoft SQL (Basic)
- NATS (OSS)
- CouchDB (OSS)



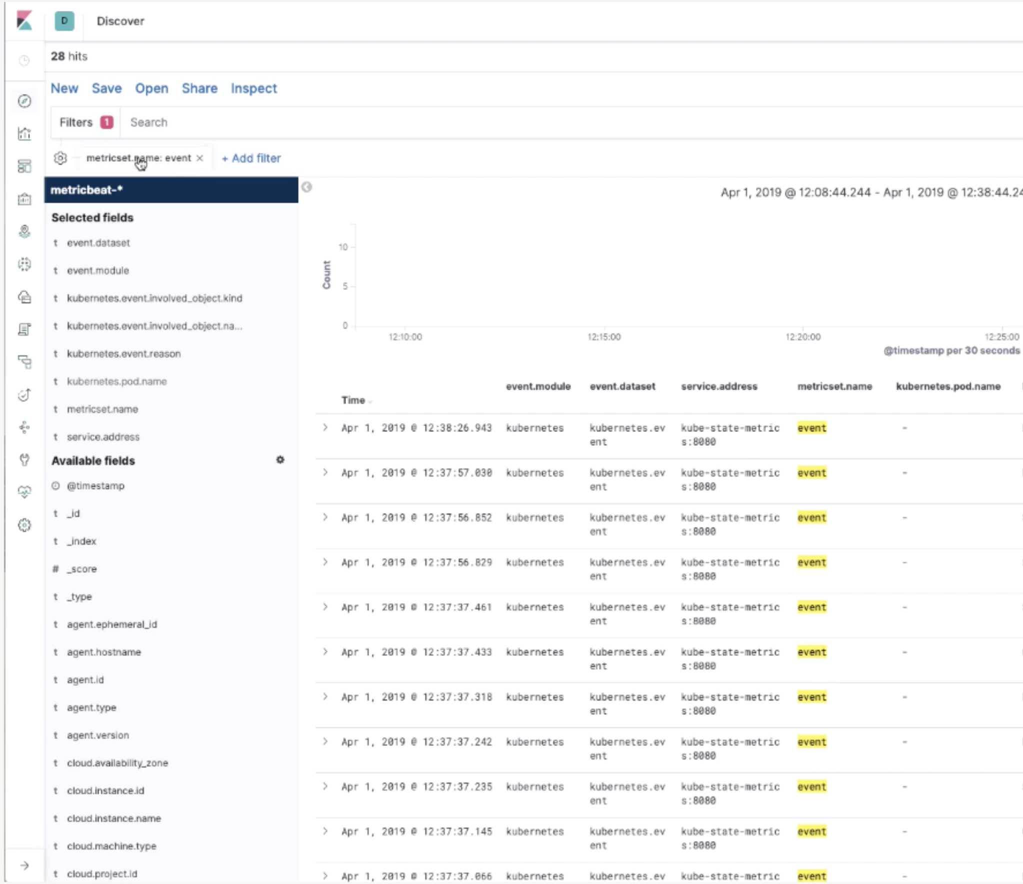
モジュールの成熟

GAになったモジュール

- Golang
- Graphite
- Munin
- Prometheus

Betaになったモジュール

- System module (Auditbeat)



28 hits

metricbeat-*

Selected fields

- t event.dataset
- t event.module
- t kubernetes.event.involved_object.kind
- t kubernetes.event.involved_object.na...
- t kubernetes.event.reason
- t kubernetes.pod.name
- t metricset.name
- t service.address

Available fields

- @timestamp
- t _id
- t _index
- # _score
- t _type
- t agent.ephemeral_id
- t agent.hostname
- t agent.id
- t agent.type
- t agent.version
- t cloud.availability_zone
- t cloud.instance.id
- t cloud.instance.name
- t cloud.machine.type
- t cloud.project.id

Apr 1, 2019 @ 12:08:44.244 - Apr 1, 2019 @ 12:38:44.24

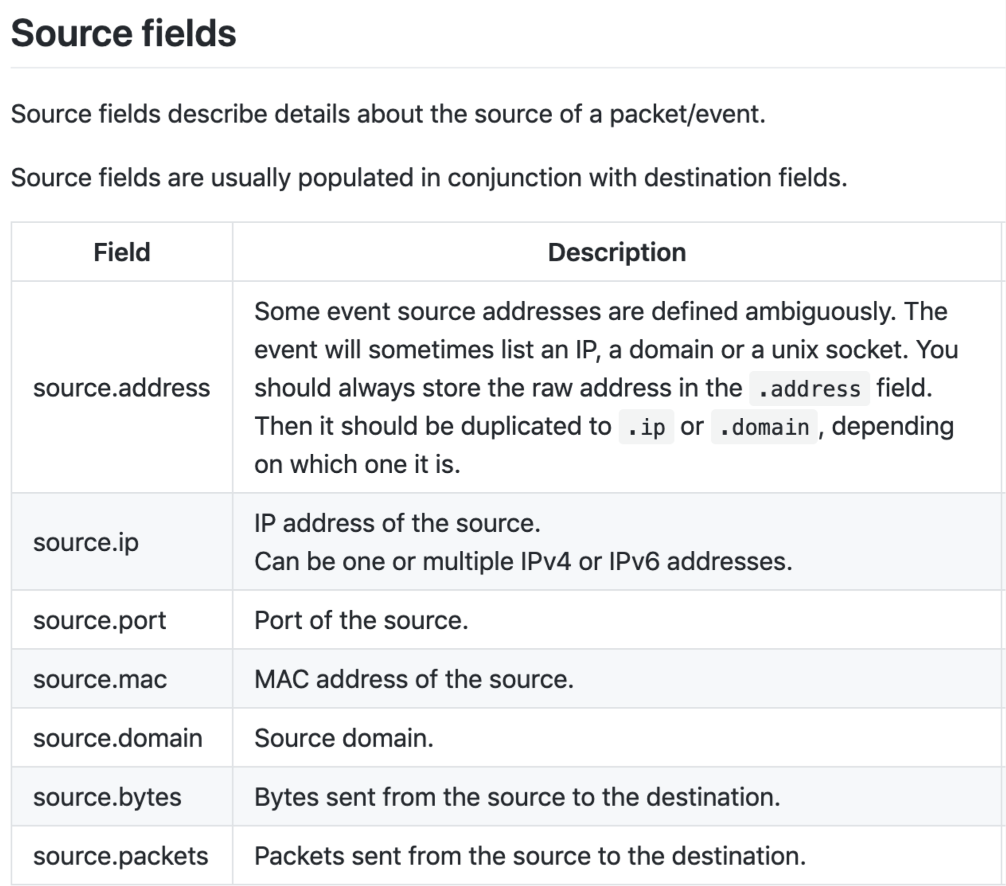
Time	event.module	event.dataset	service.address	metricset.name	kubernetes.pod.name
Apr 1, 2019 @ 12:38:26.943	kubernetes	kubernetes.event	kube-state-metrics:8080	event	-
Apr 1, 2019 @ 12:37:57.030	kubernetes	kubernetes.event	kube-state-metrics:8080	event	-
Apr 1, 2019 @ 12:37:56.852	kubernetes	kubernetes.event	kube-state-metrics:8080	event	-
Apr 1, 2019 @ 12:37:56.829	kubernetes	kubernetes.event	kube-state-metrics:8080	event	-
Apr 1, 2019 @ 12:37:37.461	kubernetes	kubernetes.event	kube-state-metrics:8080	event	-
Apr 1, 2019 @ 12:37:37.433	kubernetes	kubernetes.event	kube-state-metrics:8080	event	-
Apr 1, 2019 @ 12:37:37.318	kubernetes	kubernetes.event	kube-state-metrics:8080	event	-
Apr 1, 2019 @ 12:37:37.242	kubernetes	kubernetes.event	kube-state-metrics:8080	event	-
Apr 1, 2019 @ 12:37:37.235	kubernetes	kubernetes.event	kube-state-metrics:8080	event	-
Apr 1, 2019 @ 12:37:37.145	kubernetes	kubernetes.event	kube-state-metrics:8080	event	-
Apr 1, 2019 @ 12:37:37.066	kubernetes	kubernetes.ev	kube-state-metric	event	-

ECSのサポート

ほとんどのBeatsとモジュールはECSフォーマットでデータを生成

Beatsのadd_*プロセッサがECSをサポート

**WinlogbeatとFunctionbeat、JournalbeatはECSを限定的にサポート



Source fields

Source fields describe details about the source of a packet/event.

Source fields are usually populated in conjunction with destination fields.

Field	Description
source.address	Some event source addresses are defined ambiguously. The event will sometimes list an IP, a domain or a unix socket. You should always store the raw address in the `.address` field. Then it should be duplicated to `.ip` or `.domain`, depending on which one it is.
source.ip	IP address of the source. Can be one or multiple IPv4 or IPv6 addresses.
source.port	Port of the source.
source.mac	MAC address of the source.
source.domain	Source domain.
source.bytes	Bytes sent from the source to the destination.
source.packets	Packets sent from the source to the destination.

Elastic Common Schema (ECS)

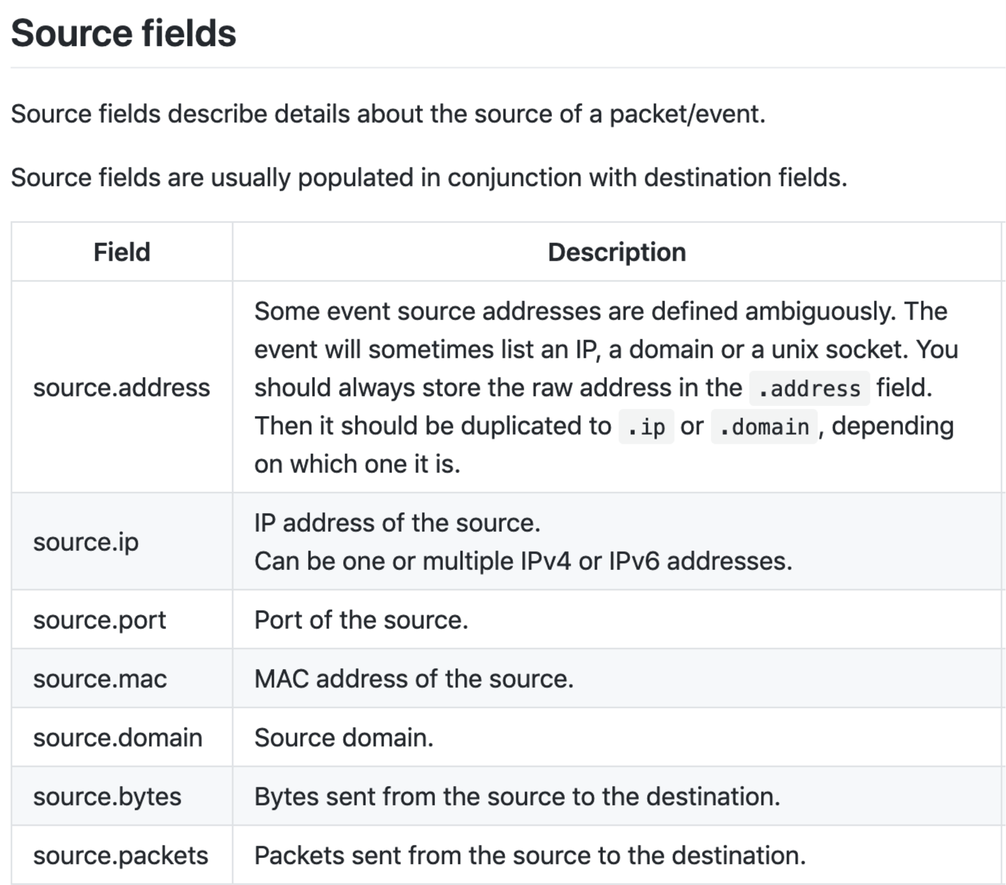
合理的な分析のための正規化

Elastic Common Schema (ECS)を、2019年3月に公開

Elasticsearchへのデータ投入に際して、フィールドとオブジェクトの共通セットを定義

多様なデータの**横断分析**を可能にする**拡張可能**な設計

https://github.com/elastic/ecs
貢献とフィードバックを歓迎します



Source fields

Source fields describe details about the source of a packet/event.

Source fields are usually populated in conjunction with destination fields.

Field	Description
source.address	Some event source addresses are defined ambiguously. The event will sometimes list an IP, a domain or a unix socket. You should always store the raw address in the `.address` field. Then it should be duplicated to `.ip` or `.domain`, depending on which one it is.
source.ip	IP address of the source. Can be one or multiple IPv4 or IPv6 addresses.
source.port	Port of the source.
source.mac	MAC address of the source.
source.domain	Source domain.
source.bytes	Bytes sent from the source to the destination.
source.packets	Packets sent from the source to the destination.

Beatsコア機能

Add_* プロセッサがECSフィールドを含める

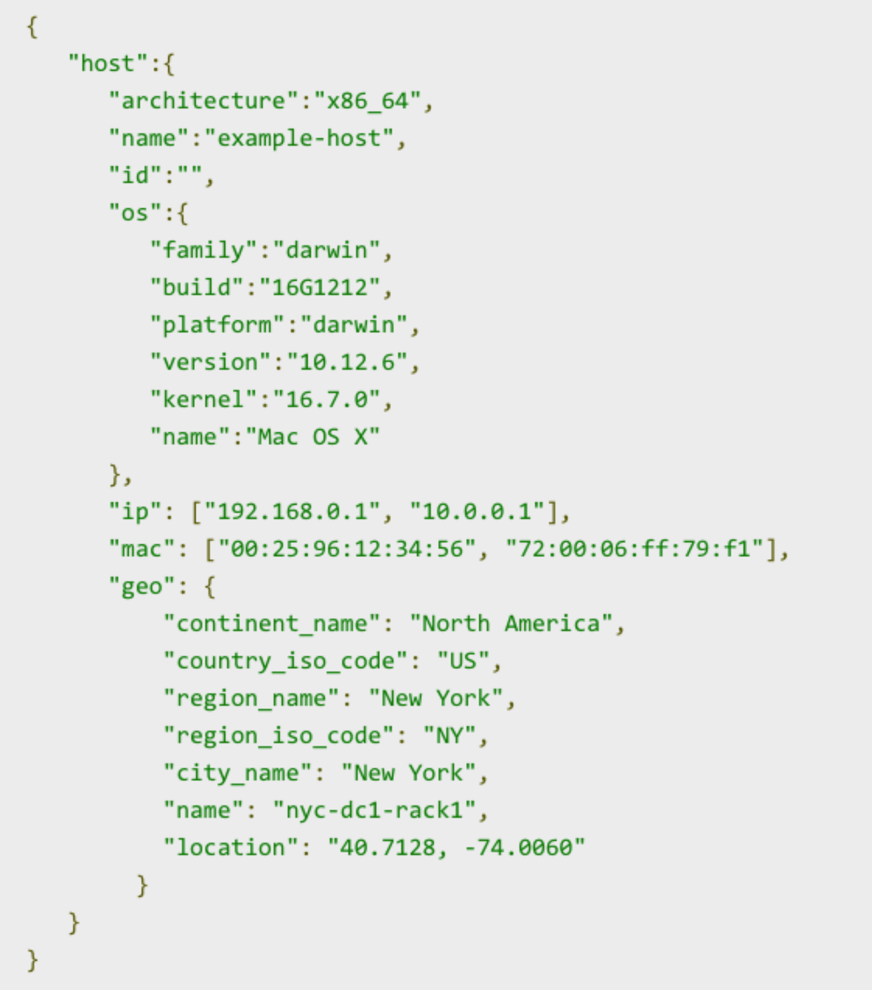
- Geo info
- OS name

新しい**Beatsプロセッサ**が利用可能

- **Add_fields**
- **Add_labels**
- **Add_tags**

新しい**Filebeatエンコーディング**

- Latin
- IBM codepages
- Cyrillic
- Macintosh
- Windows



```
{
   "host":{
      "architecture":"x86_64",
      "name":"example-host",
      "id":"",
      "os":{
         "family":"darwin",
         "build":"16G1212",
         "platform":"darwin",
         "version":"10.12.6",
         "kernel":"16.7.0",
         "name":"Mac OS X"
      },
      "ip": ["192.168.0.1", "10.0.0.1"],
      "mac": ["00:25:96:12:34:56", "72:00:06:ff:79:f1"],
      "geo": {
          "continent_name": "North America",
          "country_iso_code": "US",
          "region_name": "New York",
          "region_iso_code": "NY",
          "city_name": "New York",
          "name": "nyc-dc1-rack1",
          "location": "40.7128, -74.0060"
        }
   }
}
```

Logstash

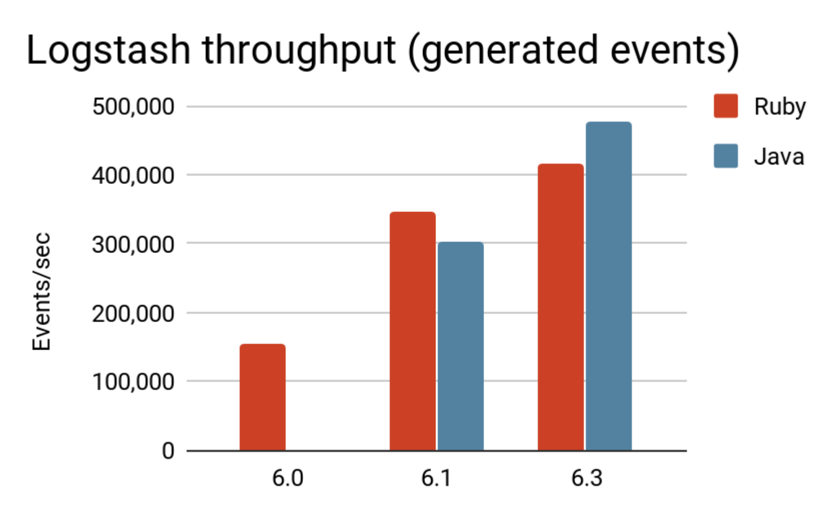
LogstashのエンハンスLogstashのエンハンス

初期値でJava Execution Engineを使用

⇒ 高パフォーマンス、短い起動時間、メモリ使用量の削減

新しいプラグインのサポート

- CIDR filter
- Clone filter
- Prune filter

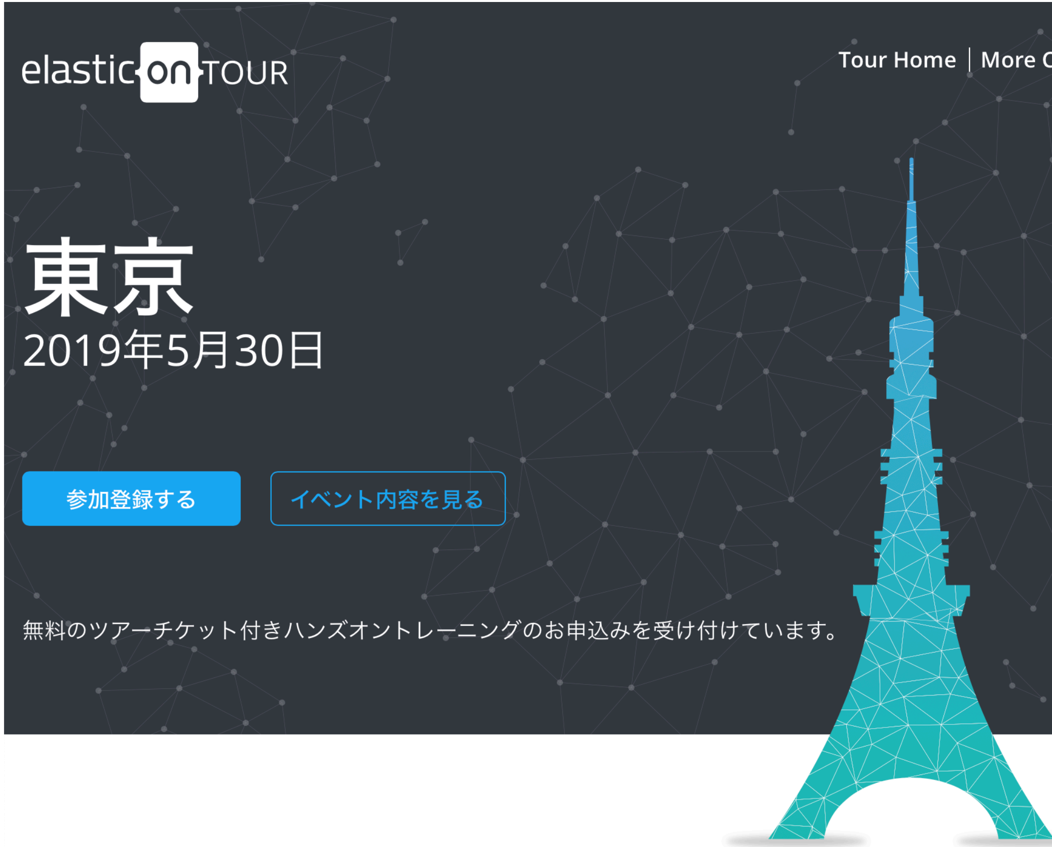


elastic

Thank You

Elastic{ON} Tour Tokyo

- Elasticsearch、Kibana、Beats、そしてLogstashの最新ロードマップが公開されます。Elasticのエキスパートや Elastic Stackユーザーから活用のヒントを得る機会にもなります。



東京

2019年5月30日

無料のツアーチケット付きハンズオントレーニングのお申込みを受け付けています。