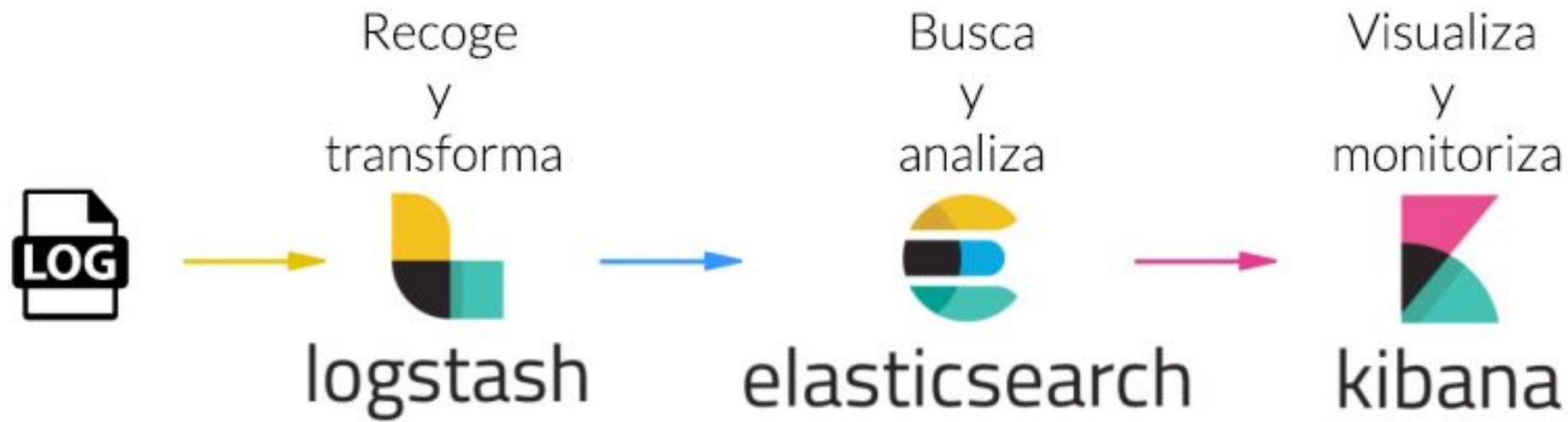




elasticsearch

logstash

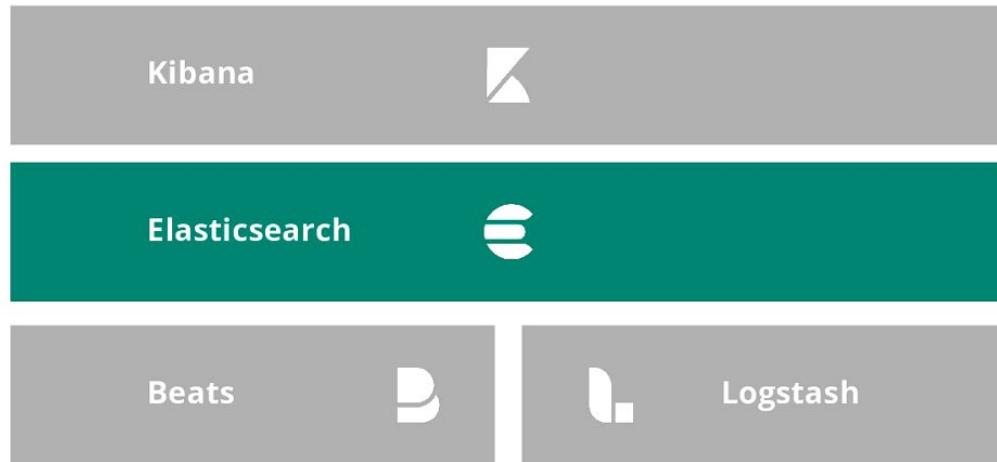
kibana



Elasticsearch



Stack Elastic



Visualizar y
Gestionar

Almacenar,
Buscar, Analizar

Ingestar

INTRODUCCIÓN A ELASTICSEARCH

- Introducción y conceptos
- Preparación del entorno
- Arquitectura

Introducción y conceptos

<https://db-engines.com/en/ranking/search+engine>

Rank			DBMS	Database Model	Score		
Sep 2021	Aug 2021	Sep 2020			Sep 2021	Aug 2021	Sep 2020
1.	1.	1.	Elasticsearch	Search engine, Multi-model 	160.24	+3.16	+9.74
2.	2.	2.	Splunk	Search engine	91.61	+1.01	+3.71
3.	3.	3.	Solr	Search engine, Multi-model 	49.81	-1.26	-1.81
4.	4.	4.	MarkLogic 	Multi-model 	9.60	+0.45	-2.34
5.	5.	↑ 7.	Sphinx	Search engine	7.64	-0.11	+1.17
6.	6.	↓ 5.	Algolia	Search engine	7.30	+0.24	+0.53
7.	7.	↓ 6.	Microsoft Azure Search	Search engine	6.85	+0.71	+0.15
8.	8.	8.	ArangoDB 	Multi-model 	4.79	+0.53	-1.01
9.	9.	↑ 10.	Virtuoso 	Multi-model 	4.42	+0.19	+1.86
10.	10.	↓ 9.	Amazon CloudSearch	Search engine	2.20	+0.02	-0.41

Introducción y conceptos

Elasticsearch es un motor de búsqueda:

- Desarrollado en java.
- Open source.
- Distribuido.
- Escalable.
- Basado en lucene.



Introducción y conceptos

Lucene es una librería de búsqueda de texto:

- Desarrollado en java.
- Open source.
- Escalable.
- Con alto rendimiento.
- Basada en **índices invertidos**.



Introducción y conceptos



- Generación de índices invertidos
- Búsqueda sobre índices invertidos
- Características adicionales:
 - Analizadores de texto
 - Ordenaciones
 - Resaltado de coincidencias
 - Corrector ortográfico
 - ...

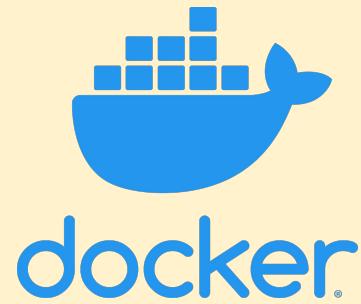


- Interfaz REST
- Distribución
- Alta disponibilidad
- Funcionalidades adicionales:
 - Herramientas de análisis
 - Múltiples índices
 - Gestión de cluster
 - ...

Introducción y conceptos

Clúster	Nodos	Índices	Tipos
Conjunto de instancias de ES que comparten el mismo nombre (<code>cluster.name</code>)	Instancia de ElasticSearch	Colección de varios documentos (objeto JSON)	Colección de varios documentos de similar estructura
		Comparable a esquemas de bases de datos	Comparable a tablas de bases de datos
		No confundir con índices de bases de datos	

Preparación del entorno



<https://github.com/jmortega/curso-elk>

<https://elk-docker.readthedocs.io/>

<https://github.com/deviantony/docker-elk>

Preparación del entorno

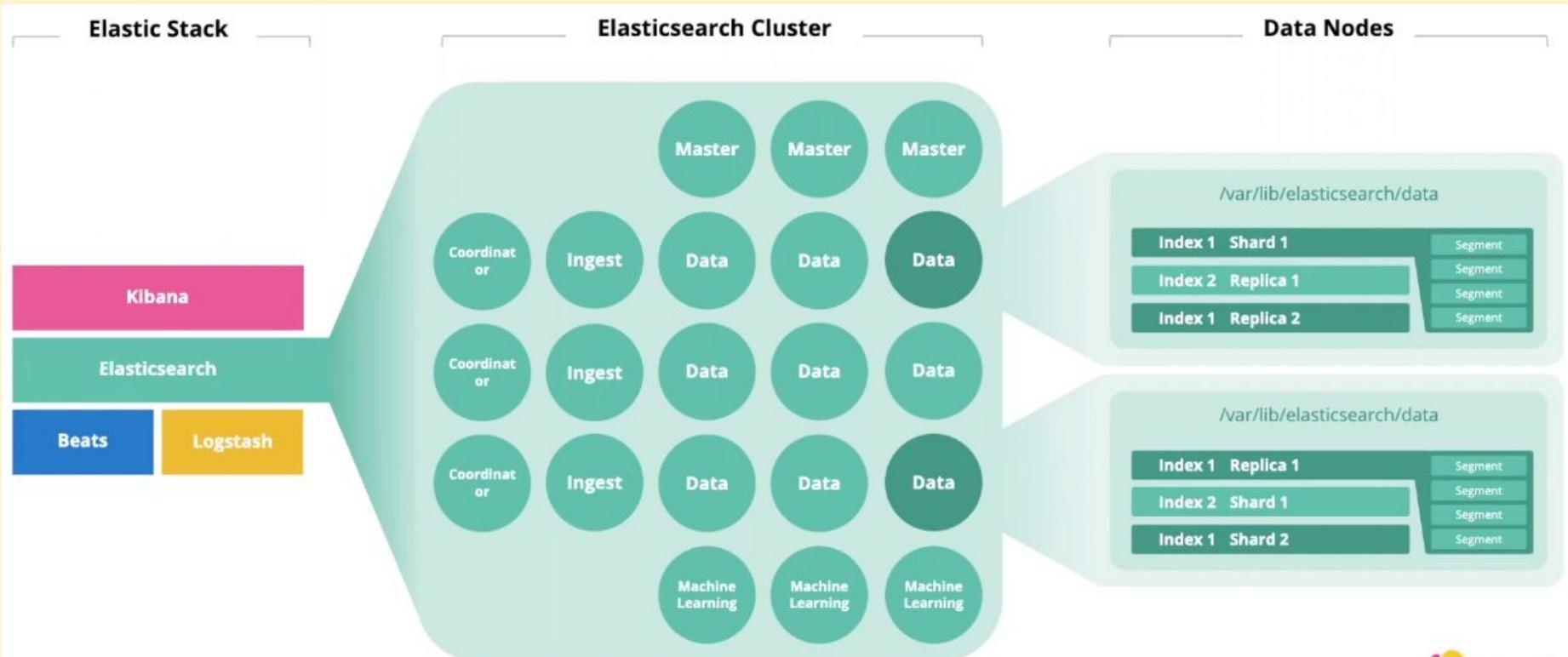
```
$ sudo sysctl -w vm.max_map_count=262144  
$ sudo sysctl -p  
$ cat /proc/sys/vm/max_map_count  
262144
```



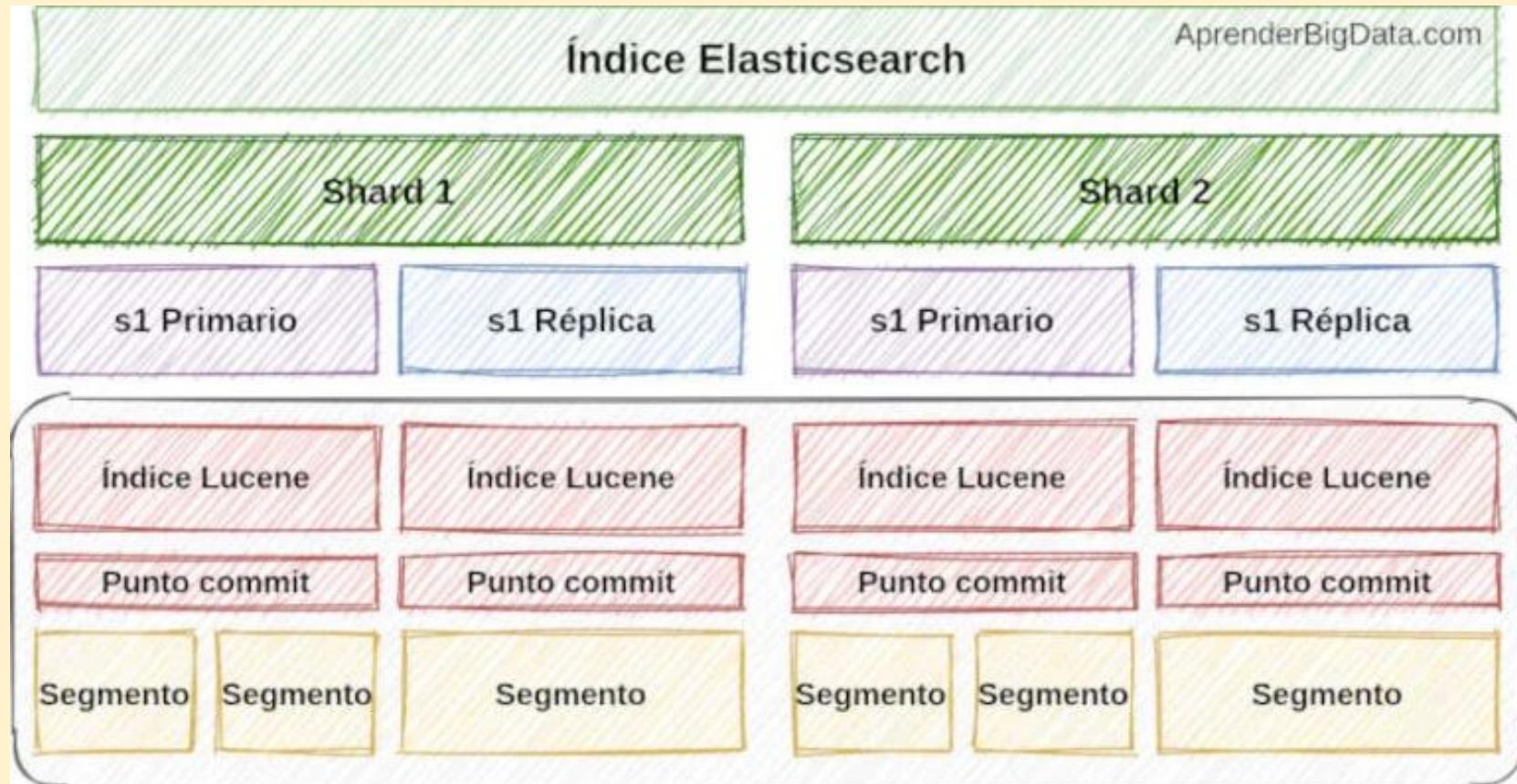
```
wsl -d docker-desktop  
sysctl -w vm.max_map_count=262144
```



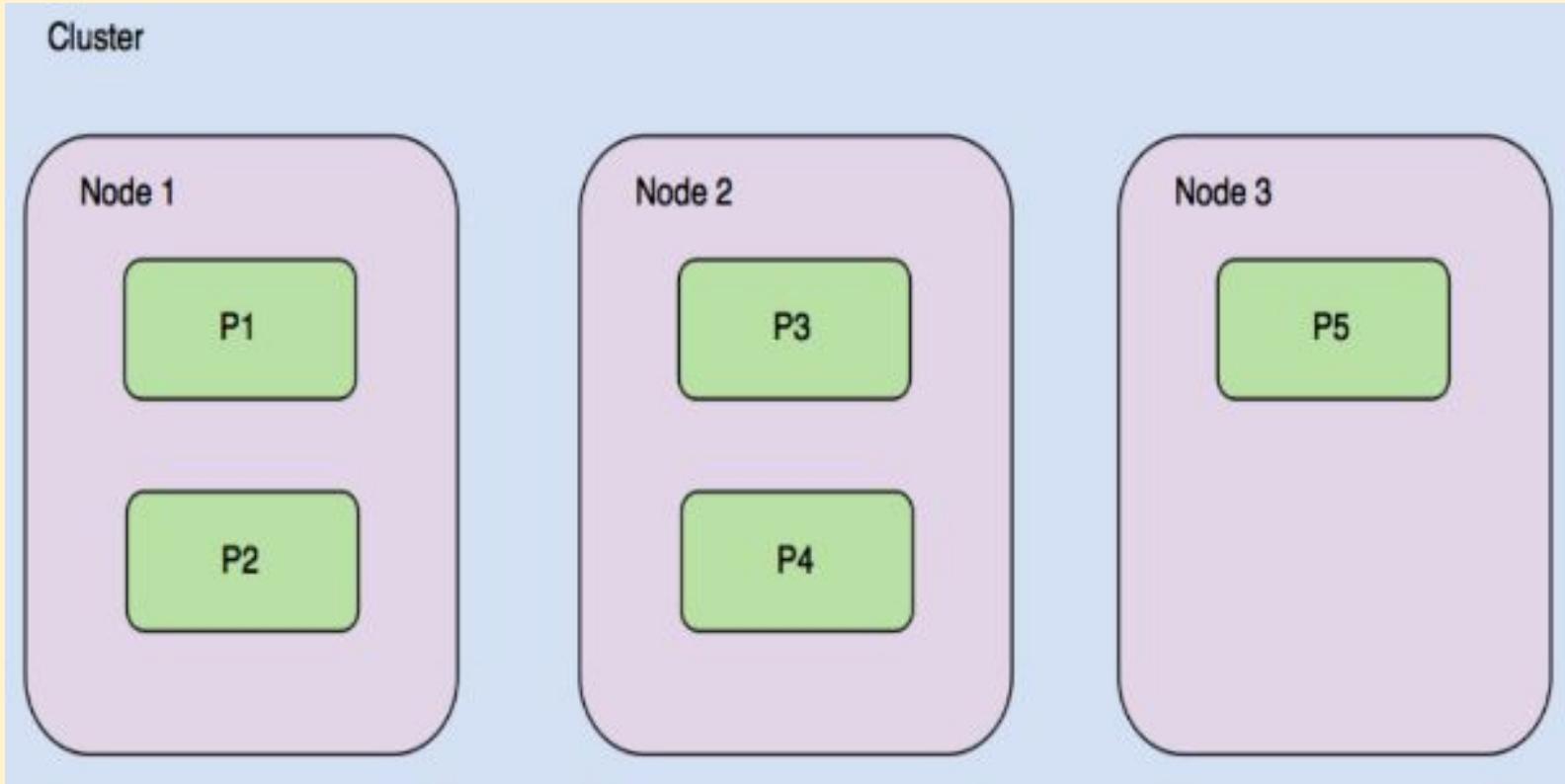
Arquitectura



Shards

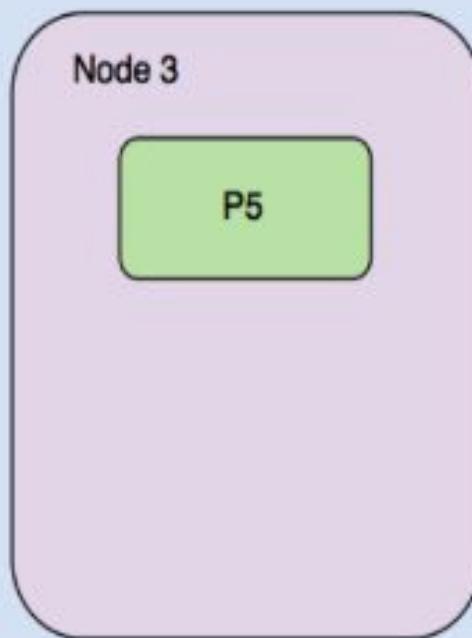
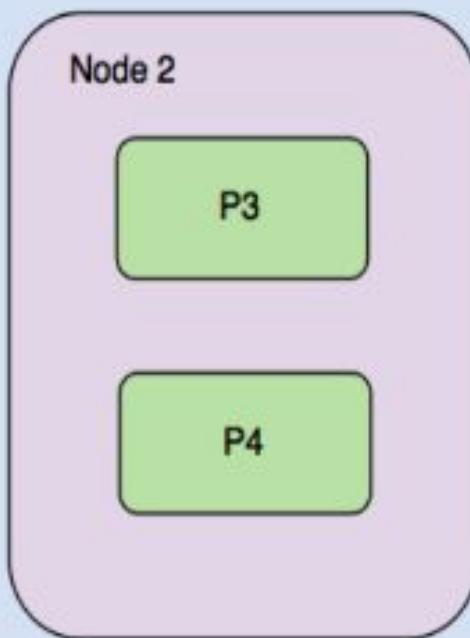
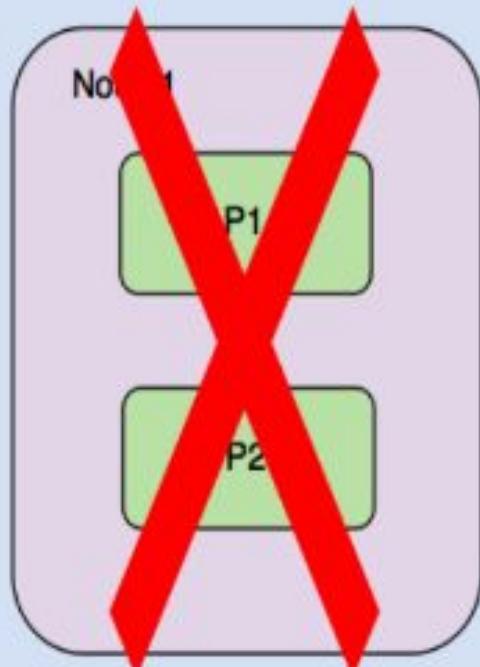


Shards y réplicas

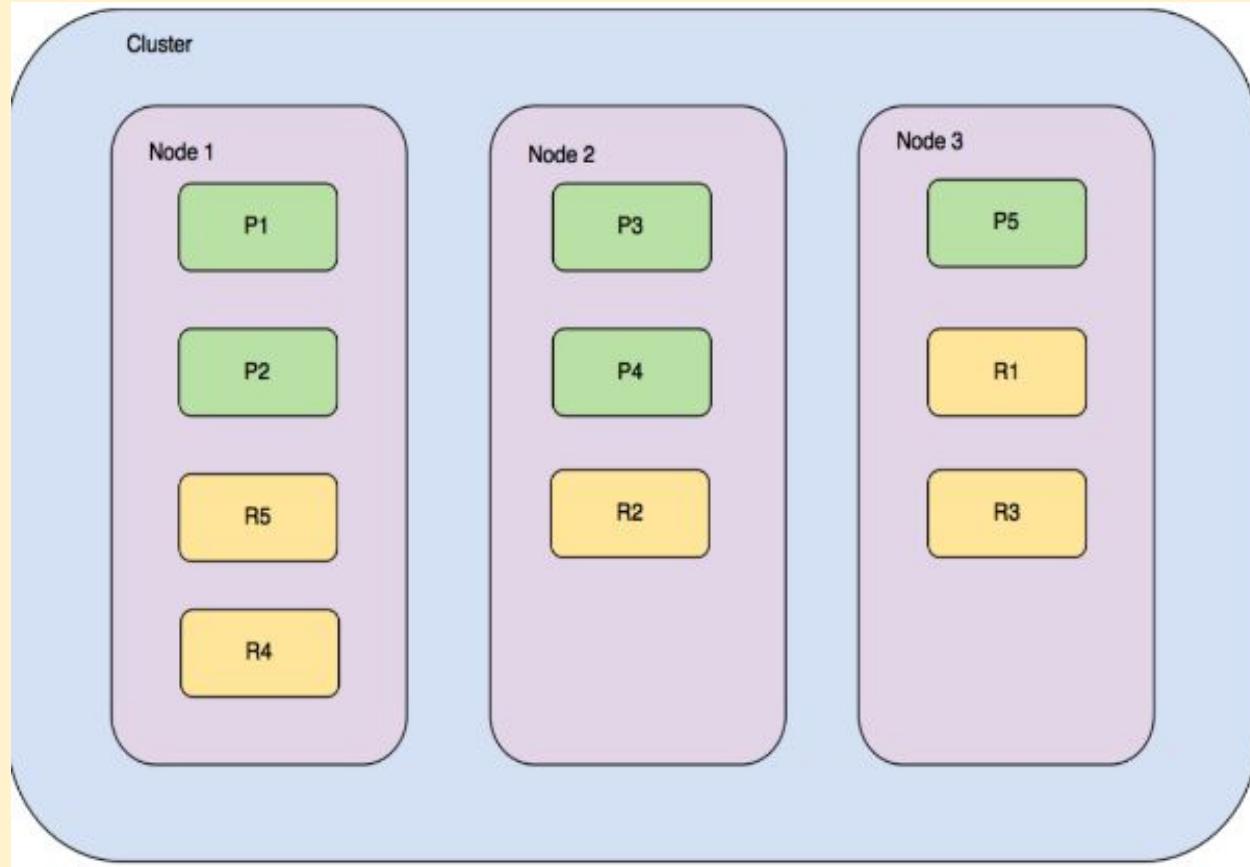


Shards y réplicas

Cluster



Shards y réplicas



Estados del cluster



Verde: Todo OK. Todos los shards y sus réplicas se han asignado a los nodos del cluster.

Amarillo: Todos los shards primarios se han asignado a nodos del cluster. Una o más réplicas no se han podido asignar a ningún nodo.

Rojo: Uno o más shards primarios no se han podido asignar a ningún nodo.

Estados del cluster

GET /_cluster/health

```
{  
  "cluster_name" : "docker-cluster",  
  "status" : "yellow",  
  "timed_out" : false,  
  "number_of_nodes" : 1,  
  "number_of_data_nodes" : 1,  
  "active_primary_shards" : 38,  
  "active_shards" : 38,  
  "relocating_shards" : 0,  
  "initializing_shards" : 0,  
  "unassigned_shards" : 6,  
  "delayed_unassigned_shards" : 0,  
  "number_of_pending_tasks" : 0,  
  "number_of_in_flight_fetch" : 0,  
  "task_max_waiting_in_queue_millis" : 0,  
  "active_shards_percent_as_number" : 86.36363636363636  
}
```

CRUD

- Crear índices
- Añadir documentos
- Leyendo documentos
- Actualizando documentos
- Actualización con script
- Upserts
- Borrar documentos
- Borrar índices
- Importando datos con cURL
- Actualizaciones en lote

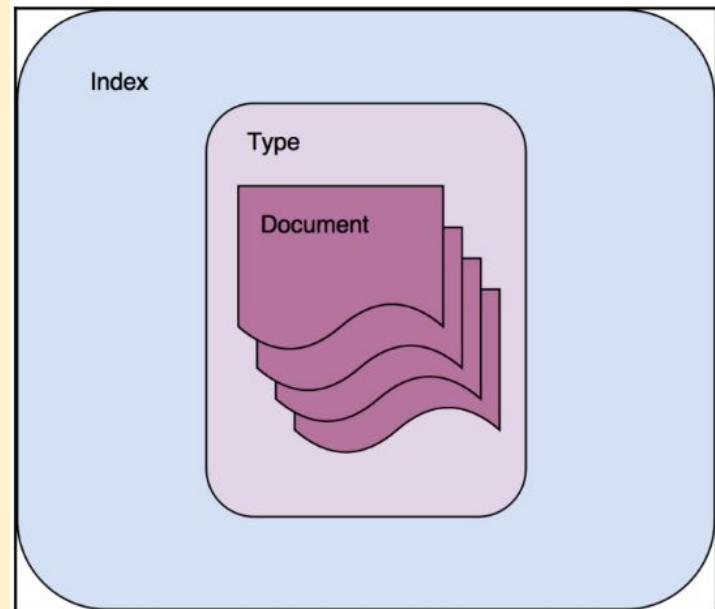
Crear índices

```
$ curl -XPUT 'http://elastic:elastic@localhost:9200/mi_indice/'
```

```
$ curl -XGET  
'http://elastic:elastic@localhost:9200/mi_indice?pretty'
```

Crear índices

```
PUT /my_index
{
  "settings": { ... settings ... },
  "mappings": {
    "type": { ... mappings ... }
  }
}
```



Crear índices

PUT /catalog

```
{  
  "settings": {  
    "index": {  
      "number_of_shards": 1,  
      "number_of_replicas": 1,  
      "codec": "best_compression"  
    }  
  }  
}
```

Crear índices

PUT /catalog

```
{  
  "mappings": {  
    "properties": {  
      "texto": {  
        "type": "text"  
      },  
      "keyword": {  
        "type": "keyword"  
      }  
    }  
  }  
}
```



Elasticsearch

¿Cómo almacenar los datos?

- Datos en origen en distintos formatos: BBDD, CSV, etc.

FlightNum	Origin	Dest	timestamp	FlightDelayMin	Cancelled
652J760	Amsterdam Airport Schiphol	Jorge Chavez International Airport	Mar 8, 2020 @ 21:08:25.000	225	false
09N00WNY	Shanghai Hongqiao International Airport	Ukrainka Air Base	Mar 8, 2020 @ 21:03:31.000	45	false
7ZS3250	Guangzhou Baiyun International Airport	Munich Airport	Mar 8, 2020 @ 20:21:16.000	285	true

- Se deben convertir a **objetos JSON** para enviar al **API REST** de Elasticsearch. Los documentos se almacenan en **índices**, agrupaciones lógicas de **Lucene shards**.

```
{  
  "FlightNum": "652J760",  
  "Origin": "Amsterdam Airport Schiphol",  
  "Dest": "Jorge Chavez International Airport",  
  "timestamp": "Mar 8, 2020 @ 21:08:25.000",  
  "FlightDelayMin": 225,  
  "Cancelled": false  
}
```

Añadir documentos

PUT <nombre_index>/_doc/<identificador>

PUT kibana_sample_data_flights/_doc/1

```
{  
  "FlightNum": "652J760",  
  "Origin": "Amsterdam Airport Schiphol",  
  "Dest": "Stockholm-Arlanda Airport",  
  "FlightDelayMin": 0,  
  "Cancelled": false,  
  "timestamp": "2020-05-01T05:21:34"  
}
```

Añadir documentos

```
POST <nombre_index>/_doc/
```

```
POST kibana_sample_data_flights/_doc/
```

```
{  
  "FlightNum": "652J760",  
  "Origin": "Amsterdam Airport Schiphol",  
  "Dest": "Stockholm-Arlanda Airport",  
  "FlightDelayMin": 0,  
  "Cancelled": false,  
  "timestamp": "2020-05-01T05:21:34"  
}
```

Añadir documentos

```
{  
  "_index": "kibana_sample_data_flights",  
  "_type": "_doc",  
  "_id": "AVrASKqgaBGmnAMj1SBe",  
  "_version": 1,  
  "result": "created",  
  "_shards": {  
    "total": 2,  
    "successful": 1,  
    "failed": 0  
  },  
  "created": true  
}
```

Obtener documentos

```
GET kibana_sample_data_flights/_doc/1
```

```
GET kibana_sample_data_flights/_doc/1/_source
```

```
GET kibana_sample_data_flights/_doc/1?_source=Origin
```

Obtener documentos

```
GET /kibana_sample_data_flights/_doc/_mget
{
  "ids" : [ "1", "AVbA4WNg7uqRWQFJiJSn" ]
}
```

```
GET /kibana_sample_data_flights/_mget
{
  "docs": [
    {
      "_id": "1"
    },
    {
      "_id": "AVbA4WNg7uqRWQFJiJSn"
    }
  ]
}
```

Obtener documentos de diferentes índices

```
GET /_mget
{
  "docs": [
    {
      "_index": "kibana_sample_data_flights",
      "_id": "1"
    },
    {
      "_index": "kibana_sample_data_log",
      "_id": "1"
    }
  ]
}
```

Actualizando documentos

```
PUT kibana_sample_data_flights/_doc/1
```

```
{  
  "FlightNum": "652J760",  
  "Origin": "Amsterdam Airport Schiphol",  
  "Dest": "Stockholm-Arlanda Airport",  
  "FlightDelayMin": 100,  
  "Cancelled": false,  
  "timestamp": "2020-05-01T05:21:34"  
}
```

Actualizando documentos de forma parcial

```
POST kibana_sample_data_flights/_update/1
{
  "doc": {
    "FlightDelayMin": 260,
    "tags": ["flight1", "delayed"]
  }
}
```

Actualización con script

```
POST kibana_sample_data_flights/_doc/1/_update
{
  "script": "ctx._source.FlightDelayMin = 200"
}
```

Actualización con script

```
POST kibana_sample_data_flights/_update/1
{
  "script": {
    "source": "if (ctx._source.Origin.contains(params.origin)) {
ctx._source.FlightDelayMin = 100 } else {
ctx._source.FlightDelayMin = 300 }",
    "lang": "painless",
    "params": {
      "origin": "Amsterdam Airport Schiphol"
    }
  }
}
```

Actualización con script

```
POST kibana_sample_data_flights/_doc/1/_update
{
  "script": {
    "source
```

Upserts

```
POST kibana_sample_data_flights/_update/1
{
  "script": {
    "source": "ctx._source.FlightDelayMin += params.count",
    "lang": "painless",
    "params": {
      "count": 4
    }
  },
  "upsert": {
    "origin": "new airport"
  }
}
```

Upserts

```
POST kibana_sample_data_flights/_update/1
{
  "scripted_upsert": true,
  "script": {
    "source": "ctx._source.FlightDelayMin += params.count",
    "lang": "painless",
    "params": {
      "count": 4
    }
  },
  "upsert": {
    "origin": "new airport",
    "FlightDelayMin":200
  }
}
```

Upserts

```
POST kibana_sample_data_flights/_update/1
```

```
{  
  "doc": {  
    "FlightDelayMin": 260,  
    "tags": [  
      "flight1",  
      "delayed"  
    ],  
    "doc_as_upsert": true  
  }  
}
```

Borrar documentos

```
DELETE kibana_sample_data_flights/_doc/<id>
```

```
POST kibana_sample_data_flights/_delete_by_query
{
  "query": {"match": {"_id": "<id>"}}
}
```

Borrar índices

```
curl -XDELETE 'http://localhost:9200/<indice>'
```

```
curl -XPOST 'http://localhost:9200/<indice>/_open'
```

```
curl -XPOST 'http://localhost:9200/<indice>/_close'
```

Importando datos con cURL

```
$ cat data.json
```

```
{ "create": { "_index": "my_index", "_id": "1" }}  
{"name": "my name", "date": "20021-01-01", "country": "Spain" }  
{ "create": { "_index": "my_index", "_id": "2" }}  
{"name": "my name1", "date": "20021-01-01", "country": "Italy" }  
{ "create": { "_index": "my_index", "_id": "3" }}  
{"name": "my name2", "date": "20021-01-02", "country": "France" }
```

```
$ curl -s -H "Content-Type: application/json" -XPOST http://elastic:elastic@localhost:9200/_bulk  
--data-binary "@data.json"; echo
```

```
{"took":2009,"errors":false,"items":[{"create":{"_index":"my_index","_type":"_doc","_id":"1","_version":1,"r  
esult":"created","_shards":{"total":2,"successful":1,"failed":0},"_seq_no":0,"_primary_term":1,"status":201  
}},{"create":{"_index":"my_index","_type":"_doc","_id":"2","_version":1,"result":"created","_shards":{  
"total":2,"successful":1,"failed":0},"_seq_no":1,"_primary_term":1,"status":201}},{"create":{"_index":  
"my_index","_type":"_doc","_id":"3","_version":1,"result":"created","_shards":{  
"total":2,"successful":1,"failed":0},"_seq_no":2,"_primary_term":1,"status":201}}]}
```

Actualizaciones en lote

POST /_bulk?pretty

```
{"delete":{"_index":"kibana_sample_data_flights","_type":"_doc","_id":"1"}}
{"create":{"_index":"kibana_sample_data_flights","_type":"_doc","_id":"1"}}
{"FlightNum":"652J760","Origin":"Amsterdam Airport
Schiphol","Dest":"Stockholm-Arlanda
Airport","FlightDelayMin":100,"Cancelled":false,"timestamp":"2020-05-01T05
:21:34"}
{"update":{"_index":"kibana_sample_data_flights","_type":"_doc","_id":"1"}}
{"doc":{"FlightDelayMin":200}}
```

Estadísticas de un índice

GET http://localhost:9200/<nombre_indice>/_stats

```
},
  "indexing" : {
    "index_total" : 18,
    "index_time_in_millis" : 21,
    "index_current" : 0,
    "index_failed" : 0,
    "delete_total" : 2,
    "delete_time_in_millis" : 1,
    "delete_current" : 0,
    "noop_update_total" : 1,
    "is_throttled" : false,
    "throttle_time_in_millis" : 0
},
```



Elasticsearch

¿Cómo almacenar los datos?

- Operaciones **CRUD - API REST**

```
POST kibana_sample_data_flights/_doc/1
{
  "FlightNum": "652J760",
  "Origin": "Amsterdam Airport Schiphol",
  "Dest": "Stockholm-Arlanda Airport",
  "FlightDelayMin": 0,
  "Cancelled" : false,
  "timestamp" : "2020-05-01T05:21:34"
}
```

```
GET kibana_sample_data_flights/_doc/1
```

```
POST kibana_sample_data_flights/_update/1
{
  "doc": {
    "FlightDelayMin": 260
  }
}
```

```
DELETE kibana_sample_data_flights/_doc/1
```

MAPPING

- Tipos de datos
- Mapping dinámico
- Meta fields
- Añadiendo mappings a índices existentes
- Parámetros de Mapping
- Formatos personalizados para fechas

Tipos de datos

<https://www.elastic.co/guide/en/elasticsearch/reference/current/mapping-types.html>

- **text**
- **keyword**
- **byte, short, integer, and long**
- **bool**
- **date**
- **Tipo de datos array**
- **Tipo de datos de objeto:** permite objetos dentro de documentos JSON.
- **Tipo de datos anidado:** útil para admitir arrays de objetos

Tipos de datos

- **text:** Se utiliza para indexar campos con valores de texto completo (FullText), como por ejemplo, el contenido de un correo electrónico o la descripción de un producto. Estos campos se analizan, es decir, se pasan a través de un analizador para convertir la cadena en una lista de términos individuales antes de indexarse.
- **keyword:** Se utiliza para los campos con valores que no pueden ser analizados para descomponerse en una serie de términos individuales, es decir, que solamente pueden ser buscados por su valor exacto.

Tipos de datos

- **Tipo de datos de puntos geográficos (Geo-point):** permite almacenar puntos geográficos con longitud y latitud. El tipo de datos de puntos geográficos permite realizar consultas, como buscar en todos los documentos a una distancia de 2 km de un punto.
- **Tipo de datos de forma geográfica (Geo-shape):** permite almacenar formas geométricas como polígonos y mapas. Geo-shape permite consultas como buscar todos los elementos dentro de una forma.
- **Tipo de datos IP:** permite almacenar direcciones IPv4 e IPv6.

Tipos de datos geo point

```
PUT my-index-geo
{
  "mappings": {
    "properties": {
      "location": {
        "type": "geo_point"
      }
    }
  }
}
```

```
PUT my-index-geo/_doc/1
{
  "text": "Geo-point as an object",
  "location": {
    "lat": 41.12,
    "lon": -71.34
  }
}
```

Tipos de datos objeto

```
POST my_index_object_type/_doc/1
{
  "person": {
    "name": {
      "firstname": "Martin",
      "lastname": "Fowler"
    }
  }
}
```

Tipos de datos nested

```
PUT /library/_mapping
{
  "properties": {
    "title": {
      "type": "text"
    },
    "review": {
      "type": "nested",
      "properties": {
        "nickname": {
          "type": "text"
        }
      }
    }
  }
}
```

Mapping dinámico

```
PUT /catalog/_doc/1
{
  "sku": "SP000001",
  "title": "Elasticsearch for Hadoop",
  "description": "Elasticsearch for Hadoop",
  "author": "Author",
  "ISBN": "1785288997",
  "price": 26.99
}
```

Mapping dinámico

```
PUT /catalog/_doc/2
{
  "sku": "SP000002",
  "title": "Google Pixel Phone 32GB - 5 inch
display",
  "description": "Google Pixel Phone 32GB",
  "price": 400.0,
  "long":100,
  "resolution": "1440 x 2560 pixels",
  "os": "Android 7.1"
}
```

Meta fields

- **_id:** este es el identificador único del documento dentro del tipo, al igual que una clave principal en una tabla de base de datos. Puede ser autogenerado o especificado por el usuario.
- **_type:** este campo contiene el tipo de documento.
- **_index:** este campo contiene el nombre de índice del documento.

Añadiendo mappings a índices existentes

```
PUT /<indice>/_mapping
```

```
{  
  "properties": {  
    "name": {  
      "type": "text"  
    }  
  }  
}
```

Formatos personalizados para fechas

<https://www.elastic.co/guide/en/elasticsearch/reference/current/mapping-date-format.html>

```
PUT my-index-date
{
  "mappings": {
    "properties": {
      "date": {
        "type": "date"
      }
    }
  }
}
```

Formatos personalizados para fechas

```
PUT my-index-date/_doc/1
{
  "date": "2021-01-01"
}
```

```
PUT my-index-date/_doc/2
{
  "date": "2021-01-01T12:10:30Z"
}
```

```
GET my-index-date/_search
{
  "sort": { "date": "asc" }
}
```

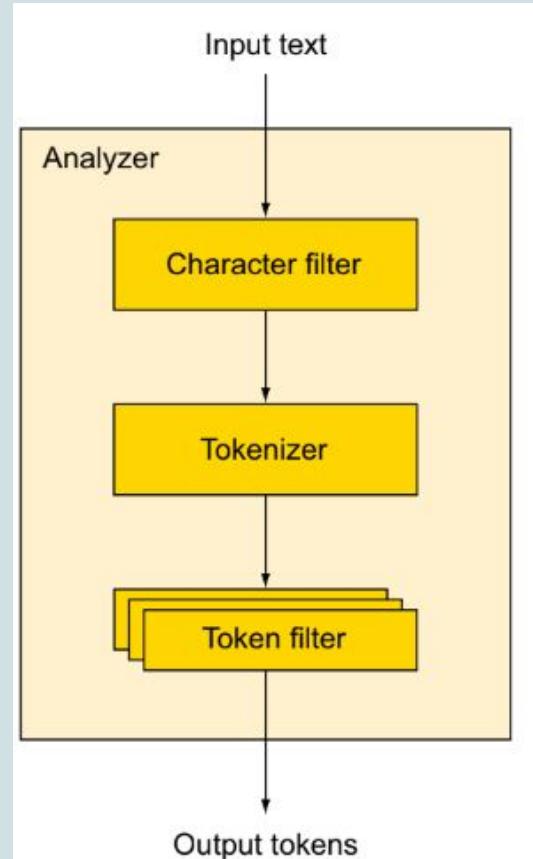
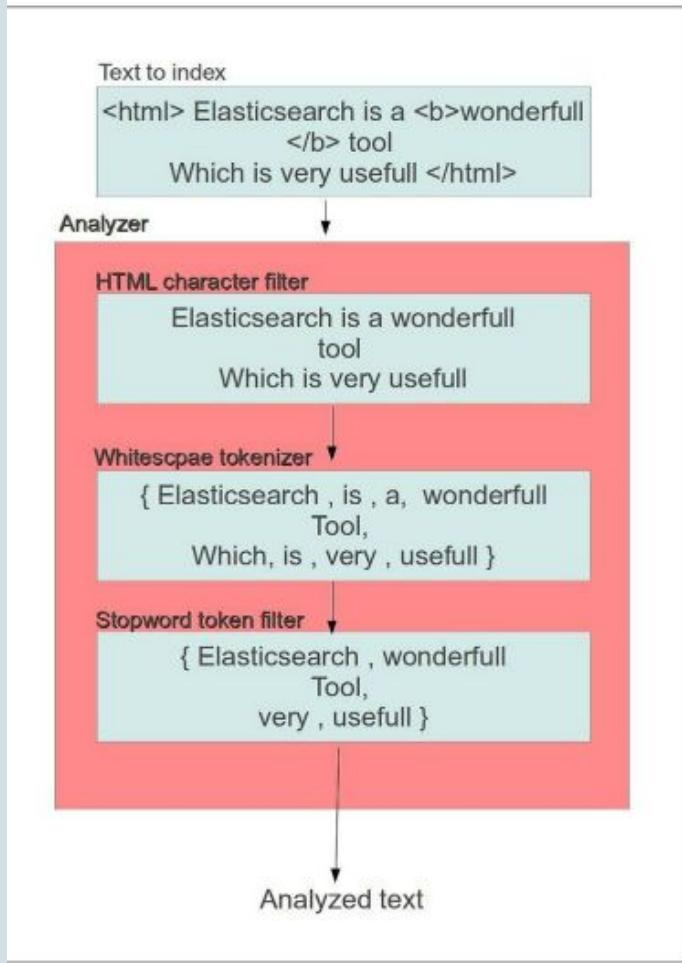
Formatos personalizados para fechas

```
PUT my-index-date/_mapping
{
  "properties": {
    "date2": {
      "type": "date",
      "format": "yyyy-MM-dd HH:mm:ss||yyyy-MM-dd"
    }
  }
}
```

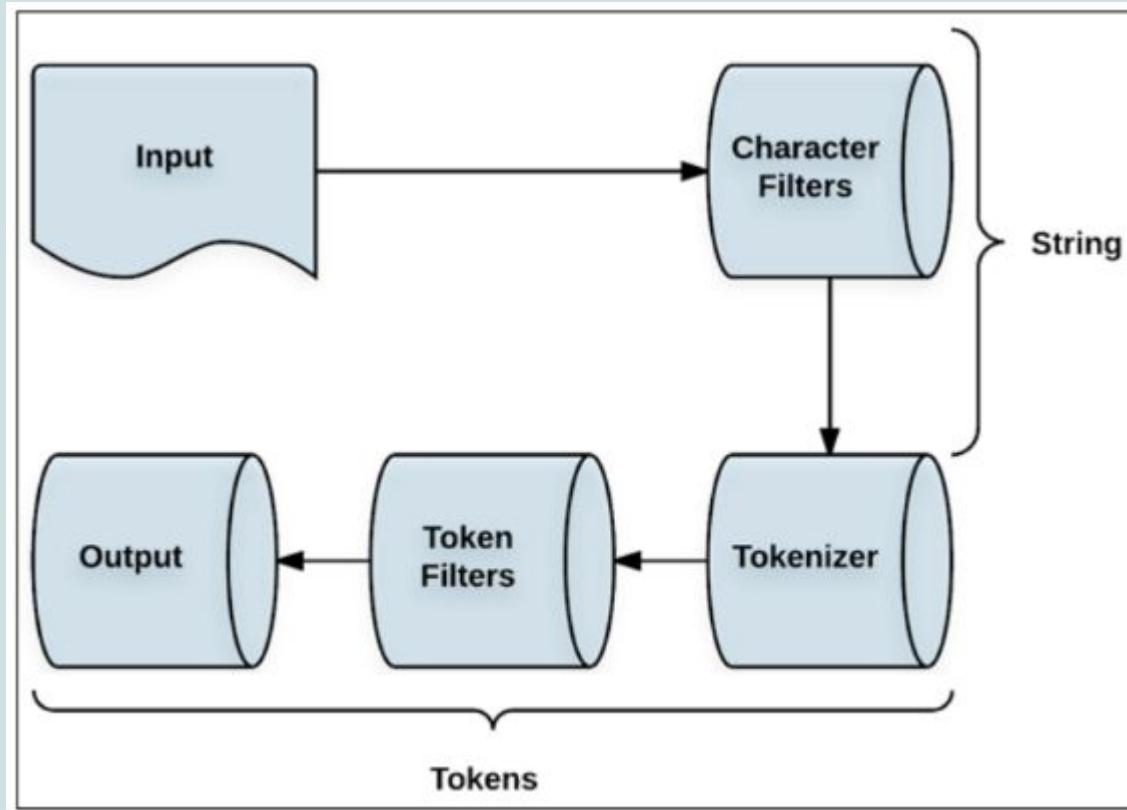
ANÁLISIS

- Introducción
- Filtros de caracteres
- Tokenizers
- Filtros de Tokens
- API de Analyze
- Analizadores de sistema
- Analizadores personalizados

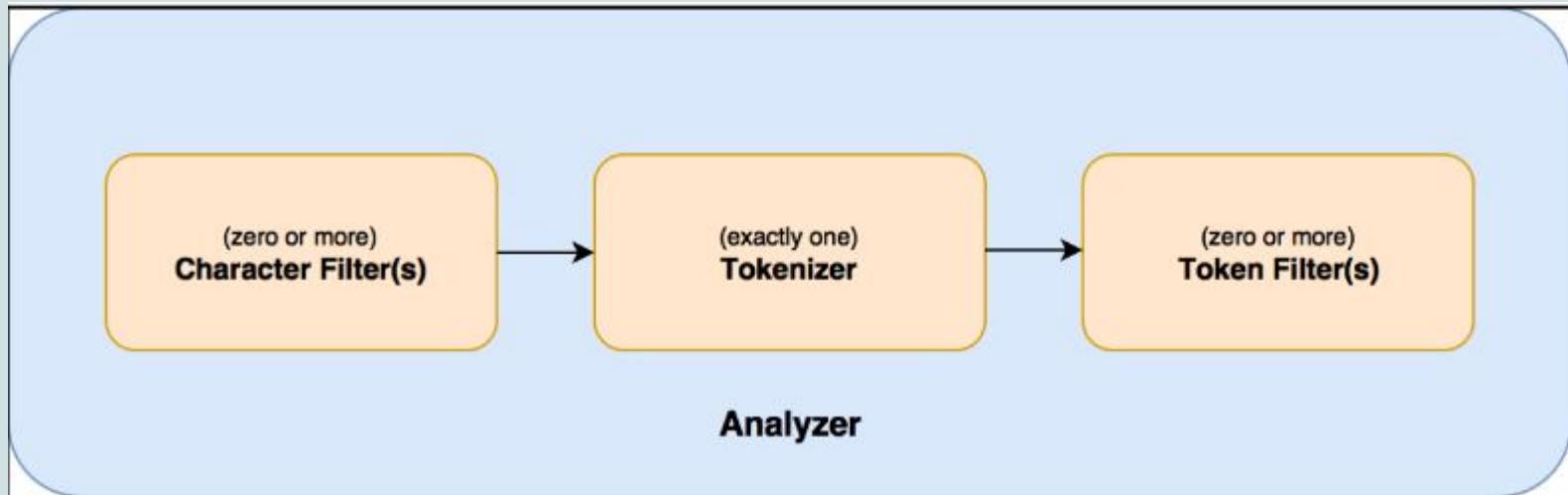
Introducción



Analizadores



Composición de un analizador



Filtros de caracteres

```
GET /_analyze
{
  "tokenizer": "keyword",
  "char_filter": [
    {
      "type": "mapping",
      "mappings
```

Filtros de caracteres

```
GET /_analyze
{
  "tokenizer": "keyword",
  "char_filter": [
    "html_strip",
    {
      "type": "mapping",
      "mappings": [
        "I'll => I will"
      ]
    }
  ],
  "text": "<p>I'll <b>learn elastic </b>!</p>"
}
```

Tokenizers

```
POST _analyze
{
  "tokenizer": "standard",
  "text": "Tokenizer breaks characters into tokens!"
}
```

Filtros de Tokens

- **Lowercase Token Filter:** reemplaza todos los tokens en la entrada con sus versiones en minúsculas.
- **Stop Token Filter:** Elimina las palabras vacías, es decir, las palabras que no añaden más significado al contexto.

Stop words

- <https://www.ranks.nl/stopwords>

Default English stopwords list

This list is used in our [Page Analyzer](#) and [Article Analyzer](#) for English text, when you let it use the default stopwords list.

a	ourselves
about	out
above	over
after	own
again	same
against	shan't
all	she
am	she'd
an	she'll

Analizadores de sistema

- **Standard Analyzer:** es el analizador predeterminado en Elasticsearch. Si no se reemplaza por ningún otro analizador de nivel de campo, nivel de tipo o nivel de índice, todos los campos se analizan utilizando este analizador.
- **Analizadores de idiomas:** los diferentes idiomas tienen diferentes reglas gramaticales.
- **Analizador de espacios en blanco:** el analizador de espacios en blanco divide la entrada en tokens siempre que encuentre un token de espacio en blanco, como un espacio, tabulación, nueva línea o retorno de carro.

Standard Analyzer

- **Standard Tokenizer:** un tokenizador que divide los tokens en caracteres de espacio en blanco.
- **Standard Token Filter:** el filtro de token estándar se utiliza como un filtro de token de marcador de posición dentro del analizador estándar.
- **Lowercase Token Filter::** convierte todos los tokens en minúsculas.
- **Stop Token Filter::** elimina las stop words especificadas.

Standard Analyzer

```
PUT index_standard_analyzer
{
  "settings": {
    "analysis": {
      "analyzer": {
        "std": {
          "type": "standard"
        }
      }
    }
  },
  "mappings": {
    "properties": {
      "my_text": {
        "type": "text",
        "analyzer": "std"
      }
    }
  }
}
```

Standard Analyzer

```
POST index_standard_analyzer/_analyze
{
  "field": "my_text",
  "text": "The Standard Analyzer works this way."
}
```

Standard Analyzer stop words

```
PUT index_standard_analyzer_english_stopwords
{
  "settings": {
    "analysis": {
      "analyzer": {
        "std": {
          "type": "standard",
          "stopwords": "_english_"
        }
      }
    }
  },
  "mappings": {
    "properties": {
      "my_text": {
        "type": "text",
        "analyzer": "std"
      }
    }
  }
}
```

Standard Analyzer stop words

```
POST  
index_standard_analyzer_english_stopwords/_analyze  
{  
  "field": "my_text",  
  "text": "The Standard Analyzer works this way."  
}
```

Analizadores personalizados

```
PUT /my_index
{
  "settings": {
    "analysis": {
      "char_filter": { ... custom character filters ... },
      "tokenizer": { ... custom tokenizers ... },
      "filter": { ... custom token filters ... },
      "analyzer
```

Analizadores personalizados

```
"keyword_tokenizer": {  
    "type": "custom",  
    "filter": [  
        "lowercase",  
        "asciifolding"  
    ],  
    "tokenizer": "keyword"  
}
```

Analizadores personalizados

```
PUT /custom_analyzer_index
{
  "settings": {
    "index": {
      "analysis": {
        "analyzer": {
          "custom_analyzer": {
            "type": "custom",
            "tokenizer": "standard",
            "filter": [
              "lowercase",
              "custom_edge_ngram"
            ]
          }
        }
      }
    }
  }
}
```

```
"filter": {
  "custom_edge_ngram": {
    "type": "edge_ngram",
    "min_gram": 2,
    "max_gram": 10
  }
}
}
}
}
},
"mappings": {
  "my_type": {
    "properties": {
      "product": {
        "type": "text",
        "analyzer": "custom_analyzer",
        "search_analyzer": "standard"
      }
    }
  }
}
}
}
}
```

Analizadores personalizados

```
POST /custom_analyzer_index/my_type
{
  "product": "Learning Elastic Stack 7"
}
```

```
POST /custom_analyzer_index/my_type
{
  "product": "Mastering Elasticsearch"
}
```

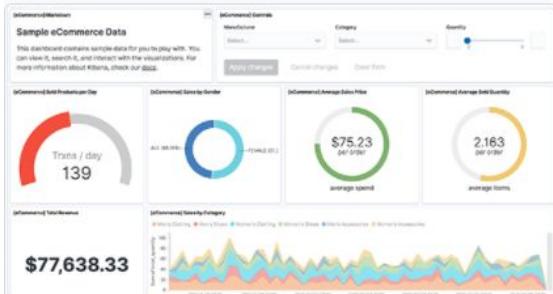
Analizadores personalizados

```
GET /custom_analyzer_index2/_search
```

```
{  
  "query": {  
    "match": {  
      "product": "Elastic"  
    }  
  }  
}
```

Analizadores personalizados

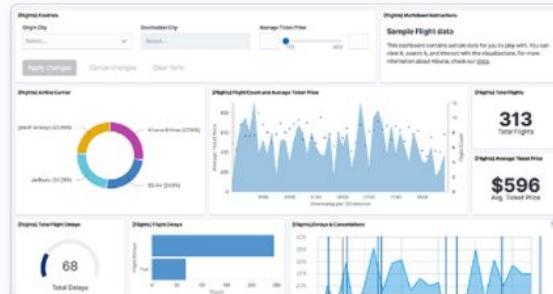
```
PUT my-index-tokenizer_max_length
{
  "settings": {
    "analysis": {
      "analyzer": {
        "my_analyzer": {
          "tokenizer": "my_tokenizer"
        }
      },
      "tokenizer": {
        "my_tokenizer": {
          "type": "standard",
          "max_token_length": 5
        }
      }
    }
  }
}
```



Sample eCommerce orders

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

Add data



Sample flight data

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

Add data



Sample web logs

Sample data, visualizations, and dashboards for monitoring web logs.

Add data



Elasticsearch

Dos tipos de búsquedas

- **Queries**
 - ¿Qué vuelos tienen **origen** en “Amsterdam”?
 - ¿Qué vuelos se **retrasaron 60 minutos o más**?
- **Agregaciones**
 - ¿Cuáles son los **top 3** aeropuertos **origen**?
 - ¿Qué **retraso** tienen en **media** los vuelos en los **2 top** aeropuertos **origen**?



Elasticsearch

Búsquedas - Queries

Query DSL (Domain Specific Language) basado en JSON para definir queries

¿Qué vuelos
tienen origen
en
“Amsterdam”?

```
GET kibana_sample_data_flights/_search
{
  "size": 1,
  "query": {
    "match": {
      "origin.text": "amsterdam"
    }
  }
}
```

```
{
  "took": 1,
  "timed_out": false,
  "_shards": {
    "total": 1,
    "successful": 1,
    "skipped": 0,
    "failed": 0
  },
  "hits": {
    "total": {
      "value": 1,
      "relation": "eq"
    },
    "max_score": 0.2876821,
    "hits": [
      {
        "_index": "flights",
        "_type": "_doc",
        "_id": "1",
        "_score": 0.2876821,
        "_source": {
          "FlightNum": "652J760",
          "Origin": "Amsterdam Airport Schiphol",
          "Dest": "Jorge Chavez International Airport",
          "timestamp": "Mar 8, 2020 @ 21:08:25.000",
          "FlightDelayMin": 225,
          "Cancelled": false
        }
      }
    ]
  }
}
```

¿Qué vuelos
se retrasaron
60 minutos o
más?

```
GET kibana_sample_data_flights/_search
?filter_path=hits.hits._source
{
  "size": 10,
  "_source": ["FlightDelayMin", "FlightNum"],
  "query": {
    "range": {
      "FlightDelayMin": {
        "gte": 60
      }
    }
  }
}
```

```
{
  "hits": {
    "hits": [
      {
        "_source": {
          "FlightNum": "EAYQW69",
          "FlightDelayMin": 180
        }
      },
      {
        "_source": {
          "FlightNum": "EVARI8I",
          "FlightDelayMin": 300
        }
      },
      {
        "_source": {
          "FlightNum": "RBFKZBX",
          "FlightDelayMin": 120
        }
      },
      {
        "size": 10,
        "_source": {
          "FlightNum": "R43CELD",
          "FlightDelayMin": 300
        }
      },
      {
        "_source": {
          "FlightNum": "1TJKW8F",
          "FlightDelayMin": 90
        }
      }
    ]
  }
}
```

Búsquedas

```
# Vuelos con origen en Amsterdam
GET kibana_sample_data_flights/_search
{
  "query": {
    "match": {
      "origin.text": "amsterdam"
    }
  }
}
```

Búsquedas

```
# Vuelos que se retrasan 60 minutos o mas
GET kibana_sample_data_flights/_search?filter_path=hits.hits._source
{
  "size": 10,
  "_source": ["FlightDelayMin","FlightNum"],
  "query": {
    "range": {
      "FlightDelayMin": {
        "gte": 60
      }
    }
  }
}
```

AGREGACIONES

- Buckets
- Agregaciones por métricas
- Agregaciones buckets
- Agregaciones estadísticas
- Agregaciones anidadas
- Agregaciones por rango
- Histogramas

Buckets

```
"aggs": {  
    "name_of_aggregation": {  
        "type_of_aggregation": {  
            "field": "document_field_name"  
        }  
    }  
}
```

Buckets

```
{  
    "aggs": {  
        "categoria": {  
            "terms": { "field": "category_id" },  
            "aggs": {  
                "tipo_publicacion" : {  
                    "terms" : { "field" : "listing_type_id"},  
                    "aggs" : {  
                        "prom_precio" : { "avg": { "field": "price"} },  
                        "min_precio" : { "min": { "field": "price"} },  
                        "max_precio" : { "max": { "field": "price"} }  
                    }  
                }  
            }  
        }  
    }  
}
```

The diagram illustrates the structure of an Elasticsearch search query, specifically focusing on the 'buckets' section. Annotations are used to identify different parts of the code:

- Nombre**: Points to the first nested aggregation, "categoria".
- Buckets**: Points to the second nested aggregation, "tipos_publicacion".
- Metrics**: Points to the three metrics defined under the "tipos_publicacion" aggregation: "prom_precio", "min_precio", and "max_precio".

Agregaciones por métricas

```
SELECT avg(column) FROM table;
```

Esta consulta calcula la puntuación media para una determinada columna.

```
"aggs" : {  
  "<name_of_aggregation>" : {  
    "avg" : {  
      "field" : "column"  
    }  
  }  
}
```



Elasticsearch

Búsquedas - Agregaciones

¿Cuáles son los top 3 aeropuertos origen?

```
GET kibana_sample_data_flights/_search
{
  "size": 0,
  "aggregations": {
    "top_aeropuertos_origen": {
      "terms": {
        "field": "origin",
        "size": 3
      }
    }
  }
}
```

¿Qué retraso tienen en media los vuelos en los 2 top aeropuertos origen?

```
GET kibana_sample_data_flights/_search
{
  "size": 0,
  "aggregations": {
    "top_aeropuertos_origen": {
      "terms": {
        "field": "origin",
        "size": 3
      },
      "aggregations": {
        "media_retraso": {
          "avg": {
            "field": "FlightDelayMin"
          }
        }
      }
    }
  }
}
```

```
{
  "took": 44,
  "timed_out": false,
  "_shards": {
    "total": 1,
    "successful": 1,
    "skipped": 0,
    "failed": 0
  },
  "hits": [
    {
      "total": {
        "value": 10000,
        "relation": "gte"
      },
      "max_score": null,
      "hits": []
    }
  ],
  "aggregations": {
    "top_aeropuertos_origen": {
      "doc_count_error_upper_bound": 0,
      "sum_other_doc_count": 12262,
      "buckets": [
        {
          "key": "Mariscal Sucre International Airport",
          "doc_count": 285
        },
        {
          "key": "Ministro Pistarini International Airport",
          "doc_count": 258
        },
        {
          "key": "El Dorado International Airport",
          "doc_count": 254
        }
      ]
    }
  }
}
```

Agregaciones por métricas

```
GET kibana_sample_data_flights/_search
{
  "aggregations": {
    "km_average": {
      "avg": {
        "field": "DistanceKilometers"
      }
    }
  },
  "size": 0
}
```

Agregaciones buckets

```
SELECT column, count(*) FROM table GROUP BY column;
```

Esta consulta divide la tabla por los diferentes valores de la columna y devuelve un recuento de documentos dentro de cada valor de la columna.

Agregaciones buckets

```
GET kibana_sample_data_flights/_search?size=0
{
  "query": {
    "match_all": {}
  },
  "aggregations": {
    "OriginCityName": {
      "terms": {
        "field": "OriginCityName",
        "size": 10
      }
    }
  }
}
```

Agregaciones buckets

```
"aggregations": {  
    "aggregation_name": {  
        "buckets": [  
            {  
                "key": value,  
                "doc_count": value  
            }  
        ]  
    }  
}
```

Agregaciones por término

```
GET kibana_sample_data_flights/_search
{
  "aggs": {
    "by_Origin": {
      "terms": {
        "field": "Origin"
      }
    }
  },
  "size": 0
}
```

Agregaciones estadísticas

```
GET kibana_sample_data_flights/_search
{
  "aggregations": {
    "km_stats": {
      "stats": {
        "field": "DistanceKilometers"
      }
    }
  },
  "size": 0
}
```

Agregaciones estadísticas

```
GET kibana_sample_data_flights/_search
{
  "aggregations": {
    "km_extended_stats": {
      "extended_stats": {
        "field": "DistanceKilometers"
      }
    }
  },
  "size": 0
}
```

Agregación cardinalidad

```
GET kibana_sample_data_flights/_search
{
  "aggregations": {
    "unique_origin": {
      "cardinality": {
        "field": "Origin"
      }
    }
  },
  "size": 0
}
```

Histogramas

```
GET kibana_sample_data_flights/_search
{
  "aggs": {
    "by_km": {
      "histogram": {
        "field": "DistanceKilometers",
        "interval": 5000
      }
    }
  },
  "size": 0
}
```

Agregaciones por rango

```
GET kibana_sample_data_flights/_search
{
  "aggs": {
    "by_km": {
      "range": {
        "field": "DistanceKilometers",
        "ranges": [
          { "to": 1000 },
          { "from": 1000, "to": 5000 },
          { "from": 5000 }
        ]
      }
    },
    "size": 0
  }
}
```

Agregaciones por rango

```
GET kibana_sample_data_flights/_search
{
  "aggs": {
    "by_km": {
      "range": {
        "field": "DistanceKilometers",
        "ranges": [
          { "key": "Upto 1000 km", "to": 1000 },
          { "key": "From 1000 to 5000 km", "from": 1000, "to": 5000 },
          { "key": "More than 5000 km", "from": 5000 }
        ]
      }
    },
    "size": 0
  }
}
```

Agregaciones por geolocalización

```
GET kibana_sample_data_flights/_search
{
  "aggs": {
    "within_radius": {
      "geo_distance": {
        "field": "OriginLocation",
        "origin": {"lat": 50.033333,"lon": 8.570556},
        "ranges": [{"from": 1000,"to": 150000}]
      }
    }
  },
  "size": 0
}
```

Agregaciones por filtro

```
GET kibana_sample_data_flights/_search
{
  "aggs": {
    "filter_origin": {
      "filter": {
        "term": {
          "Origin": "Frankfurt am Main Airport"
        }
      }
    }
  },
  "size": 0
}
```

Agregación ip_range

```
GET /kibana_sample_data_logs/_search?size=0
{
  "aggs": {
    "ip_ranges": {
      "ip_range": {
        "field": "clientip",
        "ranges": [
          {
            "from": "223.87.60.0"
          },
          {
            "to": "223.87.60.255"
          }
        ]
      }
    }
  }
}
```

Agregaciones anidadas

```
PUT nested_aggregation
{
  "mappings": {
    "properties": {
      "employee": {
        "type": "nested",
        "properties" : {
          "first_name" : { "type" : "text" },
          "last_name" : { "type" : "text" },
          "salary" : { "type" : "double" }
        }}}
```

```
GET /nested_aggregation/_search
{
  "aggs": {
    "nested_aggregation" : {
      "nested": {
        "path": "employee"
      },
      "aggs": {
        "avg_salary": {
          "avg": {
            "field": "employee.salary"
          }
        }
      }
    }}}
```

KIBANA

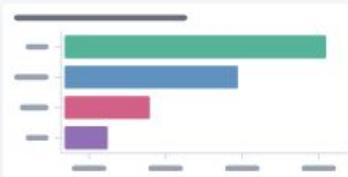
- Análisis con Kibana
- Configuración y administración
- Kibana Discover (Análisis de Logs)
- Visualizaciones: Histograma
- Visualizaciones: Pie
- Visualizaciones: Gauge
- Visualizaciones: Mapas
- Visualizaciones: Controles
- Visualizaciones: Metric
- Visualizaciones: Tablas de datos
- Otras visualizaciones
- Kibana Dashboards

Análisis con Kibana



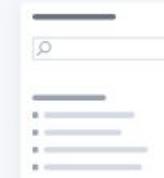
Kibana

Add data



Dashboard

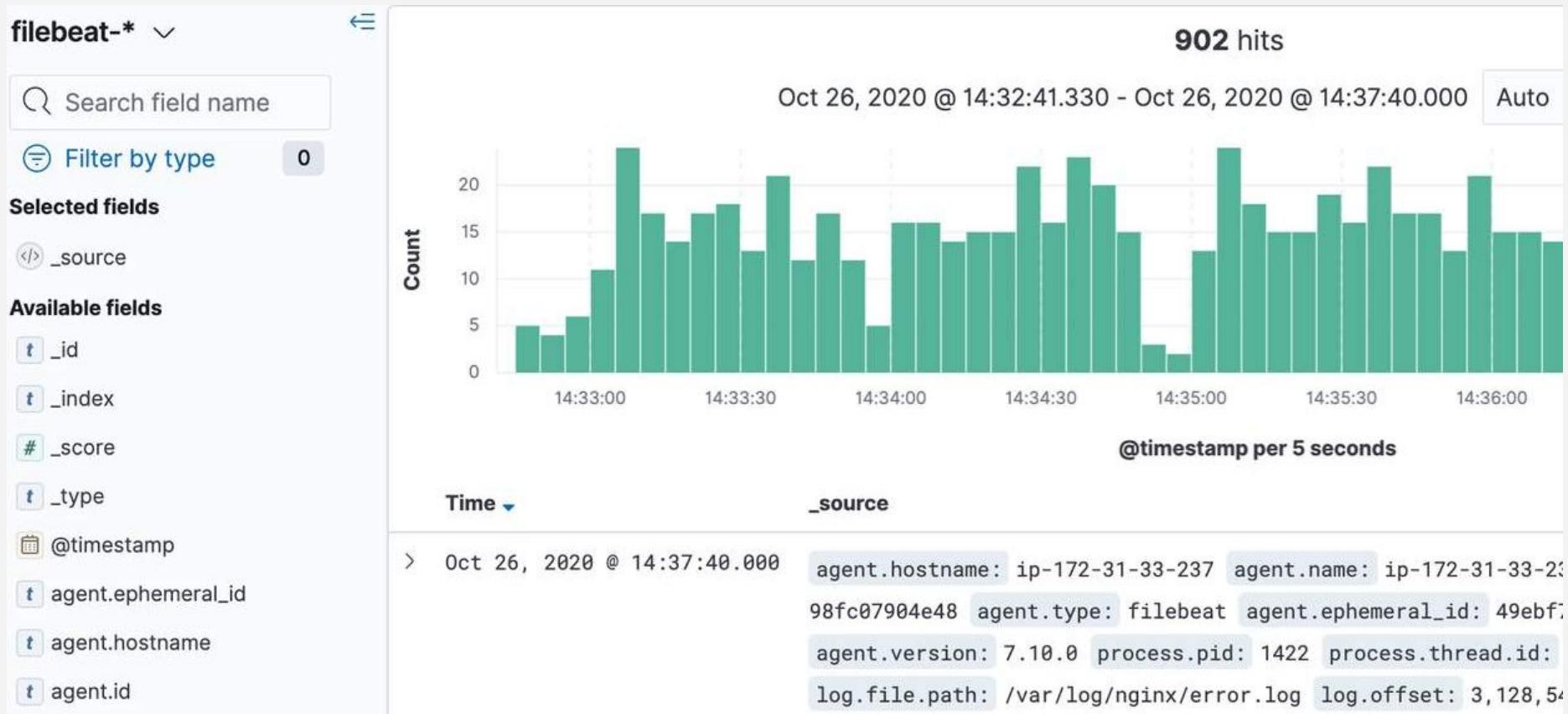
Analyze data in dashboards.



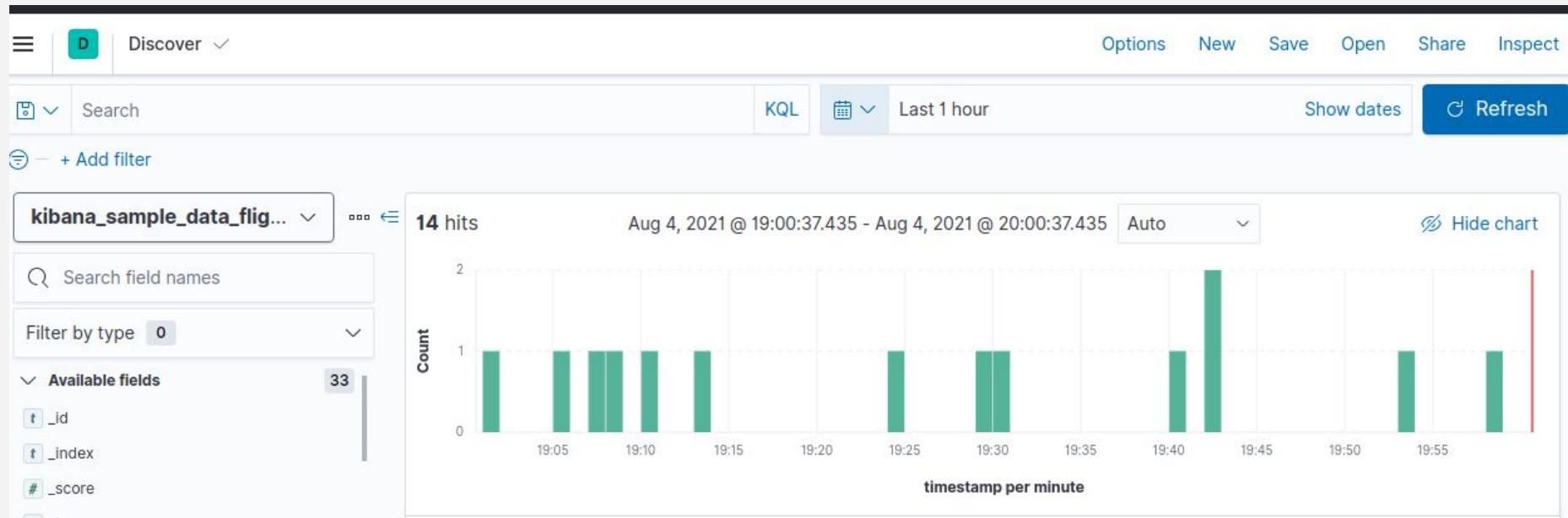
Discover

Search and find insights.

Kibana Discover (Análisis de Logs)



Kibana Discover (Análisis de Logs)



Kibana Discover (Análisis de Logs)

Quick select < >

Last 24 hours Apply

Commonly used

Today	Last 24 hours
This week	Last 7 days
Last 15 minutes	Last 30 days
Last 30 minutes	Last 90 days
Last 1 hour	Last 1 year

Recently used date ranges

Last 24 hours
<u>Aug 8, 2021 @ 22:31:39.090 to Aug 9, 2021 @ 00:31:39.090</u>
~ 7 months ago to ~ in 5 months
Last 24 months
<u>10 days ago to in 21 days</u>

Kibana Discover (Análisis de Logs)

EDIT FILTER

Edit as Query DSL

Field

host

Operator

is

@timestamp

_id

_index

_type

agent

agent.keyword

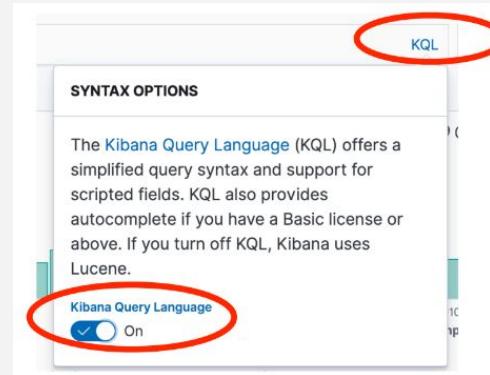
bytes

ancel

Save

Jul 17, 2021 @ 13:26:52.15

Kibana Discover (Análisis de Logs)



The screenshot displays the Kibana Discover interface. At the top, there is a search bar containing the query "agent.name :". Below the search bar is a sidebar with various filter options, including "agent.name" with four items listed, "Filebeat" with three items, and "Filter" with two items. The main area shows a table with columns for "_id", "_index", and "_score". The first row of the table is visible, showing values for each column.

	_id	_index	_score
1	c339a8202278		

Kibana Discover (Análisis de Logs)

Discover ✓

bd New Open Share Inspect

{ "bool": { "should": [{ "match": { "clientip": "84.3.192.254" } }] } }

Lucene Last 30 days Show dates Update

+ Add filter

kibana_sample_data_logs

Search field names

Filter by type 0

Available fields 30

- _id
- _index
- _score
- _type
- @timestamp
- agent
- bytes
- clientip

10 hits Jul 19, 2021 @ 23:21:05.689 - Aug 18, 2021 @ 23:21:05.689 Auto Hide chart

Count

2021-07-21 00:00 2021-07-25 00:00 2021-07-29 00:00 2021-08-03 00:00 2021-08-07 00:00 2021-08-11 00:00 2021-08-15 00:00 timestamp per 12 hours

event.dataset: sample_web_logs extension: css extension.keyword: css geo.coordinates: {"coordinates": [-95.62525583, 36.72092222], "type": "Point" } geo.dest: MM geo.src: CN

> Aug 18, 2021 @ 15:36:39.319 @timestamp: Aug 18, 2021 @ 15:36:39.319 agent: Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1; .NET CLR 1.1.4322) agent.keyword: Mozilla/4.0 (compatible; MSIE 6.0; Windows NT

Kibana Discover (Análisis de Logs)

Discover / mi_busqueda ✓

Options New Save Open Share Inspect

origin : "Huntsville International Carl T Jones Field"

+ Add filter

kibana_sample_data_flag... ▾

Search field names

Filter by type 0

Available fields 33

- _id
- _index
- _score
- _type
- AvgTicketPrice
- Cancelled
- Carrier
- dayOfWeek
- Dest

2 hits [Reset search](#) Aug 4, 2021

Count

Time ▾ Document

Time	Document
> Aug 4, 2021 @ 19:42:54.000	origin: "Huntsville International Carl T Jones Field"

mi_busqueda

View: Requests

1 request was made

Request: data

This request queries Elasticsearch to fetch the data for the search.
Search session id: 411fd8c1-b1ba-4d77-a7e4-3a476a356378

✓ 453ms

Statistics	Request	Response
② Hits	2	
② Hits (total)	2	
② Index pattern		kibana_sample_data_flights
② Index pattern ID		d3d7af60-4c81-11e8-b3d7-01146121b73d
② Query time	55ms	
② Request timestamp		2021-08-04T18:02:14.426Z

Kibana Visualize

New Visualization

Filter

Select a visualization type

Start creating your visualization by selecting a type for that visualization.

Try Lens, our new, intuitive way to create visualizations.

[Go to Lens](#)

Lens	Area	Controls	Data Table
Gauge	Goal	Heat Map	Horizontal Bar
Line	Maps	Markdown	Metric
Pie	TSVB	Tag Cloud	Timelion
Vega	Vertical Bar		

Kibana visualizaciones

The screenshot shows the Kibana interface with the following components:

- Top Bar:** Includes a search bar, KQL button, time range selector (Last 15 minutes), Show dates button, and Refresh button.
- Left Panel:** Shows a dropdown for "filebeat-*", a search field for "Search field names", a "Field filters" section (0), and a "Records" section. Below these are sections for "Available fields" (169) and specific fields: @timestamp, agent.ephemeral_id, agent.hostname, agent.id, agent.name, and agent.type.
- Middle Panel:** Displays a large value "6.3MB" and the text "Sum of http.response.body.bytes".
- Right Panel:** Titled "Metric configuration", it shows a "Select a function" dropdown with "Average" and "Median" options, and a "Records" option under "Available functions". It also lists "Available fields" including event.duration, http.response.body.bytes, http.response.status_code, log.offset, and process.pid. A search bar at the bottom contains "http.response.body.bytes".

Kibana lens

Visualize Library / Create Download as CSV Save

Search KQL Last 1 hour Show dates Refresh

+ Add filter

kibana_sample_data_logs

Search field names

Field filters 0

- agent.keyword
- # bytes
- clientip
- event.dataset
- extension.keyword
- geo.dest

Stacked bar

Drop some fields here to start

Vertical axis

Required dimension

Break down by

Top values of event.
dataset

Reset layer

Métricas y agregaciones

[Add filter](#)



Count

kibana_sample_data_flights

[Data](#) Metrics & axes Panel settings

Metrics

> Y-axis Count

+ Add

Buckets

+ Add

Buckets

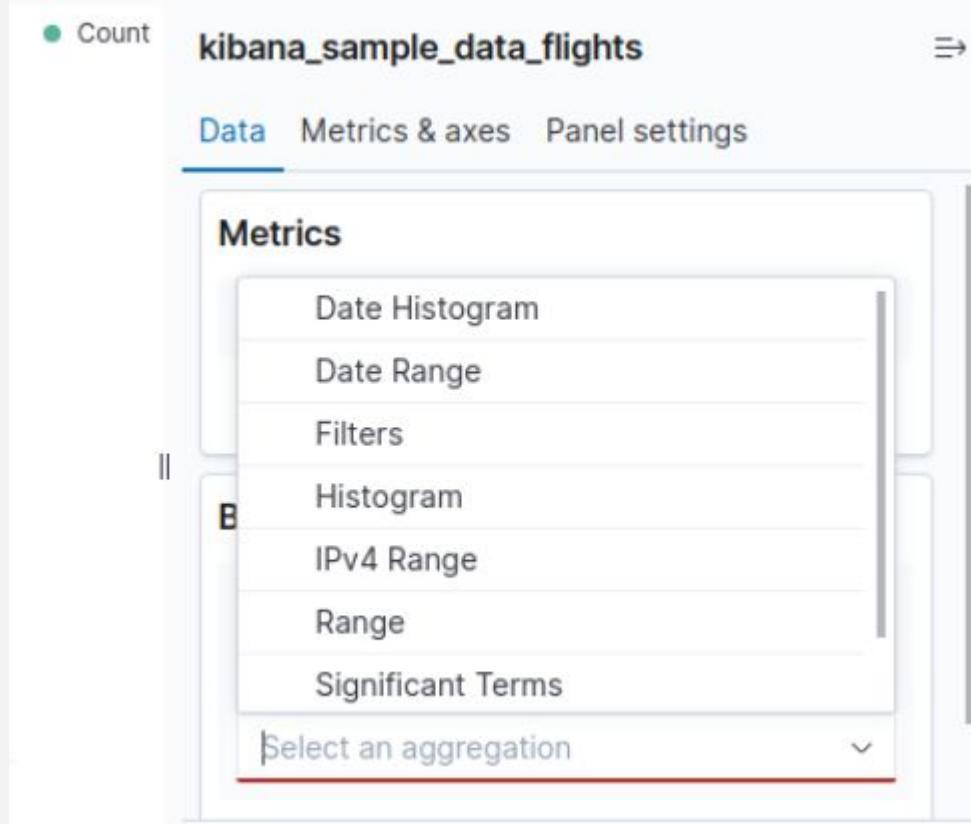
● Count kibana_sample_data_flights 

Data Metrics & axes Panel settings

Metrics

- Date Histogram
- Date Range
- Filters
- Histogram
- IPv4 Range
- Range
- Significant Terms

Select an aggregation 



Buckets

Buckets

Aggregation: Date Histogram

Field: @timestamp

Minimum interval: Auto

Select an option or create a custom value. Examples: 30s, 20m, 24h, 2d, 1w, 1M

Drop partial buckets

ADD SUB-BUCKET

X-axis

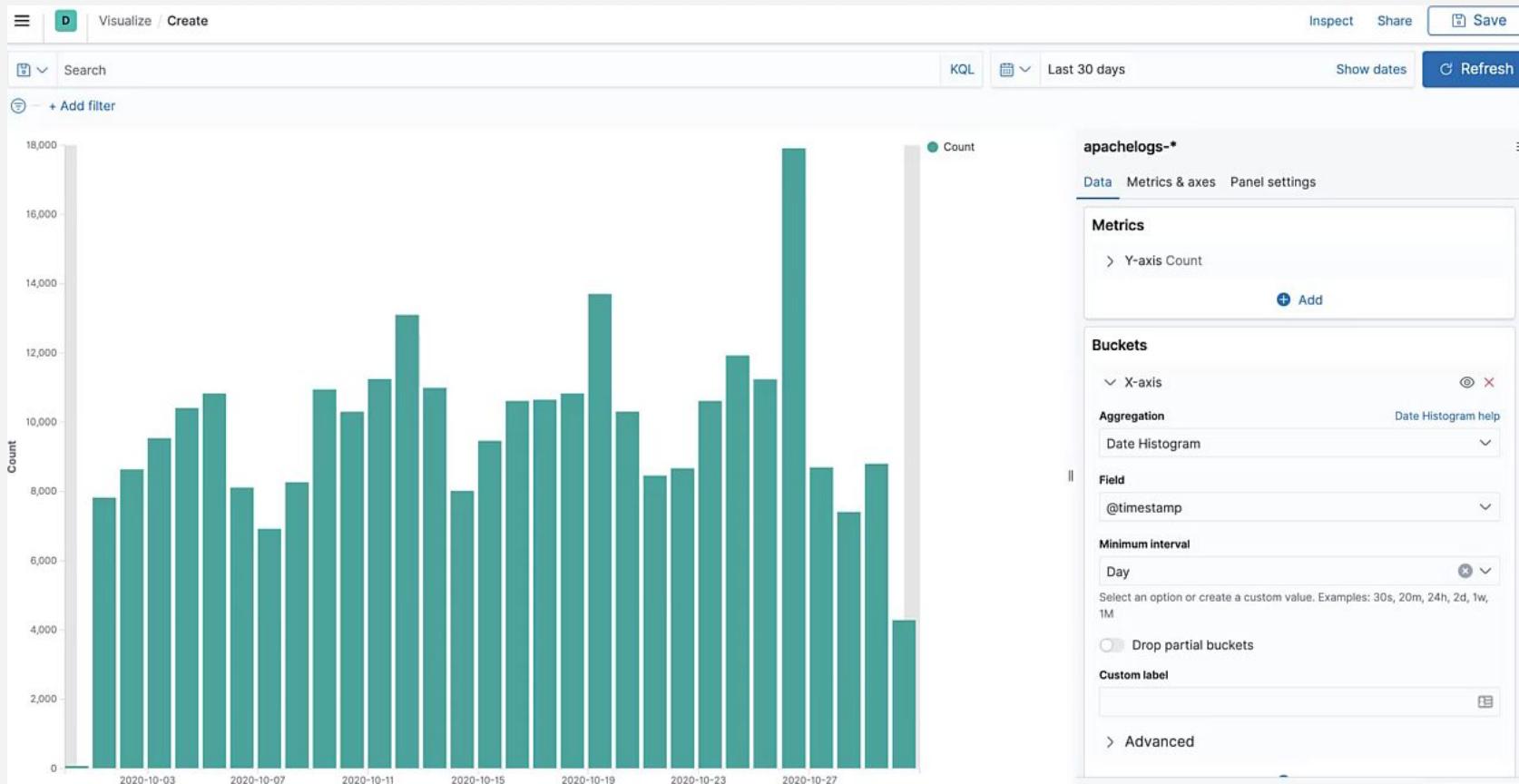
Split series

Split chart

+ Add

Date Histogram help

Buckets



Buckets

> Advanced

⊖ Split series ⊖ = X

Sub aggregation Terms help

Terms ▾

Field

response ▾

Order by

Metric: Count ▾

Order Size

Descending ▾ 5

Group other values in separate bucket

Show missing values

Custom label

> Advanced

+ Add

Métricas

A screenshot of a data visualization tool's interface. On the left, there is a vertical sidebar with a green dot icon followed by the word "Count". Below this, the word "kit" is partially visible. To the right of the sidebar, the word "Data" is underlined in blue, indicating it is the active tab. A dropdown menu is open, titled "Metric Aggregations". The menu contains the following options: Average, Count (which has a checked checkbox icon to its left), Max, Median, Min, and Percentile Ranks. At the bottom of the dropdown, there is a text input field containing the text "Count |" with a dropdown arrow icon to its right. The background of the interface shows some blurred data structures.

Count

kit

Data

Metric Aggregations

- Average
- ✓ Count
- Max
- Median
- Min
- Percentile Ranks

Count |

Buckets



Buckets

apachelogs-*

Data Metrics & Axes Panel Settings D X

Buckets

X-axis

Aggregation Date Histogram help

Date Histogram

Field @timestamp

Minimum interval Auto

Select an option or create a custom value.
Examples: 30s, 20m, 24h, 2d, 1w, 1M

ADD SUB-BUCKET

Custom label X-axis

Split series

> Advanced Split chart

Add

This screenshot shows the 'Buckets' panel for the 'apachelogs-*' index pattern. It displays settings for a Date Histogram aggregation. The 'Field' dropdown is set to '@timestamp'. The 'Add' button at the bottom of the sub-panel is highlighted with a red circle.

apachelogs-*

Data Metrics & Axes Panel Settings D X

Split series

Sub aggregation Terms help

Terms

Field response

Order by Metric: Count

Order Descending Size 5

Group other values in separate bucket

Show missing values

Custom label

This screenshot shows the 'Buckets' panel for the 'apachelogs-*' index pattern. It displays settings for a Terms aggregation. The 'Field' dropdown is set to 'response'. The 'Add' button at the bottom of the sub-panel is highlighted with a red circle.

Buckets

Visualize Library / Create 6d Inspect Share

Search KQL Last 24 hours Show dates Refresh

+ Add filter

request.keyword: Descending	Count
/apm	15
/beats/metricbeat/metricbeat-6.3.2-i686.rpm	14
/apm-server/apm-server-6.3.2-windows-x86.zip	12
/beats/metricbeat/metricbeat-6.3.2-amd64.deb	12
/kibana/kibana-6.3.2-darwin-x86_64.tar.gz	12

kibana_sample_data_logs

Data Options

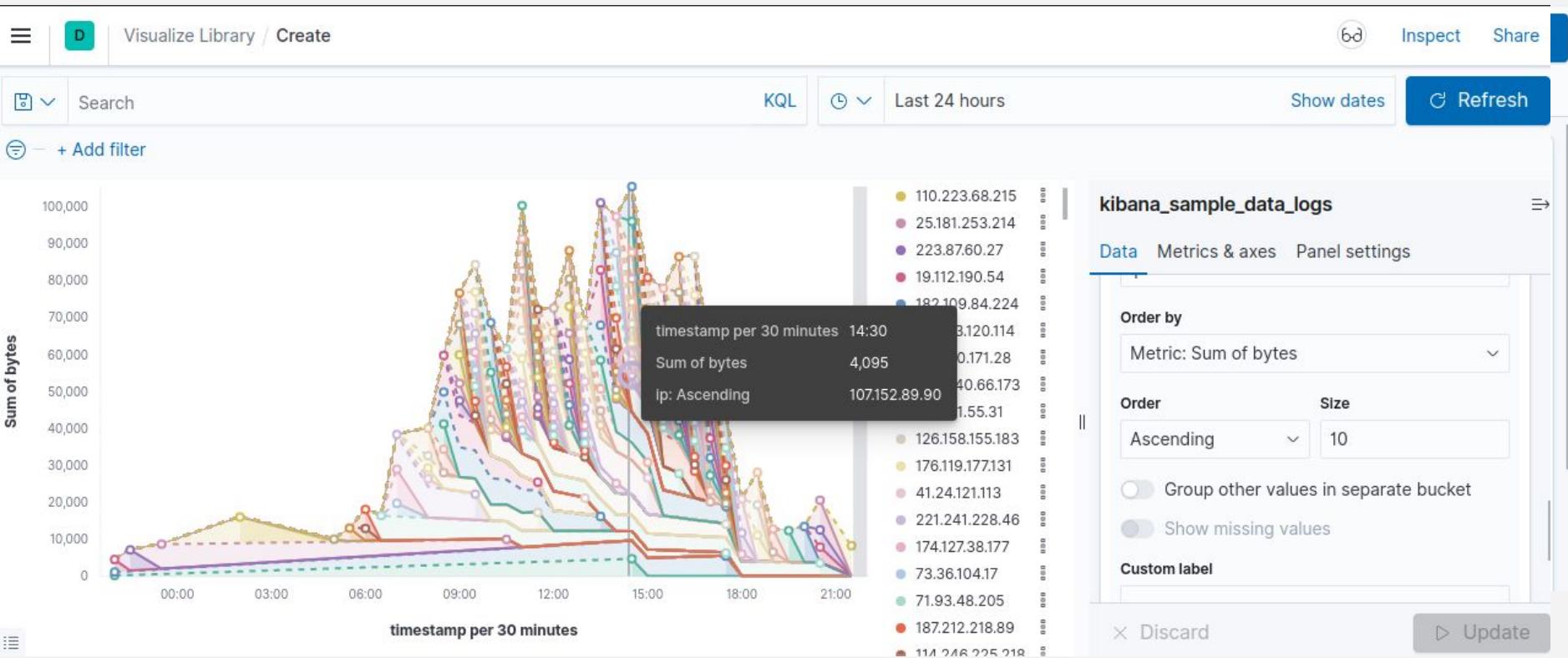
Order by Metric: Count

Order Size
Descending 5

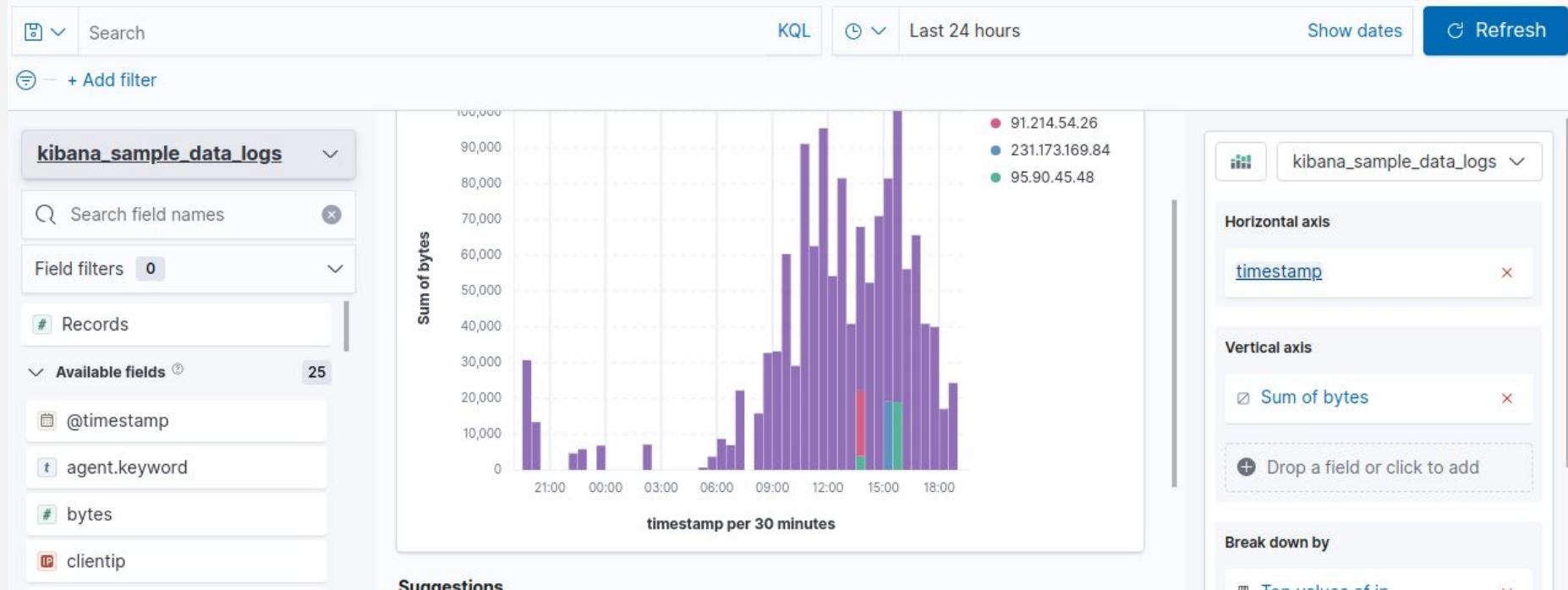
Group other values in separate bucket
 Show missing values

Custom label

Buckets



Buckets



Buckets

Visualize Library / Create

Search KQL Last 7 days Show dates Refresh

+ Add filter

cdn.elastic-elasticsearch.org
artifacts.elastic.co
www.elastic.co
elastic-elasticsearch.org

host.keyword: Descending - Count

kibana_sample_data_logs

Data Options

Buckets

Tags

Aggregation Terms help

Terms

Field host.keyword

Order by

X Discard ▷ Update

The screenshot shows the Kibana interface with a visualization titled 'Buckets'. The visualization displays three hosts: 'cdn.elastic-elasticsearch.org' in blue, 'artifacts.elastic.co' in green, and 'www.elastic.co' in orange. Below these, the URL 'elastic-elasticsearch.org' is shown in red. At the bottom, a histogram is displayed with the title 'host.keyword: Descending - Count'. To the right, a configuration panel for the 'kibana_sample_data_logs' index pattern is visible, showing settings for 'Buckets', 'Aggregation' (set to 'Terms'), 'Field' (set to 'host.keyword'), and 'Order by'. There are also 'Data' and 'Options' tabs, and buttons for 'Discard' and 'Update'.

Kibana time series

New Visualization

Start creating your visualization by selecting a type for that visualization.

Try **Lens**, our new, intuitive way to create visualizations.

[Go to Lens](#)

The visualization types shown are:

- Lens
- Area
- Controls
- Data Table
- Gauge
- Goal
- Heat Map
- Horizontal Bar
- Line
- Maps
- Markdown
- Metric
- Pie
- TSVB
- Tag Cloud
- Timelion
- Vega
- Vertical Bar

Kibana time series

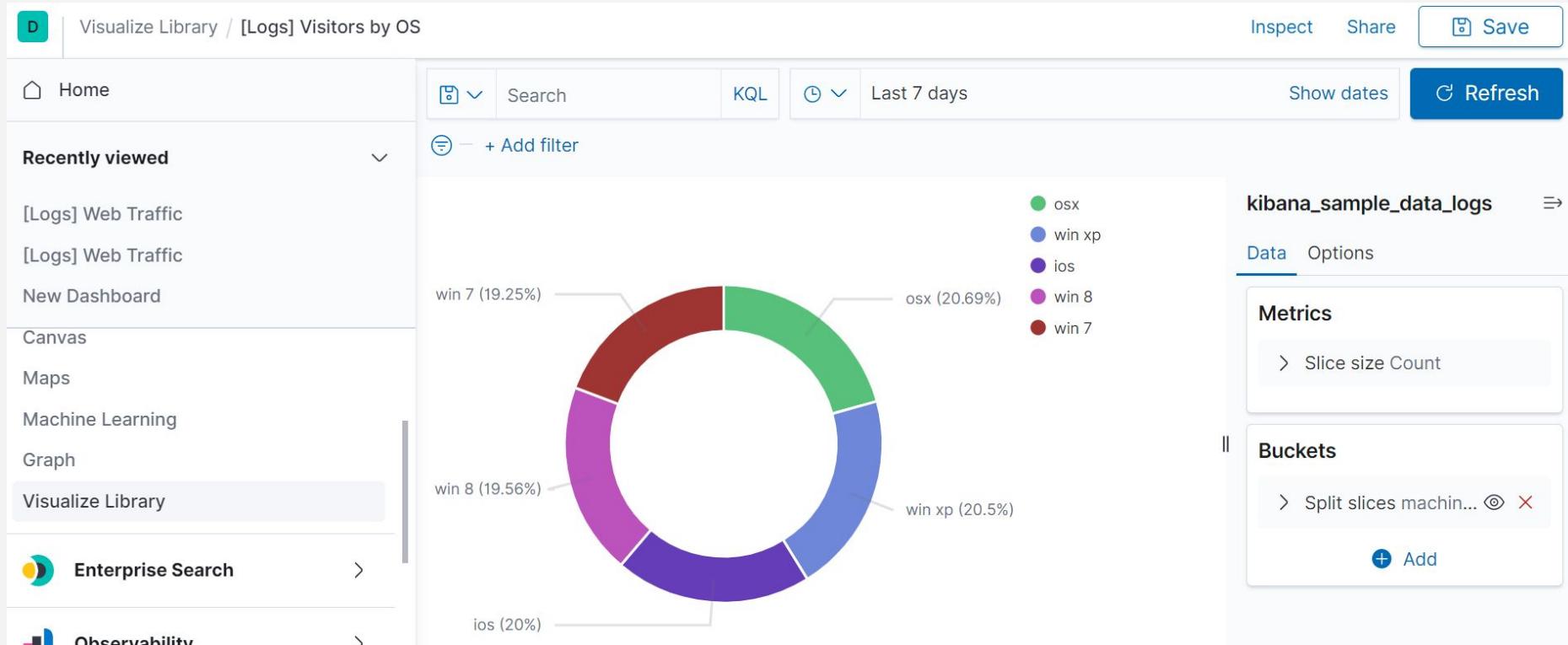


Kibana Dashboards

The screenshot shows the Kibana Dashboards interface. At the top, there is a dark header bar with the Elastic logo, a search bar labeled "Search Elastic", and some icons. Below the header, the navigation bar includes a menu icon, a dashboard icon, and the text "Dashboards". The main content area is titled "Dashboards" and contains a search bar. A blue button on the right says "+ Create dashboard". Below the search bar is a table listing ten dashboards. Each row in the table has a checkbox, a title, a description, and an "Actions" column with a pencil icon.

<input type="checkbox"/>	Title	Description	Actions
<input type="checkbox"/>	ML HTTP Access: Explorer (ECS)		
<input type="checkbox"/>	[Filebeat AWS] CloudTrail	Summary of events from AWS CloudTrail.	
<input type="checkbox"/>	[Filebeat AWS] ELB Access Log Overview	Filebeat AWS ELB Access Log Overview Dashboard	
<input type="checkbox"/>	[Filebeat AWS] S3 Server Access Log Overview	Filebeat AWS S3 Server Access Log Overview Dashboard	
<input type="checkbox"/>	[Filebeat AWS] VPC Flow Log Overview	Filebeat AWS VPC Flow Log Overview Dashboard	
<input type="checkbox"/>	[Filebeat ActiveMQ] Application Events	This dashboard shows application logs collected by the ActiveMQ filebeat module.	
<input type="checkbox"/>	[Filebeat ActiveMQ] Audit Events	This dashboard shows audit logs collected by the ActiveMQ filebeat module.	
<input type="checkbox"/>	[Filebeat Apache] Access and error logs ECS	Filebeat Apache module dashboard	
<input type="checkbox"/>	[Filebeat Auditd] Audit Events ECS	Dashboard for the Auditd Filebeat module	
<input type="checkbox"/>	[Filebeat Azure] Alerts Overview	This dashboard provides expanded alerts overview for Azure cloud	

Visualizaciones: Pie



Visualizaciones: Tablas de datos

D Visualize Library / [Logs] Host, Visits and Bytes Table

Inspect Share Save

Home

Recently viewed

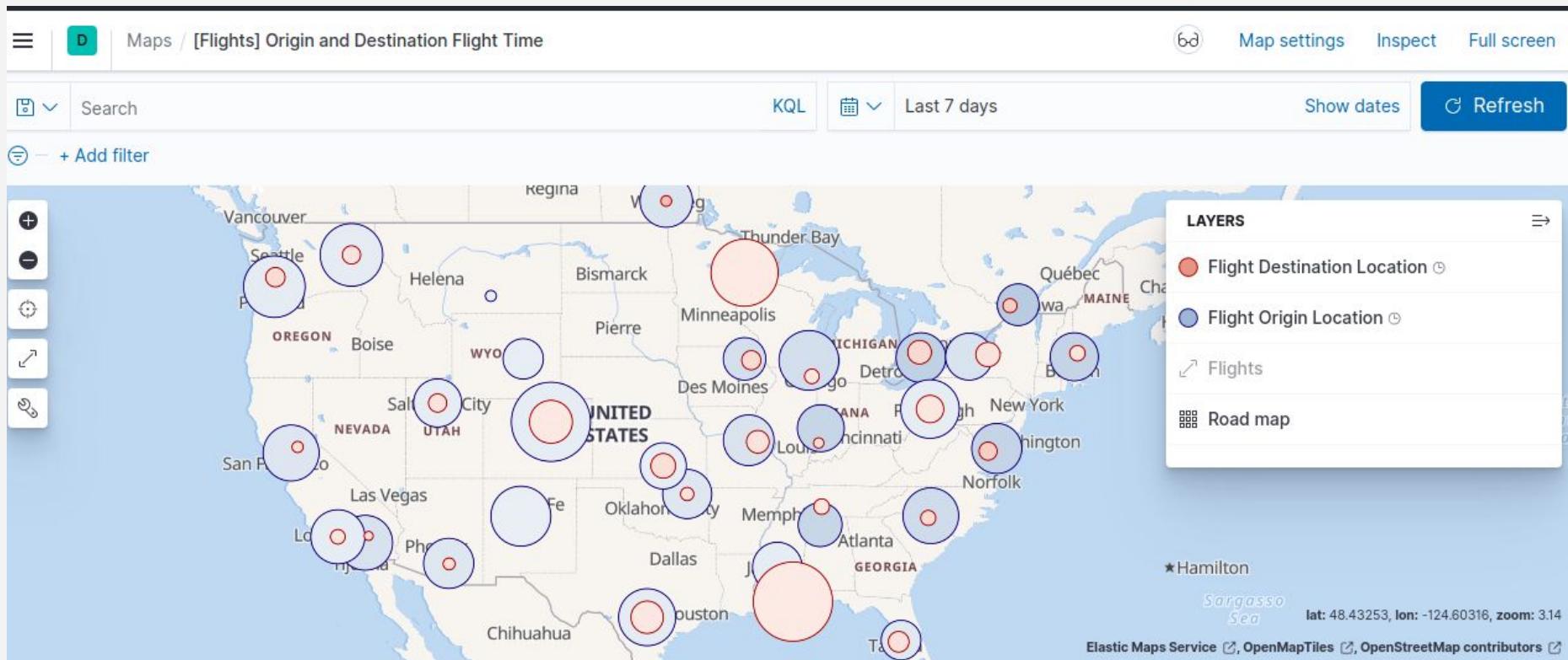
- [Logs] Total Requests and Bytes
- [Logs] Web Traffic
- [Logs] Web Traffic
- New Dashboard
- Canvas
- Maps
- Machine Learning
- Graph
- Visualize Library

Last 1 week rounded to the week Show dates Refresh

Type ↑	Bytes (Total)	Bytes (Last Hour)	Unique Visits (Total)	Unique Visits (Last Hour)
(empty)	6MB	29.3KB	1,173 ↓	8 ↑
gz	3.3MB	28.9KB	571 ↓	5 ↑
css	2.7MB	25.5KB	506 ↓	4 ↓
zip	2.3MB	7KB	384 ↓	2 ↑
deb	2.2MB	13.4KB	342 ↓	2 ↓
rpm	762.1KB	6.5KB	129 ↓	1 ↓

Auto apply The changes will be automatically applied.

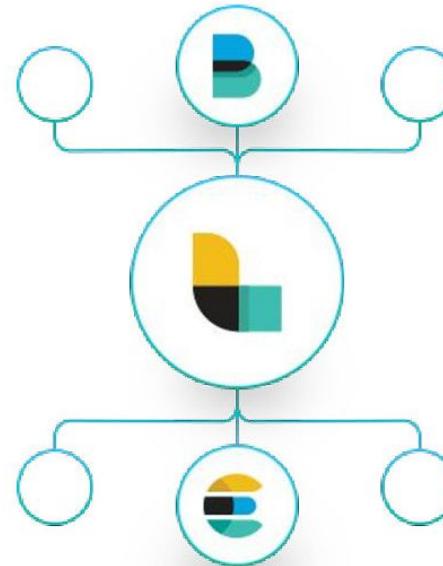
Visualizaciones: maps





Logstash

ETL para Elasticsearch



Ingestar datos de distintos tipos y fuentes

Parsear y transformar dinámicamente datos

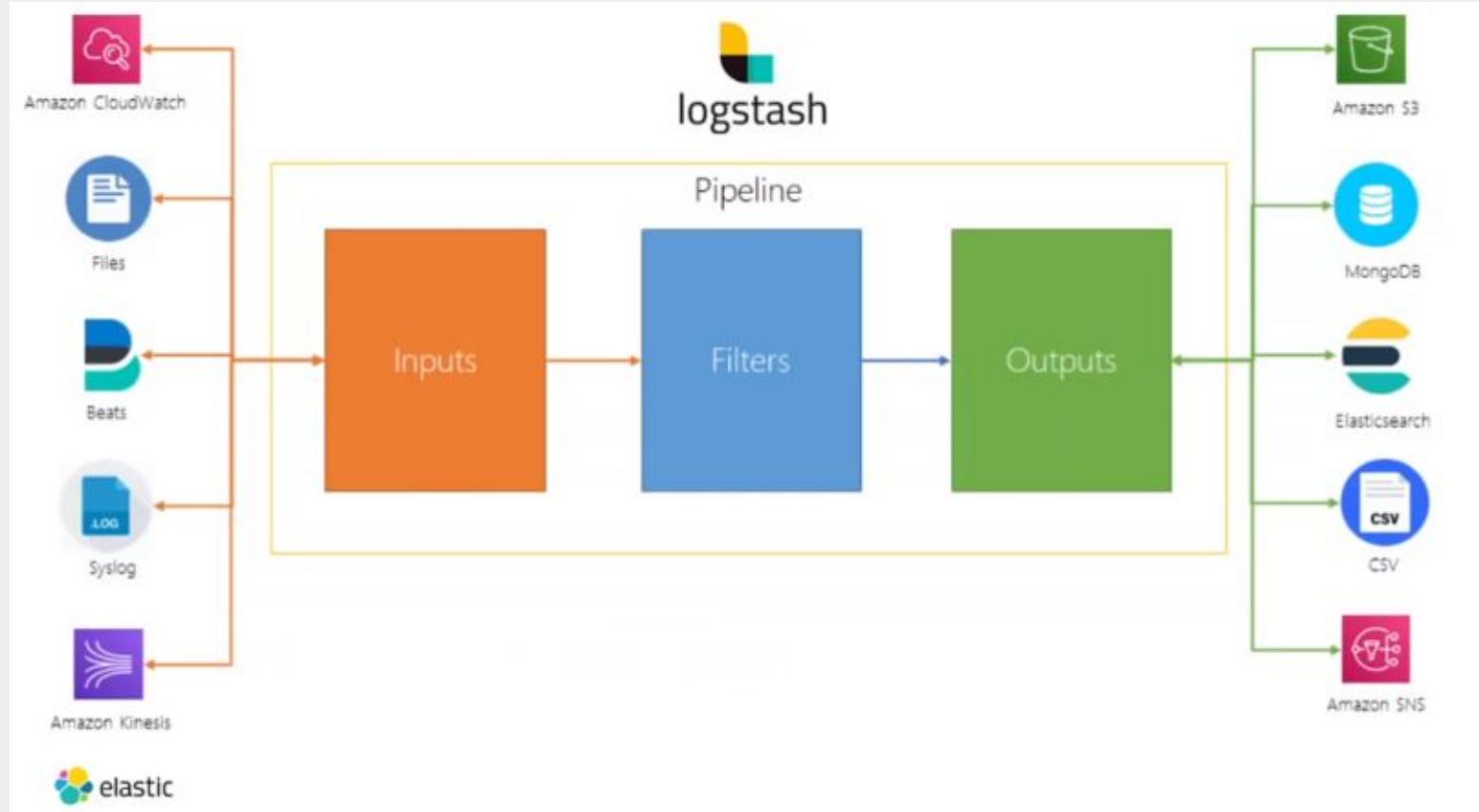
Transportar datos a varias salidas

Securizar y encriptar entradas de datos

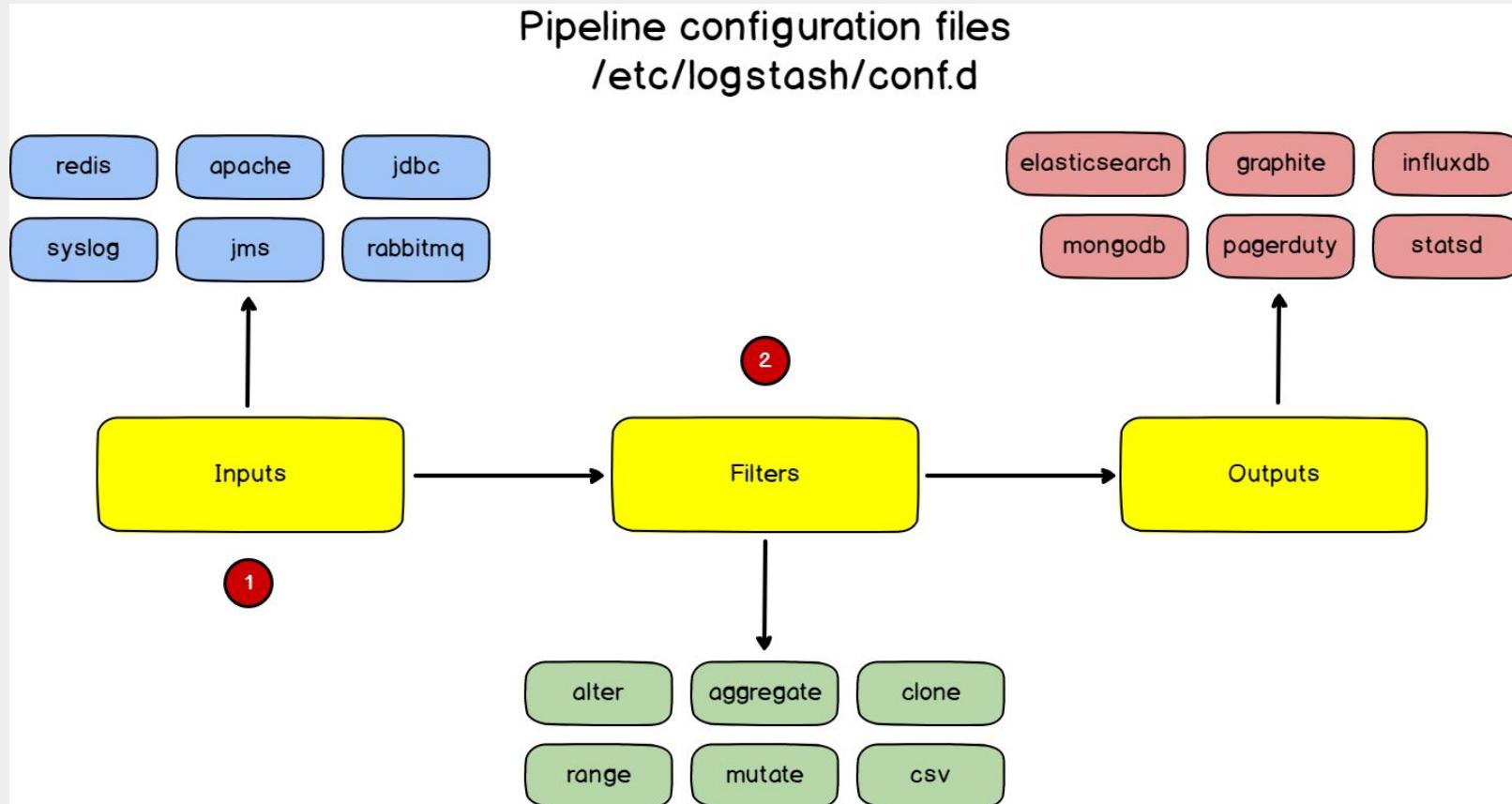
Construye tus propias pipelines

Gran variedad de plugins

Diseño de Logstash



Arquitectura de Logstash



Instalación de Logstash

```
sudo apt-get install logstash
```

```
sudo systemctl start logstash
```

```
][logstash.outputs.elasticsearch][main] ES Output version determined  
][logstash.outputs.elasticsearch][main] Detected a 6.x and above clus  
][logstash.outputs.elasticsearch][main] New Elasticsearch output {:cl  
][logstash.javapipeline    ][main] Starting pipeline {:pipeline_id=>"  
][logstash.javapipeline    ][main] Pipeline Java execution initializa  
][logstash.inputs.beats    ][main] Beats inputs: Starting input liste  
][logstash.javapipeline    ][main] Pipeline started {"pipeline.id"=>"  
][logstash.agent           ] Pipelines running {:count=>1, :running_p  
][org.logstash.beats.Server][main][0e3b303d763998eaa8de60184a2986df70  
][logstash.agent           ] Successfully started Logstash API endpoi
```

```
bin/logstash -e 'input { stdin {} } output { stdout {} }'
```

```
bin/logstash --verbose -f sample.conf
```

```
bin/logstash -e 'input { stdin {} } output {  
stdout { codec => rubydebug} }'
```

```
{  
  "message" => "Hello Logstash",  
  "@timestamp" => "2021-01-01T23:48:05.335Z",  
  "@version" => "1",  
  "host" => "127.0.0.1"  
}
```

Input Stdin (stdin_simple.conf)

```
input {
    stdin { }
}

output {
    stdout {
        codec => rubydebug
    }
}
```

Input Stdin

```
{  
  "message" => "Hello logstash",  
  "@timestamp" => "2021-09-20T23:48:05.335Z",  
  "@version" => "1",  
  "host" => "host"  
}
```

Input Stdin (stdin_json.conf)

```
input {  
    stdin{}  
}  
  
output {  
    stdout{  
        codec=>json_lines  
    }  
}
```

Input file

```
input {
  file {
    path => [ "/home/logstash/testdata.log" ]
    sincedb_path => "/dev/null"
    start_position => "beginning"
  }
}
```

Input file (file_json.conf)

```
input {
  file {
    path => "/home/linux/Descargas/elk/logstash/test.json"
    start_position => "beginning"
    codec => "json"
  }
}

filter {

}

output {
  stdout { codec => rubydebug }
}
```

Input file (test_logs.conf)

```
input {
  file {
    path => "/home/linux/Descargas/elk/logstash/logs_apm.log"
    type => "logs"
    sincedb_path => "/dev/null"
    start_position => "beginning"
  }
}

filter{}

output {
  stdout{
    codec => rubydebug
  }
}
```

Input beats (logstash-beats-sample.conf)

```
input {
  beats {
    port => 5044
  }
}

output {
  elasticsearch {
    hosts => ["http://localhost:9200"]
    index =>
    "%{@metadata}[beat]}-%{@metadata}[version]}-%{+YYYY.MM.dd}"
  }
}
```

Logstash plugins

	coralogix-resources Update twitter.conf	4bc316c on 17 Jul	 63 commits
	dead-letter-queue Update dlq.conf		3 months ago
	exec exec modifications		4 months ago
	generator adding a few more configs		4 months ago
	heartbeat updating index heartbeat epoch		4 months ago
	http-input external request		3 months ago
	http-poller test		3 months ago
	twitter Update twitter.conf		3 months ago
	unix edit conf socket		4 months ago
	websockets restructuring files and folders		4 months ago

Twitter plugin (twitter-input.conf)

```
input{
  twitter{
    consumer_key => "90Bg8UeX8uLOLn3leGS4Z3cB"
    consumer_secret => "Pzkh169mTzPXKX0FKypxLhSQGFhLerPMTA0596mYtCeCMupEXV"
    oauth_token => "201832916-39XbHLtdlinPZPVq4IkbuwgjVs0HyoHgAnqLE1QP"
    oauth_token_secret => "Q6P2clvCpKhjkIKnpKeXc95eC9B5SnmgY1bLRy9IqgcQY"
    keywords=>["ELK","ElasticSearch","Logstash","Kibana"]
  }
}

output{
  file{
    path=>"tweets.txt"
  }
  stdout{
    codec => rubydebug
  }
}
```

Input

- file, stdin, lumberjack, twitter etc.

Filter

- grok, grep, mutate, drop, date etc.

Output

- elasticsearch, stdout, mongodb etc.

LOGSTASH FILTERS

- Filter: Grok. Procesamiento básico de los logs
- Filter: Grok. Procesamiento avanzado de logs
- Filter Grok. Usos avanzados
- Filter: Mutate
- Filter: Date
- Filter: Translate
- Filter: Geolp
- Filter: Ruby

Filter: Grok. Procesamiento básico de los logs

% {PATTERN: FIELDNAME}

%{NUMBER:bytes_transferred}

Filter: Grok. Procesamiento básico de los logs

54.3.245.1 GET /index.html 14562 0.056

```
%{IP:client_ip} %{WORD:request_method }  
%{URIPATHPARAM:uri_path}  
%{NUMBER:bytes_transferred}  
%{NUMBER:duration}
```

Filter: Grok. Procesamiento básico de los logs

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

```
input {  
  file {  
    path => "/var/log/http.log"  
  }  
}  
filter {  
  grok {  
    match => { "message" => "%{IP:client} %{WORD:method} %  
{URI PATH PARAM:request} %{NUMBER:bytes} %{NUMBER:duration}" }  
  }  
}
```

Filter: Grok. Procesamiento básico de los logs

```
filter{
grok{
match => { "message"
=>"%{IP:client_ip} %{WORD:request_method} %{URIPATHPARAM:uri_path}
%{NUMBER:bytes_transferred}
%{NUMBER:duration}"}
}
}
```

- client_ip : 54.3.245.1
- request_method : GET
- uri_path :/index.html
- bytes_transferred :14562
- duration :0.056

Filter: Grok. Procesamiento avanzado de logs

```
filter {
  if [type] == "syslog" {
    grok {
      match => { "message" => "%{SYSLOGTIMESTAMP:syslog_timestamp}
%{SYSLOGHOST:syslog_hostname}
%{DATA:syslog_program}(?:\[%{POSINT:syslog_pid}\])?:
%{GREEDYDATA:syslog_message}" }
      add_field => [ "received_at", "%{@timestamp}" ]
      add_field => [ "received_from", "%{host}" ]
    }
    syslog_pri { }
    date {
      match => [ "syslog_timestamp", "MMM d HH:mm:ss", "MMM dd HH:mm:ss" ]
    }
  }
}
```

Filter: Grok. Procesamiento avanzado de los logs

- <https://github.com/logstash-plugins/logstash-patterns-core/tree/master/patterns>
- <http://grokdebug.herokuapp.com/>
- <http://grokconstructor.appspot.com/>

Filter: Grok. Procesamiento avanzado de los logs

Grok Debugger Debugger Discover Patterns

```
Aug 17 12:00:57 linux-HP-EliteBook-8470p rsyslogd: [origin software="rsyslogd" swVersion="8.32.0" x-pid="991" x-info="http://www.rsyslog.com"]  
rsyslogd was HUPed
```

```
%{SYSLOGTIMESTAMP:syslog_timestamp} %{SYSLOGHOST:syslog_hostname} %{DATA:syslog_program}(?:\[%  
{POSINT:syslog_pid}\])?: %{GREEDYDATA:syslog_message}
```

Add custom patterns Keep Empty Captures Named Captures Only Singles Autocomplete Go

```
{  
    "syslog_timestamp": [  
        [  
            "Aug 17 12:00:57"  
        ]  
    ],  
    "MONTH": [  
        [  
            "Aug"  
        ]  
    ],  
    "YEAR": [  
        [  
            "2017"  
        ]  
    ]  
}
```

Filter Mutate

```
filter {  
  mutate {  
    convert => # hash of field and data type (optional)  
    join => # hash of fields to be joined (optional)  
    lowercase => # array of fields to be converted (optional)  
    merge => # hash of fields to be merged (optional)  
    rename => # hash of original and rename field (optional)  
    replace => # hash of fields to replaced with (optional)  
    split => # hash of fields to be split (optional)  
    strip => # array of fields (optional)  
    uppercase => # array of fields (optional)  
  }  
}
```

Filter Mutate (logstash_mutate.conf)

```
input {
    stdin { }
}

filter {
    mutate {
        uppercase => [ "message" ]
    }
}

output {
    stdout {
        codec => rubydebug
    }
}
```

Filter Mutate (file_json_filter.conf)

```
filter {
    mutate {
        remove_field => [ "@version" ]
        add_field => { "tipoUsuario" => "cliente" }
        gsub => ["surname", " - ", ""]
    }
}
```

Filter Date

```
date{  
  match => ["date_field", "yyyy-MM-dd"]  
  target => "@timestamp"  
}
```

```
match => ["date_field", "MMM dd YYYY HH: mm:  
ss", "ISO8601", "MMddYYYY", "MMM d AAAA  
HH: mm: ss"]
```

Filter translate (logstash_translate.conf)

```
input {
  # The generator creates an input event
  generator {
    lines => [
      {"my_msg": "testing1234", "lookup_id": "1234"}
    ]
    count => 1
    codec => "json"
  }
}

filter {
  # Enrich the event using the lookup_id
  translate {
    field => "[lookup_id]"
    destination => "[enrichment_data]"
    fallback => "not_found"
    dictionary => {
      "1234" => "1234 found in the dictionary"
      "5678" => "5678 found in the dictionary"
    }
  }
}

output {
  stdout { codec => "rubydebug" }
}
```

Filter geoip (logstash_geoip.conf)

```
$ echo "logstash 19.1.193.230 $(date --iso-8601=seconds)" >> ~/input.txt
```

```
input {
  file {
    path => "/home/linux/Descargas/elk/logstash/input_geoip.txt"
    start_position => "beginning"
    sincedb_path => "/dev/null"
  }
}
filter {
  grok {
    match => { "message" => "%{WORD:name} %{IP:ip} %{TIMESTAMP_ISO8601:date}" }
    remove_field => [ "message", "path", "@version", "host" ]
  }
  geoip {
    source => "ip"
  }
}
output {
  stdout {
    codec => rubydebug
  }
}
```

Filter geoip

```
clientip => "83.149.9.216"
```

```
geoip{  
    source => clientip  
}
```

```
geoip:  
  •{timezone: "Europe/Moscow",  
  •ip: "83.149.9.216",  
  •latitude: 55.7485,  
  •continent_code: "EU",  
  •city_name: "Moscow",  
  •country_name: "Russia",  
  •country_code2: "RU",  
  •country_code3: "RU",  
  •region_name: "Moscow",  
  •location:  
    •{lon: 37.6184,  
    •lat: 55.7485  
    •},  
  •postal_code: "101194",  
  •region_code: "MOW",  
  •longitude: 37.6184  
  }
```

Filter geoip

```
{  
    "@timestamp" => 2021-09-11T19:57:45.958Z,  
        "ip" => "19.1.193.230",  
        "date" => "2021-09-11T21:13:17+02:00",  
    "geoip" => {  
        "longitude" => -97.822,  
        "location" => {  
            "lon" => -97.822,  
            "lat" => 37.751  
        },  
        "timezone" => "America/Chicago",  
    "country_code2" => "US",  
    "continent_code" => "NA",  
        "ip" => "19.1.193.230",  
    "country_code3" => "US",  
    "country_name" => "United States",  
        "latitude" => 37.751  
    },  
    "name" => "logstash"  
}
```

Filter ruby (logstash_ruby_filter.conf)

```
input { # ... }

filter {
    ruby {
        code => 'size = event.get("message").size;
                  event.set("message_size", size)'
    }
}

output { # ... }
```

Filter csv

```
filter{  
    csv {  
        columns => ["date","open_price","close_price"]  
        path => "/path/to/file.csv",  
        separator => ","  
    }  
}
```

Filter csv (read_csv.conf)

```
input {
  file {
    path => "/home/linux/Descargas/elk/logstash/file.csv"
    start_position => "beginning"
    sincedb_path => "/dev/null"
  }
}

filter {
  csv {
    separator => ","
    skip_header => "true"
    columns =>
    ["id","timestamp","paymentType","name","gender","ip_address","purpose","country","age"]
  }
}
output {
  stdout { codec => rubydebug }
}
```

LOGSTASH OUTPUTS

- Output: Stdout
- Output: Elasticsearch
- Otros plugins de salida
- Combinando configuraciones
- Monitorización de Logstash
- Configuración avanzada
- Múltiples pipelines
- Uso de pipelines

Output: Stdout

```
output {  
  stdout {  
    codec => rubydebug  
    workers => 2  
  }  
}
```

Output: Elasticsearch

```
output {  
  elasticsearch {  
    hosts => ["localhost:9200"]  
    index => "logstash"  
    user => "elastic"  
    password => "elastic"  
    doc_as_upsert => true  
  }  
}
```

Otros plugins de salida

```
output {  
    file {  
        create_if_deleted => true  
        file_mode => 777  
        filename_failure => "failedpath_file"  
        flush_interval => 0  
        path => "/home/linux/Descargas/elk/logstash/file_output.txt"  
    }  
}
```

Procesar logs nginx

```
input {
  file {
    type => "nginx"
    path => "/home/linux/Descargas/elk/logstash/nginx_logs.txt"
    start_position => "beginning"
    sincedb_path => "/dev/null"
  }
}
filter {
  if [type] == "nginx" {
    grok {
      patterns_dir => "/home/linux/Descargas/logstash-7.14.1/patterns"
      match => { "message" => "%{NGINX_ACCESS}" }
      remove_tag => ["_grokparsefailure"]
      add_tag => ["nginx_access"]
    }
    geoip {
      source => "clientip"
    }
  }
}
output {
  stdout{
    codec => rubydebug
  }
  elasticsearch {
    hosts => ["localhost:9200"]
    user=>"elastic"
    password=>"elastic"
  }
}
```

Procesar logs nginx

```
80.91.33.133 - - [17/May/2015:08:05:24 +0000] "GET  
/downloads/product_1 HTTP/1.1" 304 0 "-" "Debian  
APT-HTTP/1.3(0.8.16~exp12ubuntu10.17)"
```

```
NGINX_ACCESS %{IPORHOST:clientip} (?:-|(%{WORD}.%{WORD}))  
%{USER:ident} \[%{HTTPDATE:timestamp}\] "(?:%{WORD:verb}  
%{NOTSPACE:request}(?:  
HTTP/%{NUMBER:httpversion})?|%{DATA:rawrequest})"  
%{NUMBER:response} (?:%{NUMBER:bytes}|-) %{QS:referrer}  
%{QS:agent}
```

Monitorización de Logstash

```
GET /logstash-*/_search
{
  "query": {
    "match_all": {}
  },
  "sort": {
    "@timestamp": {
      "order": "desc"
    }
  }
}
```

Monitorización de Logstash

```
curl -X GET http://localhost:9600?pretty
```

```
curl -X GET http://localhost:9600/\_node/?pretty
```

```
curl -X GET http://localhost:9600/\_node/os?pretty
```

```
curl -X GET http://localhost:9600/\_node/jvm?pretty
```

```
curl -X GET http://localhost:9600/\_node/pipelines?pretty
```

```
curl -X GET http://localhost:9600/\_node/plugins?pretty
```

```
curl -X GET http://localhost:9600/\_node/stats?pretty
```

```
curl -X GET http://localhost:9600/\_node/stats/jvm?pretty
```

```
curl -X GET http://localhost:9600/\_node/stats/process?pretty
```

```
curl -X GET http://localhost:9600/\_node/stats/pipelines?pretty
```

```
curl -X GET http://localhost:9600/\_node/hot\_threads?pretty
```

Configuración avanzada

```
$ bin/logstash -w 12  
$ bin/logstash --pipeline.workers 12
```

```
$ bin/logstash -b 50  
$ bin/logstash --pipeline.batch.size 50
```

Configuración avanzada

```
$ bin/logstash -I PATH  
$ bin/logstash --path.logs <PATH>
```

```
$ bin/logstash --log.level <LEVEL>
```

- fatal
- error
- warn
- info
- debug
- trace

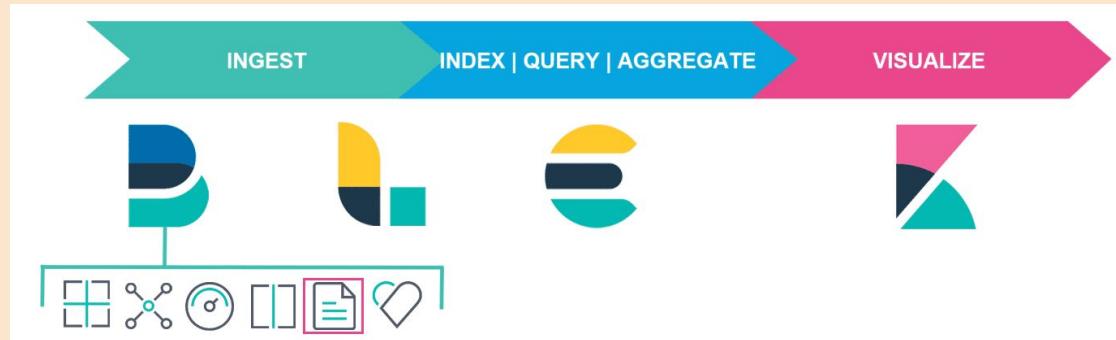
Pipelines

```
{  
  "description" : "...",  
  "processors" : [ ... ]  
}
```

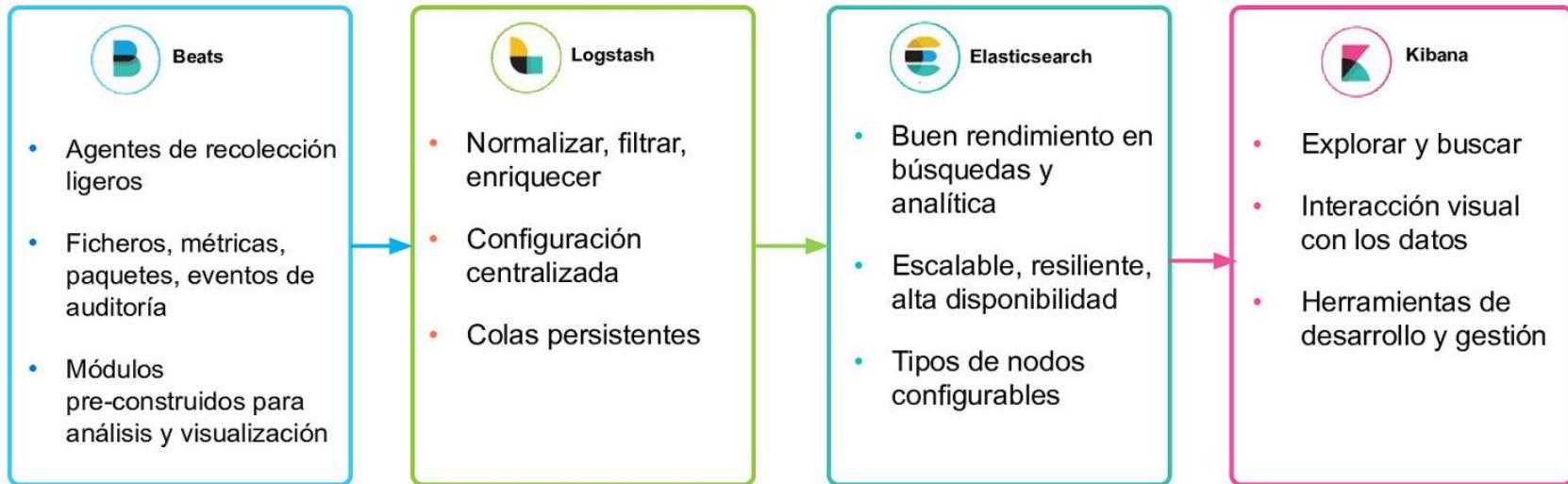
http://localhost:9200/_ingest/pipeline

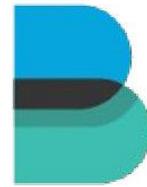
BEATS

- ¿Qué son los Beats?
- MetricBeat
- PacketBeat
- FileBeat
- FileBeat: configuración manual
- FileBeat: configuración avanzada
- FileBeat: uso de módulos preconfigurados
- Análisis con FileBeat
- WinlongBeat
- Heartbeat



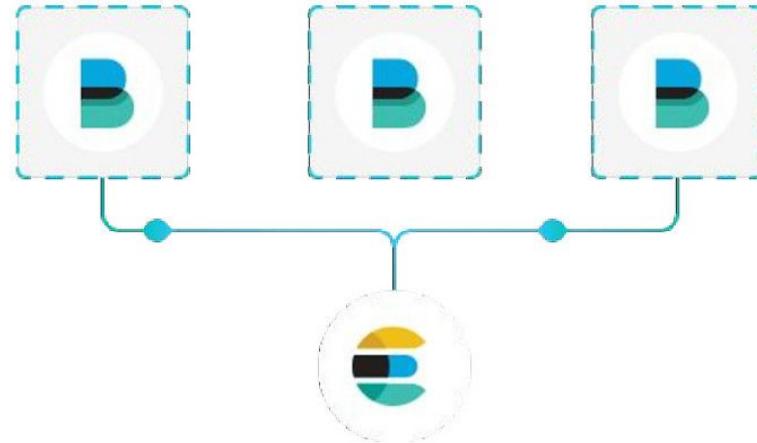
Pipeline Lógica de Ingesta





Beats

Agentes ligeros



Enviar datos desde
las fuentes

Enviar y centralizar los
datos en Elasticsearch

Enviar datos a Logstash si se
requieren transformaciones

Enviar al Cloud de Elastic

Libbeat: API framework para
construir beats a medida

70+ Beats de comunidad

¿Qué son los Beats?



Beats

Todos los módulos



FileBeat
Log Files



MetricBeat
Metrics



PacketBeat
Network Data



WinLogBeat
Window Events



HeartBeat
Uptime Monitoring

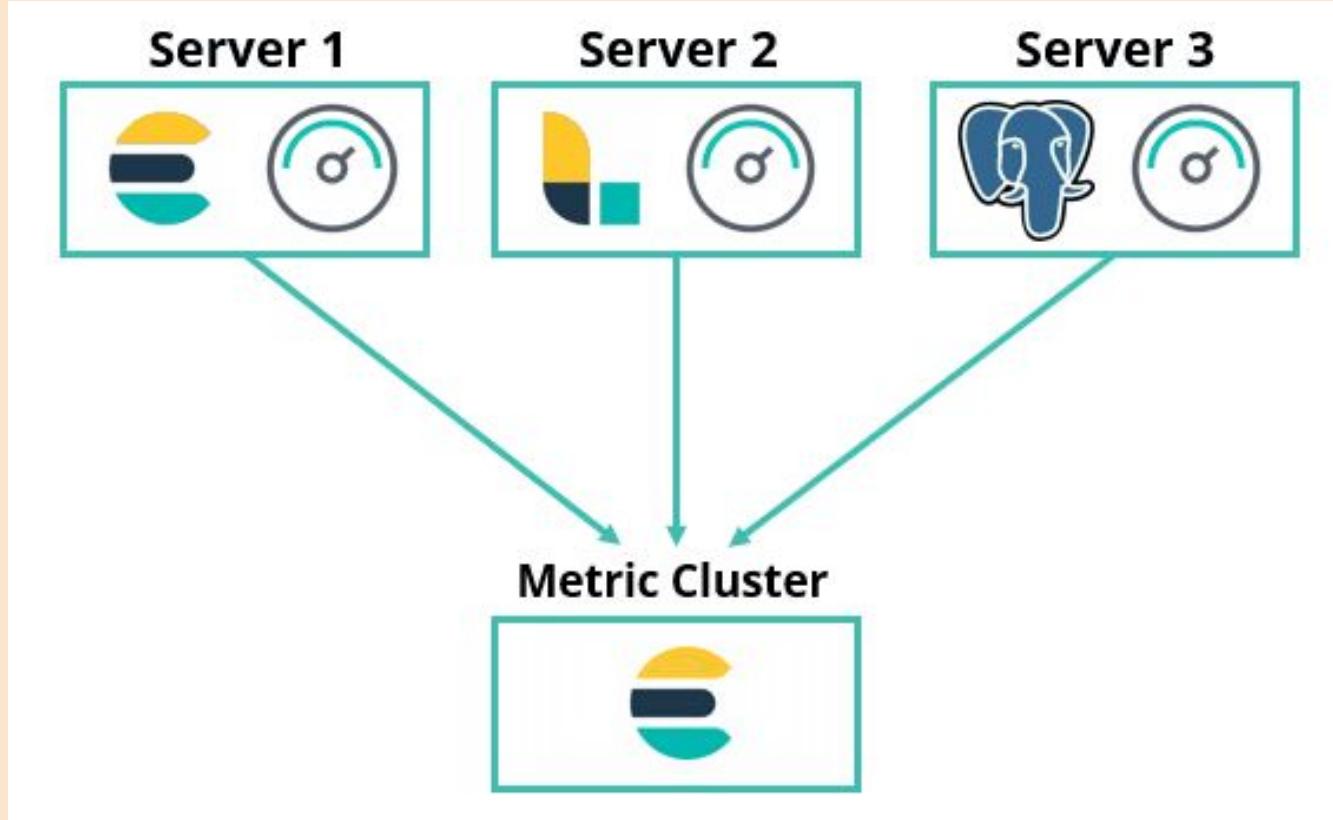


AuditBeat
Audit Data



FunctionBeat
Serverless Shipper

MetricBeat



MetricBeat

```
$ curl -L -O  
https://artifacts.elastic.co/downloads/beats/metricbeat/metricbeat-7.10.0-linux-x86\_64.tar.gz  
  
$ tar -xzf metricbeat-7.10.0-linux-x86_64.tar.gz  
  
$ cd metricbeat-7.10.0-linux-x86_64
```

MetricBeat

```
$./metricbeat modules list
```

```
$./metricbeat modules enable system  
Module system is already enabled
```

Metricbeat

```
ls modules.d
```

```
cat modules.d/system.yml
```

MetricBeat módulo system

- **core**: proporciona estadísticas de uso para cada núcleo de CPU.
- **cpu**: proporciona estadísticas de CPU.
- **diskio**: proporciona métricas de E/S a nivel de disco.
- **filesystem**: proporciona estadísticas del sistema de archivos. Para cada sistema de archivos, se crea un evento.
- **process**: proporciona estadísticas a nivel de proceso. Se crea un evento para cada proceso.
- **process_summary**: recopila estadísticas de alto nivel sobre los procesos en ejecución.
- **fsstat**: este conjunto de métricas proporciona estadísticas generales del sistema de archivos.
- **load**: proporciona estadísticas del uso de la memoria.
- **network**: proporciona métricas de E/S de red recopiladas del sistema operativo. Se crea un evento para cada interfaz de red.
- **socket**: este conjunto de métricas informa un evento para cada nuevo socket TCP. Este conjunto de métricas está disponible solo en Linux y requiere el kernel 2.6.14 o posterior.

MetricBeat metricbeat.yml

#Modules

metricbeat.config.modules:

```
# Glob pattern for configuration loading
path: ${path.config}/modules.d/*.yml
# Set to true to enable config reloading
reload.enabled: false
# Period on which files under path should be checked for changes
#reload.period: 10s
```

#Dashboards

setup.dashboards.enabled: true

Kibana

setup.kibana:

```
host: "localhost:5601"
username: "xxxx"
password: "xxxx"
```

Outputs

output.elasticsearch:

```
hosts: ["localhost:9200"]
username: "xxxx"
password: "xxxx"
```

Logging

logging.level: info

MetricBeat system.yml

```
#system.yml
- module: system
period: 10s
metricsets:
- cpu
- load
- memory
- network
- process
- process_summary
#- core
#- diskio
#- socket
processes: ['.*']
process.include_top_n:
by_cpu: 5 # include top 5 processes by CPU
by_memory: 5 # include top 5 processes by memory
- module: system
period: 1m
metricsets:
- filesystem
- fsstat
processors:
- drop_event.when.regexp:
system.filesystem.mount_point: '^/(sys|cgroup|proc|dev|etc|host|lib)(\$|/)'
```

Metrics

```
$ sudo ./metricbeat setup -e
```

```
$ sudo ./metricbeat -e -c metricbeat.yml  
--strict.perms=false
```

Dashboards MetricBeat

Dashboards

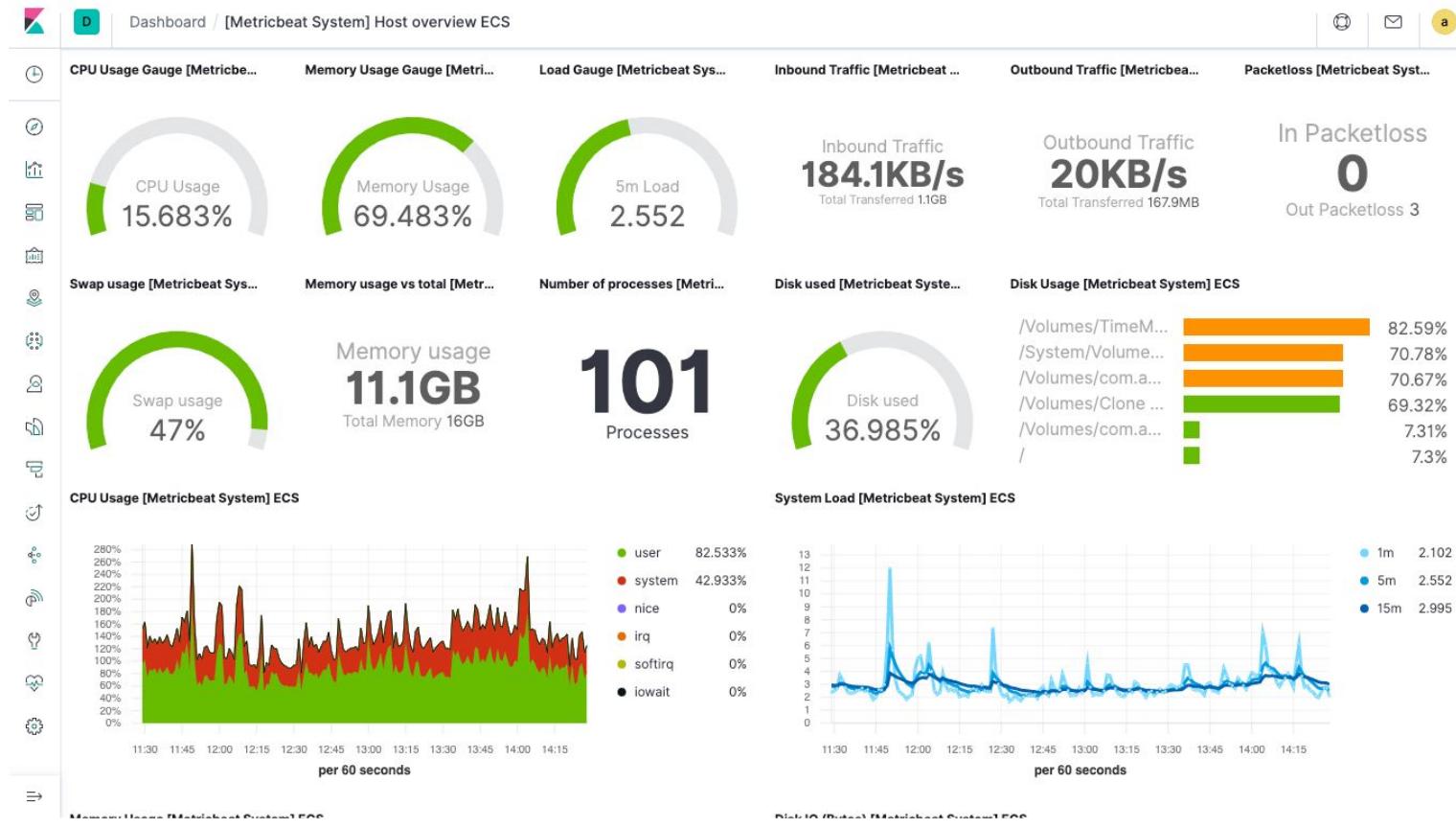
 Create dashboard

<input type="checkbox"/>	Title	Description	Actions
<input type="checkbox"/>	[Metricbeat System] Overview ECS	Overview of system metrics	
<input type="checkbox"/>	[Metricbeat System] Host Services Overview	Overview of services on an individual host.	
<input type="checkbox"/>	[Metricbeat System] Containers overview ECS	Overview of container metrics	
<input type="checkbox"/>	[Metricbeat System] Host overview ECS	Overview of host metrics	

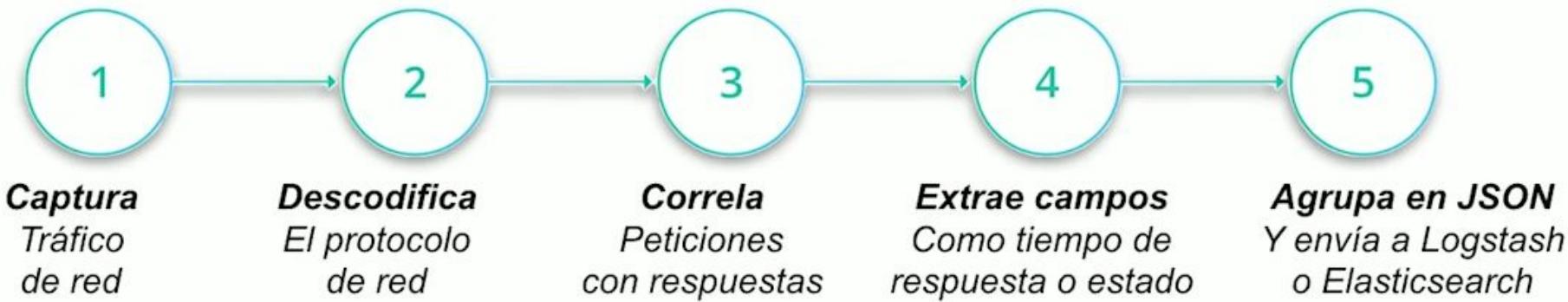
Rows per page: 20 ▾

< 1 >

Dashboards MetricBeat



PacketBeat



PacketBeat

```
$ sudo ./packetbeat devices
0: wlo1 (No description available) (192.168.18.21 fe80::946a:652b:b840:f73c)
1: veth2f3bca6 (No description available) (fe80::64a3:d4ff:fe1e:bb04)
2: br-5a6c4e7c1a9a (No description available) (172.24.0.1 fe80::42:6bff:fee1:d547)
3: any (Pseudo-device that captures on all interfaces) (Not assigned ip address)
4: lo (No description available) (127.0.0.1 ::1)
5: docker0 (No description available) (172.17.0.1)
6: br-f765bf712ace (No description available) (172.19.0.1)
7: enp0s25 (No description available) (Not assigned ip address)
8: br-57c204f68b44 (No description available) (172.18.0.1)
9: br-7916e18c4b81 (No description available) (172.22.0.1)
10: br-7193c1fd827a (No description available) (172.23.0.1)
11: br-dfb23b7f8475 (No description available) (172.21.0.1)
12: br-46a262ff8557 (No description available) (172.20.0.1)
13: nflog (Linux netfilter log (NFLOG) interface) (Not assigned ip address)
14: nfqueue (Linux netfilter queue (NFQUEUE) interface) (Not assigned ip address)
```

PacketBeat

```
packetbeat.interfaces.device: <interface>|any
```

```
packetbeat.protocols.http:
```

```
ports: [80, 8080]
```

```
output.elasticsearch:
```

```
# Array of hosts to connect to.
```

```
hosts: ["localhost:9200"]
```

PacketBeat

```
$ ./packetbeat test output
elasticsearch: http://localhost:9200...
    parse url... OK
connection...
    parse host... OK
    dns lookup... OK
    addresses: 127.0.0.1
$ ./packetbeat test config
```

PacketBeat

```
$ sudo tcpdump -i <interface> -s 65535 -w  
http-capture-packetbeat.pcap
```

```
$ ./packetbeat run -I  
http-capture-packetbeat.pcap
```

Búsquedas en PacketBeat

#Tráfico HTTP/S

event.category: network

event.type: connection

event.kind: event

network.protocol: http | tls

#Tráfico dns

event.category: network

event.type: connection

event.kind: event

network.protocol: dns

FileBeat

- Se trata de una herramienta ligera, que no consume mucha memoria.
- Puede monitorizar cualquier directorio específico de logs.



Apache



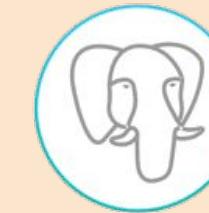
Nginx



MongoDB



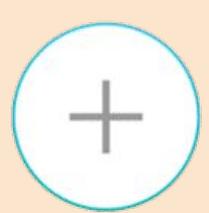
MySQL



PostgreSQL

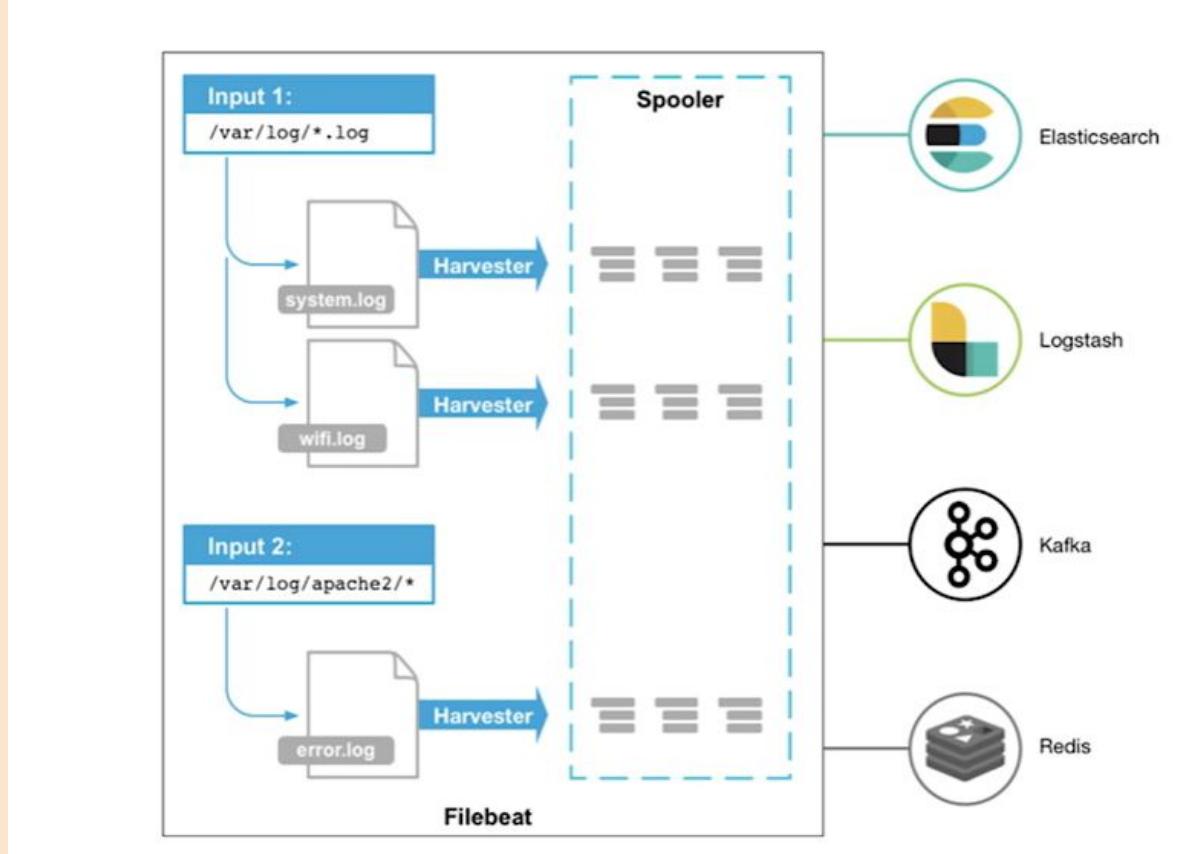


Redis



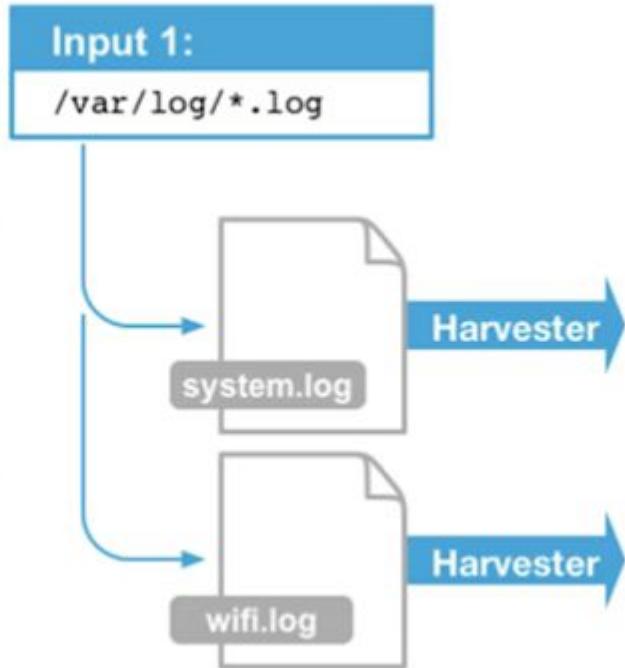
More

FileBeat

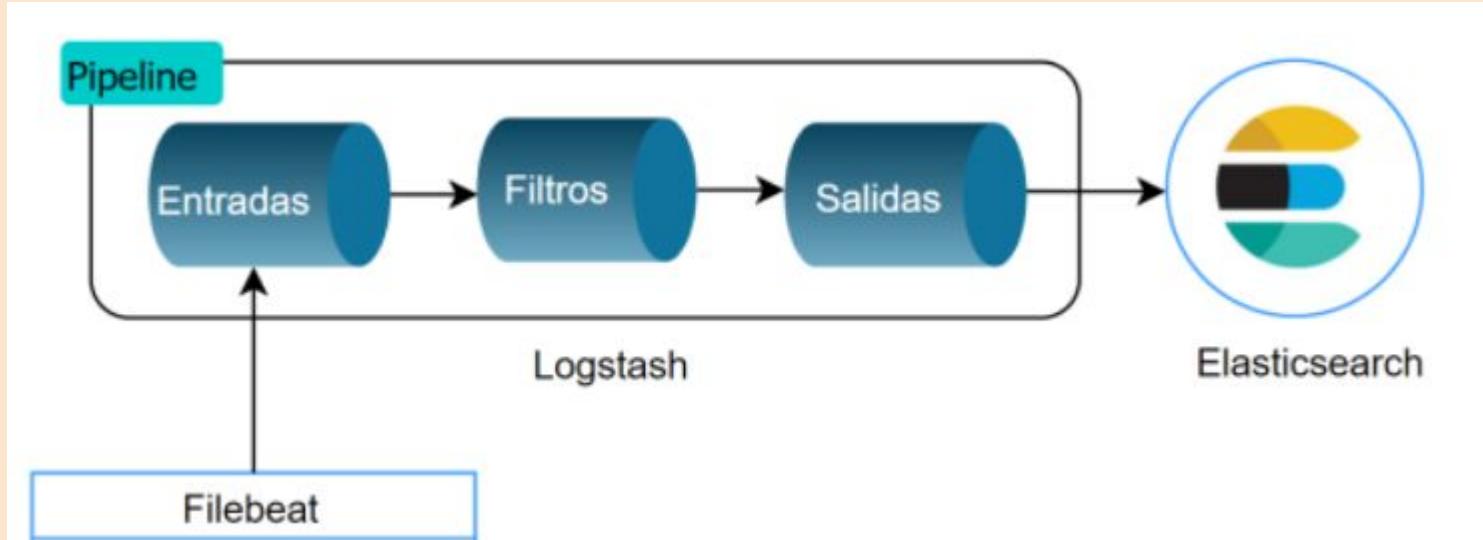


FileBeat

The new log data is aggregated and sent to the configured output by libbeat.



FileBeat



FileBeat:Instalación

1

Download and install Filebeat

First time using Filebeat? See the [Quick Start](#).

[Copy snippet](#)

```
curl -L -O https://artifacts.elastic.co/downloads/beats/filebeat/filebeat-7.13.3-darwin-x86_64.tar.gz  
tar xzvf filebeat-7.13.3-darwin-x86_64.tar.gz  
cd filebeat-7.13.3-darwin-x86_64/
```

FileBeat: configuración manual

filebeat.inputs:

- **type: log**

enabled: true

paths:

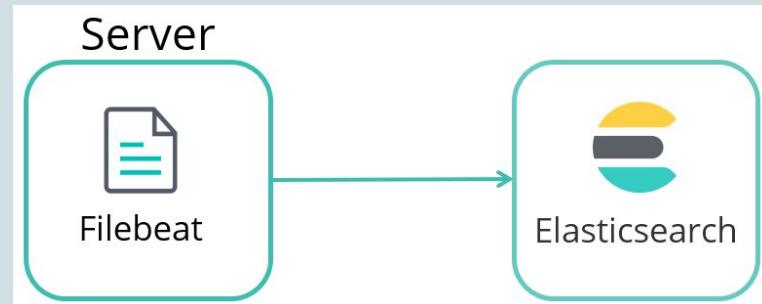
- **/var/log/*.log**

output.elasticsearch:

hosts: ["localhost:9200"]

setup.kibana:

hosts: ["localhost:5601"]



FileBeat: configuración manual

```
#----- Elasticsearch output -----
##output.elasticsearch:
#enable:true
# Array of hosts to connect to.
# hosts: ["localhost:9200"]

# Enabled ilm (beta) to use index lifecycle management instead daily indices.
#ilm.enabled: false

# Optional protocol and basic auth credentials.
#protocol: "https"
#username: "elastic"
#password: "elastic"

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

FileBeat: configuración logstash

```
input {
  beats {
    port => 5044
  }
}

output {
  elasticsearch {
    hosts => ["http://localhost:9200"]
    index =>
"%{@metadata}[beat]}-%{@metadata}[version]}-%{+YYYY.MM.dd}"
  }
}
```

FileBeat: uso de módulos preconfigurados

```
./filebeat modules list
```

```
./filebeat modules enable apache
```

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

FileBeat: uso de módulos preconfigurados

```
[elastic@server1 modules.d]$ ls
apache.yml.disabled      auditd.yml.disabled      cisco.yml.disabled
coredns.yml.disabled    .elasticsearch.yml.disabled envoyproxy.yml.disabled
googlecloud.yml.disabled haproxy.yml.disabled    icinga.yml.disabled
iis.yml.disabled         iptables.yml.disabled    kafka.yml.disabled
kibana.yml.disabled      logstash.yml.disabled   mongodb.yml.disabled
mssql.yml.disabled       mysql.yml                 nats.yml.disabled
netflow.yml.disabled     nginx.yml                osquery.yml.disabled
panw.yml.disabled        postgresql.yml.disabled rabbitmq.yml.disabled
redis.yml.disabled       santa.yml.disabled       suricata.yml.disabled
system.yml               traefik.yml.disabled    zeek.yml.disabled
```

FileBeat: uso de módulos preconfigurados

```
- module: apache
  access:
    enabled: true
    var.paths: ["/path/to/log/apache/access.log*"]
  error:
    enabled: true
    var.paths: ["/path/to/log/apache/error.log*"]
```

FileBeat: testing configuration

```
./filebeat test config
```

```
./filebeat test output
```

```
./filebeat setup -e
```

```
./filebeat -e -c filebeat.yml
```

Análisis con FileBeat

http://localhost:9200/filebeat*/_search?pretty

Create index pattern

Kibana uses index patterns to retrieve data from Elasticsearch indices for things like visualizations.



Include system indices

Step 1 of 2: Define index pattern

Index pattern

filebeat*

You can use a * as a wildcard in your index pattern.

You can't use spaces or the characters \, /, ?, ", <, >, |.

> Next step

✓ Success! Your index pattern matches 11 indices.

Análisis con FileBeat

Create index pattern

Kibana uses index patterns to retrieve data from Elasticsearch indices for things like visualizations.

X Include system indices

Step 2 of 2: Configure settings

You've defined **filebeat*** as your index pattern. Now you can specify some settings before we create it.

Time Filter field name

Refresh

@timestamp



The Time Filter will use this field to filter your data by time.

You can choose not to have a time field, but you will not be able to narrow down your data by a time range.

> Show advanced options

< Back

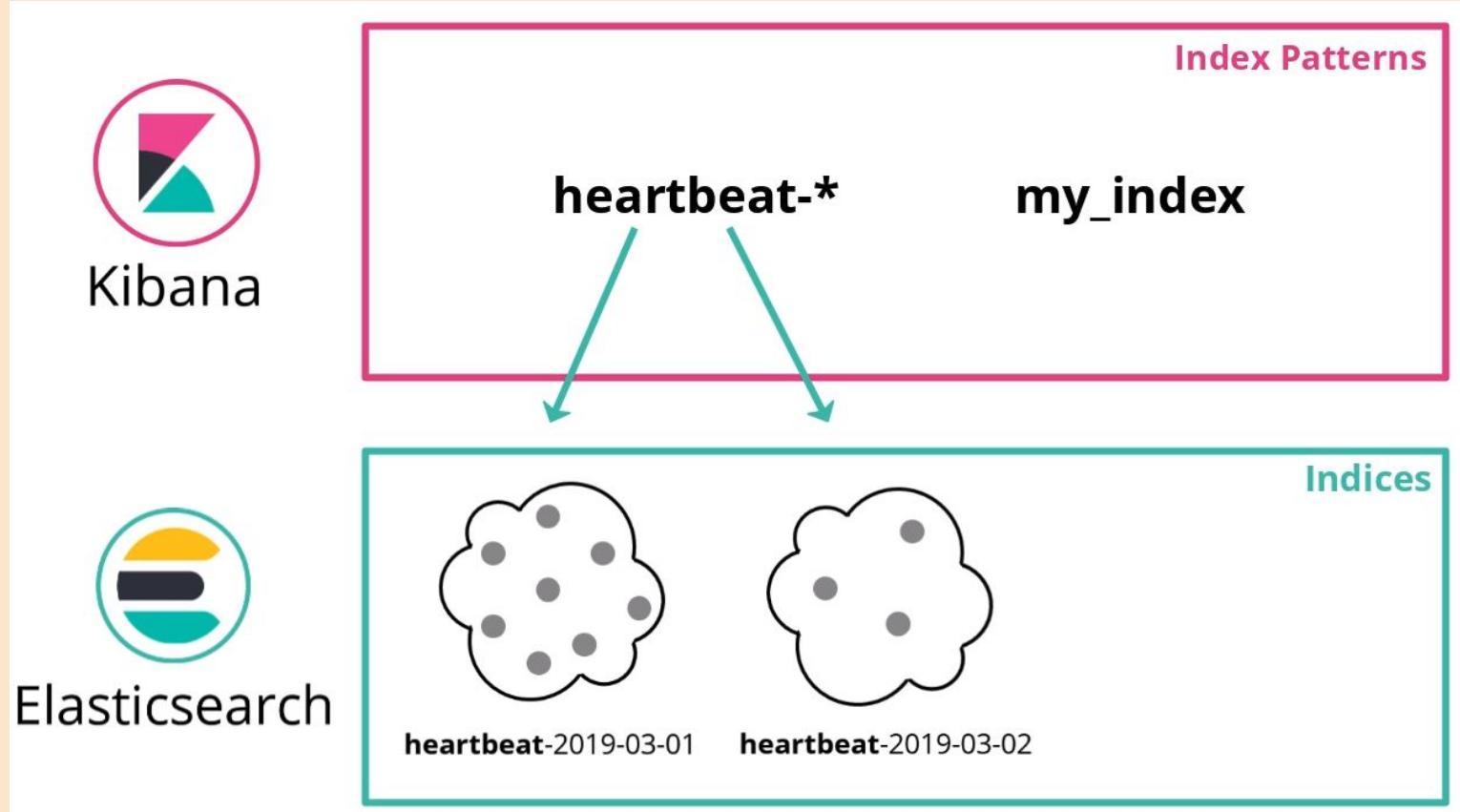
Create index pattern

Análisis con FileBeat

```
GET filebeat*/_search
```

```
{  
    "took" : 0,  
    "timed_out" : false,  
    "_shards" : {  
        "total" : 1,  
        "successful" : 1,  
        "skipped" : 0,  
        "failed" : 0  
    },  
    "hits" : {  
        "total" : {  
            "value" : 377,  
            "relation" : "eq"  
        },  
        "max_score" : 1.0,  
        "hits" : [  
            {  
                "_index" :  
                    "filebeat-7.3.1-2019.10.18-000001",  
                "_type" : "_doc",  
                "_id" :  
                    "mnVV320BHYYifCddW4a",  
                "_score" : 1.0,  
                "_source" : {  
                    "agent" : {  
                        "hostname" : ...  
                    }  
                }  
            }  
        ]  
    }  
}
```

HeartBeat



HeartBeat

heartbeat.monitors:

- **type: http**

- enabled: true**

- urls:**

- ["http://localhost:9200","http://localhost:8080","http://no_existe"]**

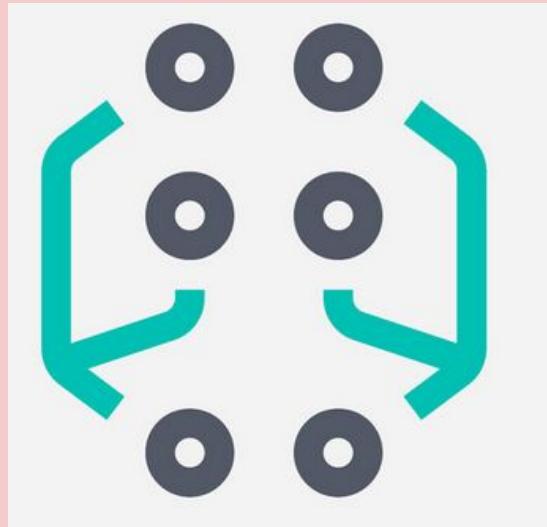
- schedule: '@every 10s'**

output.elasticsearch:

- hosts: ["localhost:9200"]**

INTRODUCCIÓN A MACHINE LEARNING

- ¿Qué es ML?
- Fundamentos esenciales de ML
- Usando Kibana para ML
- Limitaciones según la plataforma
- Limitaciones según la configuración
- Limitaciones operativas
- Limitaciones de Kibana
- Fundamentos de la API ML

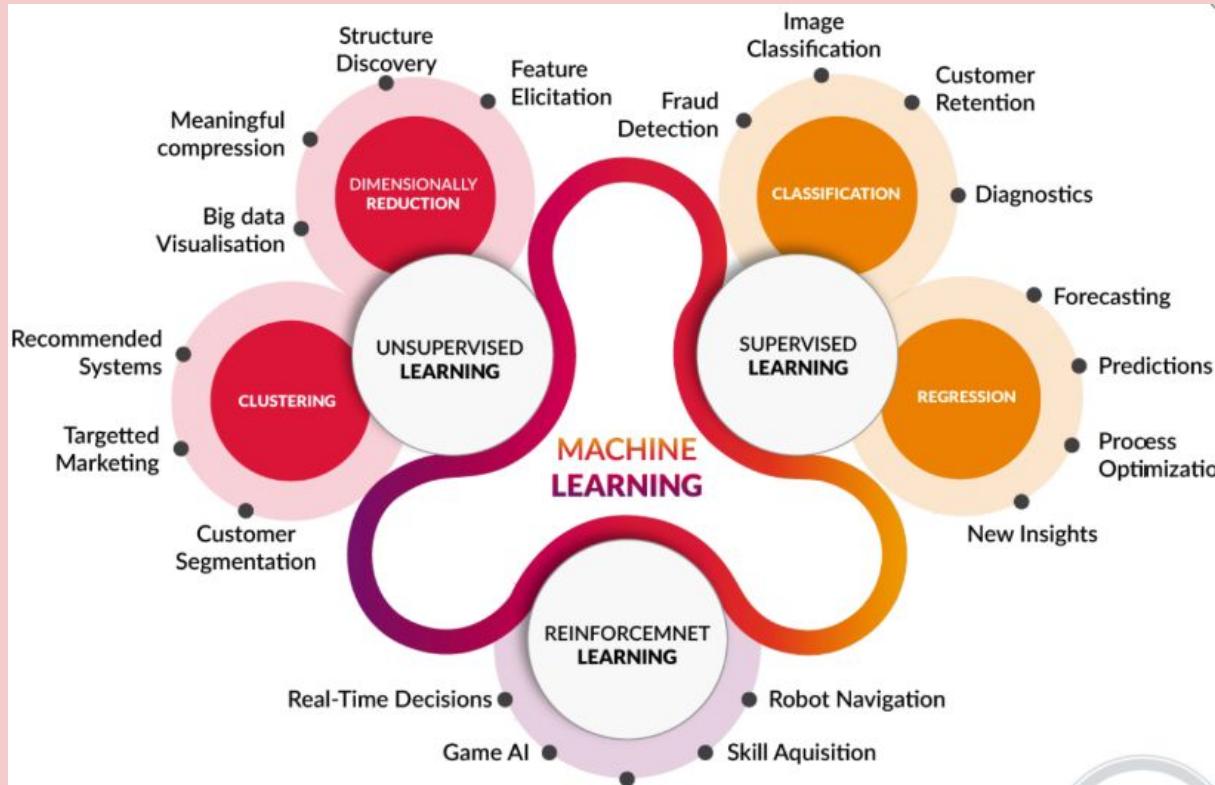


Fundamentos esenciales de ML

La **inteligencia artificial** es un término utilizado para describir un sistema que percibe su entorno y toma medidas para maximizar las posibilidades de lograr sus objetivos.

El **aprendizaje automático** es un conjunto de técnicas que permiten a las computadoras realizar tareas sin ser programadas explícitamente. Los sistemas de ML generalizan a partir de datos pasados para hacer predicciones sobre datos futuros.

Fundamentos esenciales de ML



Fundamentos esenciales de ML

El **aprendizaje supervisado** se centra en modelos que predicen las probabilidades de nuevos eventos en función de las probabilidades de eventos observados previamente. Por ejemplo: **determinar si un archivo es malware o no.**

Los modelos de **aprendizaje no supervisado** intentan encontrar patrones en datos no etiquetados. Por ejemplo : **determinar cuántas familias de malware existen en el conjunto de datos y qué archivos pertenecen a cada familia.**

Aprendizaje Supervisado

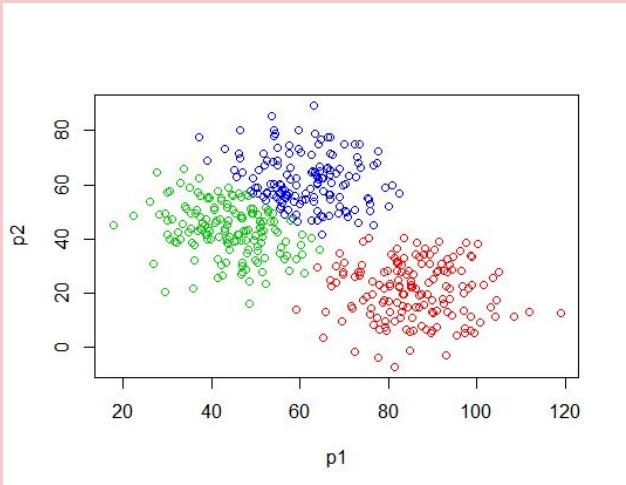
Clasificación: Los algoritmos de clasificación predicen a qué categoría pertenece una entrada en función de las probabilidades aprendidas de las entradas observadas previamente. **Por ejemplo: determinar si un archivo es malware o no.**

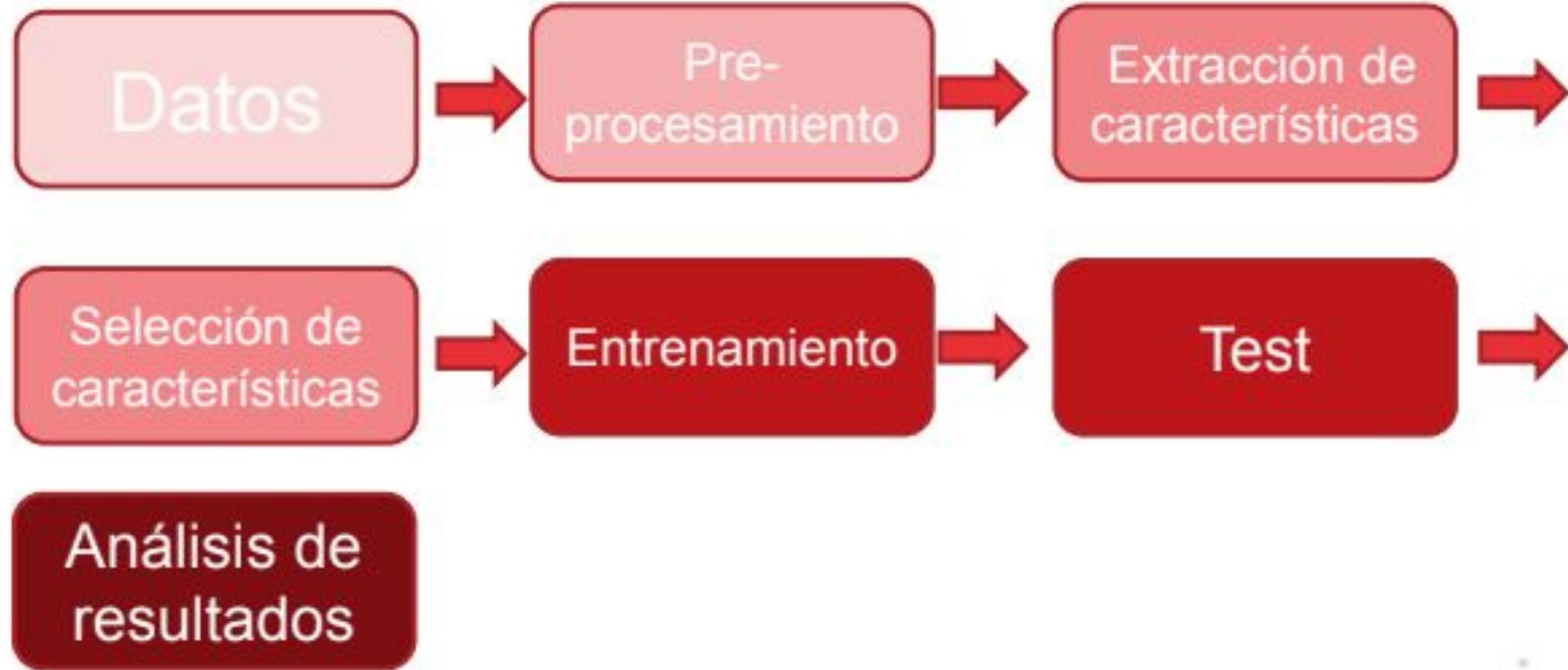
Regresión: los modelos de regresión (lineal, logística) predicen un valor de salida continuo para una entrada determinada en función de los valores de salida asociados con las entradas anteriores. **Por ejemplo: predecir cuántas muestras de malware se detectarán el próximo mes.**

Aprendizaje No Supervisado

Clustering: Consiste en agrupar un conjunto de objetos de tal manera que los objetos en el mismo grupo(cluster) sean más similares entre sí que con los de otros grupos

Detección de anomalías





- **Recopilar** muestras de datos de ambas clasificaciones para entrenar el modelo de aprendizaje automático.
- **Extraer** características de cada ejemplo de entrenamiento para representar el ejemplo numéricamente.
- **Entrenar** al sistema de aprendizaje automático para identificar elementos que sigan un patrón específico.
- **Probar** el sistema con datos que no se utilizaron durante el entrenamiento para evaluar su precisión o accuracy.

- **Conjunto de datos (data set):** El total del conjunto de datos sobre los que queremos desarrollar un algoritmo de ML con el fin de obtener un modelo que lo represente lo mejor posible. Contendrá variables independientes y dependientes
- **Variables independientes (features):** Aquellas columnas del data set que serán usadas por el algoritmo para generar un modelo que prediga lo mejor posible las variables dependientes
- **Variables dependientes (labels):** Columna del data set que responde a una correlación de las variables independientes y que debe ser predicha por el futuro modelo
- **Conjunto de datos de entrenamiento (training set):** Subconjunto del data set que será utilizado para entrenar el modelo que se pretende generar
- **Conjunto de datos de test (test set):** Subconjunto del data set que se le pasará al modelo una vez haya sido entrenado para comprobar, mediante el uso de diferentes métricas, su calidad

¿Supervisado o no supervisado?

Dos enfoques distintos para resolver distintos casos de uso

No Supervisado

- ¿Por qué este servidor está enviando muchos más datos ahora?
- ¿Qué es este nuevo proceso?
- ¿Por qué son de mayor tamaño que antes estas peticiones al DNS?
- ¿Tengo sistemas comprometidos, con algún malware?
- ¿Qué usuarios podrían ser una amenaza de seguridad interna?
- ¿Qué incidentes en el tráfico de red están causando el mayor retraso?

Supervisado

- ¿Qué transacciones son fraudulentas?
- ¿Qué productos deberíamos recomendar a los clientes?
- ¿Qué clientes es probable que abandonen?
- ¿Cómo puedo clasificar la actividad según el tipo de terminal de origen (ej. smartphone)?
- ¿En qué idioma está escrito un documento?
- ¿Qué resultados de búsqueda son más relevantes basándonos en el ratio de click-through?

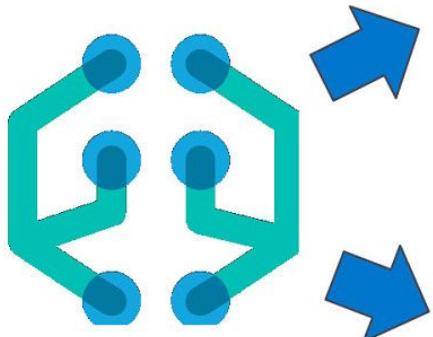
Casos de Uso en los que **construir** nuestro propio **Modelo**

¿De qué comportamiento podemos aprender para hacer predicciones?

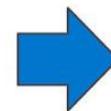
- Machine Learning en **Search**
 - ¿En qué idioma está escrito un documento?
 - ¿Cómo incrementar la *relevancia* en las búsquedas para entidades nombradas?
 - ¿Qué resultados son más relevantes basados en el ratio de click-through?
- **Seguridad**
 - ¿Cómo identificar nombres de dominio maliciosos generados por DGAs?
 - ¿Cómo clasificar la actividad según el tipo de terminal de origen (ej. smartphone)?
- **Observabilidad**
 - ¿Que usuarios o servidores son atípicos?
 - ¿Cómo clasificar alertas y enrutarlas al equipo adecuado?
 - ¿Qué clientes es probable que nos abandonen?

Expandiendo casos de uso

Machine Learning Supervisado extremo a extremo con la versión 7.6

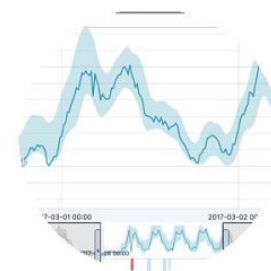


- *Detección Anomalías*
- *Detección valores atípicos*
- *Forecasting*

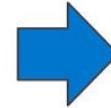


Data Driven:
Reconocimiento
de patrones

No Supervisado



- *Identificación idioma*
- *Detección fraude*
- *Clasificación usuarios*



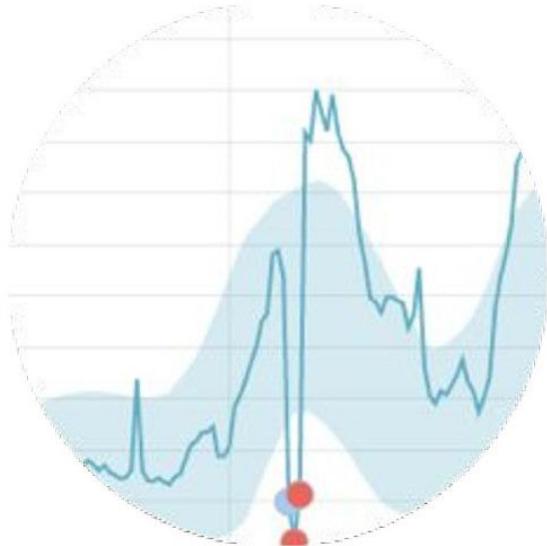
Datos
etiquetados para
aprendizaje y
predicción



Supervisado

Machine Learning en el Stack Elastic

Detección de anomalías en series temporales



Analítica con Data Frames

user_count	products.quantity.sum	products.
348	878	3078.
188	506	17829.9.
170	433	14911.722
154	301	10702.8789
158	409	14159.821
76	158	6999.76953
106	211	6895.21481
57	114	3344.9336
148	287	9453.8
84	167	5741.6
114	307	1072.1
109	292	1
21	276	

Jobs de ML **pre-configurados** para campos basados en **ECS**

Simplifica la configuración de jobs en formatos de datos habituales

<https://github.com/elastic/ecs>

The screenshot shows the Kibana interface with three tabs open:

- Create a job from the index pattern filebeat-***: This tab displays a summary of the job configuration. It includes a section for "Use a supplied configuration" which highlights "Filebeat NGINX" as a matching known configuration. A red box highlights this section.
- New job from index pattern filebeat-***: This tab shows the detailed configuration for the job. It lists "Job settings" like "job ID prefix" and "job group". Under "Job details", there are sections for "remote_ip_request_rate", "remote_ip_url_count", "visitor_rate", "response_code", "new_request_rate", and "use_full_filebeat_*_data". A "Create job" button is at the bottom.
- Job Management - Anomaly Explorer**: This tab displays a table of active jobs. The table has columns for "Job ID", "Description", "Processed records", "Memory status", "Job state", "Datafeed state", "Latest timestamp", and "Actions". It lists several jobs related to "NGINX Access Logs" such as "Detect low request rate", "Detect unusual remote_ip", "Detect unusual url_count", "Detect unusual response_code", and "Detect unusual visitor_rate". The table shows 3,743 processed records, 1,592,348 closed jobs, and 1,327,894 stopped jobs.

- <https://www.elastic.co/es/blog/introducing-the-elastic-common-schema>

```
10.42.42.42 - - [07/Dec/2018:11:05:07 +0100] "GET /blog HTTP/1.1" 200 2571 "-"  
"Mozilla/5.0 (Macintosh; Intel Mac OS X 10_14_0) AppleWebKit/537.36 (KHTML, like Gecko)  
Chrome/70.0.3538.102 Safari/537.36"
```

event.original	10.42.42.42 - - [07/Dec ...	Log completo y sin modificaciones para auditorías
http.request.method	get	
http.response.body.bytes	2571	
http.response.status_code	200	
http.version	1.1	
host.hostname	webserver-blog-prod	
message	"GET /blog HTTP/1.1" 200 2571	Representación de texto de la información significativa del evento para una visualización precisa en un visualizador de logs



Analítica con Data Frames

Outlier Detection y ML Supervisado

Analítica con Data Frames

- Permite entrenar un modelo y evaluarlo
 - **Outlier Detection** / Detección de valores atípicos (*No supervisado*)
 - **Análisis de Regresión** (*Supervisado*)
 - **Análisis de Clasificación** (*Supervisado*)
 - **Inferencia** - Predicción sobre datos no etiquetados

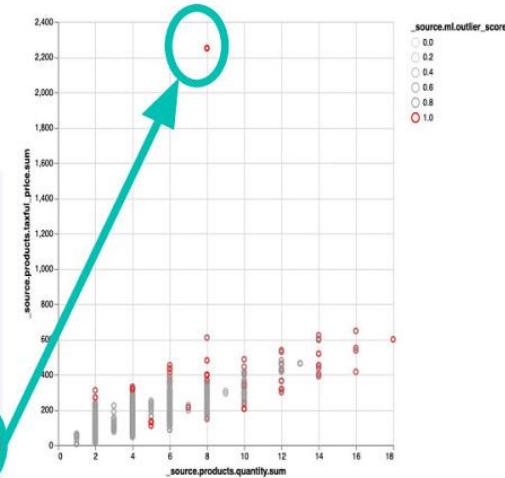


Detección de **valores atípicos**

¿Cómo los identifica Elasticsearch?

- El objetivo es identificar **puntos** de datos que **no siguen el modelo**
- Métodos utilizados basados en
 - distancia
 - densidad
- Salidas
 - **Outlier score**
 - **Feature influencer score**
- Casos de uso
 - Detección de fraude
 - Problemas médicos
 - Detección de amenazas en seguridad

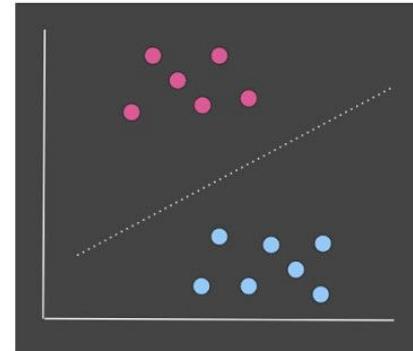
```
"customer_full_name": {  
    "keyword": "Nagdi Shaw"  
},  
"ml_id_copy": "Vyu9e08pKNaST-9TLV9p3k0AAAAAAA",  
"products": {  
    "taxful_price": {  
        "sum": 2250.0  
    },  
    "quantity": {  
        "sum": 8.0  
    }  
}  
}  
  
ml : {  
    "outlier_score" : 0.9848338961601257,  
    "feature_influence.products.quantity.sum" : 0.007586637046188116,  
    "feature_influence.products.taxful_price.sum" : 0.992413341999054
```



Modelos de **Clasificación**

¿Cómo funcionan?

- El objetivo es predecir la **categoría/clase** de un punto dentro del conjunto de datos.
- Método utilizado: **boosted tree regression**
- Dos tipos: **binaria, multi-clase**
- Casos de uso
 - Detección de cáncer
 - Clasificar música o texto
 - Predicción de riesgo en préstamos



Modelos de Clasificación

Requerimientos

- Supervisado implica que alguien tiene que **etiquetar** los datos para aprender
- Requiere
 - **Feature variables** / características
 - **Dependent variable** / variable dependiente

Características

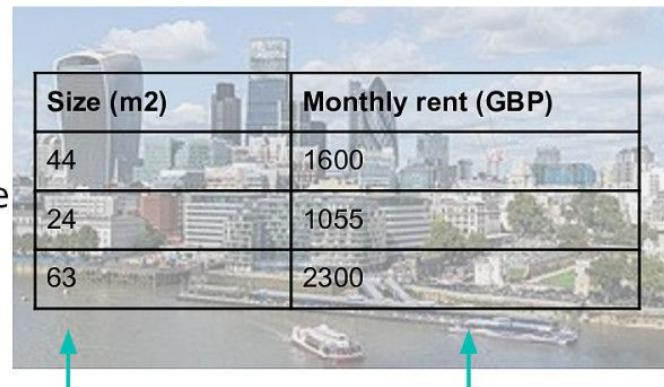
Variable dependiente

	customer a	customer b
total duration of customer sessions in last month	80:21:07	1:01:11
tv episodes watched in last month	24	1
films watched in last month	5	0
newness of titles watched in last month	9.8	1.2
change in duration of customer sessions this month	6:22:17	16:43:29
customer subscription plan	gold	platinum
customer tenure	32	26
has churned?	no	yes

Modelos de **Regresión**

¿Cómo funcionan?

- El objetivo es estimar la **relación** entre varios campos dentro del conjunto de datos.
- Método utilizado: **extreme gradient boost**
- Mismos requerimientos que clasificación
 - **Feature variables** / características
 - **Dependent variable** / variable dependiente
- Casos de uso
 - Precio del alquiler/compra
 - Tiempo que se retrasará un vuelo



Size (m ²)	Monthly rent (GBP)
44	1600
24	1055
63	2300

Característica

Variable dependiente

Validación del Modelo

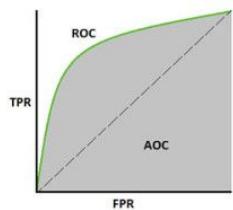
Medir la calidad del modelo con datos de prueba

Clasificación

Matriz de confusión

		Predicted label	
		0	1
Actual label	0	98%	2%
	1	9%	91%

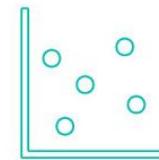
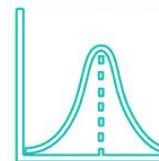
Curva ROC



Importancia
de las **feature**
variables

Regresión

Error Medio cuadrado



R-cuadrado

Predicción mediante Inferencia

Aplicar un modelo a nuevos datos en streaming

Entrenamiento/Test/Validación



```
PUT _ml/data_frame/analytics/churn
{
  "source": {
    "index": "customer_behaviour"
  },
  "dest": {
    "index": "customer_behaviour_churn"
  },
  "analysis": {
    "regression": {
      "dependent_variable": "churn_probability",
      "training_percent": 80,
      "save_model": {
        "name": "churn"
      }
    }
  }
}
POST _ml/data_frame/analytics/churn/_start
```

Inferencia del modelo



```
PUT _ingest/pipeline/predict_churn
{
  "description" : "Predict customer churn",
  "processors" : [
    {
      "inference" : {
        "model": {
          "regression": {
            "model_id": "churn",
            "target_field": "churn_probability"
          }
        }
      }
    }
  ]
}

PUT _ingest/pipeline/lang_ident
{
  "description" : "Identify language",
  "processors" : [
    {
      "inference" : {
        "model": {
          "lang_ident": {
            "target_field": "text",
            "target_language_field": "lang",
            "target_probability_field": "lang_prob"
          }
        }
      }
    }
  ]
}
```

Machine Learning Supervisado - Regresión



Machine Learning / Data Frame Analytics

Create job



Switch to json editor

1 Configuration



Outlier detection

Outlier detection identifies unusual data points in the data set.

Select



Regression

Regression predicts numerical values in the data set.

✓ Selected



Classification

Classification predicts classes of data points in the data set.

Select

Machine Learning Supervisado - Regresión

Dependent variable

FlightDelayMin



Included fields

22 fields included in the analysis

Search...

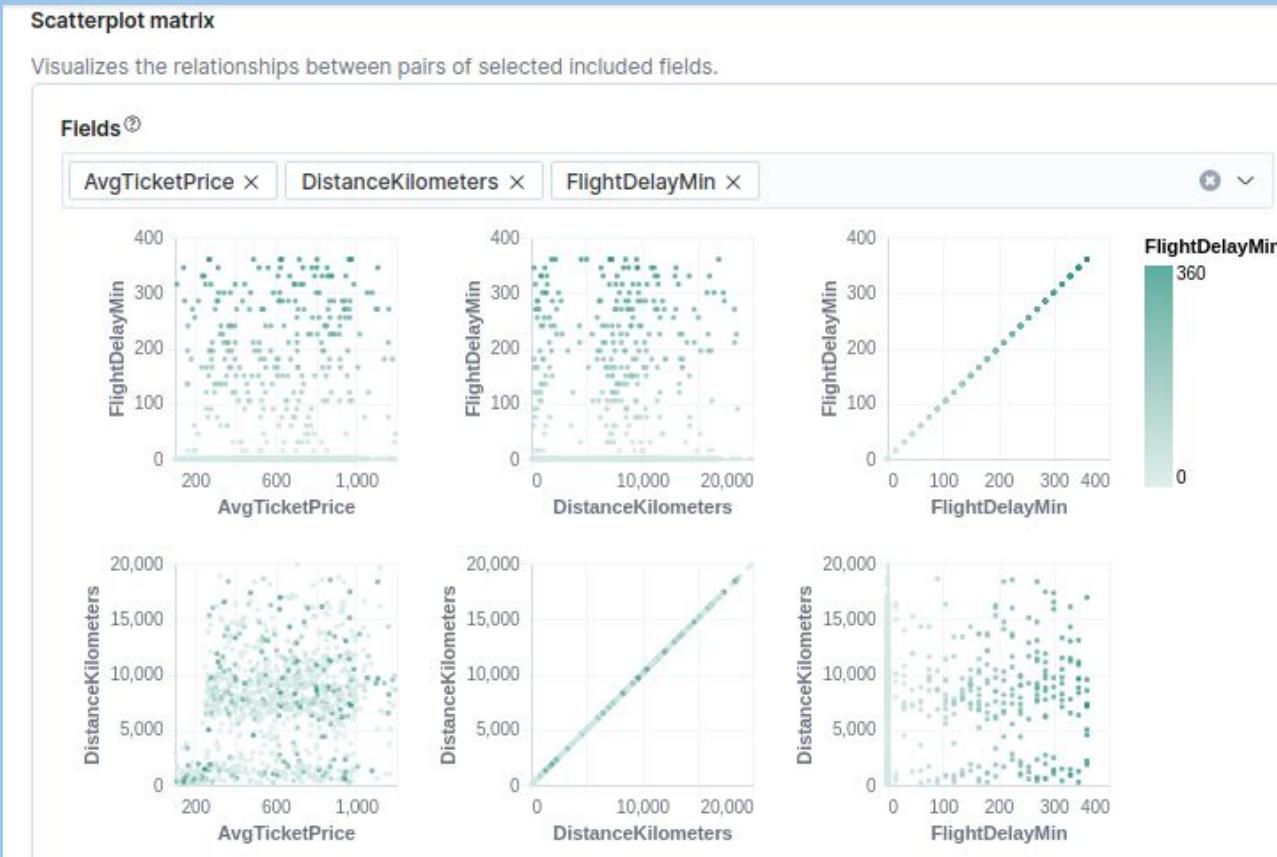
Is included Is not included

<input type="checkbox"/>	Field name	Mapping	Is included	Is required	Reason
<input checked="" type="checkbox"/>	DistanceKilometers	float	Yes	No	
<input checked="" type="checkbox"/>	DistanceMiles	float	Yes	No	
<input type="checkbox"/>	FlightDelay	boolean	No	No	field not in includes list
<input checked="" type="checkbox"/>	FlightDelayMin	integer	Yes	Yes	
<input type="checkbox"/>	FlightDelayType	keyword	No	No	field not in includes list

Machine Learning Supervisado - Regresión

Scatterplot matrix

Visualizes the relationships between pairs of selected included fields.



Machine Learning Supervisado - Regresión

2

Additional options

Advanced configuration

Feature importance values

5

Specify the maximum number of feature importance values per document to return.

Top classes

All classes

The number of categories for which the predicted probabilities are reported. If you have a large number of classes there could be a significant effect on the size of your destination index.

Prediction field name

Defines the name of the prediction field in the results. Defaults to <dependent_variable>_prediction.

Randomize seed

The seed for the random generator used to pick training data.

Model memory limit

 533mb

Use estimated model memory limit

The approximate maximum amount of memory resources that are permitted for analytical processing.

Maximum number of threads

1

The maximum number of threads to be used by the analysis. The default value is 1.

Machine Learning Supervisado - Regresión

Histogram charts 24 columns hidden Sort fields

ml.is_training	ml.FlightDelayMin_pr...	FlightDelayMin	AvgTicketPrice	Cancelled	Carrier
true	3.056	0	761.208	false	ES-Air
true	213.322	285	362.625	false	JetBeats
true	54.12	0	399.309	false	ES-Air
true	2.623	0	153.454	false	Kibana Airlines
true	117.459	105	317.185	false	ES-Air
true	5.271	0	493.243	false	Logstash Airways
true	111.118	105	418.23	true	Logstash Airways
true	72.169	0	785.814	false	JetBeats
true	87.319	195	819.186	false	Kibana Airlines
true	12.984	0	955.818	false	Logstash Airways
true	1.571	0	246.958	true	Kibana Airlines

Machine Learning Supervisado - Regresión

Model evaluation ^				 Regression evaluation docs
Job status				stopped
Generalization error				
2,492 docs evaluated		3610	0.594	Training error
Mean squared error 		R squared 	9,964 docs evaluated	Mean squared error 
NaN		35.7	0.664	R squared 
Mean squared logarithmic error 		Pseudo Huber loss function 	Mean squared logarithmic error 	Pseudo Huber loss function 

Machine Learning Supervisado - Clasificación

flights_classification ok kibana_sa... flights_cla... classification started Phase 5/8    

[Job details](#) [Job stats](#) [JSON](#) [Job messages](#)

State

id	flights_classification
state	started
data_size	{"training_docs_count":10447,"test_docs_count":0,"skipped_docs_count":0}
mem	{} {"timestamp":1627992823294,"peak_usage_bytes":17531040,"status":"ok"}
node	{"id":"AV_hbL_3SNCWJInWXAdTQA","name":"elasticsearch","ephemeral_id":"7_JvRDYTSDiAkoopSR3M1w","transport_address":"127.0.0.1:9300","attributes":{"ml.machine_memory":2147483648,"xpack.installed":true,"transform.node":true,"ml.max_open_jobs":512,"ml.max_jvm_size":1073741824}}
assigned_shards	0

Progress

Phase	Progress
reindexing	100%
loading_data	100%
feature_selection	100%
coarse_parameter_search	100%
fine_tuning_parameters	32%
final_training	0%
writing_results	0%
inference	0%

Machine Learning Supervisado - Clasificación

Histogram charts 27 columns hidden Sort fields

Queries run to fetch histogram chart data will use a sample size per shard of 5000 documents.

mi.is_training	mi.FlightDelay_prediction	FlightDelay	0.28 - 1 ml.prediction_probability	0 documents contain field. ml.feature_importance	0.28 - 1 ml.prediction_score	Chart not supported. ml.top_classes	100.15 - 119 AvgTicketPr
true	true	false		0.402 [{"feature_name": ["Dest"], ...}		0.402 [{"class_score": [0.401576...}	
true	false	false		0.727 [{"feature_name": ["AvgTic...]		0.279 [{"class_score": [0.279079...}	
true	true	false		0.338 [{"feature_name": ["AvgTic...]		0.338 [{"class_score": [0.338370...}	
true	false	false		0.966 [{"feature_name": ["Dest"], ...}		0.371 [{"class_score": [0.370994...}	
true	false	false		0.963 [{"feature_name": ["AvgTic...]		0.37 [{"class_score": [0.369760...}	
true	false	false		0.929 [{"feature_name": ["AvgTic...]		0.357 [{"class_score": [0.356730...}	
true	false	false		0.783 [{"feature_name": ["AvgTic...]		0.301 [{"class_score": [0.300654...}	
true	false	false		0.947 [{"feature_name": ["AvgTic...]		0.364 [{"class_score": [0.363603...}	
true	false	false		0.976 [{"feature_name": ["AvgTic...]		0.375 [{"class_score": [0.374901...}	
true	true	true		0.99 [{"feature_name": ["Dest"], ...}		0.99 [{"class_score": [0.989603...}	
true	false	false	0.771 [{"feature_name": ["Dest"]}]		0.296 [{"class_score": [0.296132...}		

Machine Learning Supervisado - Clasificación

Normalized confusion matrix for testing dataset [?](#)

		Predicted class	
		false	true
Actual class	false	81%	19%
	true	24%	76%

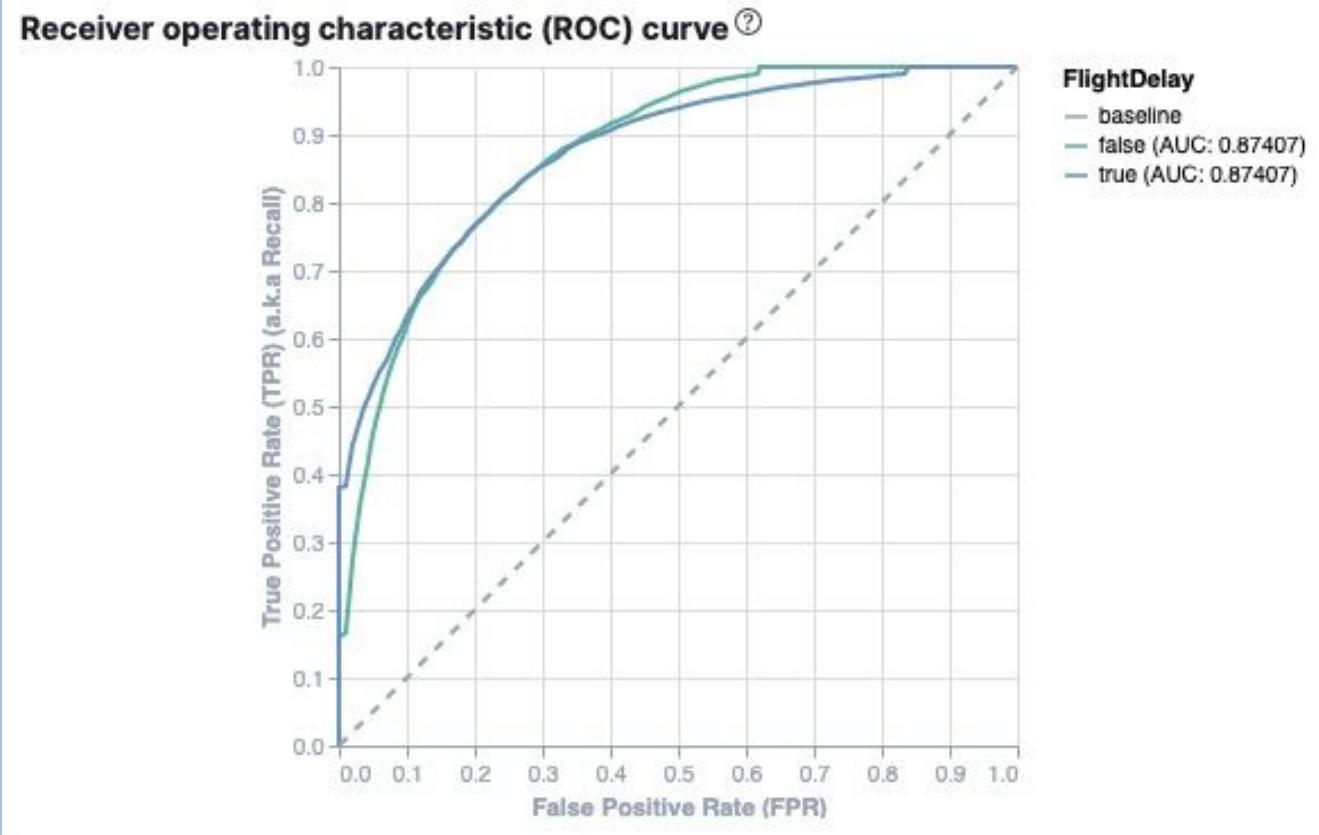
0.798

Overall accuracy [?](#)

0.786

Mean recall [?](#)

Machine Learning Supervisado - Clasificación



Machine Learning Supervisado - Identificar idioma

lang_ident_model_1 Model used for identifying language from arbitrary input text. classification built-in 2019-12-05 13:28:34

Details Config Stats

Details

description	Model used for identifying language from arbitrary input text.
tags	lang_ident prepackaged
version	7.6.0
estimated_operations	39629
estimated_heap_memory_usage_bytes	1053992
license_level	basic

Machine Learning Inferencia de modelos

GET _ml/inference

**GET
_ml/inference/?filter_path=trained_model_configs.model_id**

Limitaciones plataforma

- Sólo funciona en máquinas cuyas CPU son compatibles con SSE4.2
- Las mejoras a nivel de rendimiento de la CPU se aplican solo a Linux y MacOS

Limitaciones configuración

- El tamaño de las agregaciones de términos afecta el análisis de datos
- No puede utilizar los siguientes nombres de campo en las propiedades by_field_name o over_field_name en un job [by; count; over]
- Puede generar sólo tres forecasts por trabajo de detección de anomalías al mismo tiempo.

Limitaciones operativa

- Los datos que envíe al job deben usar el formato JSON.
- La API de obtención de jobs y la API de obtención de estadísticas pueden trabajar con un máximo de 10.000 jobs

CASO PRÁCTICO: ML PARA DETECCIÓN DE ANOMALÍAS

- Introducción al proyecto
- Indexación en ElasticSearch de valores temporales
- Estableciendo el algoritmo de detección de anomalías
- Ejecución y pruebas

Detección de Anomalías en Series Temporales

Conceptos sobre Anomalías

Detectando anomalías en los datos

Aprendizaje no supervisado

- Aprendizaje sin ejemplos etiquetados por humanos
- Basarse sólo en los datos

Detección de Anomalías

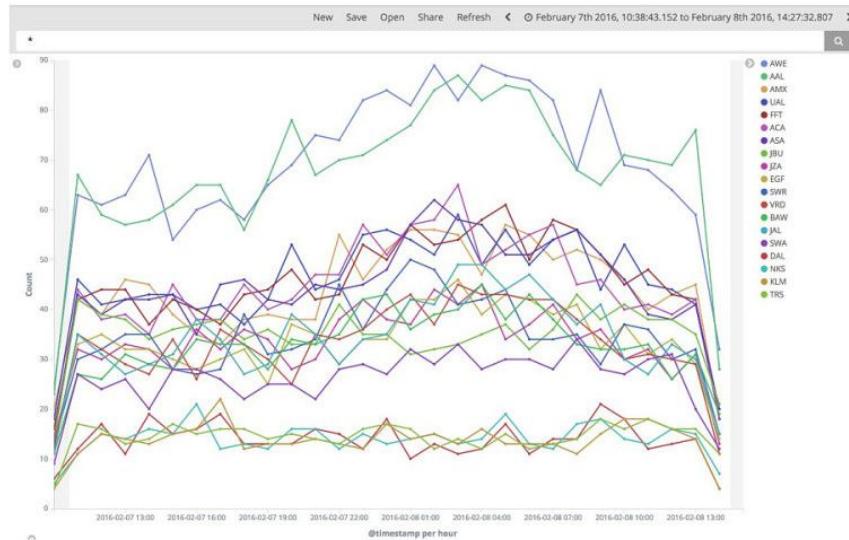
- Descubrir lo que es extraño o diferente, no necesariamente malo.

Bayesiano

- Determinar las probabilidades posteriores en función de las probabilidades anteriores y la nueva información. Cuando se recopile información adicional.

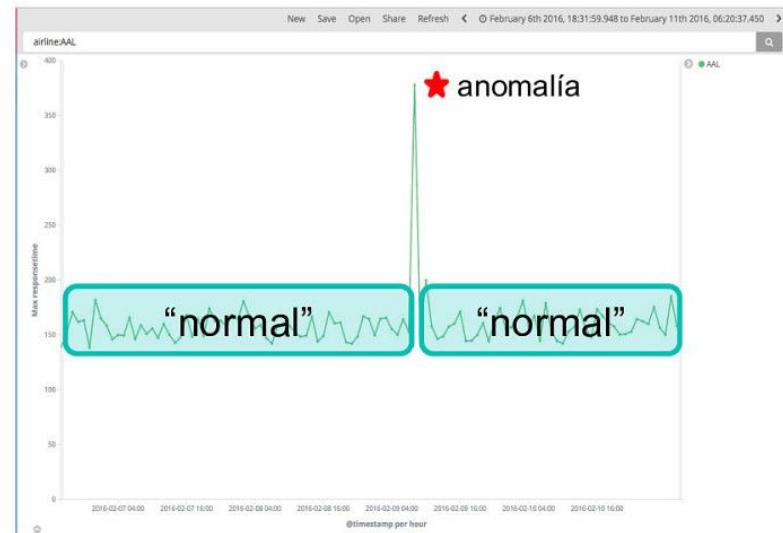
¿Cómo definimos “Normal”?

1. Algo que se comporta de manera **consistente** con **respecto a sí mismo**, en el **tiempo**
2. Algo se comporta de manera **consistente** en **comparación** con **entidades similares**



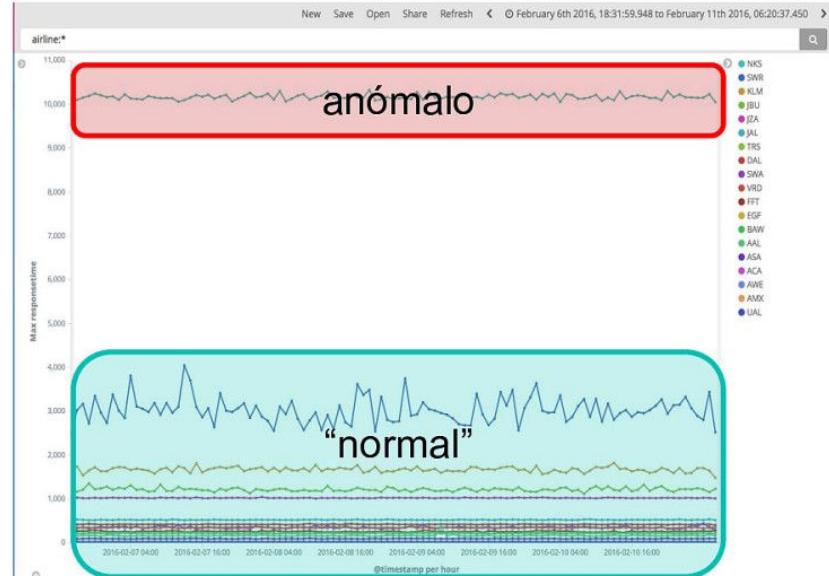
¿Qué es “Anormal”?

1. Si algo cambia su comportamiento, en comparación con su propia historia, ese cambio es **anómalo**



¿Qué es “Anormal”?

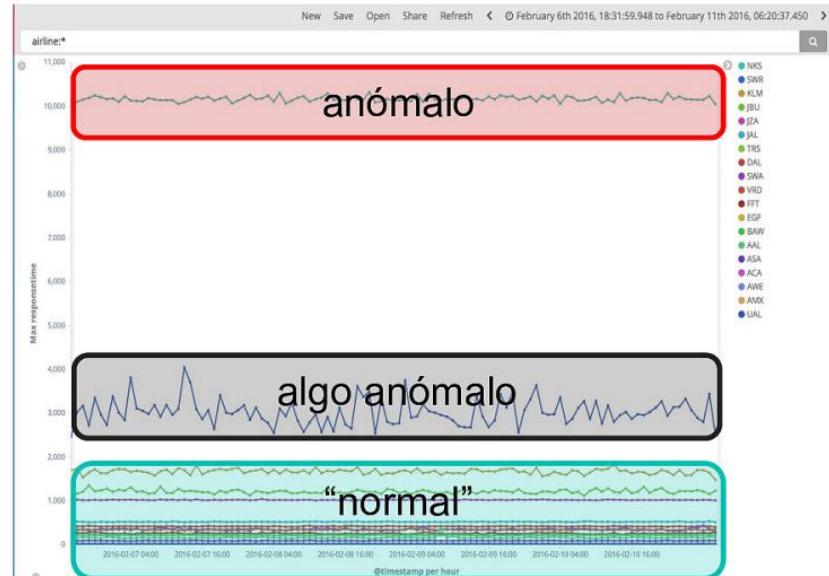
2. Si algo es drásticamente diferente que otros dentro de una población, entonces esa entidad es **anómala**



¿Qué es “**Anormal**”?

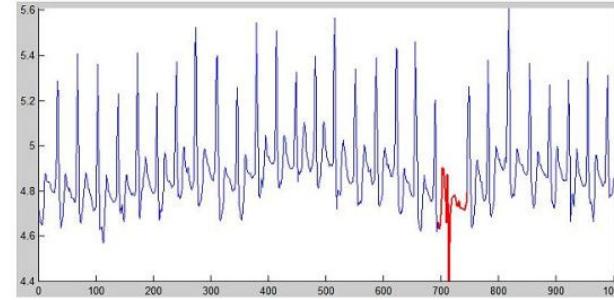
2. Si algo es drásticamente diferente que otros dentro de una población, entonces esa entidad es **anómala**

También existe el concepto de ser "**algo anómalo**"



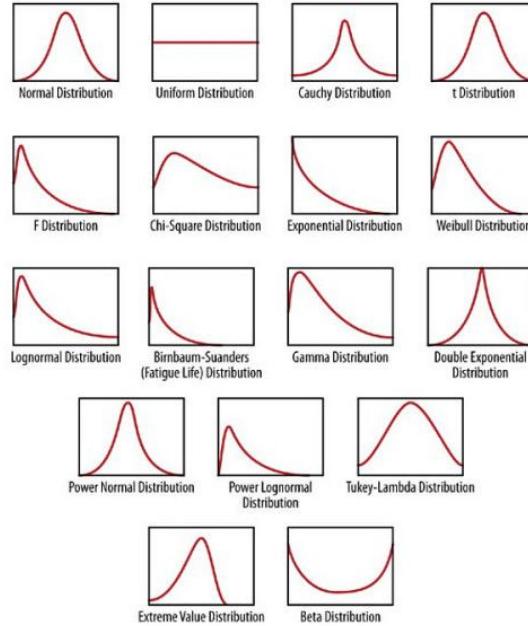
¿Cómo definimos “Anormal”?

1. Cuando el **comportamiento** de una entidad **cambia** de manera *significativa* y *repentina*
2. Cuando una entidad es **drásticamente diferente** de otras dentro de una población



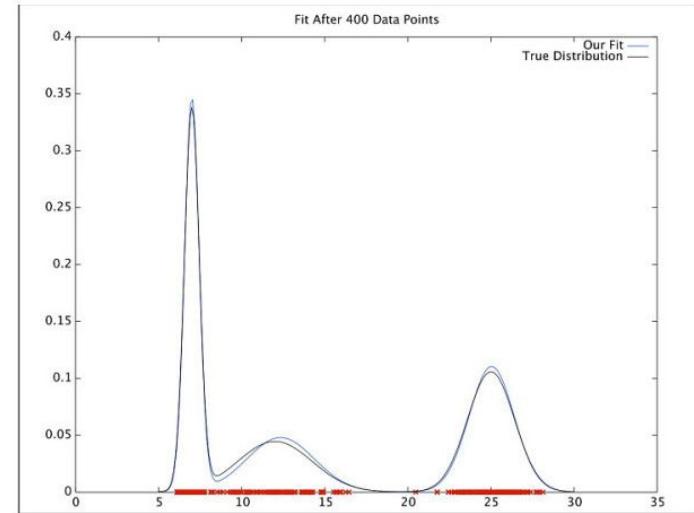
¿Cómo se escoge el “**Modelo**” que mejor se adapta a nuestros datos?

2



Machine Learning lo escoge por nosotros, **no supervisado**

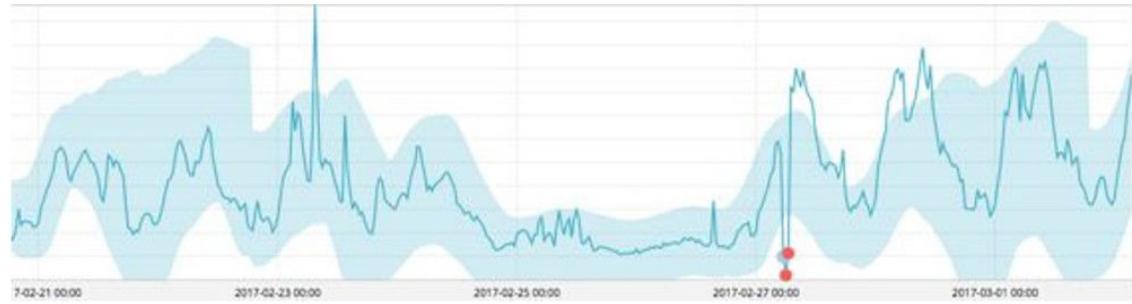
- Utiliza técnicas sofisticadas de machine-learning para ajustar mejor el modelo estadístico adecuado para nuestros datos
- Mejor modelo = mejor detección valores atípicos = menos falsas alarmas
- Las anomalías ocurren cuando estamos en zonas de baja probabilidad de esos valores



Los modelos deben tener en cuenta la **periodicidad**



Mejorar y facilitar Alertas



Predicción / Forecasting



Jobs para detección de anomalías

- **Single-metric jobs:** el análisis de datos se realiza en un solo campo de índice.
- **Multi-metric jobs:** el análisis de datos se puede realizar en varios campos de índice; sin embargo, cada campo se analiza por separado.
- **Advanced jobs:** el análisis de datos se puede realizar en varios campos de índice. Los trabajos avanzados proporcionan ajustes de configuración completos para detectores.
- **Population jobs:** análisis de datos del comportamiento de distribución para datos menos comunes, como la detección de valores atípicos en un conjunto de datos.

Single-metric jobs

Create job: Single metric

Using index pattern kibana_sample_data_logs

1
Time range

2
Pick fields

3
Job details

4
Validation

5
Summary

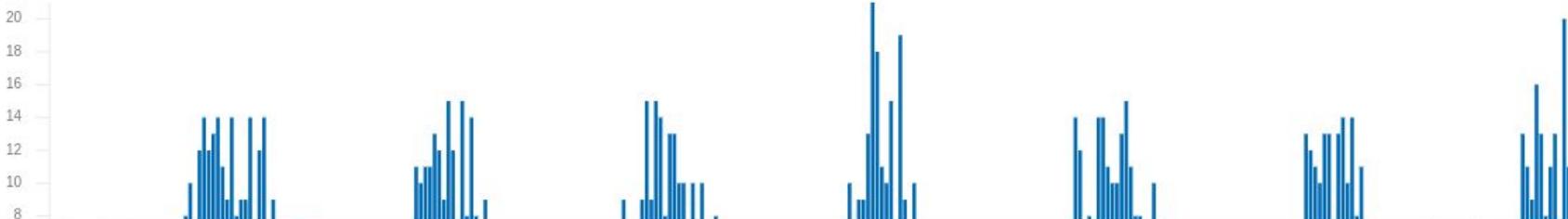
Time range

Aug 31, 2021 @ 19:28:29.527



Sep 7, 2021 @ 19:28:29.527

Use full kibana_sample_data_logs data



API ML

#obtener jobs

GET _ml/anomaly_detectors/_all

GET _ml/anomaly_detectors/<job_id>

#obtener estadísticas

GET _ml/anomaly_detectors/<job_id>/_stats

#obtener buckets

GET _ml/anomaly_detectors/<job_id>/results/buckets/

API ML

#abrir un job

POST _ml/anomaly_detectors/<job_id>/_open

#cerrar un job

POST _ml/anomaly_detectors/<job_id>/_close

#borrar un job

DELETE _ml/anomaly_detectors/<job_id>

API ML

#obtener buckets

```
GET _ml/anomaly_detectors/<job_id>/results/overall_buckets
```

#resultados del job

```
GET _ml/anomaly_detectors/<job_id>/results/records
```

```
GET _ml/anomaly_detectors/<job_id>/results/influencers
```

```
GET _ml/anomaly_detectors/<job_id>/model_snapshots/
```

```
GET _ml/anomaly_detectors/<job_id>/model_snapshots/<snapshot_id>
```

API ML

```
PUT _xpack/ml/anomaly_detectors/suspicious_login_activity
{
  "job_id": "suspicious_login_activity",
  "job_type": "anomaly_detector",
  "job_version": "6.3.0",
  "description": "suspicious login activity",
  "established_model_memory": 69458,
  "analysis_config": {
    "bucket_span": "5m",
    "detectors": [
      {
        "detector_description": "high_count",
        "function": "high_count",
        "partition_field_name": "system.auth.hostname",
        "detector_index": 0
      }
    ],
  }
}
```

Machine Learning no supervisado para detección de anomalías

Machine Learning / [Anomaly Detection](#) / Job Management

Overview [Anomaly Detection](#) Data Frame Analytics Data Visualizer Settings

Anomaly detection jobs

Refresh 30 seconds

Active ML nodes: 0 Total jobs: 0 Open jobs: 0 Closed jobs: 0 Active datafeeds: 0

[Create job](#)

Search... [Opened](#) [Closed](#) [Failed](#) [Started](#) [Stopped](#) [Group](#)

ID ↑	Description	Processed records	Memory status	Job state	Datafeed state	Latest timestamp
No jobs found						

Machine Learning no supervisado para detección de anomalías

☰ | D Machine Learning / Anomaly Detection / Create job

Create a job from the index pattern kibana_sample_data_logs

Use preconfigured jobs

The fields in your data match known configurations. Create a set of preconfigured jobs.



**Kibana sample
data web logs**
Find anomalies in
Kibana sample web
logs data.

Machine Learning no supervisado para detección de anomalías

New job from index pattern kibana_sample_data_logs

Job settings

Job ID prefix

The prefix is added to the beginning of each job ID.

Job ID prefix

demo_

Start datafeed after save

Use full kibana_sample_data_logs data

> Advanced

Create Jobs

Jobs

demo_low_request_rate 

Find unusually low request rates

kibana_sample_data

kibana_sample_web_logs

demo_response_code_rates 

Find unusual event rates by HTTP response code (high and low)

kibana_sample_data

kibana_sample_web_logs

demo_url_scanning 

Find client IPs accessing an unusually high distinct count of URLs

kibana_sample_data

kibana_sample_web_logs

Machine Learning no supervisado para detección de anomalías

Machine Learning / Anomaly Detection / Job Management

ID ↑	Description	Processed records	ory status	Job state	Datafeed state	Latest timestamp	W	E	...
<input type="checkbox"/> demo_low_request_rate	Find unusually low request rates kibana_sample_data kibana_sample_web_logs	1,216	ok	closed	stopped	2021-09-23 22:49:29			...
<input type="checkbox"/> demo_response_code_rates	Find unusual event rates by HTTP response code (high and low) kibana_sample_data kibana_sample_web_logs	14,073	ok	closed	stopped	2021-09-23 22:49:29			...
<input type="checkbox"/> demo_url_scanning	Find client IPs accessing an unusually high distinct count of URLs kibana_sample_data kibana_sample_web_logs	14,073	ok	closed	stopped	2021-09-23 22:49:29			...

Machine Learning no supervisado para detección de anomalías

Types of Time Series Anomaly Detection models

Single metric



High • Low • Count

Sum • Min • Max • Mean • Median

Multi metric



Machine Learning no supervisado para detección de anomalías



Machine Learning no supervisado para detección de anomalías

Anomaly timeline

View by

job ID

...

Overall

2021-07-25 2021-08-01 2021-08-08 2021-08-15 2021-08-22 2021-08-29 2021-09-05 2021-09-12 2021-09-19

Annotations

> 0 > 3 > 25 > 50 > 75

demo_response_code_rates

demo_url_scanning

demo_low_request_rate

2021-07-25 2021-08-01 2021-08-08 2021-08-15 2021-08-22 2021-08-29 2021-09-05 2021-09-12 2021-09-19

Rows per page: 10

< 1 >

Machine Learning no supervisado para detección de anomalías

Time	Severity	Detector	Found for	Influenced ...	Actual	Typical	Description	Job ID	Actions
> August 24th 2021, 14:00	● 95	Low request rates			0	28.4	↓ Unexpect ed zero value	demo_low_request_rate	
> August 24th 2021, 14:00	● 94	Event rate by response code	200	response.key word: 200	0	27.4	↓ Unexpect ed zero value	demo_response_code_rates	
> August 24th 2021, 13:00	● 3	Event rate by response code	200	response.key word: 200	0	22.6	↓ Unexpect ed zero value	demo_response_code_rates	
> August 24th 2021, 13:00	● 1	Low request rates			0	24.7	↓ Unexpect ed zero value	demo_low_request_rate	
> August 23rd 2021, 14:00	● < 1	High distinct count of URLs for a client IPs		clientip: 95.110.66.82	95.110.66.82	2	↑ 2x higher	demo_url_scanning	

Machine Learning no supervisado para detección de anomalías

Time	Severity	Detector	Found for	Influenced ...	Actual	Typical	Description	Job ID	Actions
September 4th 2021, 14:00	● 99	High distinct count of URLs for a client IPs	30.156.16.16 4	clientip: 30.156.16.16 4  	88	1	↑ 88x higher	demo_url_scanning	
September 4th 2021, 14:00	● 98	Event rate by response code	404	clientip: 30.156.16.16 4   response.key word: 404  	101	1.17	↑ 87x higher	demo_response_code_rates	

Predicción a futuro

X

Forecasting

Previous forecasts ②

Created	From	To	View
September 7th 2021, 20:23:47	September 7th 2021, 18:30:00	September 8th 2021, 18:30:00	W

Run a new forecast

Duration

Run

Length of forecast, up to a maximum of 3650 days. Use s for seconds, m for minutes, h for hours, d for days, w for weeks.

[Close](#)

Predicción a futuro

Single time series analysis of count

show model bounds annotations show forecast



3 soluciones con 1 stack



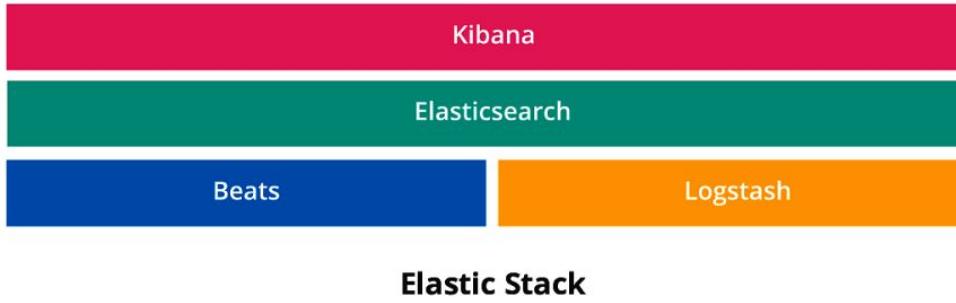
Elastic Enterprise Search



Elastic Observability



Elastic Security



ELASTIC OBSERVABILITY

- Introducción a Elastic Observability
- Centralización de Elastic Logs
- Logs de Web, Apps, bases de datos y contenedores
- Generación de métricas
- Métricas de contenedores, bases de datos, red y almacenamiento
- Rastreos de trazas con Elastic APM
- Monitorización de usuarios reales, transacciones y dependencias
- Medición de SLA con Elastic Uptime
- Medición de respuestas, corrección y validación de certificados
- Análisis detallados para la identificación de problema

Introducción a Elastic Observability

- Unificación de observaciones en único lugar

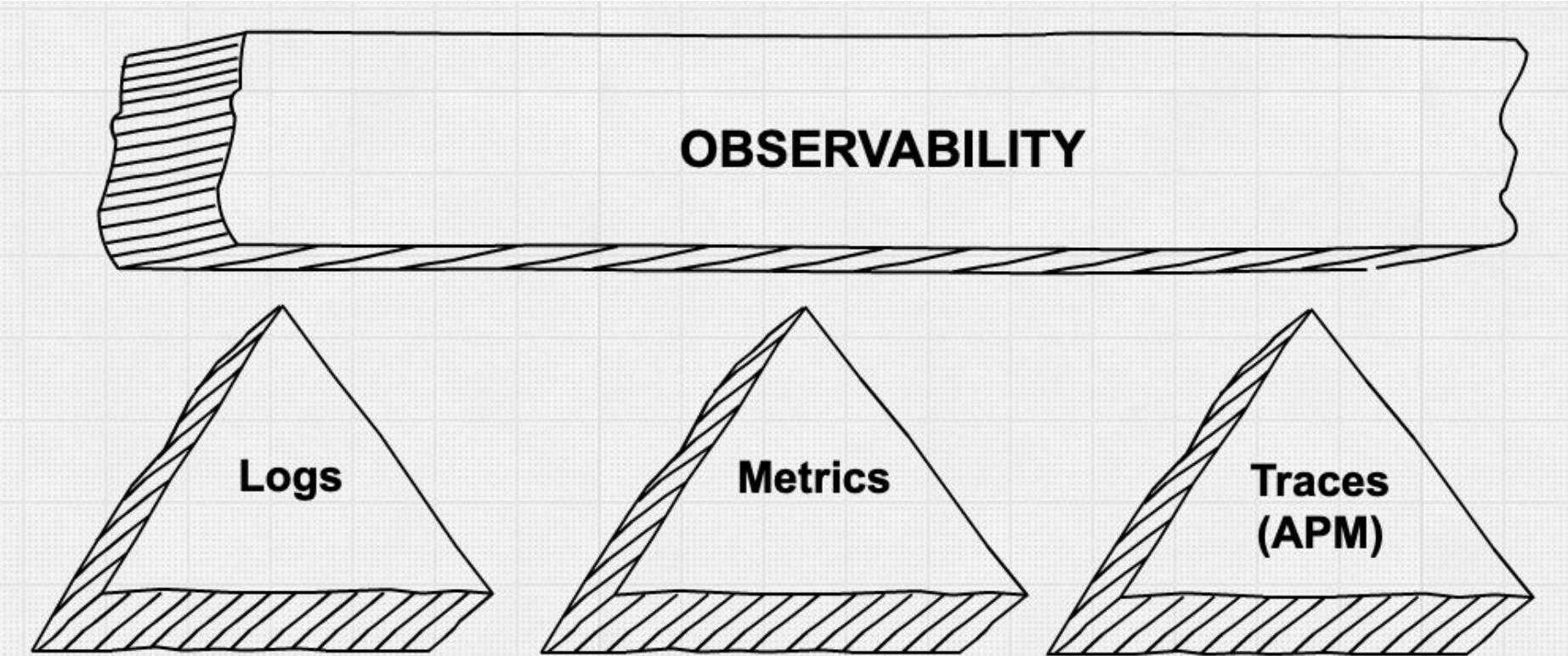
Podemos entender observabilidad por todo el conjunto de acciones que se realizan para monitorizar un sistema y reaccionar ante eventualidades. Entre esas las acciones que podemos encontrar están medir, leer y controlar.



Introducción a Elastic Observability



Introducción a Elastic Observability



Dev & Ops Teams

Filebeat

Metricbeat

Elastic APM

Heartbeat

Web Logs
App Logs
Database Logs
Container Logs

Container Metrics
Host Metrics
Database Metrics
Network Metrics
Storage Metrics

Real User Monitoring
Transaction Monitoring
Distributed Tracing
Dependency Mapping

Uptime
Response
Correctness
Certificate Validation

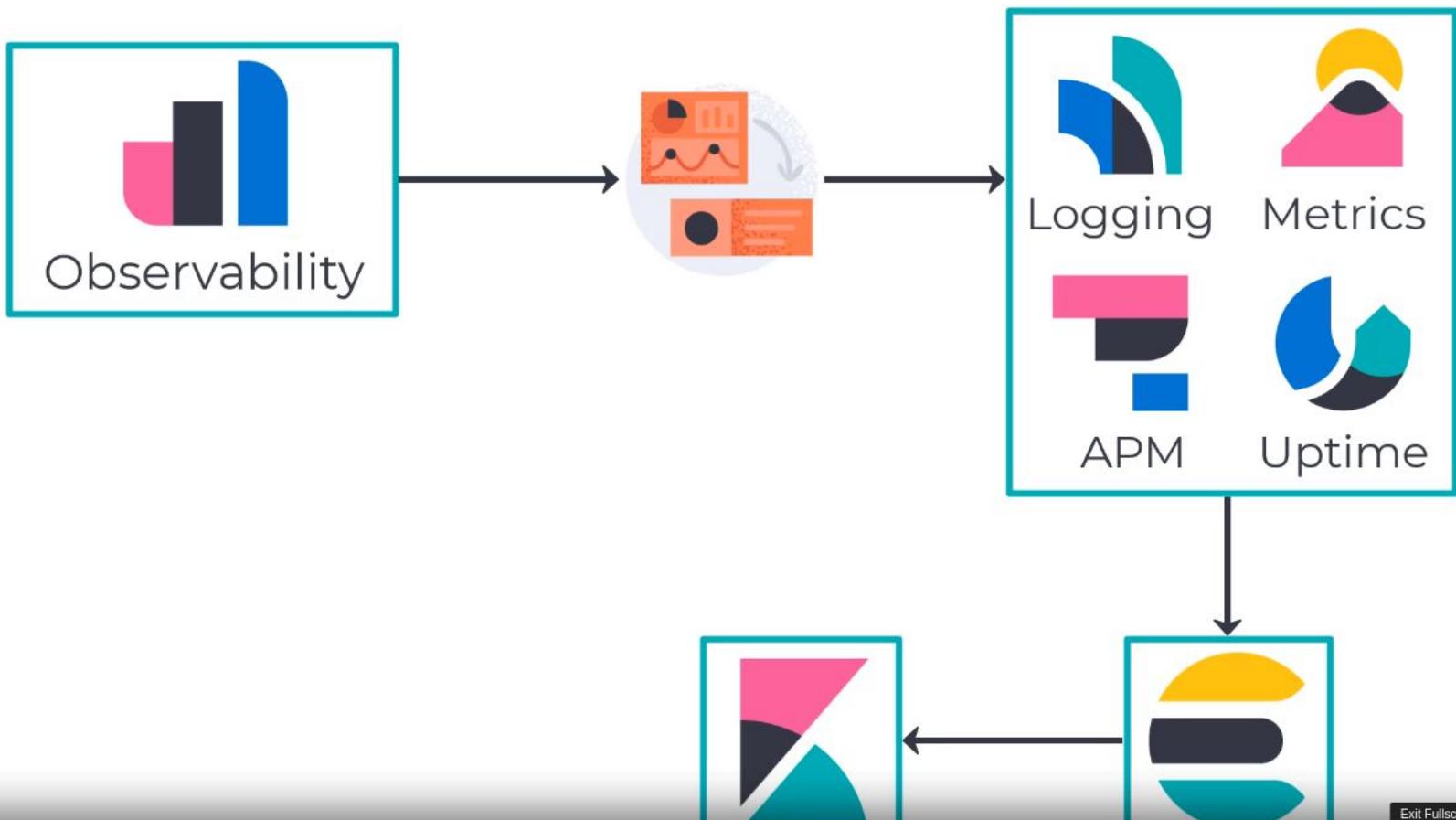
Elastic Common Schema



Elasticsearch



Kibana



Exit Fullscr

Centralización de Elastic Logs

```
66.249.73.185 -- [16/Feb/2014:09:47:54 -0500] "GET / HTTP/1.1" 200 37932 "-"
"Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
```

```
[2017-05-18 00:00:05,871][INFO ][cluster.metadata      ] [esnode-2] [.data-
es-1-2017.05.18] creating index, cause [auto(bulk api)], templates [.data-
es-1], shards [1]/[1], mappings [_default_, shards, node, index_stats,
index_recovery, cluster_state, cluster_stats, node_stats, indices_stats]
```

```
120707 0:37:09 [Note] Plugin 'FEDERATED' is disabled.
120707 0:37:09 InnoDB: The InnoDB memory heap is disabled
```

Centralización de Elastic Logs

Overview Anomaly Detection Data Frame Analytics **Data Visualizer** Settings

Visualize data from a log file EXPERIMENTAL

The File Data Visualizer helps you understand the fields and metrics in a log file. Upload your file, analyze its data, and then choose whether to import the data into an Elasticsearch index.

The File Data Visualizer supports these file formats:



- Delimited text files, such as CSV and TSV
- Newline-delimited JSON
- Log files with a common format for the timestamp

You can upload files up to 100 MB.

This feature is experimental. Got feedback? Please create an issue in [GitHub](#).



Select or drag and drop a file

Centralización de Elastic Logs

Overview Anomaly Detection Data Frame Analytics **Data Visualizer** Settings

sample.log

File contents

First 1 line

```
1 192.168.86.255 - 2020-04-15T21:43:32.020+0000 DEBUG 471 - Saving timestamp of Fundamentals of Securing Elasticsearch - Limited Offer
```

Summary

Number of lines analyzed 1

Format semi_structured_text

Grok pattern %{IP:ipaddress} .*? %{TIMESTAMP_ISO8601:timestamp} %{LOGLEVEL:loglevel} %{INT:field} .*

Time field timestamp

Time format ISO8601

[Override settings](#)

[Analysis explanation](#)

Logs vs Metrics

Logs are recorded
as events occur

[2018-09-07T07:48:00,127][INFO][o.e.x.m.MIDailyMaintenanceService] triggering scheduled [ML] maintenance tasks

[2018-09-07T07:48:00,381][INFO][o.e.x.m.a.TransportDeleteExpiredDataAction]
[_8LMCWq] Deleting expired data

[2018-09-07T07:48:00,648][INFO][o.e.x.m.MIDailyMaintenanceService] Successfully completed [ML] maintenance tasks

Metrics are recorded
based on a schedule

[2018-09-07T06:00:00,000][filesystem] 50085941248 overlay / 67371577344

[2018-09-07T06:05:00,000][filesystem] 50085917352 overlay / 67371577344

[2018-09-07T06:10:00,000][filesystem] 50075903715 overlay / 67371577344

Metrics

All Logs **Metrics** Security Sample data



ActiveMQ metrics

Fetch monitoring metrics from ActiveMQ instances.



Aerospike metrics

Fetch internal metrics from the Aerospike server.



Apache metrics

Fetch internal metrics from the Apache 2 HTTP server.



AWS metrics

Fetch monitoring metrics for EC2 instances from the AWS APIs and Cloudwatch.



Azure metrics

Fetch Azure Monitor metrics.



Ceph metrics

Fetch internal metrics from the Ceph server.



CockroachDB metrics

Fetch monitoring metrics from the CockroachDB server.



Consul metrics

Fetch monitoring metrics from the Consul server.



CoreDNS metrics

Fetch monitoring metrics from the CoreDNS server.



Couchbase metrics

Fetch internal metrics from Couchbase.



CouchDB metrics

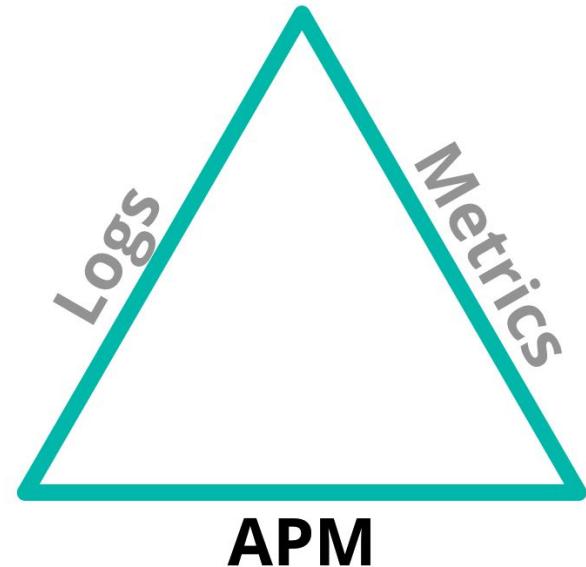
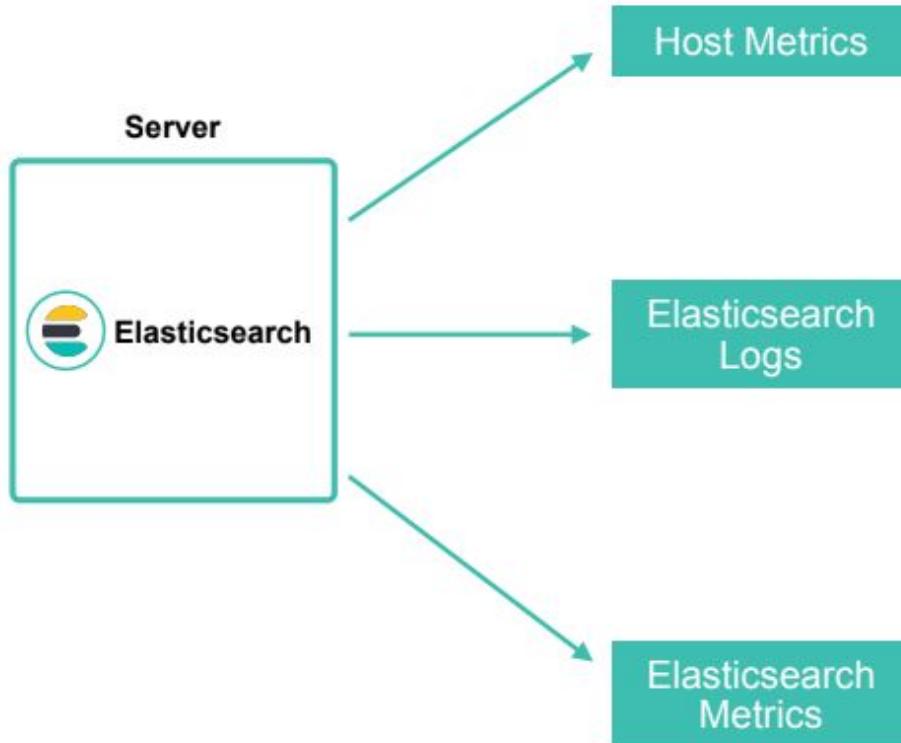
Fetch monitoring metrics from the CouchDB server.



Docker metrics

Fetch metrics about your Docker containers.

Elastic APM (Application Performance Monitoring)



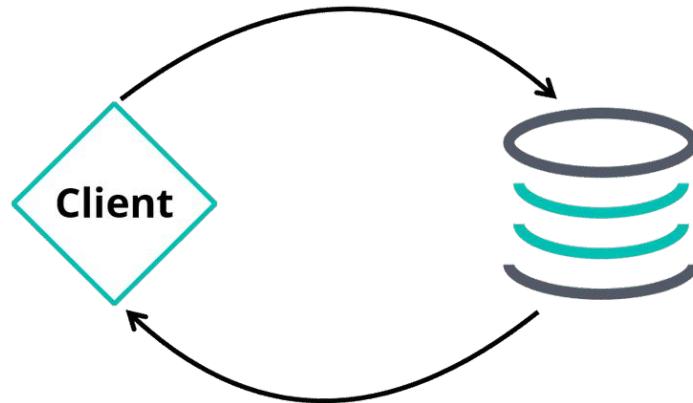
Elastic APM (Application Performance Monitoring)

Añadir APM a tu aplicación te permite lo siguiente:

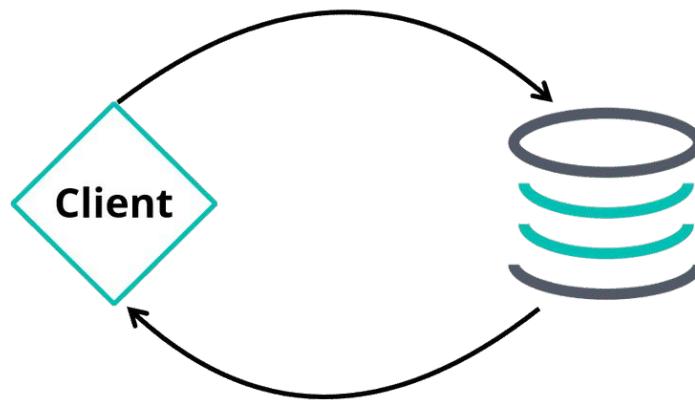
- Comprender el comportamiento de los servicios.
- Ver cómo los servicios interactúan entre ellos y visualizar cuellos de botella.
- Descubrir y corregir cuellos de botella y errores de rendimiento.
- Mantener un seguimiento de la experiencia del usuario final con la aplicación

Elastic APM (Application Performance Monitoring)

17:36:30 Request /top/10



17:36:30 Request /top/10



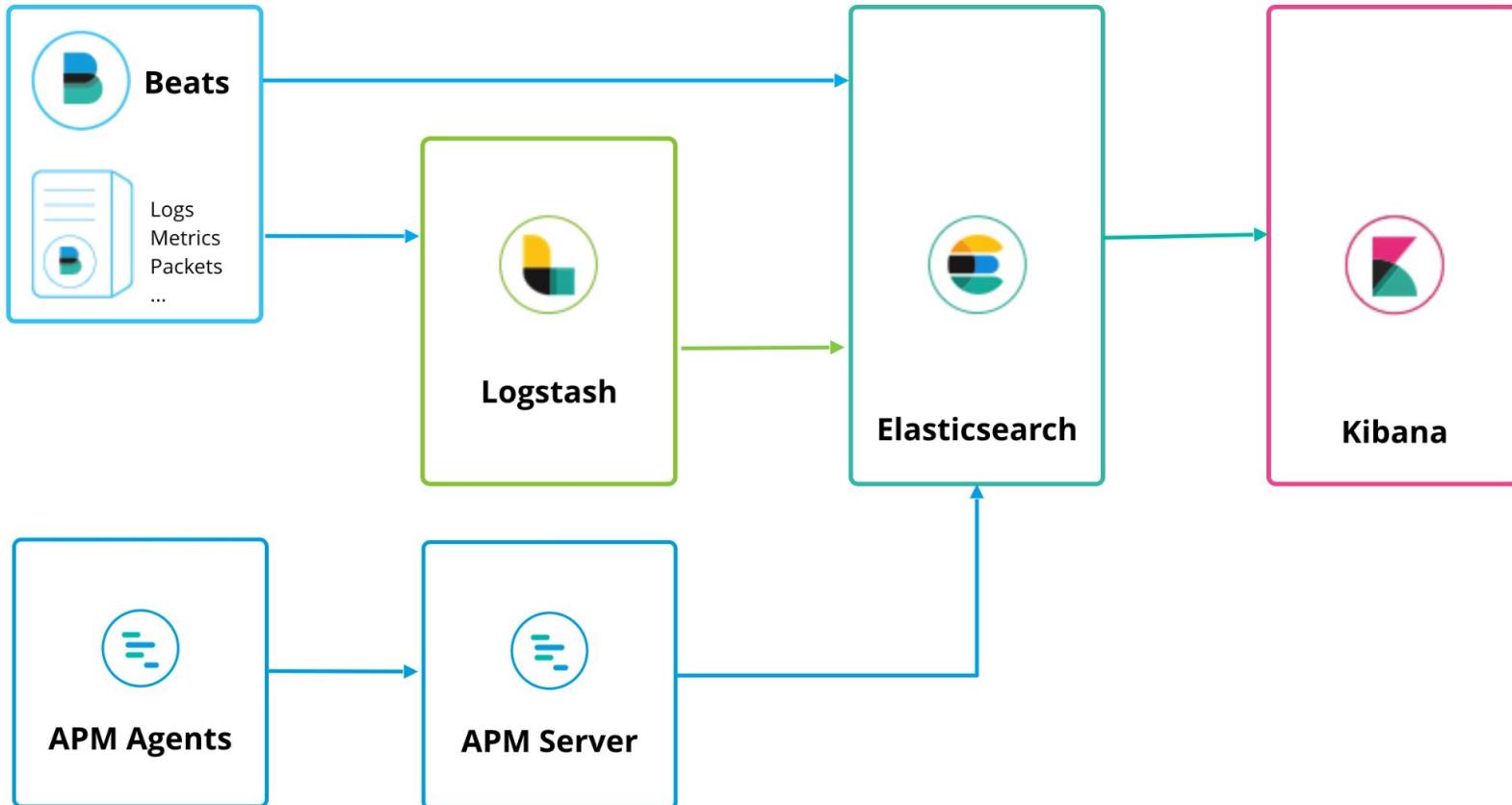
17:36:38 Response /top/10 200 OK



17:36:31 Response /top/10 500 ERROR



Elastic APM (Application Performance Monitoring)



Elastic APM (Application Performance Monitoring)

Edge Machines



Elastic Stack



Elastic APM (Application Performance Monitoring)

1. Iniciar un clúster de Elasticsearch con Kibana (**versión > 5.6**)
2. Iniciar el servidor APM
3. Configurar el agente APM en sus aplicación
4. Comprobar que los datos se indexan en Elasticsearch

Elastic APM (Application Performance Monitoring)

<https://www.elastic.co/es/downloads/apm>

Release date: August 03, 2021

License: [Elastic License](#)

Downloads:

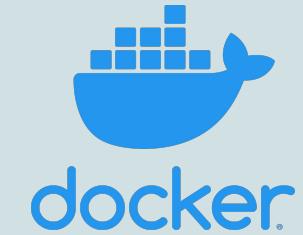
DEB 32-BIT shaasc	DEB 64-BIT shaasc
RPM 32-BIT shaasc	RPM 64-BIT shaasc
LINUX 32-BIT shaasc	LINUX 64-BIT shaasc
MAC shaasc	WINDOWS 32-BIT shaasc
WINDOWS 64-BIT shaasc	LINUX AARCH64 shaasc
DEB AARCH64 shaasc	RPM AARCH64 shaasc

Elastic APM (Application Performance Monitoring)

<https://www.elastic.co/guide/en/apm/server/current/running-on-docker.html>

```
$ docker pull docker.elastic.co/apm/apm-server:7.14.0
```

```
$ docker run -d \
-p 8200:8200 \
--name=apm-server \
--user=apm-server \
--volume="$(pwd)/apm-server.docker.yml:/usr/share/apm-server/apm-server.yml:ro" \
docker.elastic.co/apm/apm-server:7.14.0 \
--strict.perms=false -e \
-E output.elasticsearch.hosts=["elasticsearch:9200"]
```



Elastic APM (Application Performance Monitoring)

APM Agents



Java

RUM (JS)

Node.js

Django

Flask

Ruby on Rails

Rack

Go

.NET

PHP

4 Download the APM agent

Download the agent jar from [Maven Central](#). Do **not** add the agent as a dependency to your application.

5 Start your application with the javaagent flag

Add the `-javaagent` flag and configure the agent with system properties.

- Set the required service name (allowed characters: a-z, A-Z, 0-9, -, _, and space)
- Set the custom APM Server URL (default: `http://localhost:8200`)
- Set the APM Server secret token
- Set the service environment

Elastic APM (Application Performance Monitoring)

```
java -javaagent:/path/to/elastic-apm-agent-<version>.jar \
    -Delastic.apm.service_name=my-application \
    -Delastic.apm.server_url=http://localhost:8200 \
    -Delastic.apm.application_packages=org.example \
    -jar my-application.jar
```

Elastic APM (Application Performance Monitoring)

☰ D APM / Services 6d Settings Alerts Anomaly detection Add data

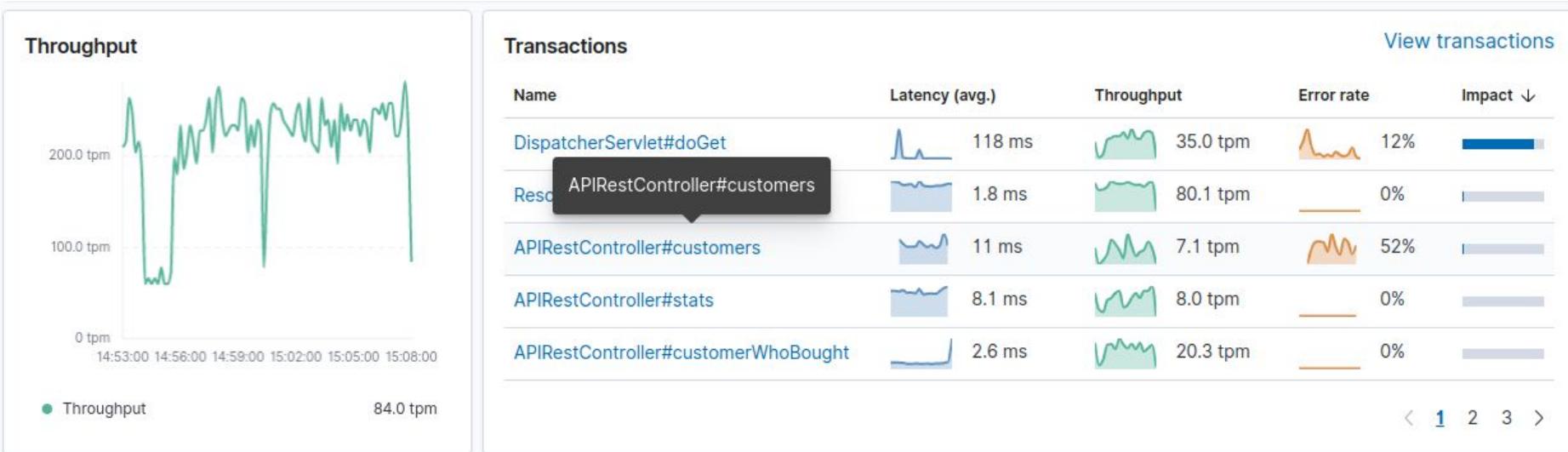
Services Traces Service Map

Search transactions, errors and metrics (E.g. transaction.duration.us > 300000 AND h) Last 15 minutes Show dates Refresh

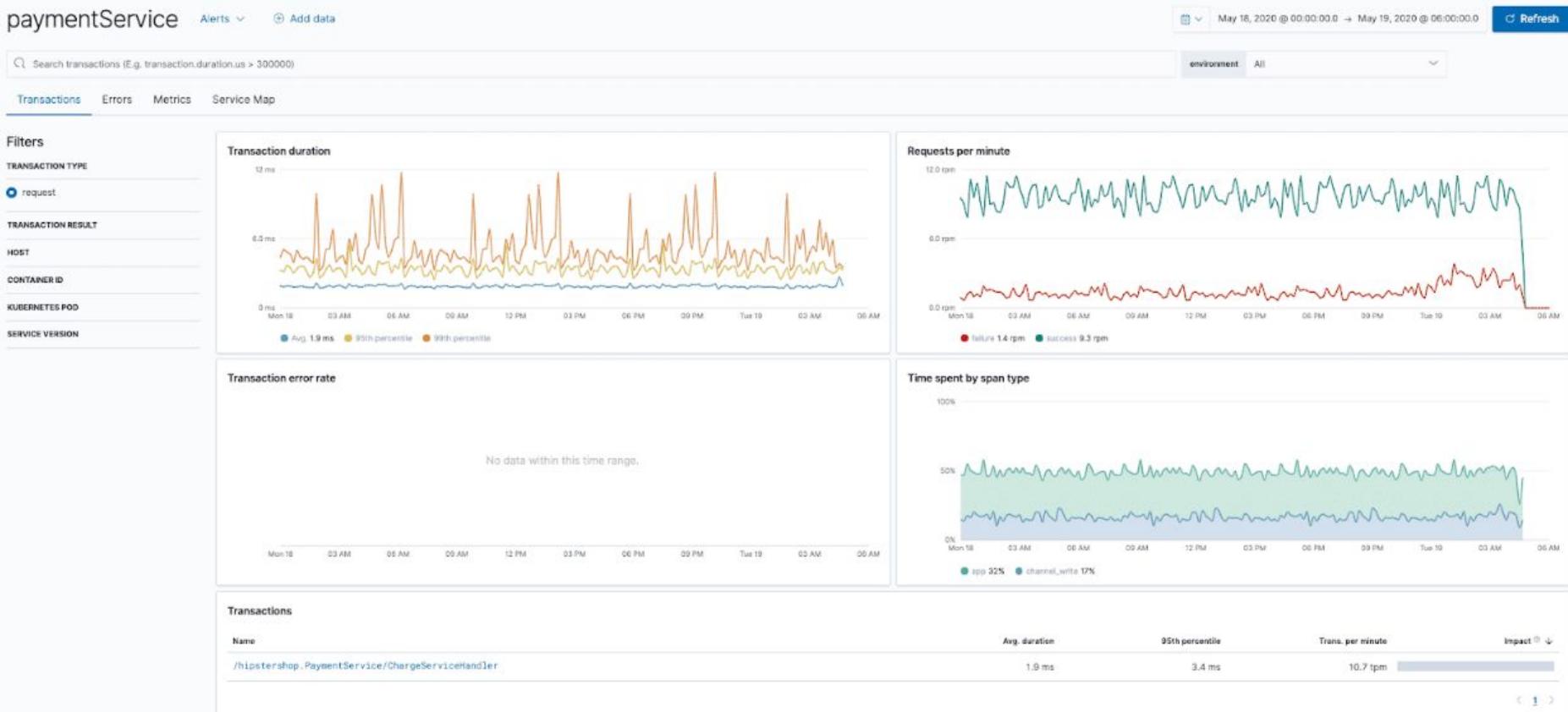
② What are these metrics?

Name	Environment	Latency (avg.)	Throughput ↓	Error rate %
⬢ opbeans-node	production	28 ms	215.3 tpm	4.2%
☕ opbeans-java	production	22 ms	211.7 tpm	8.6%
🥩 opbeans-ruby	production	39 ms	210.1 tpm	3.8%
🐹 opbeans-golang	production	30 ms	209.5 tpm	4.4%
🐍 opbeans-python	production	132 ms	203.1 tpm	2.7%

Elastic APM (Application Performance Monitoring)

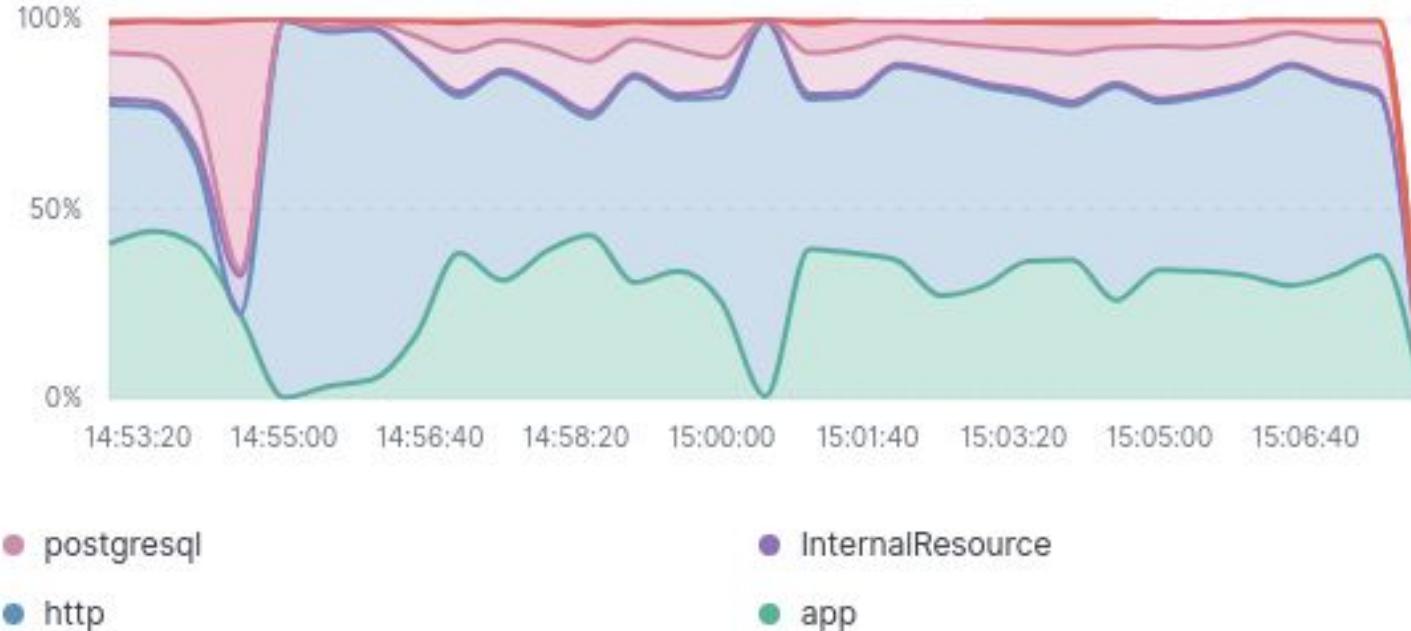


Elastic APM (Application Performance Monitoring)

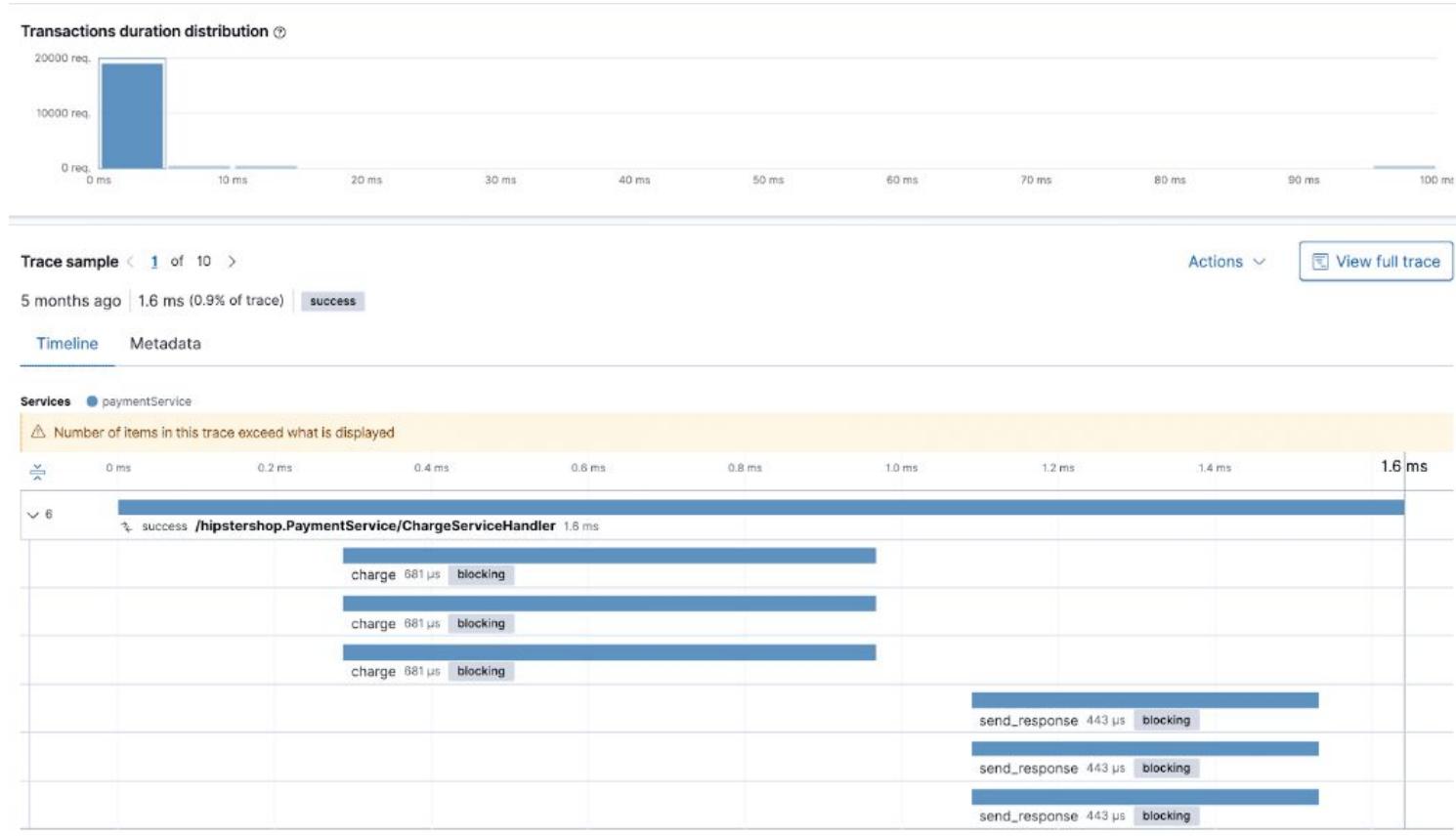


Elastic APM (Application Performance Monitoring)

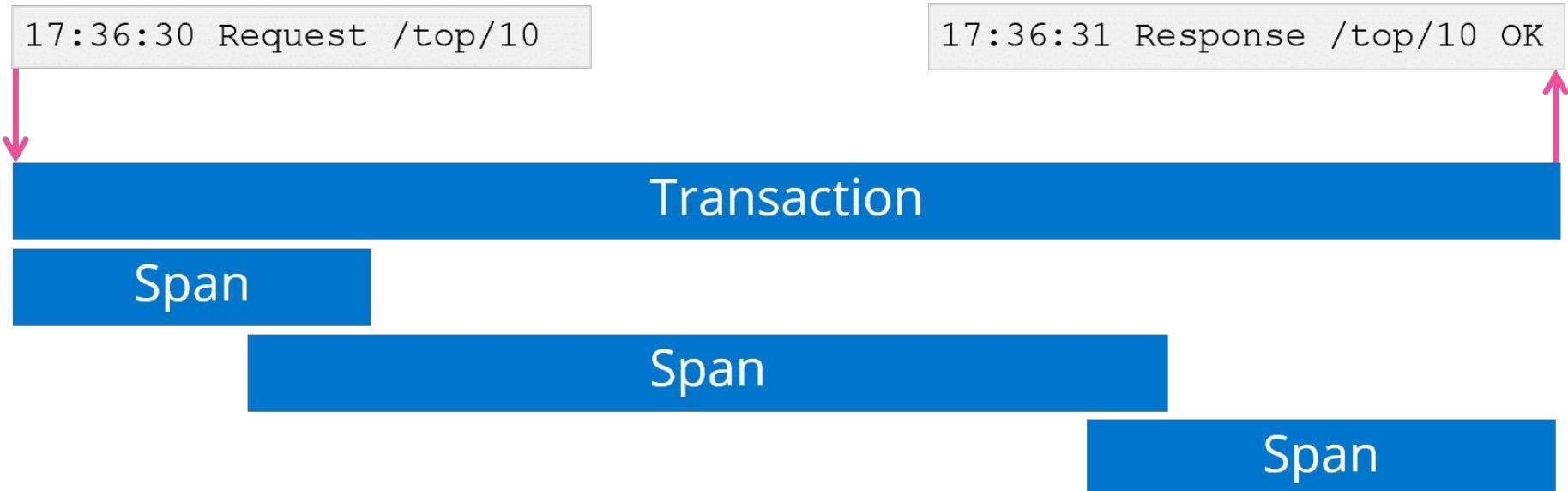
Time spent by span type



Elastic APM (Application Performance Monitoring)



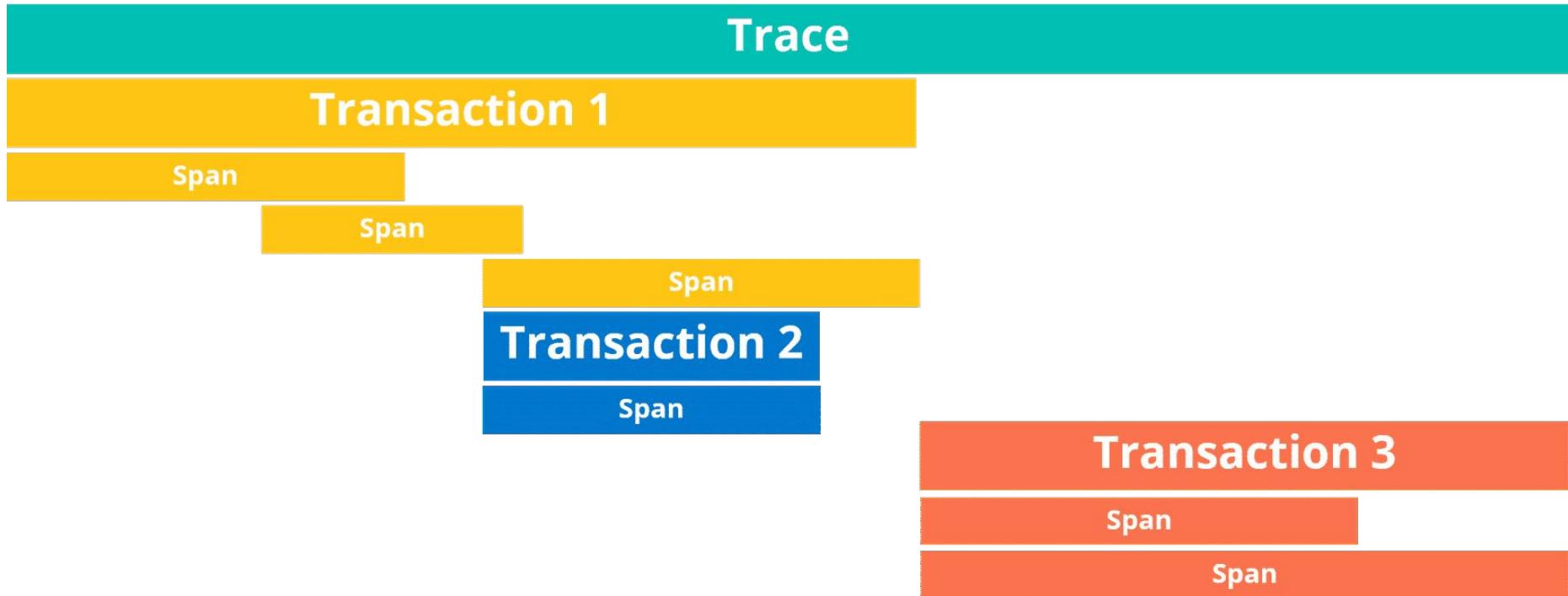
Elastic APM (Application Performance Monitoring)



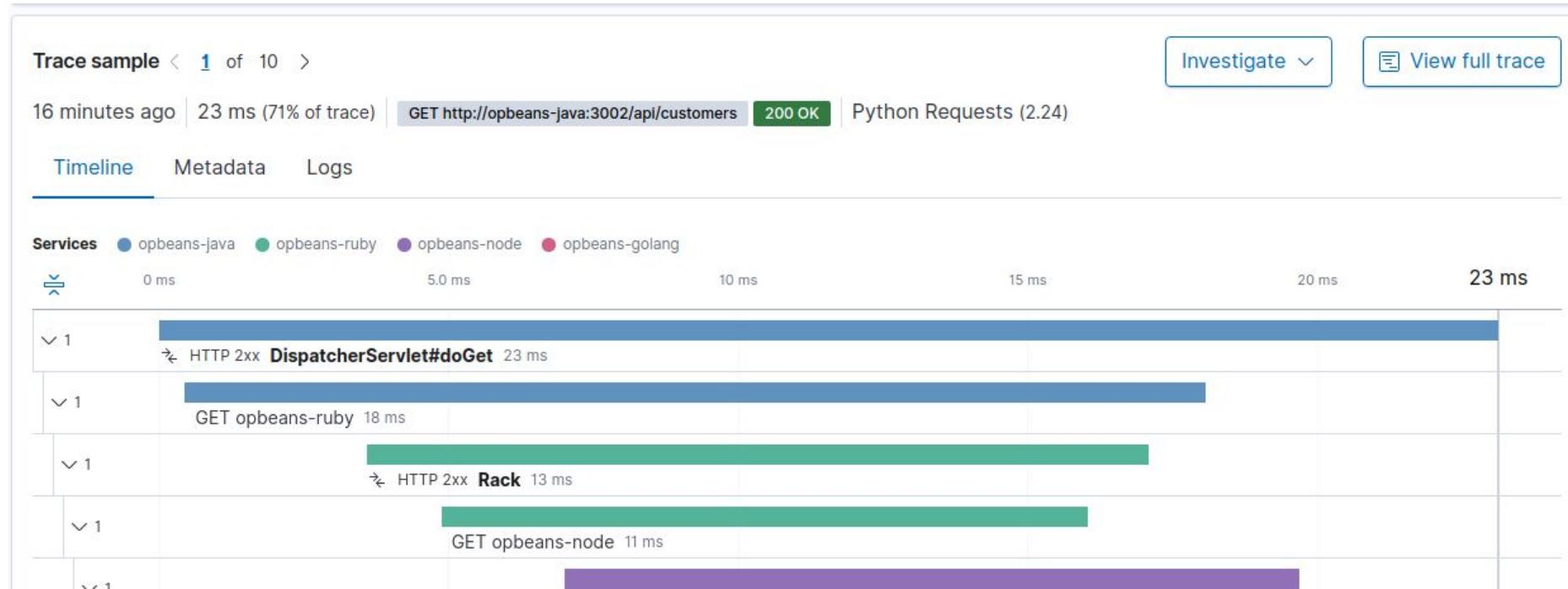
Elastic APM (Application Performance Monitoring)



Elastic APM (Application Performance Monitoring)



Elastic APM (Application Performance Monitoring)



Elastic APM (Application Performance Monitoring)

Span details

[View span in Discover](#)

Name	Service	Transaction
SELECT FROM customers	opbeans-golang	GET /api/customers

44 minutes ago | 319 µs (1.4% of transaction) | [DB](#) [postgresql](#) [query](#)

Database statement

```
SELECT
  customers.id, full_name, company_name, email,
  address, postal_code, city, country
FROM customers
```

Elastic APM (Application Performance Monitoring)

APM / Services / opbeans-java / Errors

6d Settings Alerts ▾ Anomaly detection Add data

Errors

Group ID	Type	Error message and culprit	Occurrences	Latest occurrence
cc927	org.springframework.web.bind.support.converter.ConverterNotFoundException	No converter found for return value of type: class com.sun.prism.Graphics2D	156	4 minutes ago
d16d3	org.springframework.web.bind.annotation.ResourceNotFoundException	Response status 404 co.elastic.apm.opbeans.controllers.APIRestController.lambda\$notFound\$0()	148	4 minutes ago
25fae	java.io.IOException	Connection reset by peer co.elastic.apm.opbeans.controllers.DTInterceptor.preHandle(DTInterceptor.java:44)	49	4 minutes ago
3bb34	org.apache.catalina.connector.ClientAbortException	java.io.IOException: Connection reset by peer N/A	44	4 minutes ago
23756	org.springframework.web.bind.annotation.ResourceNotFoundException	Response status 404 co.elastic.apm.opbeans.controllers.APIRestController.lambda\$notFound\$0()	21	4 minutes ago
...	...	Request method 'POST' not supported	10	4 minutes ago

Elastic APM (Application Performance Monitoring)

Error occurrence

 View 66 occurrences in Discover.

5 minutes ago

GET http://10.15.240.127:3002/api/customers

500 Internal Server Error

Python aiohttp (3.3.2)

APIRestController#customers

Exception stack trace

Metadata

java.io.IOException: Connection reset by peer

```
at org.apache.catalina.connector.OutputBuffer.realWriteBytes(OutputBuffer.java:356)
at org.apache.catalina.connector.OutputBuffer.flushByteBuffer(OutputBuffer.java:825)
at org.apache.catalina.connector.OutputBuffer.append(OutputBuffer.java:730)
at org.apache.catalina.connector.OutputBuffer.writeBytes(OutputBuffer.java:391)
at org.apache.catalina.connector.OutputBuffer.write(OutputBuffer.java:369)
at org.apache.catalina.connector.CoyoteOutputStream.write(CoyoteOutputStream.java:96)
at com.fasterxml.jackson.core.json.UTF8JsonGenerator._flushBuffer(UTF8JsonGenerator.java:2085)
at com.fasterxml.jackson.core.json.UTF8JsonGenerator.writeString(UTF8JsonGenerator.java:457)
at com.fasterxml.jackson.databind.ser.std.StringSerializer.serialize(StringSerializer.java:41)
at com.fasterxml.jackson.databind.ser.BeanPropertyWriter.serializeAsField(BeanPropertyWriter.java:727)
at com.fasterxml.jackson.databind.ser.std.BeanPropertyWriter.serializeFields(BeanPropertyWriter.java:710)
```

Elastic APM (Application Performance Monitoring)

The screenshot shows the Elastic APM interface for monitoring application performance. At the top, there's a navigation bar with the elastic logo, a search bar, and various management icons like Settings, Alerts, Anomaly detection, and Add data. Below the navigation, the path APM / Services / opbeans-ruby / Transactions / Rack is displayed. The main content area is titled "Trace sample" and shows a single trace entry from 15 minutes ago. The trace details a GET request to http://10.15.249.213:3004/api/orders, which was successful (200 OK) using Python aiohttp (3.3.2). The "Logs" tab is selected, showing a table with columns for Timestamp, event.dataset, and Message. The first log entry is timestamped at 16:07:52.000 and is categorized as nginx.error. The message contains a detailed log entry about loading a REDIS Cache and handling a specific API request.

Timestamp	event.dataset	Message
16:07:52.000	nginx.error	[nginx][info] [lua] ip_blacklist.lua:15: Loading REDIS Cache, client: 10.12.1.29, server: blue.demo.elastic.co, request: "GET /api/apm/services/opbeans-ruby/transactions/charts/error_rate?environment=ENVIRONMENT_ALL&start=2021-08-08T13%3A52%3A00.000Z&end=2021-08-08T14%3A07%3A42.097Z&transactionType=request&transactionName=Rack&uiFilters=%7B%22environment%22%3A%22ENVIRONMENT_ALL%22%7D HTTP/1.1", host: "demo.elastic.co", referrer: "https://demo.elastic.co/app/apm/services/opbeans-ruby/transactions/view?rangeFrom=now-15m&rangeTo=now&traceId=c553ab037638b0d299ab33a3d18d0dab&transactionId=38f326def11d8efd&transactionName=Rack&transactionType=request&latencyAggregationType=avg&environment=ENVIRONMENT_ALL"

Elastic APM (Application Performance Monitoring)

≡ | D | APM / Services / paymentService / Metrics

paymentService

Alerts ▾ + Add data

May 18, 2020 @ 00:00:00.0 → May 19, 2020 @ 06:00:00.0 Refresh

Search metrics (E.g. process.pid = "1234") environment All

Transactions Errors Metrics Service Map

Filters

HOST

CONTAINER ID

KUBERNETES POD

SERVICE VERSION

CPU usage

Mon 18 06 AM 12 PM 06 PM Tue 19 06 AM

System max 88% System average 34% Process max 0.2%
Process average 0.1%

System memory usage

Mon 18 06 AM 12 PM 06 PM Tue 19 06 AM

Max 33% Average 30%

Elastic APM (Application Performance Monitoring)

☰ | D APM / Services / advertService / JVMs

advertService

Alerts ▾ [+ Add data](#) May 18, 2020 @ 00:00:00.0 → May 19, 2020 @ 06:00:00.0 [⟳ Refresh](#)

Search metrics (E.g. process.pid = "1234") environment All

[Transactions](#) [Errors](#) [JVMs](#) [Service Map](#)

Filters

	Name	CPU avg ↓	Heap memory avg	Non-heap memory avg	Thread count max
HOST	advertiservice-pod	0.5%	113.7 MB	79.8 MB	33
CONTAINER ID					
KUBERNETES POD					

< 1 >

Elastic APM (Application Performance Monitoring)

☰ | D APM / Traces

APM Settings Alerts Anomaly detection ⚠️ + Add data May 18, 2020 @ 00:00:00.0 → May 19, 2020 @ 09:00:00.0 Refresh

Search transactions (E.g. transaction.duration.us > 300000) environment All

Services Traces Service Map

Filters

TRANSACTION RESULT

HOST

CONTAINER ID

KUBERNETES POD

Name	Originating service	Avg. response time	Traces per minute	Impact ⚪ ↓
productHandler	frontend	331 ms	123.237 tpm	<div style="width: 100%;"></div>
homeHandler	frontend	305 ms	49.192 tpm	<div style="width: 50%;"></div>
viewCartHandler	frontend	65 ms	57.149 tpm	<div style="width: 10%;"></div>
placeOrderHandler	frontend	164 ms	9.71 tpm	<div style="width: 5%;"></div>
addToCartHandler	frontend	25 ms	28.622 tpm	<div style="width: 2%;"></div>
setCurrencyHandler	frontend	0.2 ms	18.945 tpm	<div style="width: 1%;"></div>

< 1 >

Elastic APM (Application Performance Monitoring)

Trace sample < 1 of 10 >

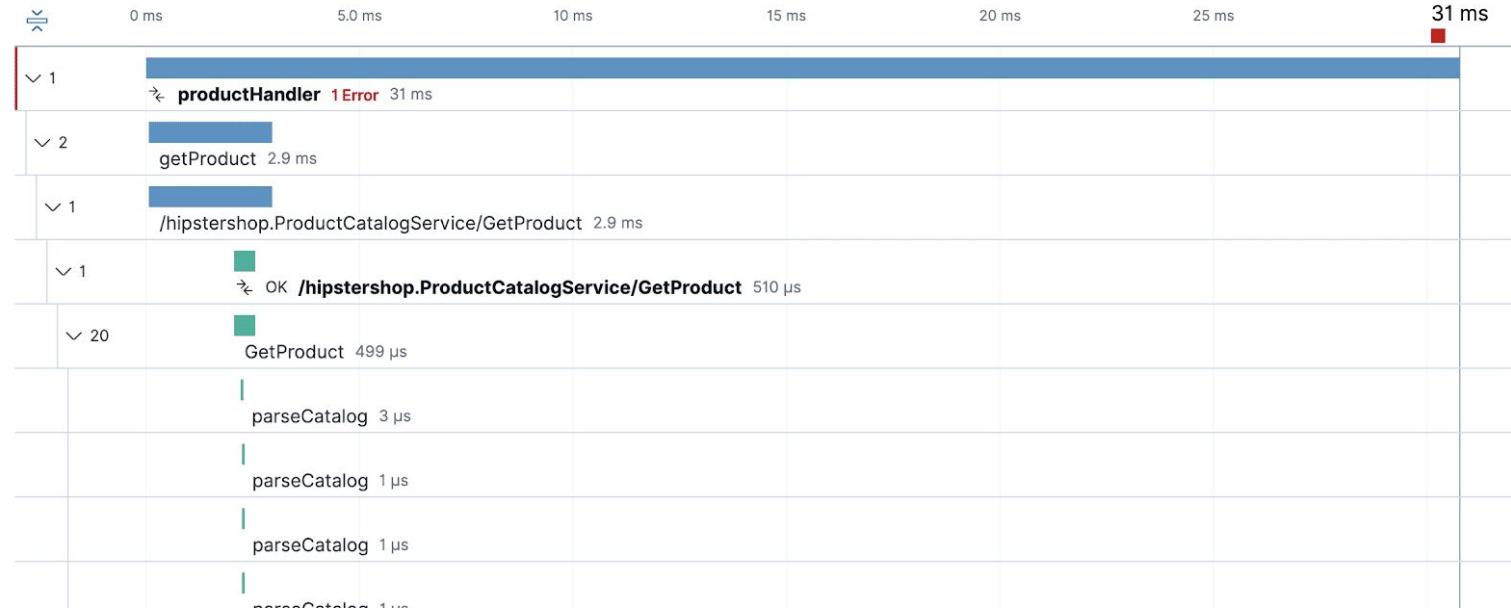
Actions ▾

[View full trace](#)

5 months ago | 31 ms (100% of trace) | **1 Error** | Chrome (51.0.848.0)

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

Services ● frontend ● productCatalogService ● currencyService ● cartService ● recommendationService



Elastic APM (Application Performance Monitoring)

APM | D APM / Service Map

APM Settings Alerts Anomaly detection Add data May 18, 2020 @ 00:00:00.0 → May 19, 2020 @ 06:00:00.0 Refresh

Search is not available here environment All

Services Traces Service Map

The Service Map displays the relationships between different services. The 'frontend' service is the central hub, with arrows pointing to it from 'advertismentService', 'checkoutService', 'recommendationService', and 'productCatalogService'. It also has arrows pointing away to 'elasticsearch', 'carService', 'currencyService', 'emailService', 'paymentService', and 'shippingService'. The 'advertismentService' and 'productCatalogService' services are shown with a yellow warning icon.

```
graph TD; frontend --> advertismentService; frontend --> checkoutService; frontend --> recommendationService; frontend --> productCatalogService; frontend --> elasticsearch; frontend --> carService; frontend --> currencyService; frontend --> emailService; frontend --> paymentService; frontend --> shippingService; advertismentService --> elasticsearch; carService --> currencyService; currencyService --> emailService; paymentService --> productCatalogService; productCatalogService --> shippingService;
```

Elastic APM (Application Performance Monitoring)

GET `_cat/indices/apm*?v&s=index`

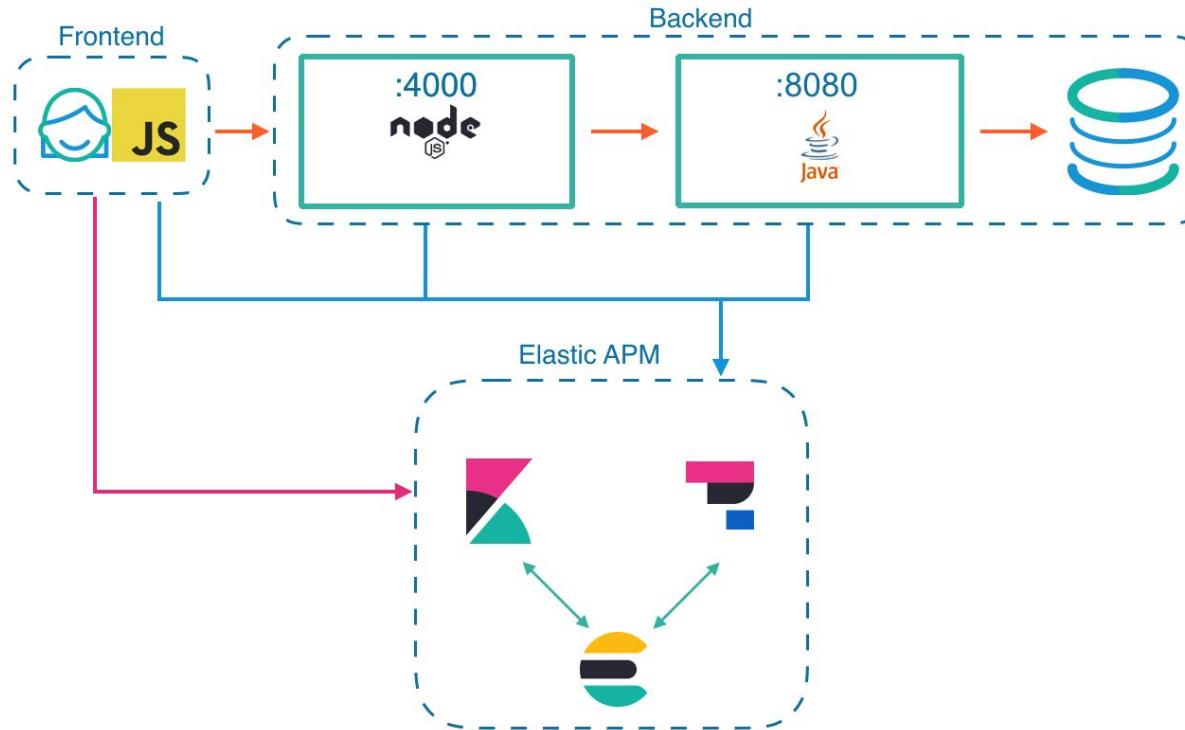
health	status	index	uuid	pri	rep	docs.count	docs.
green	open	apm-7.10.0-error-000001	jl60LSC5QtIyucIeapx7BA	1	1	0	0
green	open	apm-7.10.0-metric-000001	oRgvTpNfTI0l_Mr3XzVStA	1	1	0	0
green	open	apm-7.10.0-onboarding-2020.10.19	Af5D7ZA5QqmWbtX7exeWJQ	1	1	1	1
green	open	apm-7.10.0-profile-000001	X7_zr1XjRy0UEoSfBI4dWw	1	1	0	0
green	open	apm-7.10.0-span-000001	M8ILpHxGR2m4403Dks7zog	1	1	0	0
green	open	apm-7.10.0-transaction-000001	TlHuoIHnRVenK-4Wa5VXgQ	1	1	0	0

Elastic APM (Application Performance Monitoring)

apm-[version]-span-00000X
apm-[version]-transaction-00000X
apm-[version]-error-00000X
apm-[version]-metric-00000X

Elastic APM (Application Performance Monitoring)

<https://github.com/elastic/spring-petclinic>



ELASTIC SECURITY

- Introducción a Elastic Security
- SIEM
- Endpoint Security
- Antimalware
- Prevención contra Ransomware
- Recopilación de datos
- Reglas de detección y priorización de amenazas a escala
- Flujos de trabajo SecOps
- Detención de amenazas en frío
- Buenas prácticas buscando amenazas
- Búsqueda de amenazas automatizada
- Monitorización cloud
- Integración con Kibana Lens para visualizaciones

Introducción a Elastic Security

Screenshot of the Elastic Security Overview page.

The top navigation bar includes the elastic logo, a search bar with placeholder "Search Elastic", and several icons.

The main navigation bar shows "Security / Overview" and tabs for "Overview", "Detections", "Hosts", "Network", "Timelines", "Cases", and "Administration". The "Overview" tab is selected and highlighted with an orange border.

Below the tabs are search and date range filters: "Search" (with a dropdown arrow), "KQL" (selected), "Last 90 days" (with a calendar icon), "Show dates" (button), and "Refresh" (button).

The left sidebar contains "Data sources" (dropdown), "Recent cases" (with a file icon), and "Recent timelines" (with a star and file icon). Under "Recent cases", there is a section for "Tesla Agent Match" showing 1 alert generated by Tesla Agent.

The main content area features a chart titled "Detection alert trend" showing the number of alerts over time. The chart displays two major peaks: one around April 3rd (approx. 30 alerts) and another smaller peak around April 17th (approx. 5 alerts). The Y-axis represents the count of alerts (0 to 30), and the X-axis represents dates from March 6th to May 29th. The legend indicates alert types: Malware Detection Alert (teal), Packtpub Network Traffic Test (blue), Imported Packtpub Network Traffic Test (pink), and Malicious Indicator Match Rule (dark purple).

At the bottom of the sidebar, there is a message: "You haven't favorited any timelines yet. Get started!"

Endpoint Security

The screenshot shows the Elastic Security interface with the "Overview" tab selected. The top navigation bar includes links for Overview, Detections, Hosts, Network, Timelines, Cases, Administration, and Add data. Below the navigation is a search bar with "Search" and "KQL" options, a date range selector for "Last 90 days", and buttons for "Show dates" and "Refresh".

The main content area is divided into two sections: "Host events" and "Network events".

Host events: Shows 826,156 events. The top item is Auditbeat (0). Below it is Endpoint Security (757,917), followed by DNS (41,563), File (247,585), Image Load (18,600), Network (35,127), Process (40,990), Registry (370,118), and Security (3,934).

Event Type	Count
Auditbeat	0
Endpoint Security	757,917
DNS	41,563
File	247,585
Image Load	18,600
Network	35,127
Process	40,990
Registry	370,118
Security	3,934

Network events: Shows 171,603 events. The top item is Auditbeat (0). Below it is Packetbeat (171,603), followed by DNS (6,267), Flow (161,358), and TLS (3,978).

Event Type	Count
Auditbeat	0
Filebeat	0
Packetbeat	171,603
DNS	6,267
Flow	161,358
TLS	3,978

On the left side of the page, there is a sidebar with two blog posts:

- Detecting rare and unusual processes with Elastic machine learning** (published 2021-03-25) - This post explores identifying truly rare host process executions using anomaly detection jobs in both Elastic Machine Learning and Elastic Security.
- Hunting for Lateral Movement using Event Query Language** (published 2021-03-18) - This post explores some lateral movement techniques and leverages the capabilities of Elastic's Event Query Language (EQL) to design behavioral hunts and detections.

SIEM

- **Centralizar la vista de potenciales amenazas**
- **Determinar qué amenazas requieren resolución y cuáles son falsos positivos.**
- **Escalar información a los analistas de seguridad para que puedan tomar una acción.**
- **Documentar en un registro de auditoría los eventos detectados y cómo fueron resueltos**

SIEM

- **Blacklist.** Capacidad para detectar archivos a partir el hash.
- **Acceso a credenciales.** Capacidad para detectar o prevenir intentos de volcado de datos de credenciales de la máquina.
- **Exploits.** Instrumentación binaria dinámica para detectar o prevenir la explotación de aplicaciones vulnerables.
- **Malware.** Utilizando ML tiene la capacidad de inspeccionar archivos que pueden ser maliciosos al escribir, modificar, sobrescribir o ejecutar en el sistema de archivos.
- **Escalado de privilegios.** Capacidades de detección y prevención para la manipulación de credenciales y el escalado de privilegios.
- **Inyección de procesos.** Capacidad de prevención y detección de ataques que intentan de infectar la máquina mediante la inyección de procesos en memoria.

SIEM

SIM / Hosts

Overview Hosts Network Timeline

Untitled Timeline Description Notes 0 Year to date Show dates Refresh

e.g. host.name: "foo"

Drop here to build an OR query

AND Filter event.action:"config_change" and event.dataset:"file"

Fields @timestamp event.severity event.category event.action host.name

Jun 3, 2019 @ 19:40:15.160 -- audit-rule executed siem-es

Session # unset @ root @ siem-es in / executed > ip route is table local type flocal scope host dev eth0 proto 66 with result success

Jun 3, 2019 @ 19:40:15.160 -- audit-rule executed siem-es

Session # unset @ root @ siem-es in / executed > google_network_ with result fail

Jun 3, 2019 @ 19:40:15.160 -- audit-rule executed siem-es

Session # unset @ root @ siem-es in / executed > google_network_ with result fail

Jun 3, 2019 @ 19:40:15.160 -- audit-rule executed siem-es

Session # unset @ root @ siem-es in / executed > google_network_ with result fail

Hosts 904

User #

All Hosts Showing: 21,116 Hosts

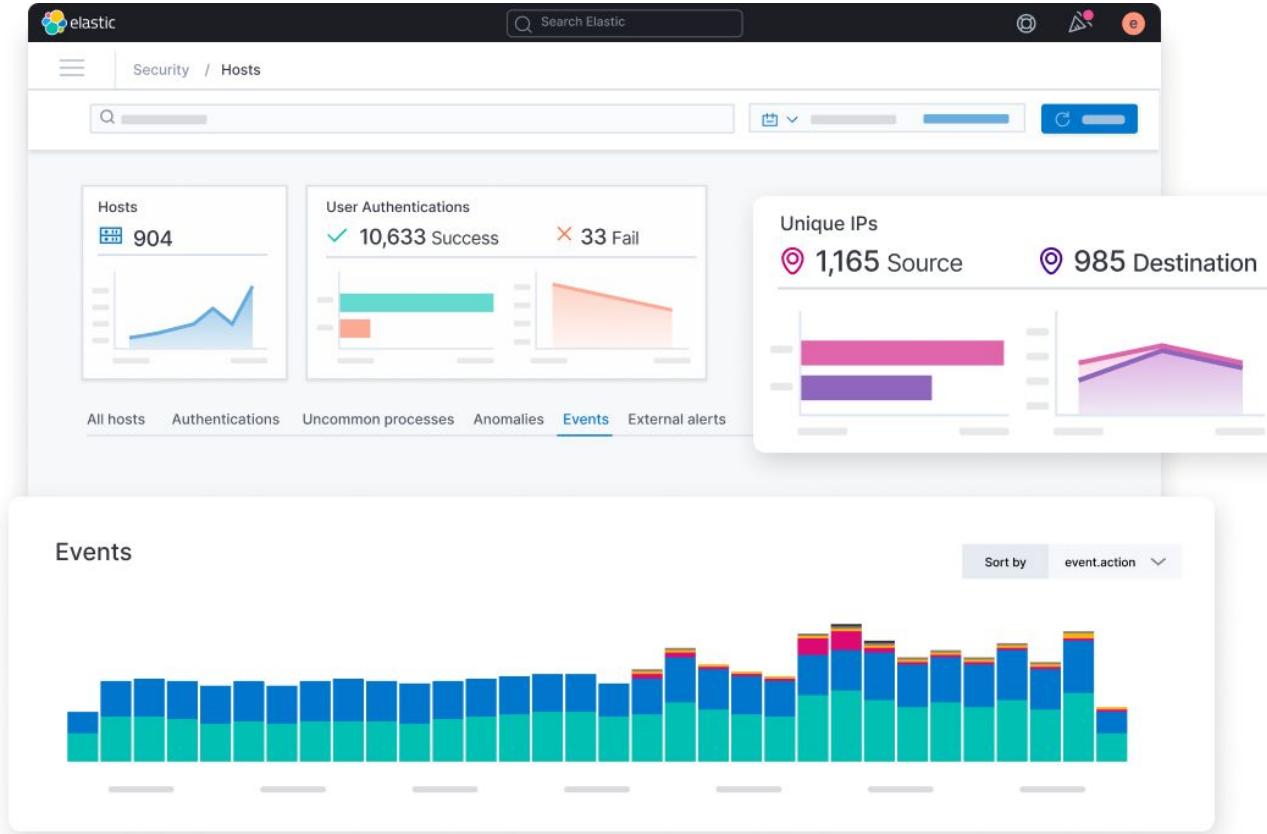
Name

heate-ci-immutable-ubuntu-1804

25 of 14846121 Events Load More Updated 4 minutes ago

The screenshot shows a SIEM application interface. At the top, there's a navigation bar with tabs for Overview, Hosts (which is selected), Network, and Timeline. Below the navigation is a search bar with placeholder text 'e.g. host.name: "foo"'. To the right of the search bar is a search builder area with a dropdown for 'OR' and a single search term 'host.name: "siem-es"'. Below this is a placeholder text 'Drop here to build an OR query' and a 'AND Filter' button followed by a complex search query involving 'event.action' and 'event.dataset'. The main content area is titled 'Hosts' and shows a summary of 904 hosts with a line graph showing host count over time from June 3, 2019. Below this is a table of audit events for host 'siem-es' on June 3, 2019, at 19:40:15.160. The table columns are Fields (@timestamp, event.severity, event.category, event.action, host.name), and the rows show multiple audit events with status 'executed'. On the left side, there's a sidebar with various icons and sections for 'All Hosts' (Showing: 21,116 Hosts) and a 'Name' filter input field containing 'heate-ci-immutable-ubuntu-1804'. At the bottom, there are pagination controls (25 of 14846121 Events), a 'Load More' button, and a timestamp indicating the data was updated 4 minutes ago.

Flujos de trabajo SecOps



Malware y Ransomware



Prevención contra Malware y Ransomware

Edit Endpoint Security integration

Modify integration settings and deploy changes to the selected agent policy.

Agent policy
Default policy

Integration settings

Choose a name and description to help identify how this integration will be used.

Integration name: endpoint

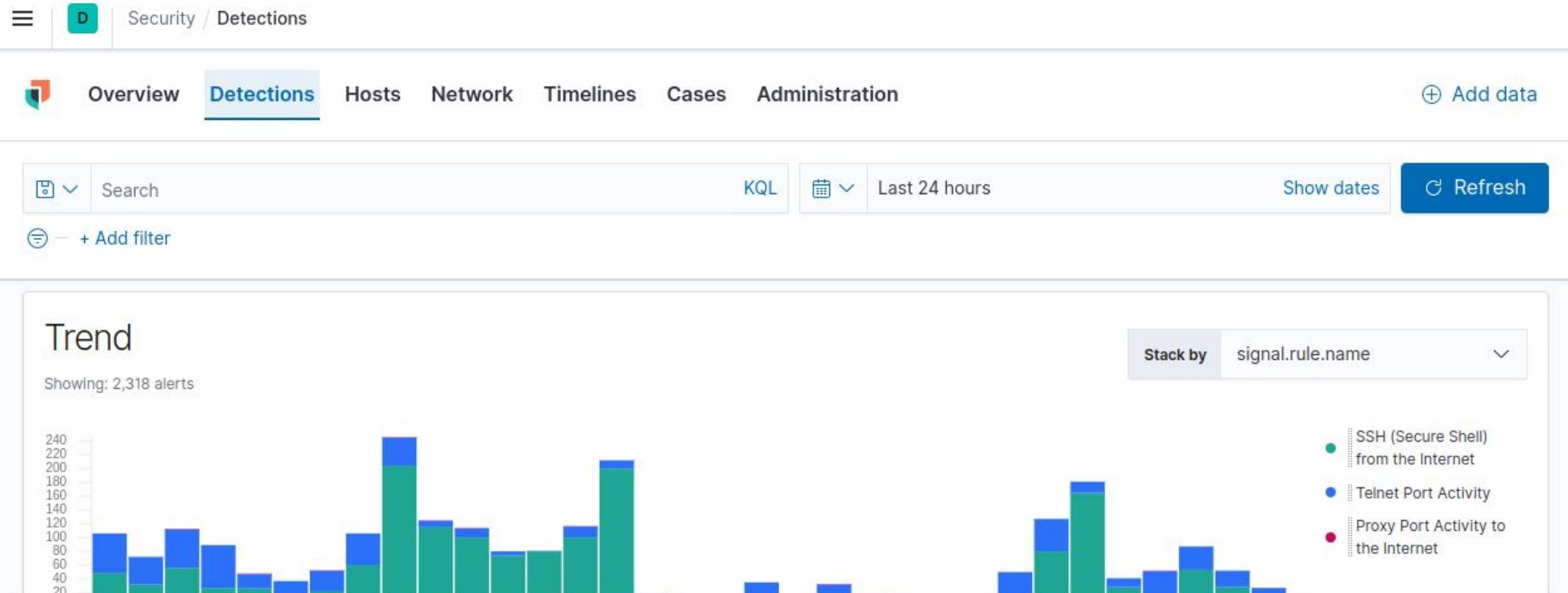
Description: (Optional)

> Advanced options

Protections

Type	Operating System	Malware protections enabled
Malware	Windows, Mac	<input checked="" type="checkbox"/> Malware protections enabled

Detection alerts



Detection alerts details

Security / Detections

Overview Detections Hosts Network Timelines Cases Administration

Search + Add filter

Showing 2,240 alerts

Icon	@timestamp	Rule	Versi...	Method
	Aug 8, 2021 @ 19:00:16.945	Telnet Port Activity	7	query
	Aug 8, 2021 @ 19:00:16.087	SSH (Secure Shell) from th...	8	query
	Aug 8, 2021 @ 19:00:16.086	SSH (Secure Shell) from th...	8	query
	Aug 8, 2021 @ 19:00:16.086	SSH (Secure Shell) from th...	8	query

Generic Endpoint Timeline 0

Alert details

Summary Table JSON View

signal.status: Open

@timestamp: Aug 8, 2021 @ 19:00:16.945

Rule: Telnet Port Activity

Severity: medium

Risk Score: 47

host.name: server-02

source.ip: 124.133.201.110

destination.ip: 178.62.77.103

Detection rules

[Overview](#)[Detections](#)[Hosts](#)[Network](#)[Timelines](#)[Cases](#)[Administration](#)[ML job settings](#)[+ Add data](#)

All rules

 e.g. rule name[Tags](#)[Elastic rules \(0\)](#) [Custom rules \(0\)](#)

Load Elastic prebuilt detection rules

Elastic Security comes with prebuilt detection rules that run in the background and create alerts when their conditions are met. By default, all prebuilt rules except the Endpoint Security rule are disabled. You can select additional rules you want to activate.

[Load Elastic prebuilt rules](#)[Create your own rules](#)

Timeline <

Detection rules

≡ | D | Security / Detections / Detection rules

[Overview](#)[Detections](#)[Hosts](#)[Network](#)[Timelines](#)[Cases](#)[Administration](#)[+ Add data](#)

All rules

⌚ Updated 50 seconds ago

 ×[Tags](#) ▼[Elastic rules \(546\)](#)[Custom rules \(0\)](#)

Showing 5 rules | Selected 0 rules [⌚ Refresh](#) [Refresh settings](#) ▾

Rule	Risk score	Severity	Last run	Last response	Last updated	Version	Tags	Activated
SSH (Secure Shell) from the Internet	47	● Medium	18 minutes ago	● failed	Jun 23, 2021 @ 02:31:54.383	8	Command and Control Elastic Host See all	<input checked="" type="checkbox"/>
SSH (Secure Shell) to the Internet	21	● Low	18 minutes ago	● failed	Jun 23, 2021 @ 02:30:54.326	8	Command and Control Elastic Host See all	<input checked="" type="checkbox"/>
Potential SSH Brute Force Detected	47	● Medium	—	● —	Jun 23, 2021 @ 02:28:16.516	2	Credential Access Elastic Host See all	<input type="checkbox"/>

+ ● Untitled timeline

Detection rules

- https://github.com/elastic/detection-rules/blob/main/rules/linux/credential_access_ssh_backdoor_log.toml

87 lines (78 sloc) | 2.54 KB

Raw Blame  

```
1 [metadata]
2 creation_date = "2020/12/21"
3 maturity = "production"
4 updated_date = "2021/03/03"
5
6 [rule]
7 author = ["Elastic"]
8 description = """
9 Identifies a Secure Shell (SSH) client or server process creating or writing to a known SSH backdoor log file.
10 Adversaries may modify SSH related binaries for persistence or credential access via patching sensitive functions to
11 enable unauthorized access or to log SSH credentials for exfiltration.
12 """
13 false_positives = ["Updates to approved and trusted SSH executables can trigger this rule."]
14 from = "now-9m"
15 index = ["auditbeat-*", "logs-endpoint.events.*"]
16 language = " eql"
17 license = "Elastic License v2"
18 name = "Potential OpenSSH Backdoor Logging Activity"
19 references = [
20     "https://github.com/eset/malware-ioc/tree/master/sshdoor",
21     "https://www.welivesecurity.com/wp-content/uploads/2021/01/ESET_Kobalos.pdf",
22 ]
23 risk_score = 73
24 rule_id = "f28e2be4-6eca-4349-bdd9-381573730c22"
25 severity = "high"
26 tags = ["Elastic", "Host", "Linux", "Threat Detection", "Persistence", "Credential Access"]
```

Detection rules

[Back to detection rules](#)

Potential SSH Brute Force Detected

Created by: elastic on Sep 26, 2021 @ 19:00:33.401 Updated by: elastic on Sep 26, 2021 @ 19:00:33.401

Last response: ● — ⏲



Edit rule settings



About

Identifies a high number (20) of macOS SSH KeyGen process executions from the same host. An adversary may attempt a brute force attack to obtain unauthorized access to user accounts.

Author Elastic



Severity Medium

Risk score 47

Reference URLs • <https://themittenmac.com/detecting-ssh-activity-via-process-monitoring/>



Definition

Index patterns

auditbeat-* logs-endpoint.events.*

Custom query

event.category:process and event.type:start and process.name:"sshd-keygen-wrapper" and process.parent.name:launchd

Rule type

Threshold

Timeline template

None

Threshold

Results aggregated by host.id >= 20

Rule monitoring

Overview **Detections** Hosts Network Timelines Cases Administration ML job settings Add data

Back to detections

Detection rules

Install 1 Elastic prebuilt rule Upload value lists Import rule Create new rule

Rules Rule Monitoring Exception Lists

All rules

Updated 28 seconds ago

e.g. rule name Tags Elastic rules (546) Custom rules (0)

Rule	Indexing Time (ms)	Query Time (ms)	Last Gap (if any)	Last run	Last response	Activated
Setgid Bit Set via chmod	—	67.48	—	19 minutes ago	● succeeded	active
SSH Authorized Keys File Modification	—	83.01	—	19 minutes ago	● succeeded	active
Sensitive Files Compression	—	45.06	—	19 minutes ago	● succeeded	active
WebProxy Settings Modification	—	58.86	—	19 minutes ago	● succeeded	active
Public IP Reconnaissance Activity	—	6.23	—	16 minutes ago	● succeeded	active
Endpoint Security	—	1.76	3 hours	15 minutes ago	● succeeded	active

Detection alerts

Security / Detections

Overview Detections Hosts Network Timelines Cases Administration Add data

Search KQL Last 24 hours Show dates Refresh

+ Add filter

Icon	Date	Action	Count	Type	Severity	Count	Icon
SSH icon	Aug 8, 2021 @ 18:35:05.568	SSH (Secure Shell) from th...	8	query	medium	47	—
SSH icon	Aug 8, 2021 @ 18:35:05.568	SSH (Secure Shell) from th...	8	query	medium	47	—
SSH icon	Aug 8, 2021 @ 18:35:05.568	SSH (Secure Shell) from th...	8	query	medium	47	—
SSH icon	Aug 8, 2021 @ 18:35:05.568	SSH (Secure Shell) from th...	8	query	medium	47	—
SSH icon	Aug 8, 2021 @ 18:35:05.568	SSH (Secure Shell) from th...	8	query	medium	47	—
SSH icon	Aug 8, 2021 @ 18:35:05.449	Telnet Port Activity	7	query	medium	47	—
SSH icon	Aug 8, 2021 @ 18:35:05.449	Telnet Port Activity	7	query	medium	47	—

Filter icons:

Detection alerts details

Query 1 Correlation Analyzer Notes Pinned

Aug 8, 2021 @ 18:51:17.343 → Aug 8, 2021 @ 19:00:17.343

All data sources

`(_id: "05a4ab275bb81c386bb7d426740984b5ea72a1b59c56bccdc357242aa5eb3962")`

OR
() + Add field

AND Filter Search KQL

+ Add filter

@timestamp ↓ 1	message	event.category	event.action	host.name	source.ip	destination.ip	user.name
Aug 8, 2021 @ 19:00:17.343	—	network_traffic	network_flow	server-02	124.133.201.110	178.62.77.103	—
		114B 2 pkts	tcp	1:edzV8NYfXWcm32jwInCxJ4shrhE=			
		33903ns					
		Jan 10, 2020 @ 05:12:41.161					
		Jan 10, 2020 @ 05:12:41.161					
		Source: 124.133.201.110 : 19373		Destination:			
		Asia CN Shandong		(50.88%) 58B	1 pkts	178.62.77.103 : 23	Europe GB England London
				< (49.12%) 56B	1 pkts		

1 of 1 events

Updated 2 minutes ago

Hosts

Overview Detections Hosts Network Timelines Cases Administration [+ Add data](#)

Search KQL Last 90 days Show dates Refresh

+ Add filter

Hosts

Last event: 1 minute ago

Data sources ▾

Hosts
3

2021-03-07 2021-04-11 2021-05-16

User authentications
✓ 11,966 success... ✗ 33 fail

Succ Fail

0 3,000 6,000 9,000

2021-03-07 2021-04-11 2021-05-16

Unique IPs
448 source 1,244 destin...

Succ Dest

0 200 400 600 800 1,000 1,200

0 200 400 600 800

2021-03-07 2021-04-11 2021-05-16

All hosts Authentications Uncommon processes Events External alerts

Hosts

Authentications



Search

KQL

Calendar

Mar 10, 2020 @ 17:28:37.66 → Mar 12, 2020 @ 17:27:50.81

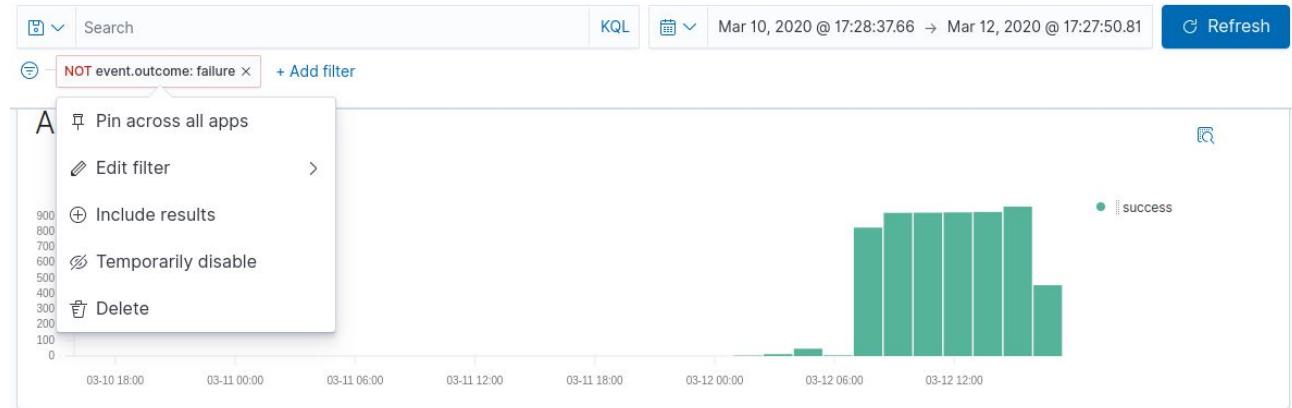
Refresh

NOT event.outcome: failure X + Add filter

Authentications

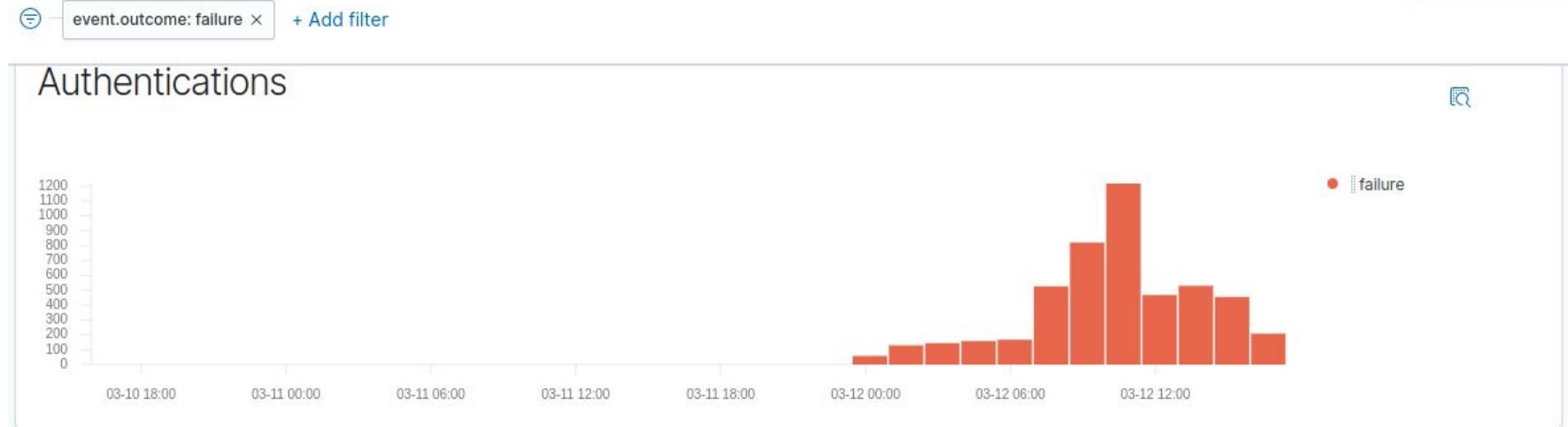


Hosts



event.outcome: failure × + Add filter

Authentications



Hosts

User	Successes	Failures	Last success	Last successful source	Last successful destin...	Last failure	Last failed source	Last failed destination
root	0	8911	—	—	—	11 minutes ago	49.88.112.112	workstation-03
admin	0	591	—	—	—	12 minutes ago	171.5.221.56	server-02
test	0	313	—	—	—	23 minutes ago	27.47.25.242	workstation-03
oracle	0	237	—	—	—	20 minutes ago	185.184.24.33	webserver-01
user	0	236	—	—	—	30 minutes ago	77.35.17.149	server-02
git	0	204	—	—	—	11 minutes ago	52.130.85.47	server-02
ftouser	0	188	—	—	—	29 minutes ago	77.35.17.149	server-02

Hosts

159.65.100.137

As Source ▾

Last event: Aug 25, 2021 @ 13:07:40.928

Data sources ▾

Location	First seen	Host ID	Whois
New York, New York	Jun 23, 2021 @ 19:40:28.199	—	iana.org ↗
Autonomous system	Last seen	Host name	Reputation
DigitalOcean, LLC / 14061	Aug 25, 2021 @ 13:07:40.928	—	virustotal.com ↗ , talosIntelligence.com ↗
Max anomaly score by job			

Hosts

Events



Showing: 6.851.232 events

@timestamp	message	host.name	event.module	event.dataset	event.action
Mar 12, 2020 @ 21:08:45.000	Process ended.	james-honeypot-logstash-d...	endgame	esensor	termination_event
	root @ james-honeypot-logstash-demo terminated process > / (30958)	/usr/sbin/unix_chkpwd	root	nonull	with exit code 7

Mar 12, 2020 @ 21:08:45.000	—	james-honeypot-logstash-d...	endgame	esensor	file_create_event
	root @ james-honeypot-logstash-demo created a file imjournal.state.tmp in /var/lib/rsyslog/imjournal.state.tmp via > rsyslogd	/var/lib/rsyslog/imjournal.state.tmp	rsyslogd	/var/lib/rsyslog/imjournal.state.tmp	via /var/lib/rsyslog/imjournal.state.tmp

Table

JSON View



Filter by Field, Value, or Description...

Field	Value	Description
<input checked="" type="checkbox"/> @timestamp	Mar 12, 2020 @ 21:08:45.000	Date/time when the event originated. This is the date/time extracted from the event, typically representing when the event was generated by the source. If the event source has no original timestamp, this value is typically populated by the first time the event was received by the pipeline. Required field for all events. Example: 2016-05-23T08:05:34.853Z
<input type="checkbox"/> _id	1Jd0HABcprL3sKu3CGj	Each document has an _id that uniquely identifies it Example: Y-6TfmcB0WOhS6qyMv3s
<input type="checkbox"/> _index	endgame-4.21.0-2020.03.12	An index is like a 'database' in a relational database. It has a mapping which defines multiple types. An index is a logical namespace which maps to one or more primary shards and can have zero or more replica shards. Example: auditbeat-8.0.0-2019.02.19-000001
<input type="checkbox"/> _score	1	

Network

Overview Detections Hosts **Network** Timelines Cases Administration + Add data

Search KQL Last 24 hours Show dates Refresh

+ Add filter

Source countries

Showing: 131 Countries

Country	Bytes in	Bytes out ↓	Flows	Source IPs	Destination IPs
United States of America	1.2GB	9.9GB	31,552	1,215	5
United Kingdom	13.5GB	4.3GB	16,132	105	59
Germany	6.1MB	95.9MB	746	60	4
China	69.1MB	45.4MB	4,533	814	3
Italy	1.5MB	18.2MB	138	71	4
Brazil	3.9MB	2.2MB	390	179	3

Destination countries

Showing: 17 Countries

Country	Bytes in ↓	Bytes out	Flows	Destination IPs	Source IPs
United States of America	4.3GB	13.4GB	5,632	19	1
United Kingdom	539.1MB	743.3MB	42,510	6	4,222
Germany	6.4MB	2.8MB	10	1	1
Netherlands	46.3KB	110.1KB	68	6	1
Australia	16.3KB	41.1KB	8	1	1
Japan	140B	2.2KB	1	1	1

Flows

Flows DNS HTTP TLS External alerts

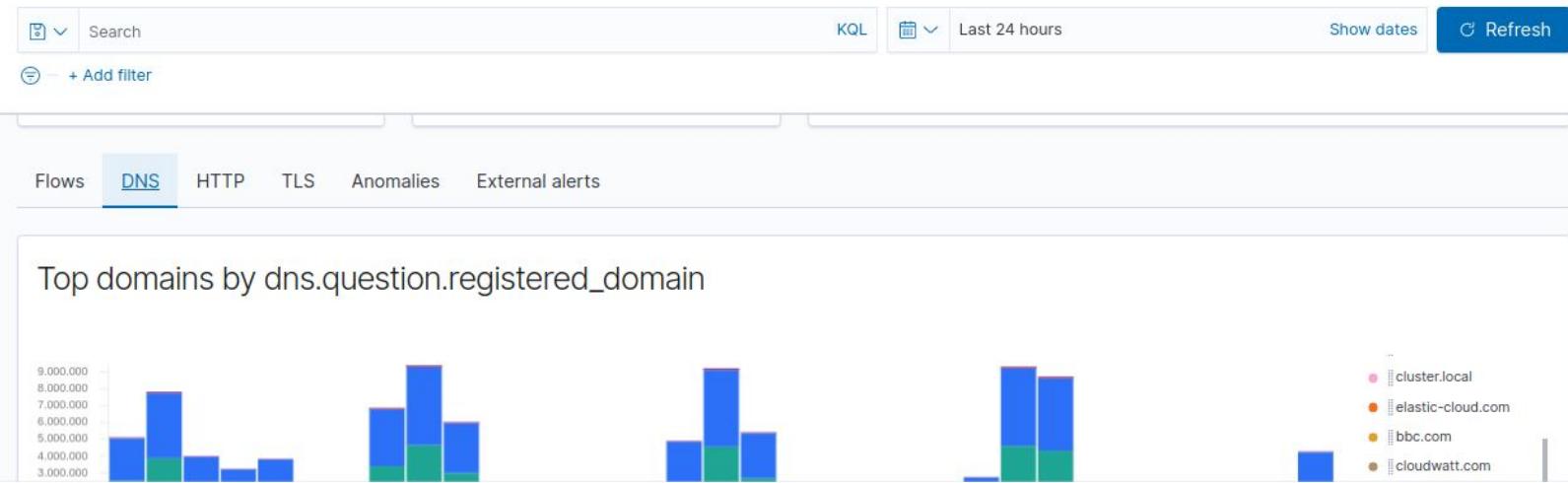
Source IPs



Showing: 277 IPs

IP	Domain	Autonomous system	Bytes in	Bytes out ↓	Flows	Destination IPs
127.0.0.1	—	—	3GB	172.7GB	37,189	2
185.255.105.40	—	—	35.8MB	425.2MB	90	1
::1	—	—	1.5MB	8.8MB	17,887	1

DNS



Top DNS domains ⓘ

Showing: 97 domains

Include PTR records

Registered domain	Total queries	Unique domains ↓	DNS bytes in	DNS bytes out
google.internal	33484506	22	2.1GB	4.4GB
elastic-product.internal	33483883	21	2.4GB	4.6GB
elastic.co	301815	7	7.7MB	23.2MB
es.io	492767	4	30.8MB	77MB

CASO PRÁCTICO: LOG DE SERVIDOR APACHE

- Introducción al proyecto
- Montando el servidor de apache
- Ingestar y transformar los logs con Logstash
- Indexarlos en ElasticSearch
- Visualizarlos en Kibana

<https://www.elastic.co/es/demos>

THE ELASTIC STACK

Elastic Demo Gallery

Little examples designed to let you explore various facets of the Elastic Stack, from Kibana dashboards and Canvas workpads to Elasticsearch SQL snippets and machine learning jobs.

Because trying is better than seeing.

[Launch Demos](#)