The Magic of Search + Location

Nathan Reese | Kibana Engineer

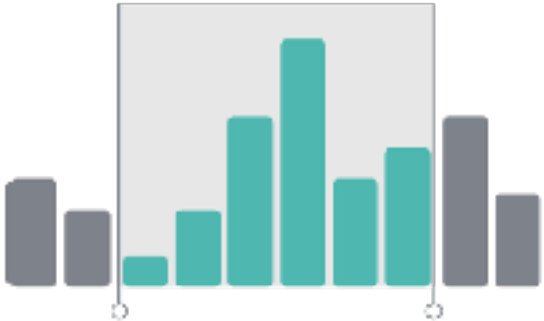
Nov 12th 2019

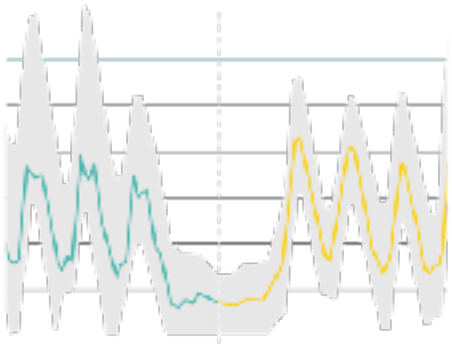
About me

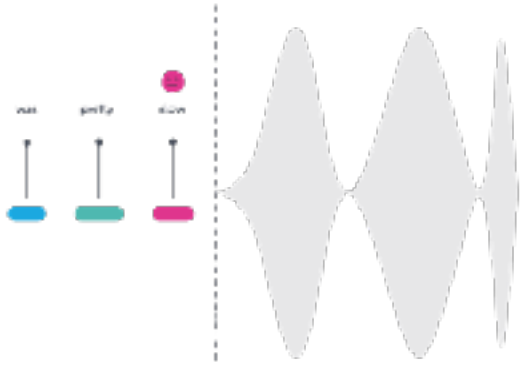
- Software Engineer at Elastic
- Kibana + Elastic Maps
- https://github.com/nreese

elastic

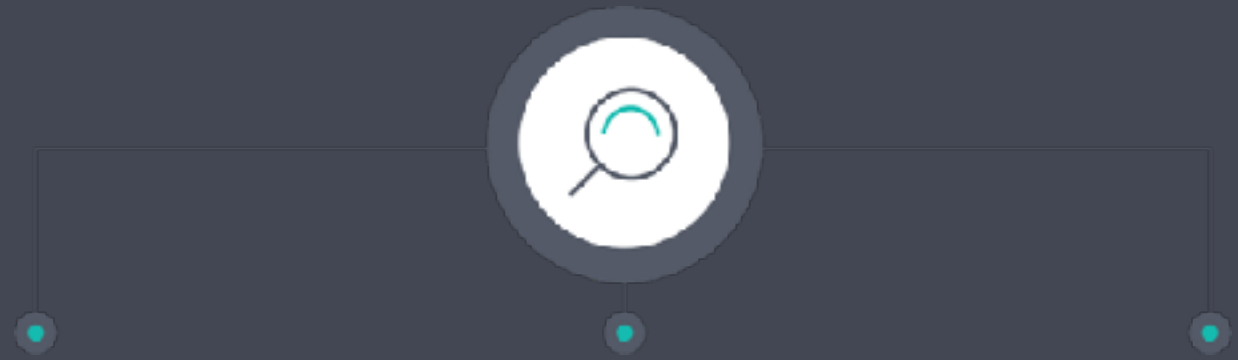
Search is a **constant/foundation**







Technology **differentiation**



SCALE

Distributed by design

SPEED

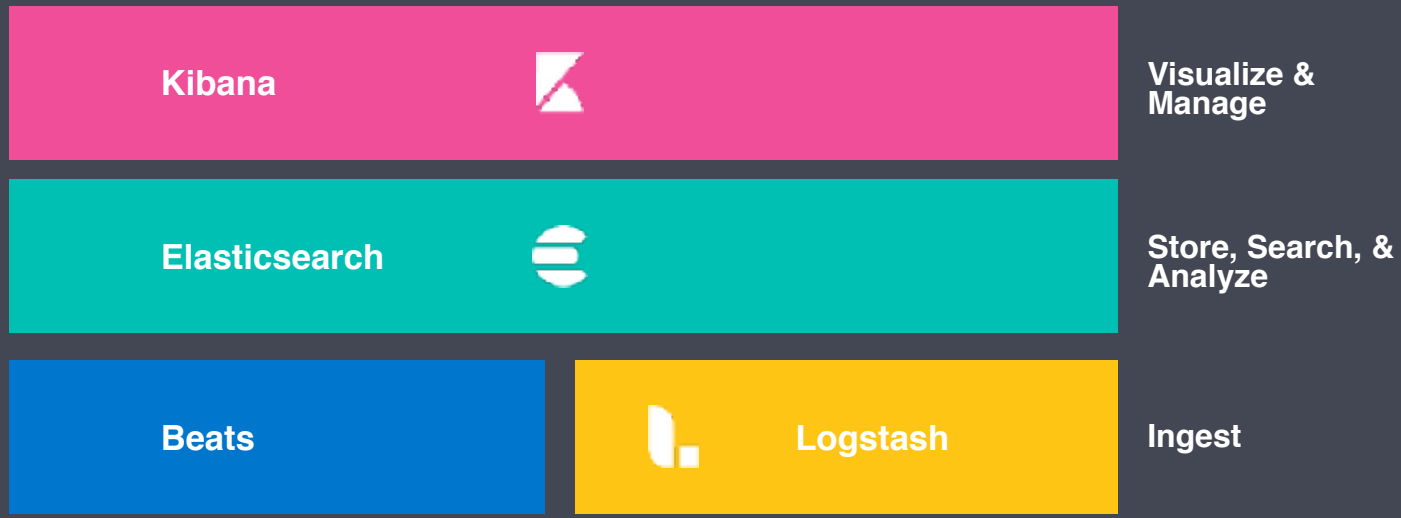
Find matches in milliseconds

RELEVANCE

Get highly relevant results

elastic

Elastic **Stack**



Elasticsearch

Elasticsearch is a distributed, open source search and analytics engine for all types of data, including textual, numerical, geospatial, structured, and unstructured.

elastic

16 AUGUST 2010

Geo Location and Search

By **Shay Banon**

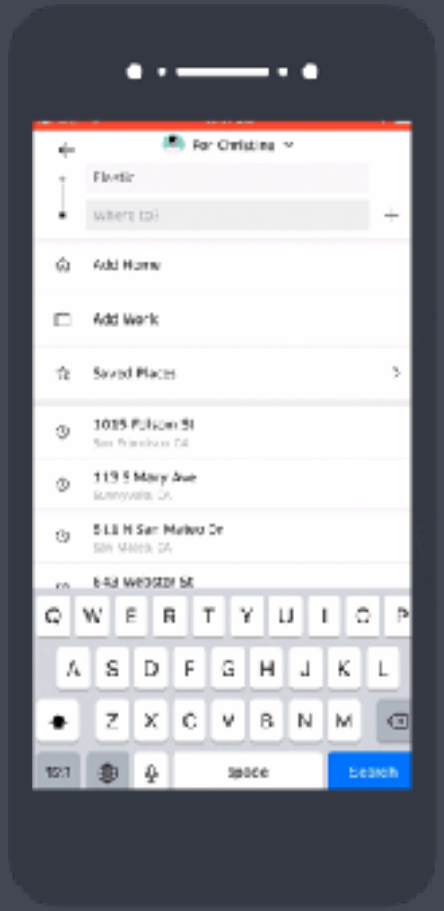
Share

One of the coolest search technology combinations out there are the ability to combine geo and search. Queries such as give me all the restaurants that serves meat ([insert your query here]) within 20 miles from me, or create a distance heat map of them, is slowly becoming a must have for any content website. This is becoming even more relevant with new browsers supporting **Geolocation API**.

Already in **master** (and in the upcoming `0.9.1` release), elasticsearch comes with rich support for geo location. Lets take a drive down the geo support path:

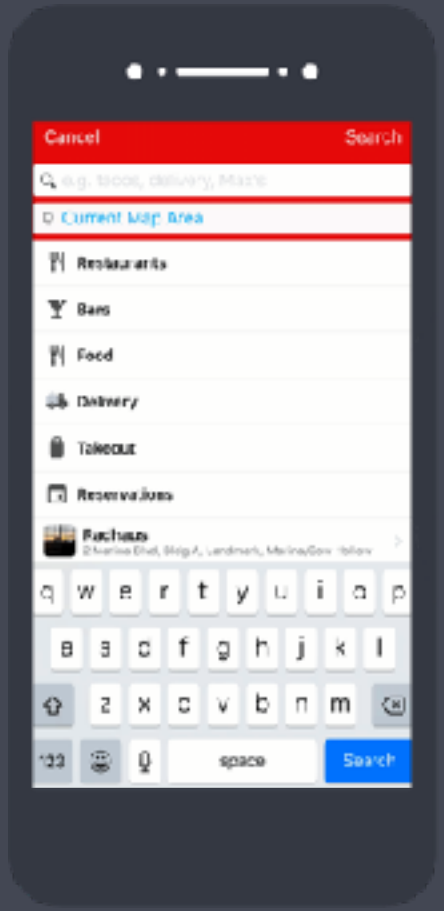
https://www.elastic.co/blog/geo-location-and-search

elastic



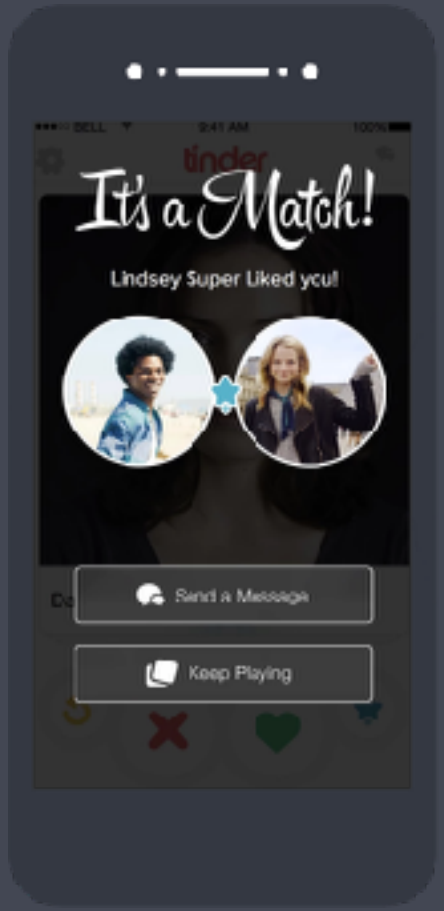
Uber

elastic



yelp

elastic



tinder

elastic

Index mapping

```
PUT /cities
{
  mappings: {
    properties: {
      center: { type: "geo_point" },
      outline: { type: "geo_shape" },
      name,
      state,
      population
    }
  }
}
```


elastic

Inserting documents

```
POST /cities/_doc
{
  center: { lat: 32.7157, lon: 117.1611 },
  outline: {
    type : "polygon",
    coordinates: [
      [lon, lat], [lon, lat], ...
    ]
  },
  name: "San Diego",
  state: "CA",
  population: 1420000,
}
```


elastic

Geo-bounding box query

```
GET cities/_search
{
  query: {
    bool: {
      must: { match_all: {} },
      filter: {
        geo_bounding_box: {
          center: {
            top_left: { lat: 40.73, lon: -74.1 },
            bottom_right: { lat: 40.01, lon: -71.12 }
          }
        }
      }
    }
  }
}
```


elastic

Geo-shape query

```
GET /cities/_search
{
  query:{
    bool: {
      must: { match_all: {} },
      filter: {
        geo_shape: {
          outline: {
            shape: {
              type: envelope,
              coordinates : [[13.0, 53.0], [14.0, 52.0]]
            },
            relation: within
          }
        }
      }
    }
```


elastic

geo-shape spatial relation operators

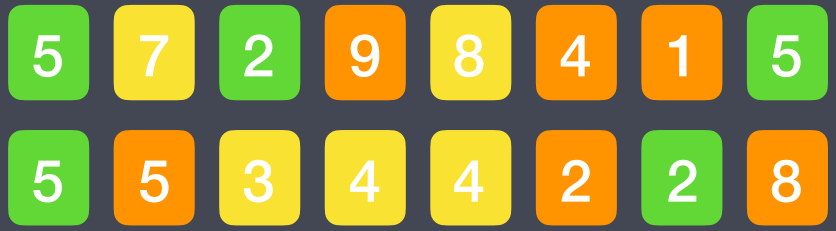
- Intersects
- Disjoint
- Within
- Contains

elastic

Aggregations

elastic

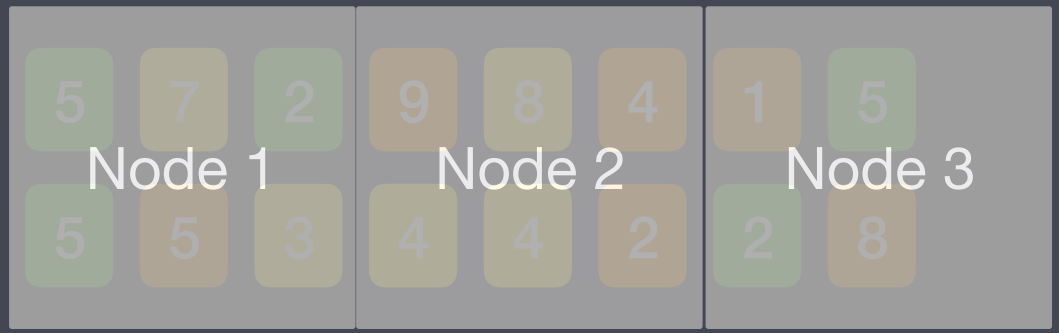
Aggregations



```
POST /my_index/_search?size=0
{
  aggs: {
    my_bucket_agg: {
      terms: {
        field: "color"
      },
      aggs: {
        my_metric_agg: {
          avg: {
            field: "my_field"
          }
        }
      }
    }
  }
}
```


elastic

Aggregations



avg: 3.8

avg: 5.2

avg: 4.8

elastic

Nested aggs

```
POST /my_index/_search?size=0
{
  aggs: {
    my_bucket_agg: {
      terms: {
        field: "color"
      },
      aggs: {
        my_nested_bucket_agg: {
          terms: {
            field: "shape"
          },
          aggs: {
            my_metric_agg: {
              avg: {
              field: "my_field"
              }
            }
          }
        }
      }
    }
  }
}
```


elastic

GeoTile Grid Aggregation

```
POST /my_index/_search?size=0
{
  aggs: {
    my_bucket_agg: {
      geotile_grid: {
        field: "location",
        precision: 13
      },
      aggs: {
        my_metric_agg: {
          avg: {
            field: "my_field"
          }
        }
      }
    }
  }
}
```

elastic

Geo Centroid Aggregation

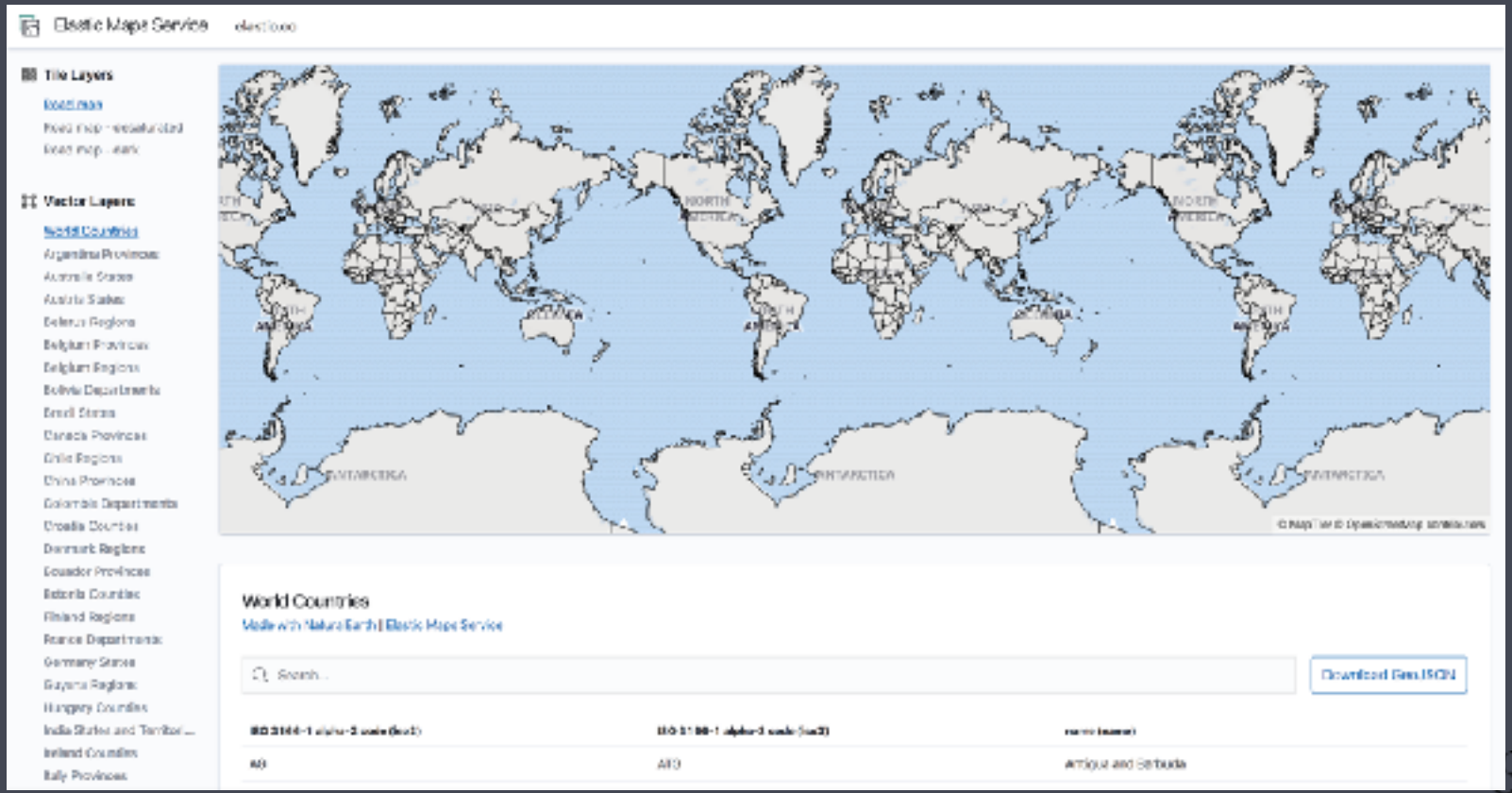
```
{
  aggs: {
    my_bucket_agg: {
      geotile_grid: {
        field: "location",
        precision: 13
      },
      aggs: {
        my_metric_agg: {
          avg: {
            field: "my_field"
          }
        },
        my_other_metric_agg: {
          geo_centroid: {
            field: "location"
          }
        }
      }
    }
  }
}
```

elastic

Choropleth maps

elastic

Elastic Maps Service
https://maps.elastic.co



GeoIP Processor

```
GET my_logs/_doc/1
{
  _index: "my_logs",
  _id: "1",
  _source : {
    client: {
      ip: "33.217.174.45",
    }
  }
}
```

elastic

GeoIP Processor

```
GET my_logs/_doc/1
{
  _index: "my_logs",
  _id: "1",
  _source : {
    client: {
      ip: "33.217.174.45",
      geo: {
        country_iso_code: “US”,
        location: {
          lat: 45.505918,
          lon: -73.614830
        }
      }
    }
```

elastic

Choropleth maps

US 5	CA 6	CA 8	MX 1	MX 5
US 7	CA 9	MX 4	MX 2	MX 3

US avg: 6

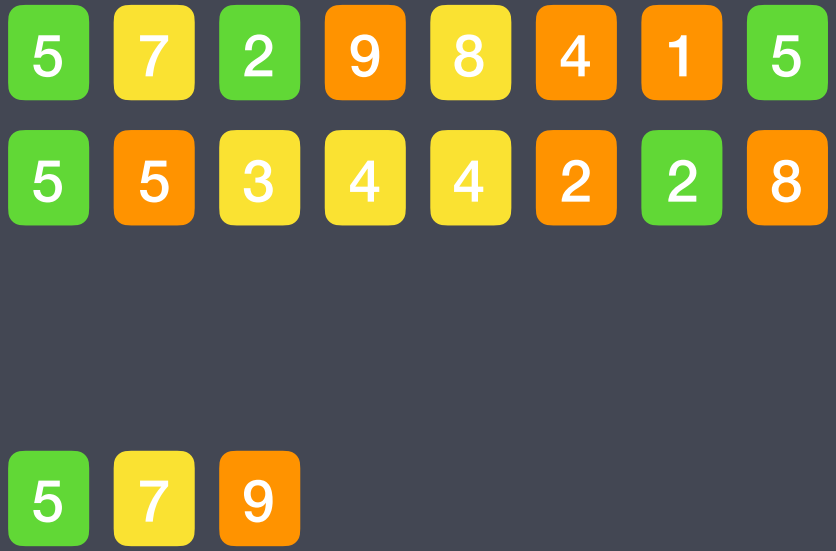
CA avg: 8

MX avg: 3

```
POST /my_logs/_search?size=0
{
  aggs: {
    my_bucket_agg: {
      terms: {
        field: "country_iso_code",
      },
      aggs: {
        my_metric_agg: {
          avg: {
            field: "my_field"
          }
        }
      }
    }
  }
}
```

elastic

Top hits aggregations




```
POST /my_index/_search?size=0
{
  aggs: {
    my_bucket_agg: {
      terms: {
        field: "color"
      },
      aggs: {
        my_metric_agg: {
          top_hits: {
            size: 1,
            sort: {
              timestamp: {
                order: "desc"
              }
            }
          }
```


elastic

Metric aggregations

https://www.elastic.co/guide/en/elasticsearch/reference/current/search-aggregations-metrics.html

- Avg Aggregation
- Weighted Avg Aggregation
- Cardinality Aggregation
- Extended Stats Aggregation
- Geo Bounds Aggregation
- Geo Centroid Aggregation
- Max Aggregation
- Min Aggregation
- Percentiles Aggregation
- Percentile Ranks Aggregation
- Scripted Metric Aggregation
- Stats Aggregation
- Sum Aggregation
- Top Hits Aggregation
- Value Count Aggregation
- Median Absolute Deviation Aggregation

Bucket aggregations

https://www.elastic.co/guide/en/elasticsearch/reference/current/search-aggregations-bucket.html

- Adjacency Matrix Aggregation
- Auto-interval Date Histogram Aggregation
- Children Aggregation
- Composite Aggregation
- Date Histogram Aggregation
- Date Range Aggregation
- Diversified Sampler Aggregation
- Filter Aggregation
- Filters Aggregation
- Geo Distance Aggregation
- GeoHash grid Aggregation
- GeoTile Grid Aggregation
- Global Aggregation
- Histogram Aggregation
- IP Range Aggregation
- Missing Aggregation
- Nested Aggregation
- Parent Aggregation
- Range Aggregation
- Rare Terms Aggregation
- Reverse nested Aggregation
- Sampler Aggregation
- Significant Terms Aggregation
- Significant Text Aggregation
- Terms Aggregation
- Subtleties of bucketing range fields

elastic

Pipeline aggregations

https://www.elastic.co/guide/en/elasticsearch/reference/current/search-aggregations-pipeline.html

- Avg Bucket Aggregation
- Derivative Aggregation
- Max Bucket Aggregation
- Min Bucket Aggregation
- Sum Bucket Aggregation
- Stats Bucket Aggregation
- Extended Stats Bucket Aggregation
- Percentiles Bucket Aggregation
- Moving Average Aggregation
- Moving Function Aggregation
- Cumulative Sum Aggregation
- Cumulative Cardinality Aggregation
- Bucket Script Aggregation
- Bucket Selector Aggregation
- Bucket Sort Aggregation
- Serial Differencing Aggregation

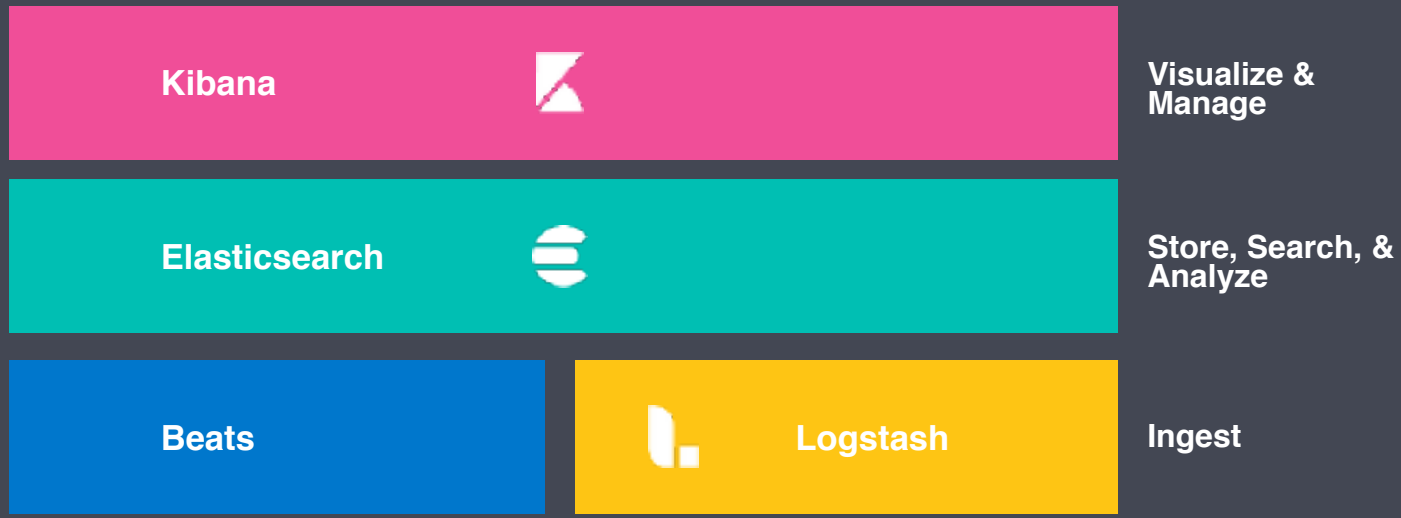
Matrix aggregations

https://www.elastic.co/guide/en/elasticsearch/reference/current/search-aggregations-matrix.html

Matrix Stats

elastic

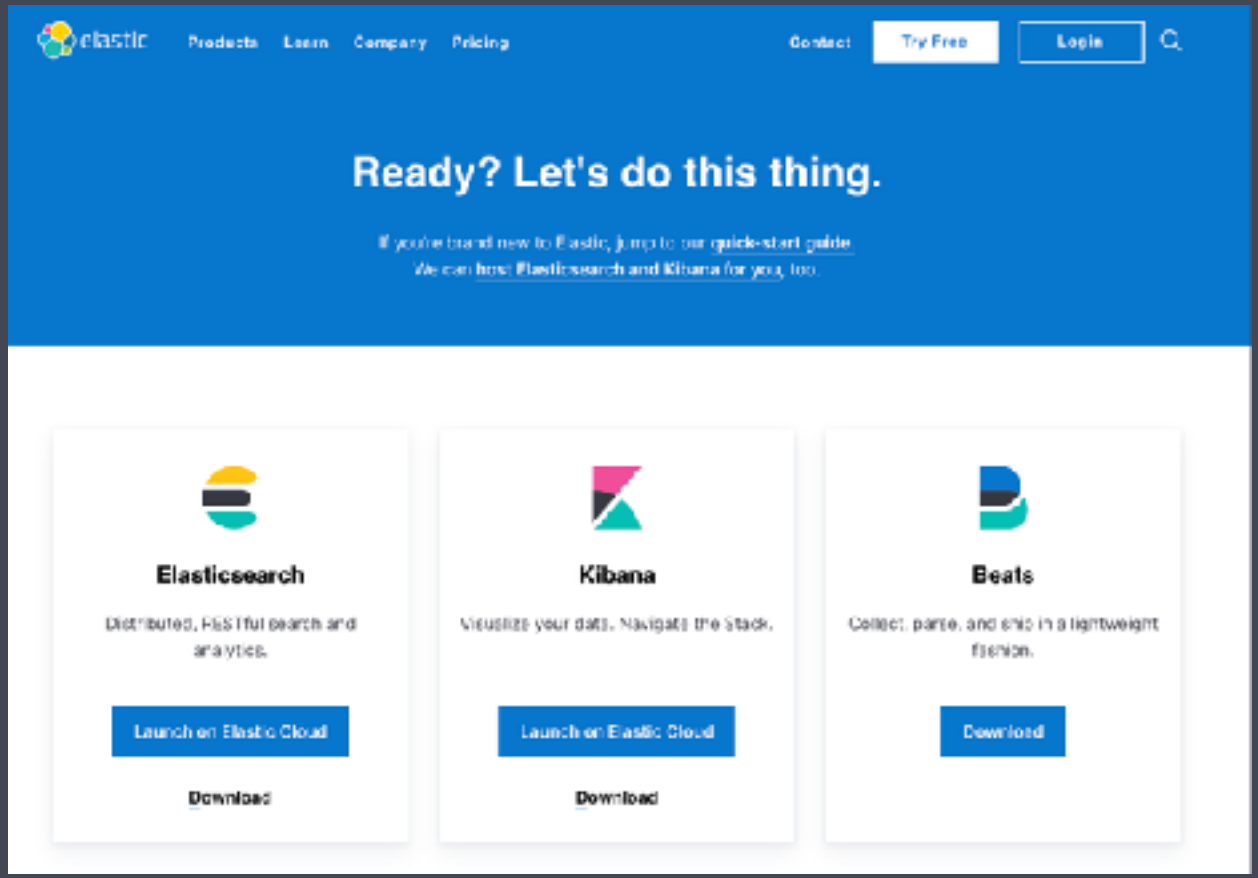
Elastic **Stack**



elastic

Try the elastic stack today

https://www.elastic.co/downloads/



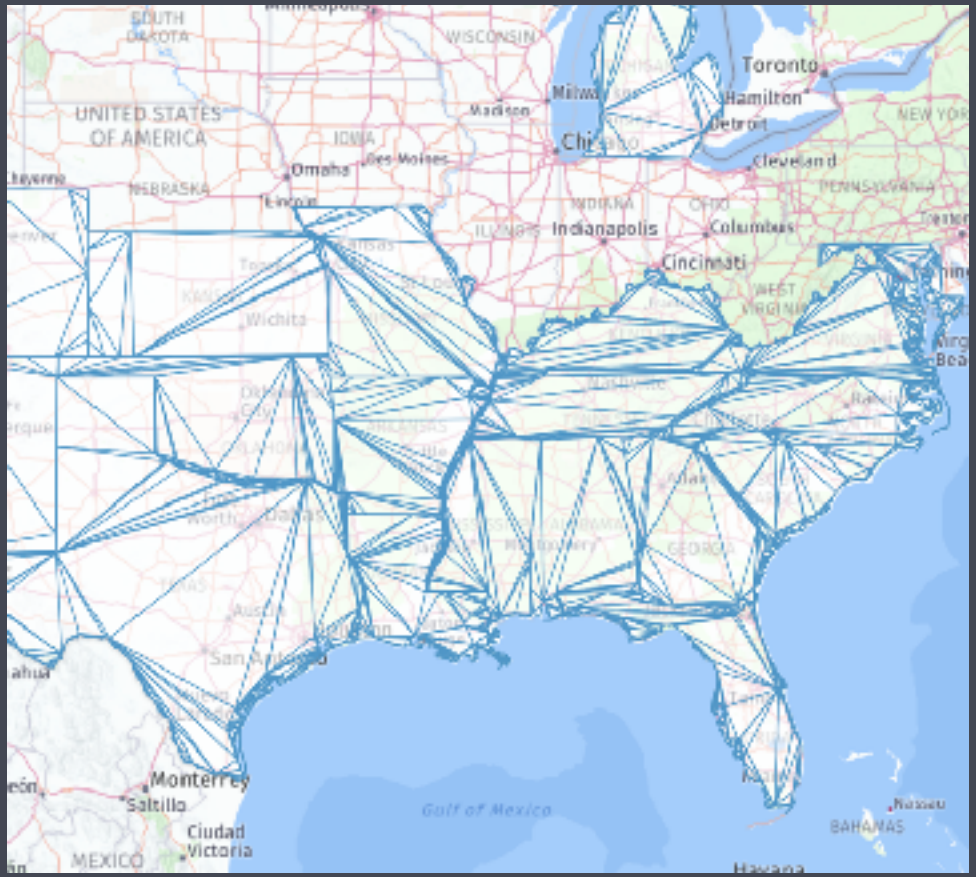
elastic

Thank you

Bkd-Backed Geoshape

Much more efficient geoshape, similar to the work on geopoints in 5.0

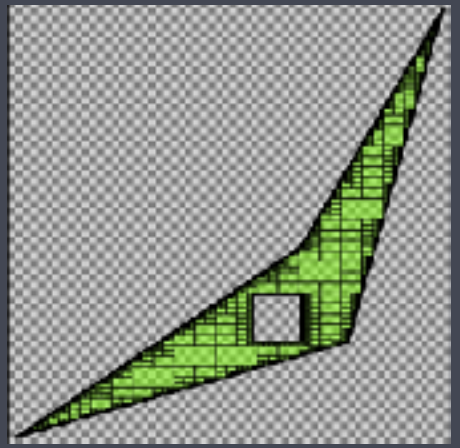
- Faster indexing
- Smaller index size
- Faster box queries



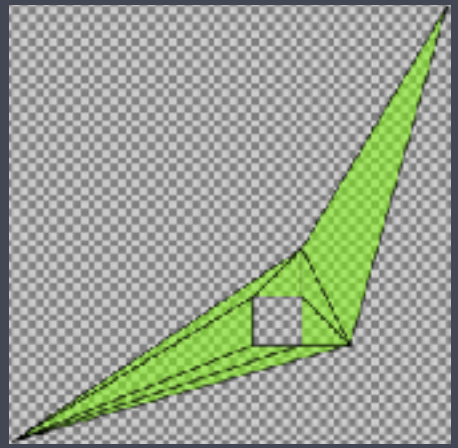
geo_shape indexing

“triangle” encoding (Bkd-tree)

- LatLonShape - new Lucene Field Type (Lucene 7.6)
  - Lat/Lon geometry represented using Triangular Irregular Network (TIN)



Quad Tree decomposition
1,105,889 quad terms at
3 meter spatial resolution



TIN decomposition
8 triangle terms at
~1.11 centimeter spatial resolution

OSS

GA

Distance Feature Query

Geo and Time Proximity: closer is frequently more relevant

- Benefits:
  - Newer data is frequently more relevant
  - Geographical proximity as a relevance indicator in non-geo searches

- Easy to use, sophisticated under the hood:
  - Designed to leverage top-k faster queries for improved performance
  - Supports normalization through Saturation

$$Score = Boost \cdot \frac{Pivot}{Pivot + Distance}$$

```
GET items/_search
{
  "query": {
    "bool": {
      "must": {
        "match": {
          "name": "chocolate"
        }
      },
      "should": {
        "distance_feature": {
          "field": "location",
          "pivot": "1000m",
          "origin": [-71.3, 41.15]
        }
      }
    }
  }
}
```

elastic

OSS

GA

Search As You Type

Get relevant results while you type the query

- Frequently requested
- Easy to configure - balance index size with num of shingles
- Mid-term: infix (in addition to prefix)
- Top k faster queries

elastic