elastic

Elastic Stack 7.0

新機能ツアー

Kosho Owa, Principal Solution Architect
Jun Ohtani, Developer | Evangelist

Elasticsearch

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

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

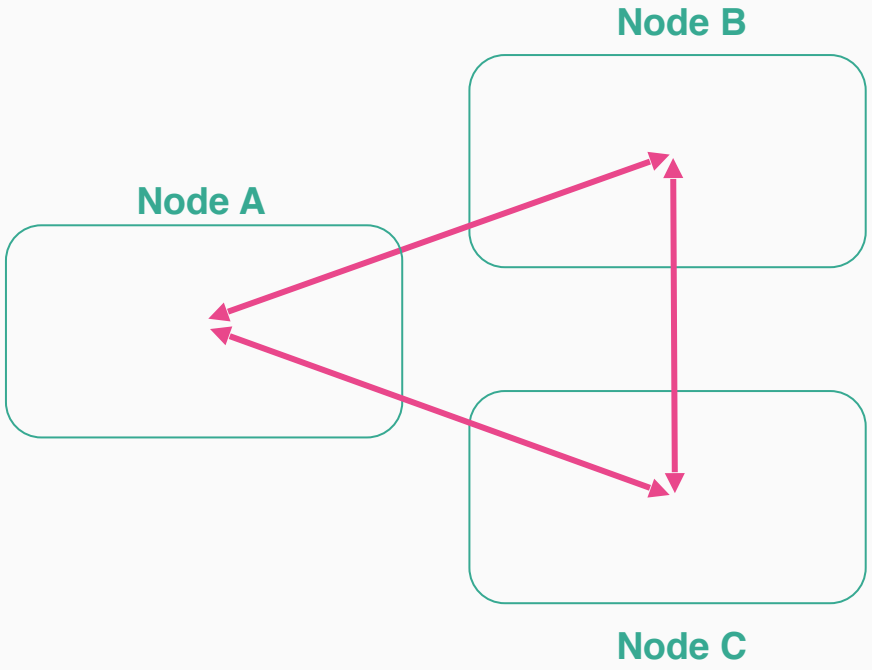
形式モデルで検証済み

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



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)

Script Score Query

Experimental

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

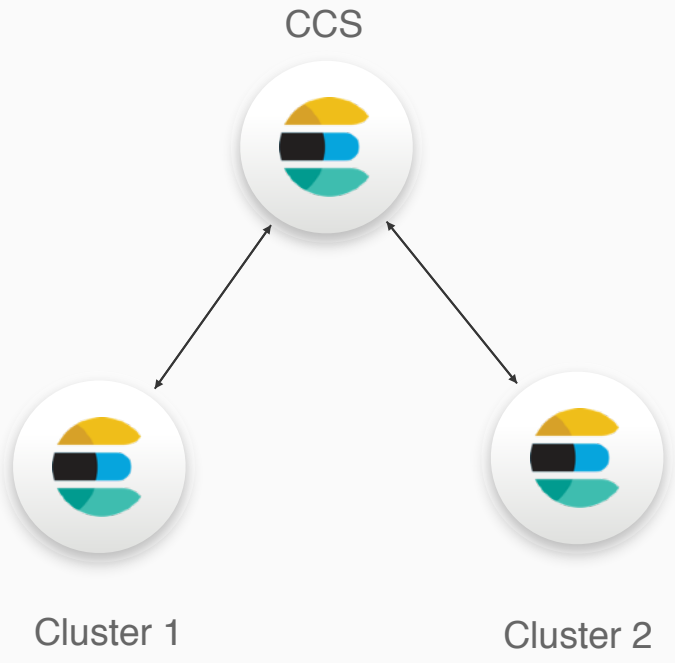
WANのために最適化された新実行モードの登場
(`ccs_minimize_roundtrips`)

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



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

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

新しい
ルック・アンド・フィール

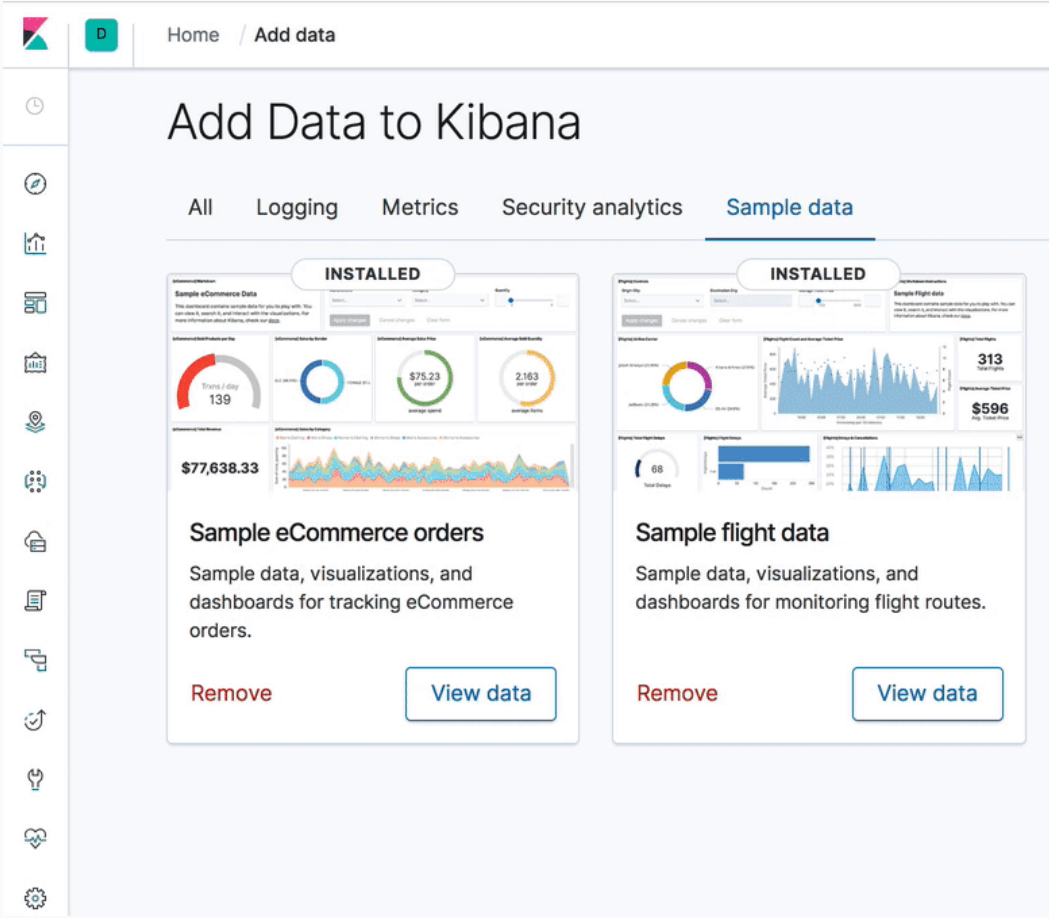
新しいグローバル・ナビゲーション

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



新しい
ルック・アンド・フィール

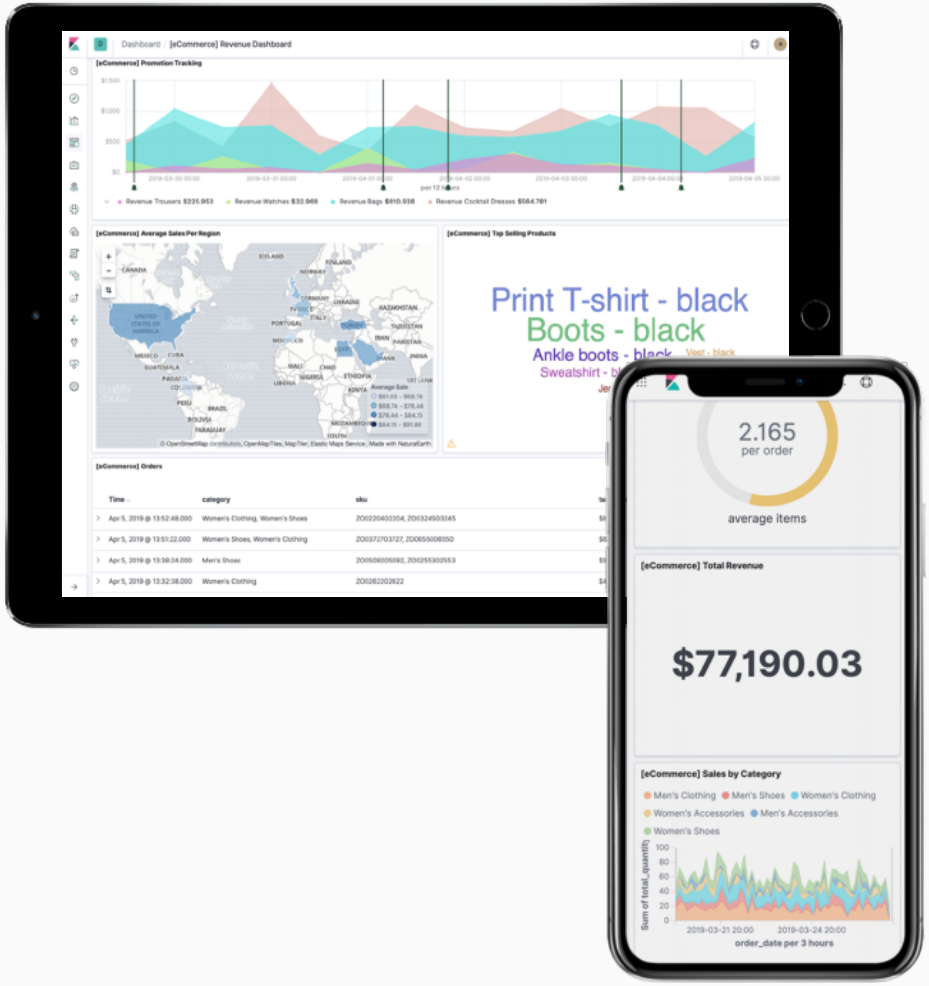
新しいグローバル・ナビゲーション

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



新しい
ルック・アンド・フィール

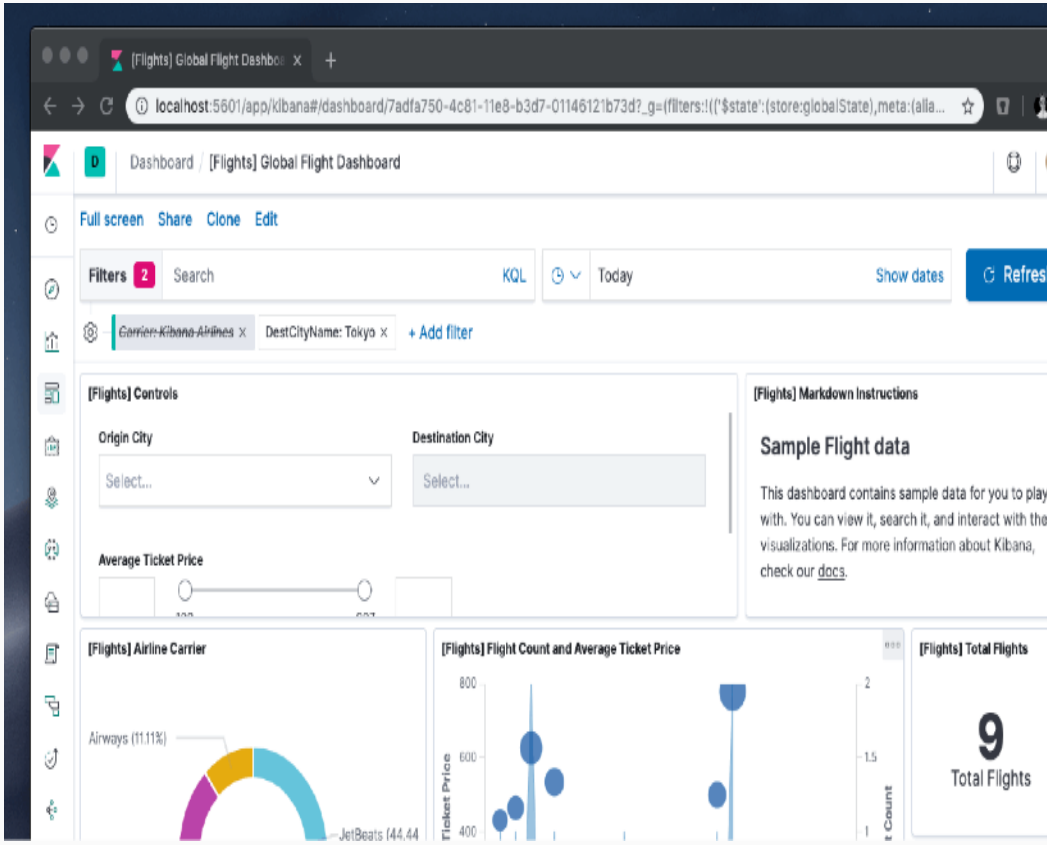
新しいグローバル・ナビゲーション

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



新しい
ルック・アンド・フィール

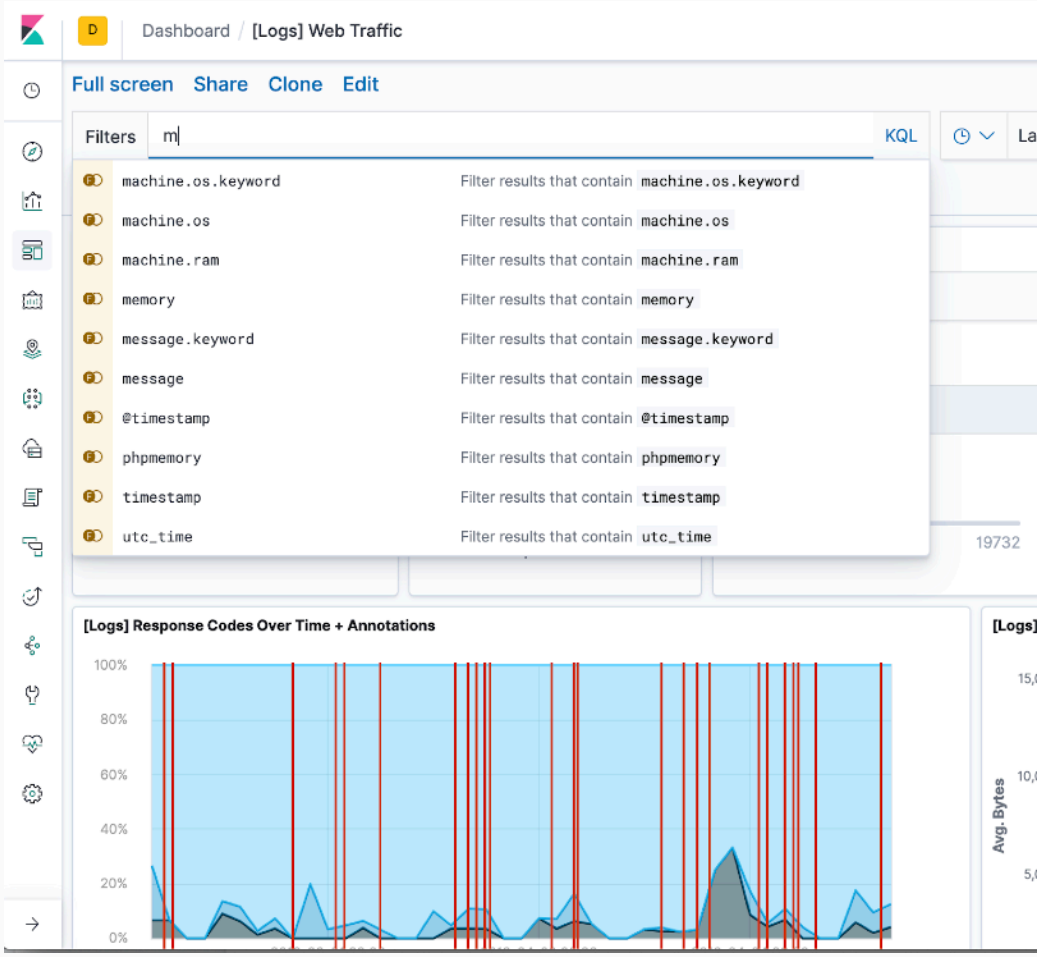
新しいグローバル・ナビゲーション

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



Beats

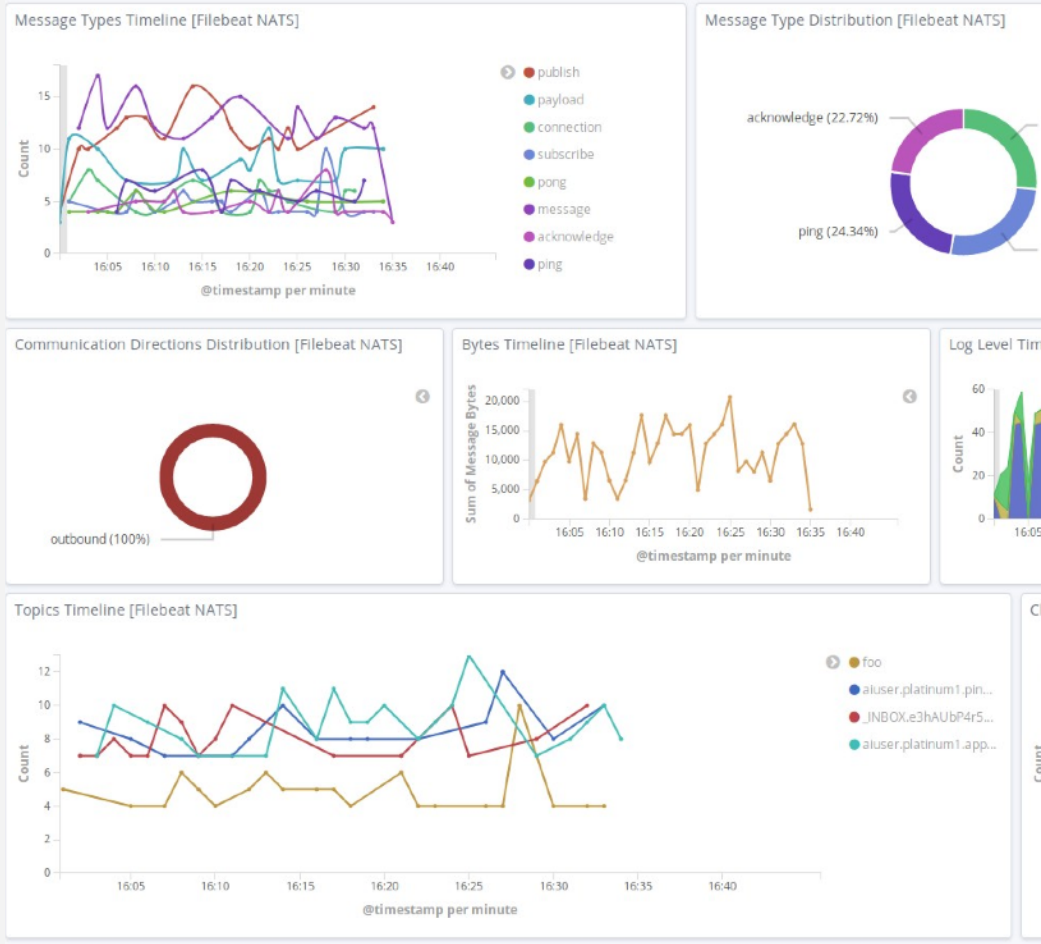
新しいBeatsモジュール

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)



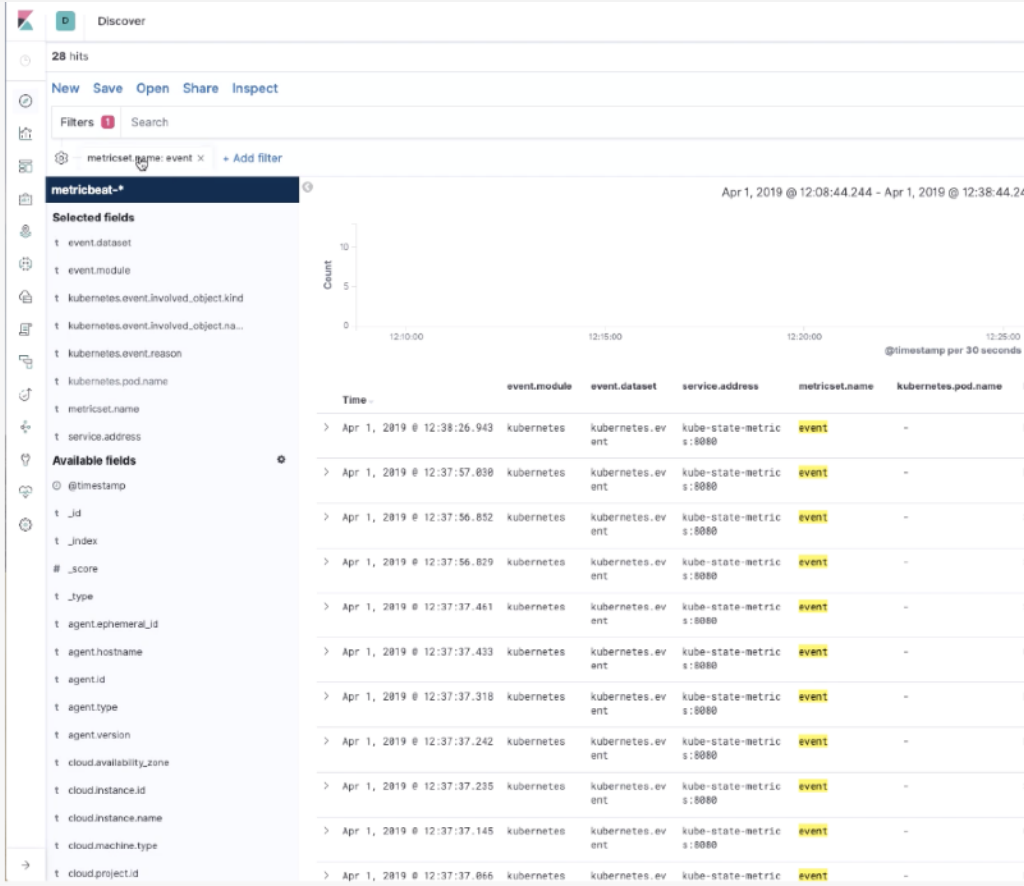
モジュールの成熟

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)



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

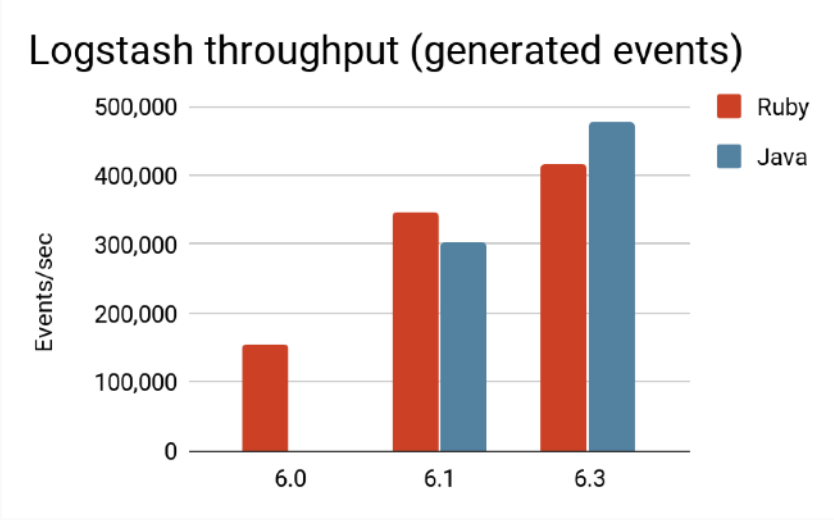
Logstashのエンハンス

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