

# ELK STACK

Meltem YILMAZ

Big Data Engineer

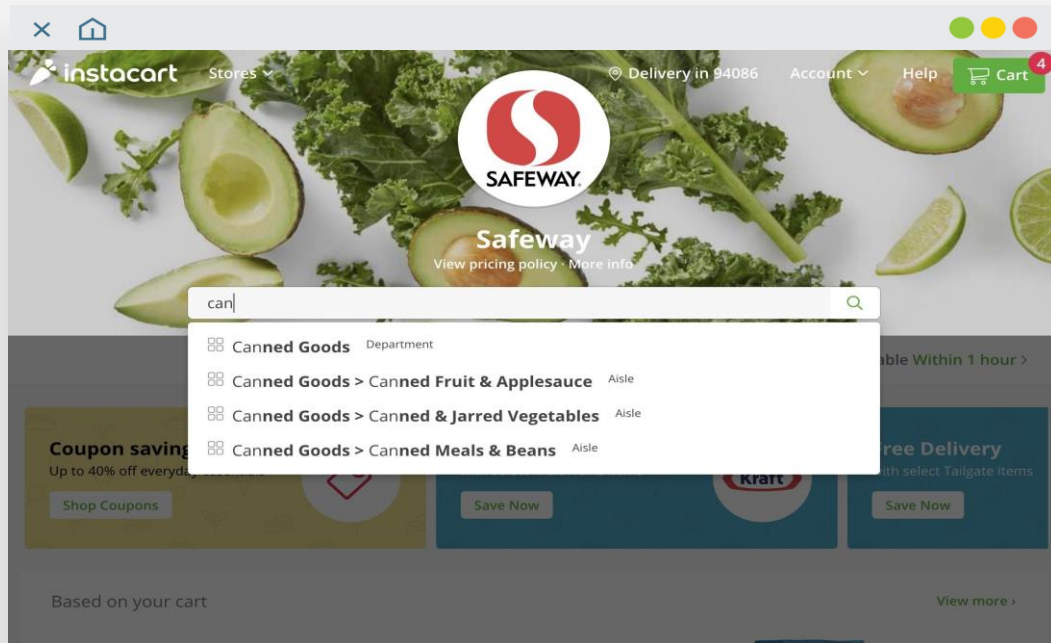
[meltemyilmaz@havelan.com.tr](mailto:meltemyilmaz@havelan.com.tr)

# The Elastic Stack

Reliably and securely take data from any source, in any format, then search, analyze, and visualize it in real time.



Great Search Experience = Get **fast and relevant results, no matter the scale.**



**Elasticsearch**

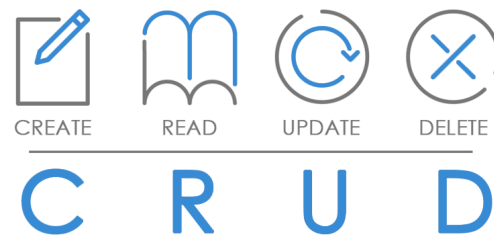
Store | Search | Analyze



elasticsearch

# Etkinliğin Sonunda Neler Yapabiliyor Olacağız?

- Elasticsearch ve Kibana'nın kullanım alanlarını anlamak
- Elasticsearch mimarisini anlamak
- Logstash ile log dosyası okumak
- Elasticsearch ve Kibana ile CRUD operasyonlarını gerçekleştirmek



**Elasticsearch**

Store | Search | Analyze



elasticsearch

Elasticsearch, **Lucene** tabanlı bir arama sunucusudur.

- Java ile geliştirilmiştir.
- Apache Lisansı ile yayınlanmıştır.
- Full-text arama yapabilmenizi sağlar.
- Apache Solr'dan sonra gelen ikinci en popüler arama motorudur.

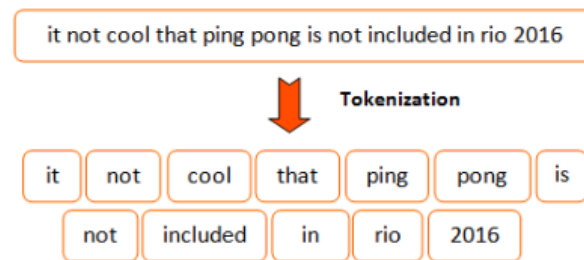
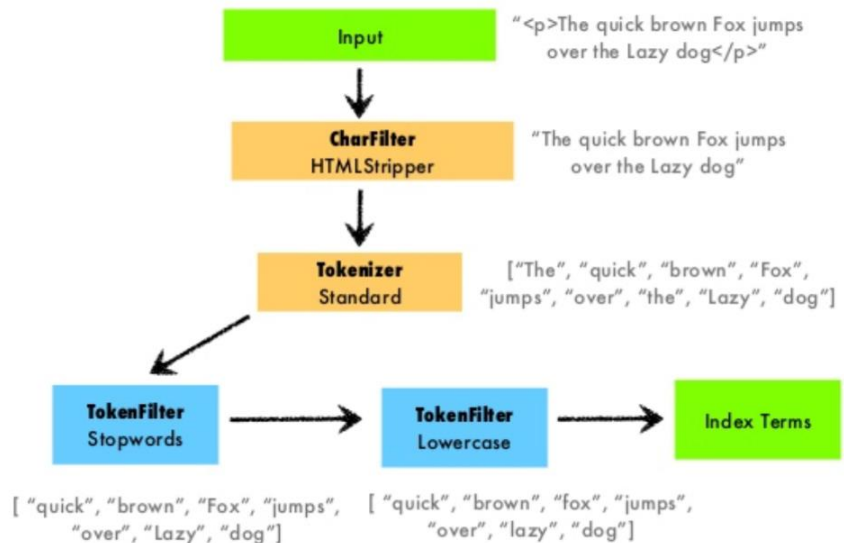


**Elasticsearch**

Store | Search | Analyze



elasticsearch



Elasticsearch'e gönderilen dokümanlar maksimum performans ve amaca hizmet edebilen bir search yapısı oluşturabilmek için, indexlenmeden önce bir dizi işlemten geçer. İşlemi yapan kısım ise **Analyzer** olarak adlandırılır.

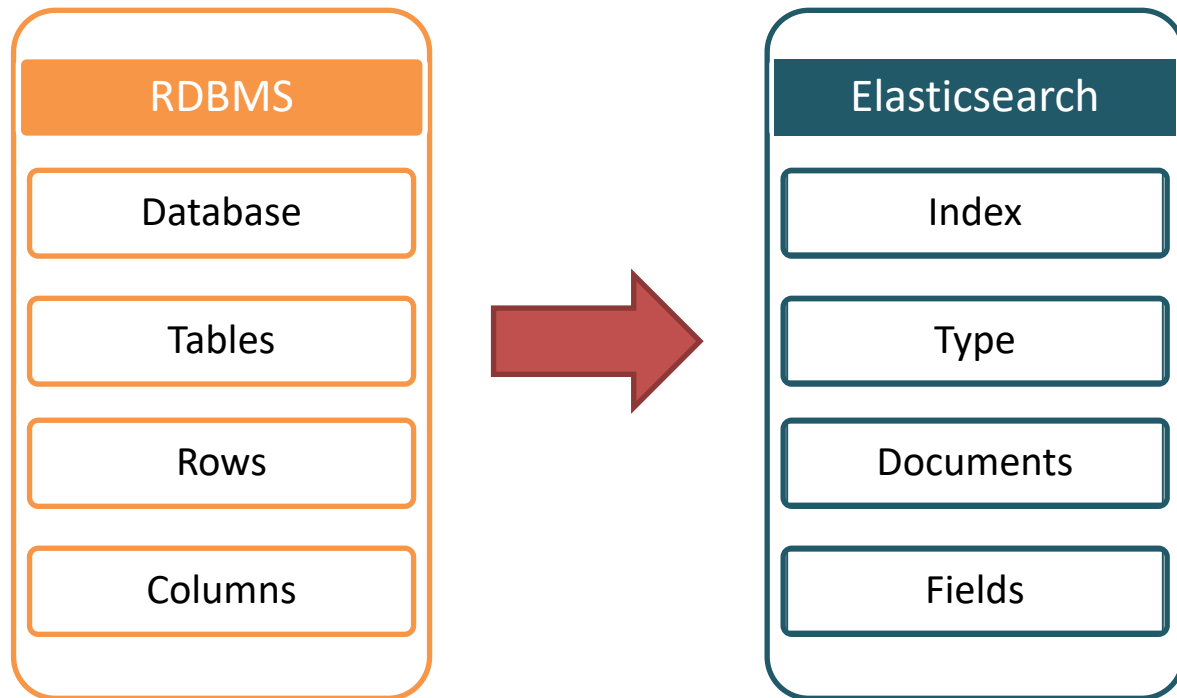
# Elasticsearch

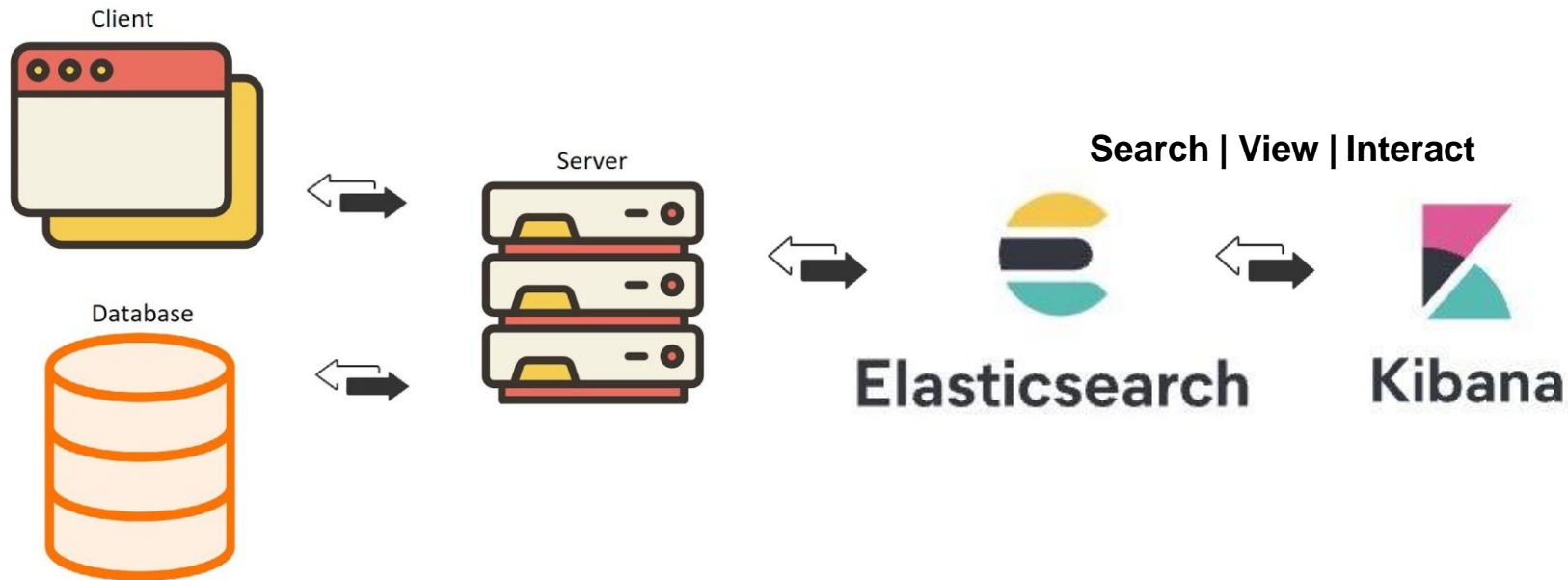
Store | Search | Analyze



elasticsearch

# RDBMS & Elasticsearch





# Elasticsearch

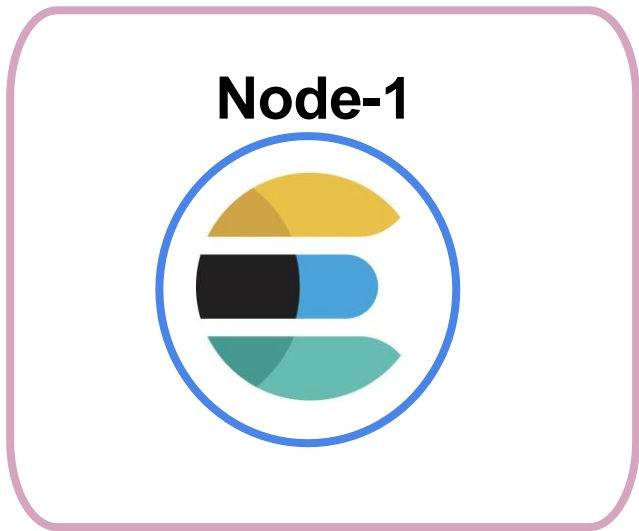
Store | Search | Analyze



elasticsearch



# Single Node Cluster



- Bir node Elasticsearch'ün bir instance'ıdır.
- Her node'un kendine has bir adı var.
- Bir node'dan oluşan Cluster'a Single Cluster denilir.

**Elasticsearch**

Store | Search | Analyze



elasticsearch

# Cluster

**Node-1**



**Node-2**



**Node-3**



**Node-4**



# Elasticsearch & Data

```
{  
  "name": "Baby Carrots(1lb bag)",  
  "category": "Vegetables",  
  "brand": "365",  
  "price": "$0.99"  
}
```

- Veriler, Elasticsearch'te doküman olarak saklanır.
- Dokümanlar ise bir JSON nesnesidir.

**Elasticsearch**

Store | Search | Analyze



elasticsearch

# Dokümanlar bir index içerisinde gruplanır.

## Produce Index

```
{
  "name": "Baby Carrots(1lb bag)",
  "category": "Vegetables",
  "brand": "365",
  "price": "$0.99"
}

{
  "name": "Clementines(3lb bag)",
  "category": "Fruits",
  "brand": "Cuties",
  "price": "$4.29"
}
```

## Wine & Beer Index

```
{
  "name": "Unanime Malbec(750ml)",
  "brand": "Mascota Vineyards",
  "country": "Argentina",
  "region": "Mendoza",
  "wine_type": "Red Wine",
  "ABV": "14%",
  "price": "$22.99"
}

{
  "name": "Hazy Little Thing IPA(750ml)",
  "country": "US",
  "state": "California",
  "beer_type": "Ale",
  "beer_style": "India Pale Ale",
  "ABV": "6.7%",
  "price": "$14.99"
}
```

Elasticsearch

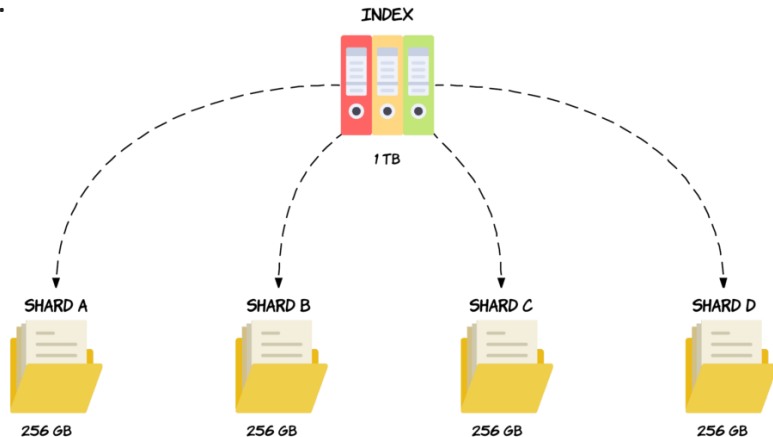
Store | Search | Analyze



elasticsearch

# Shard

- Bir *Index* tek başına bir ya da daha fazla *Shard*'dan oluşabilir.
- Genellikle *Index* in büyüyebilmesi ve bir kaç node üzerinde dağıtılabilmesi için bir kaç *Shard*'dan oluşmaktadır.
- Primary Shard bir doküman için ev sahipliği yapmaktadır.
- Bir Replica Shard ise Primary shard bir kopyasıdır.
- Primary Shard «fail over» olduğunda ya da okuma çok arttığında replica shard kullanılmaktadır.



## Elasticsearch

Store | Search | Analyze



elasticsearch

# Shard Nedir?

## Cluster

Node-1

Node-2

Node-3

Produce Index

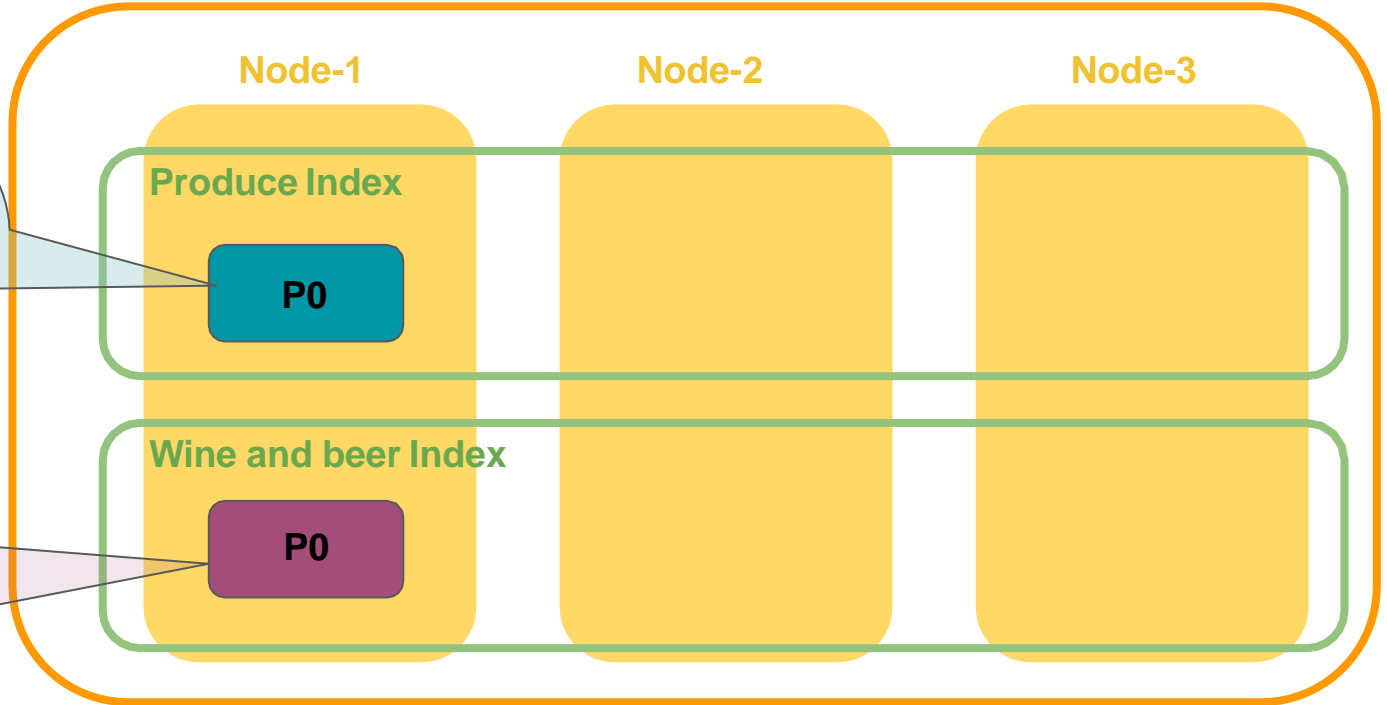
P0

Wine and beer Index

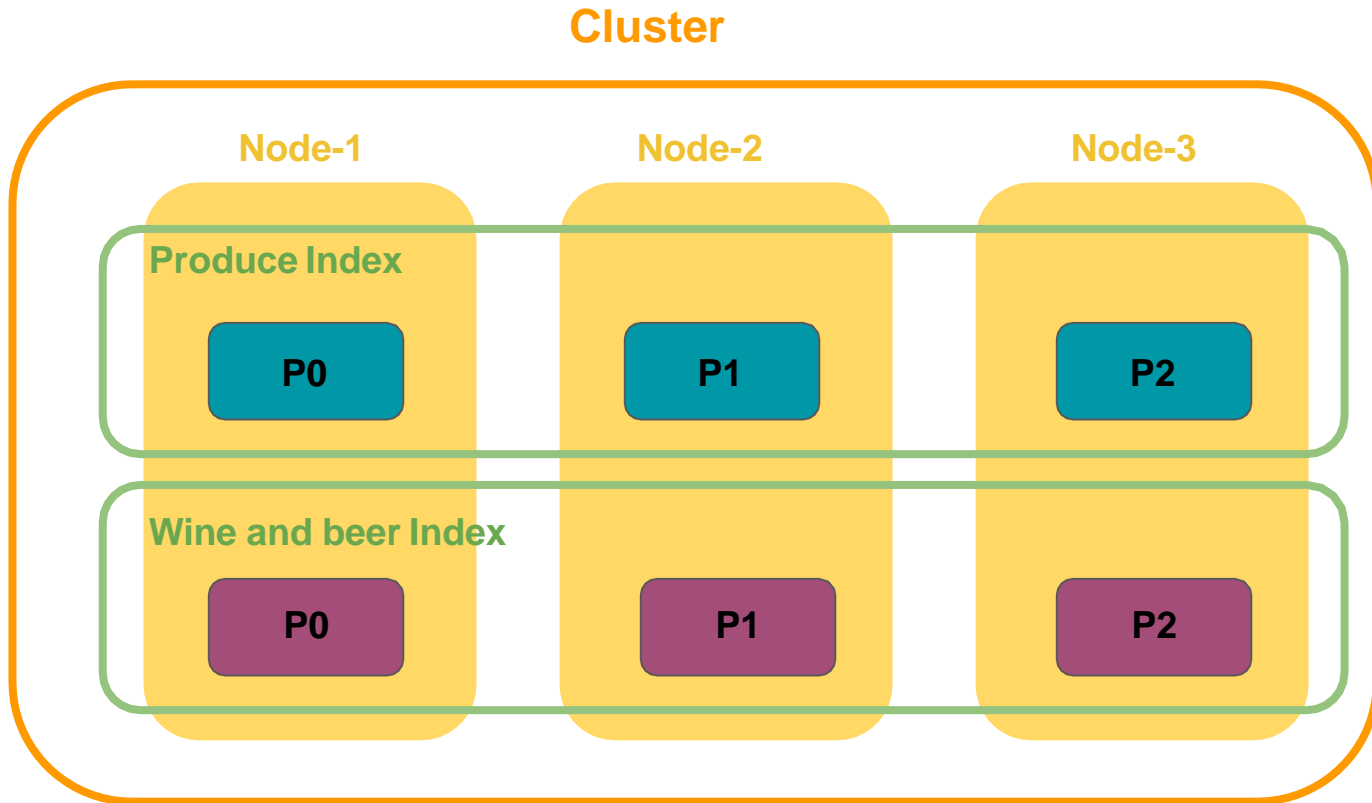
P0

Merhaba,  
Ben bir Shard'ım.  
Produce\* ile ilgili  
dokümanları  
tutarım.

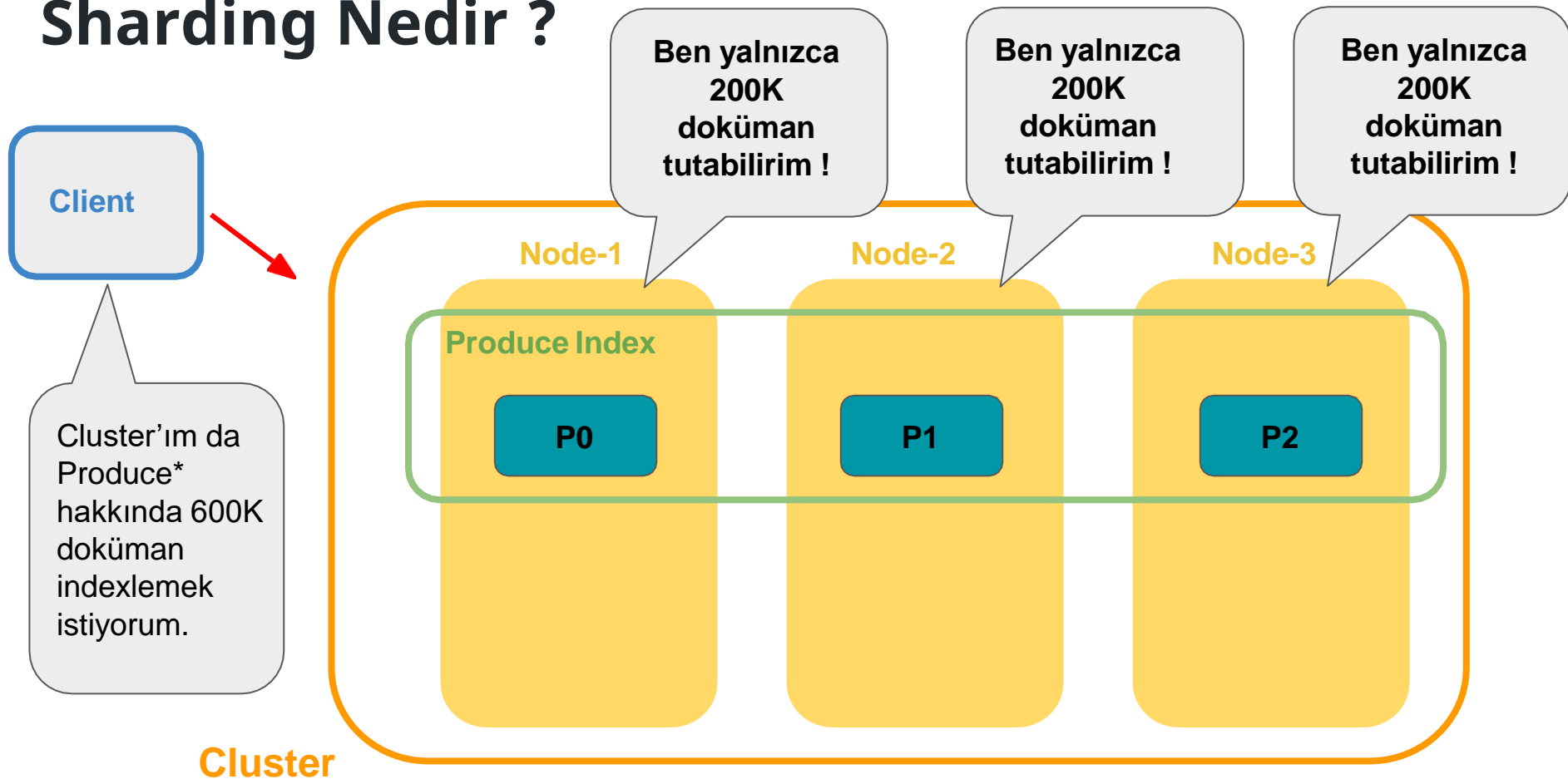
Merhaba,  
Ben bir Shard'ım.  
Beer\* ile ilgili  
dokümanları  
tutarım.



# Sharding Nedir ?



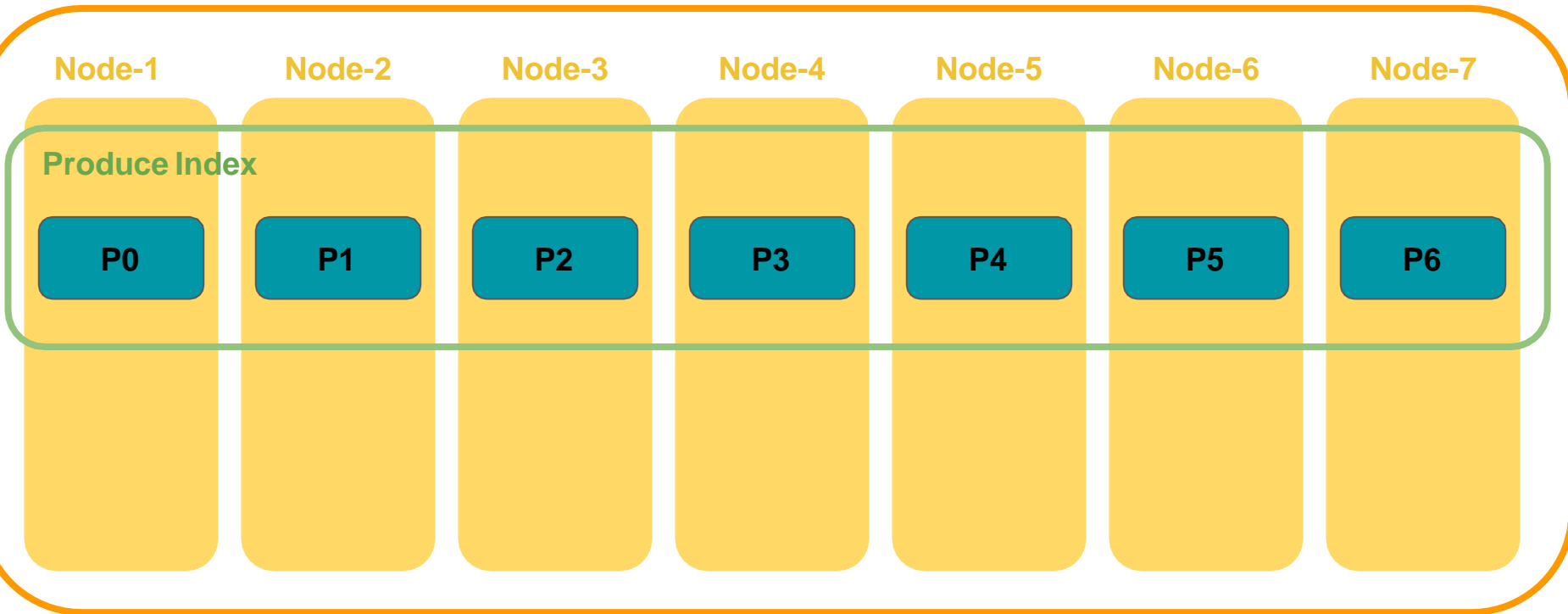
# Sharding Nedir ?



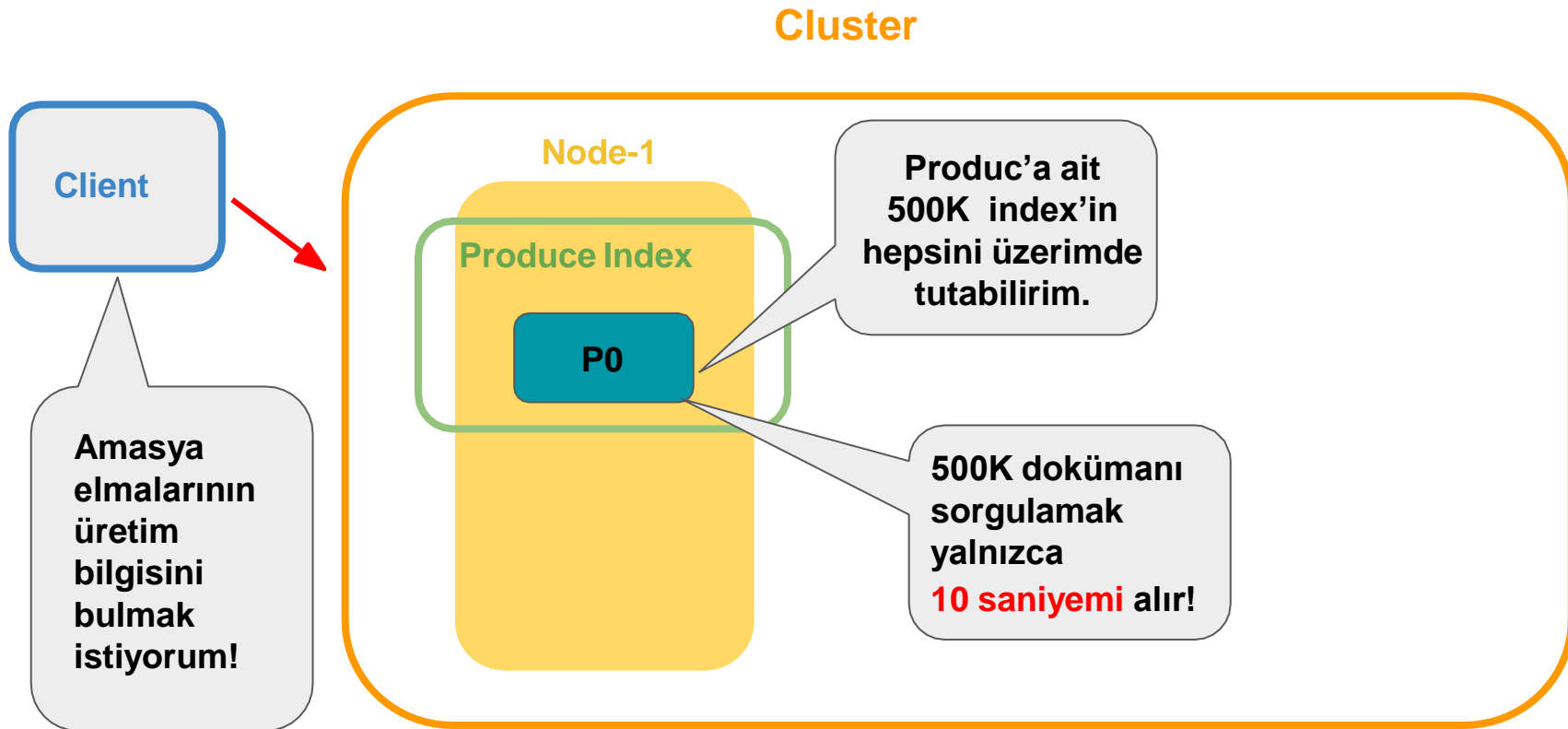


# Sharding Nedir ?

## Cluster



# Sharding Nedir ?



# Sharding arama hızınızı artırır !

Cluster

1 Saniye'de 500K doküman  
sorgulama yapabiliriz ! ⚡⚡

Node-1

Node-2

Node-3

Node-4

Node-5

Node-6

Node-7

Node-8

Node-9

Node-10

Produce Index keeps track of 500K produced documents

P0

50K

P1

50K

P2

50K

P3

50K

P4

50K

P5

50K

P6

50K

P7

50K

P8

50K

P9

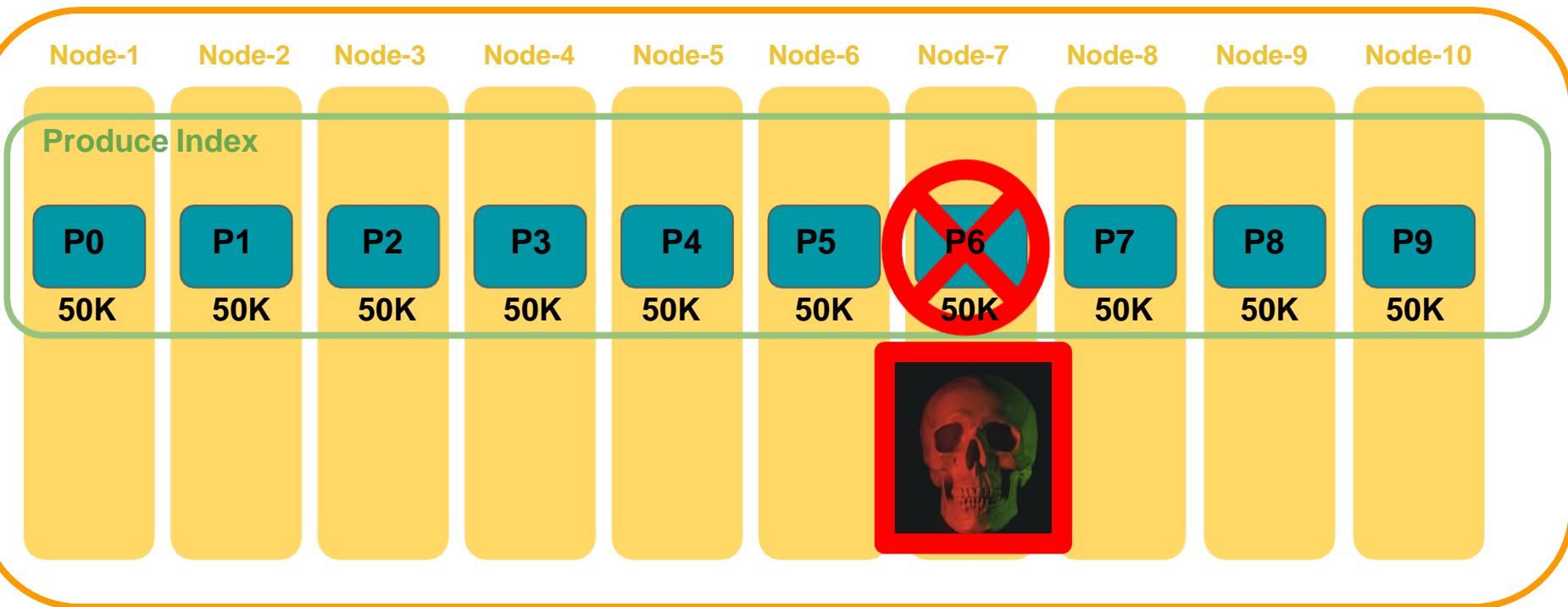
50K

50K veride aramayı  
çalıştırmam 1 saniyemi alır !

# Sharding arama hızınızı artırır !

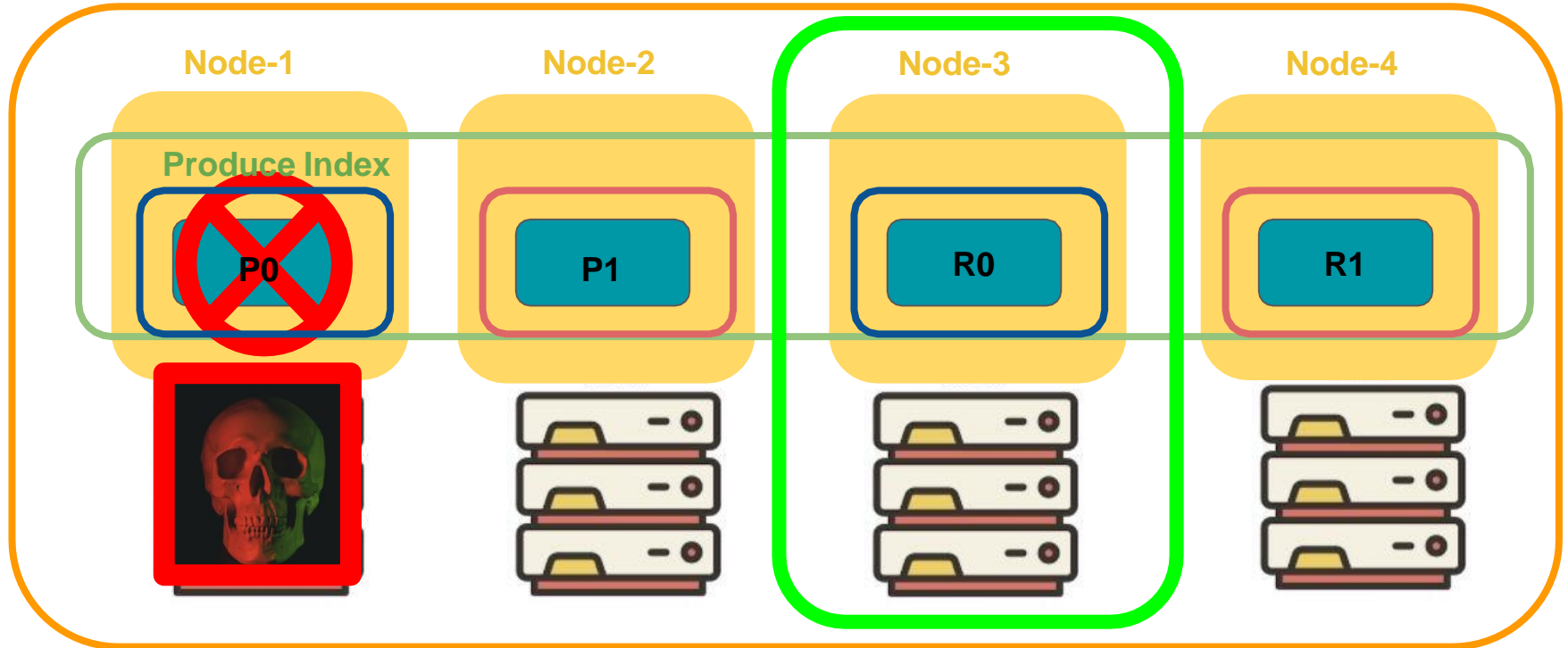
1 Saniye'de 500K doküman  
sorgulama yapabiliriz ! ⚡

Cluster

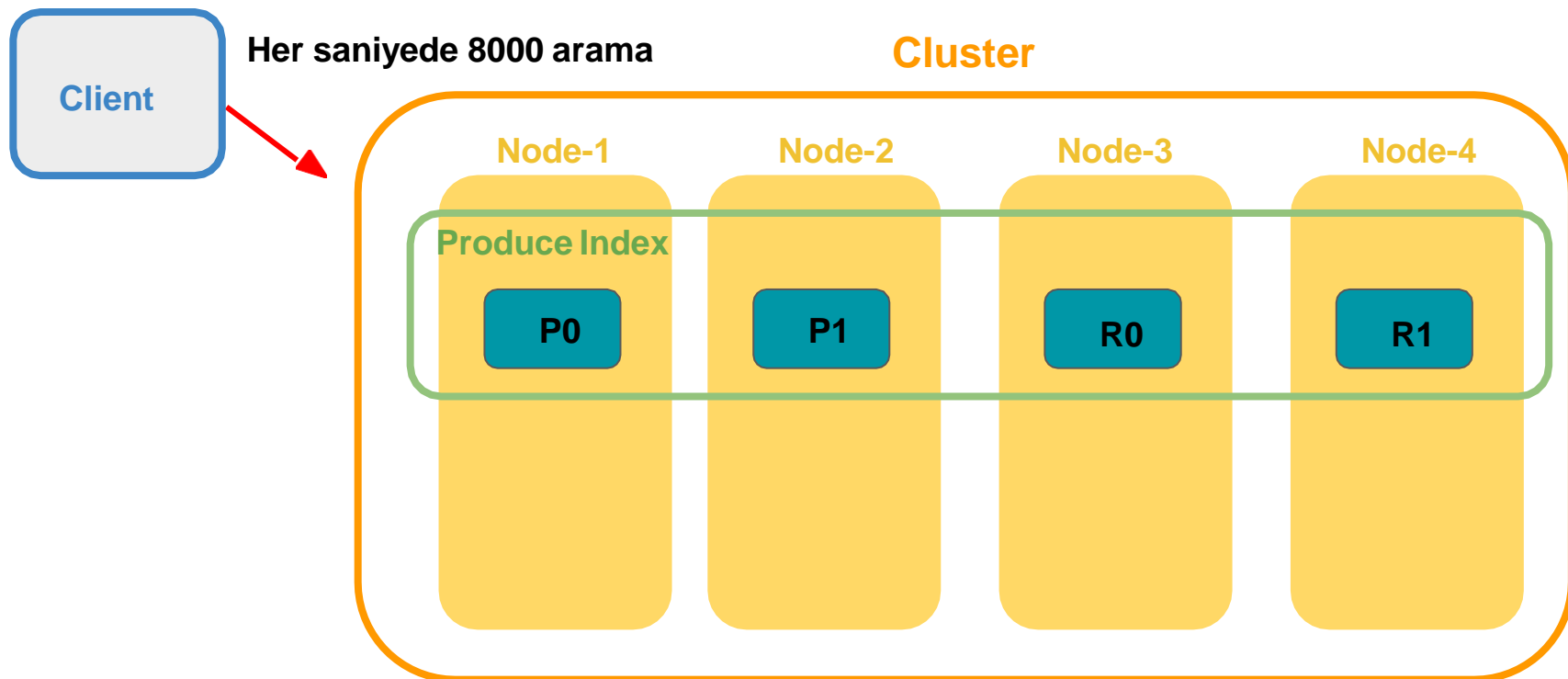


# Replica Shard Nedir?

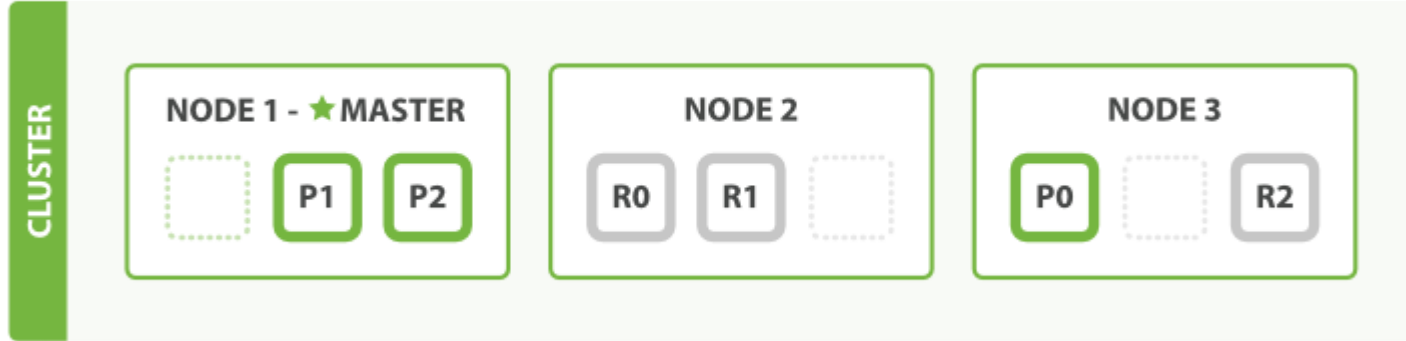
## Cluster



# Replica shard kullanarak arama performansınızı artırabilirsiniz !

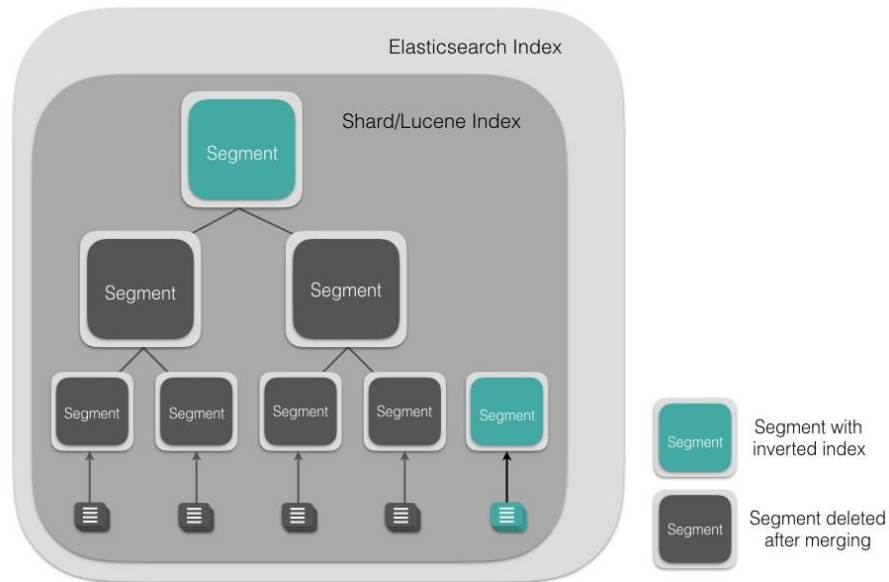


**Replica shard kullanarak  
arama performansınızı artırabilirsiniz !**



## Segment(Inverted Index)

- Segmentler bir Inverted Index'tir.
- Bir shard üzerindeki bir arama sırayla her bir segment'de aranacaktır ve sonra sonuçlar bu shard için son bir sonuçlar kümesinde toplanacaktır.
- Siz bir dökümanı indexlerken,
- Elasticsearch onları bellekte toplayacaktır (ve güvenlik için transaction log'da), sonra her saniye veya daha fazla zamanda bir bunu yapar,
- disk'e yeni bir segment oluşturur, ve aramaları refresh eder.



# Elasticsearch

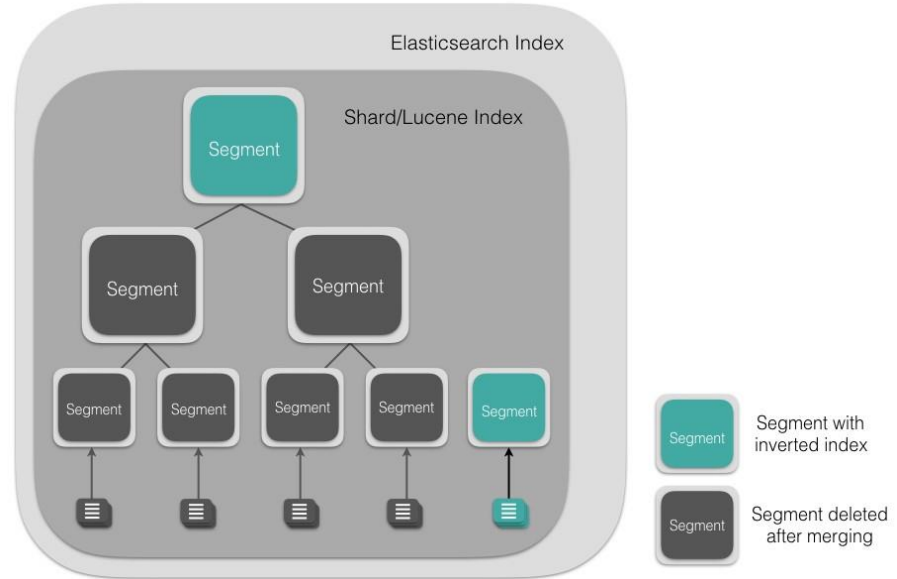
Store | Search | Analyze





## Segment(Inverted Index)

- Ne kadar çok segment varsa aramalar o kadar uzun sürer.
- Yani Elasticsearch arkaplanda çalışan birleştirici işlemler (merge process) ile benzer büyüklükteki bir kaç segmenti daha büyük bir segmente birleştirecektir.
- Yeni oluşan daha büyük segment diske yazıldıktan sonra eskileri silinecektir.
- Bu işlem bir çok aynı büyüklükte segment olduğu sürece tekrar eder.



# Elasticsearch

Store | Search | Analyze



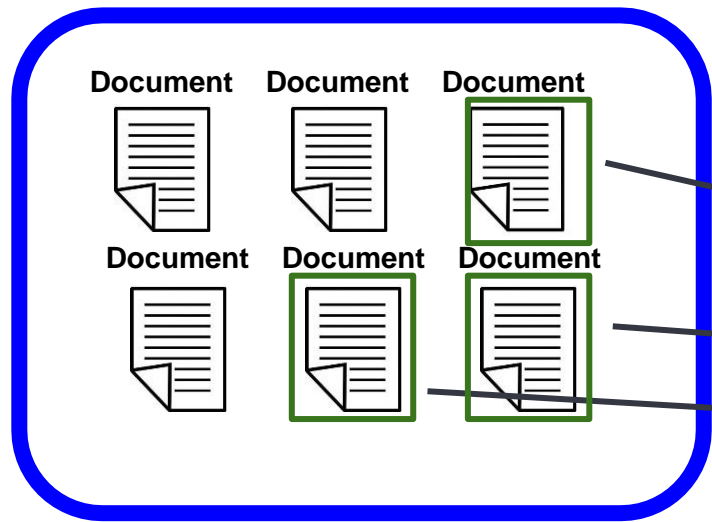
# *Speed,* **Scale,** **Relevance**

## **Elastic is a search company.**

We focus on value to users by producing fast results that operate at scale and are relevant. This is our DNA. We believe search is an experience. It is what defines us, and makes us unique.

Arama sorgusu gönderildiğinde, Elasticsearch ilgili belgeleri alır ve belgeleri arama sonuçları\* olarak sunar.

## Index

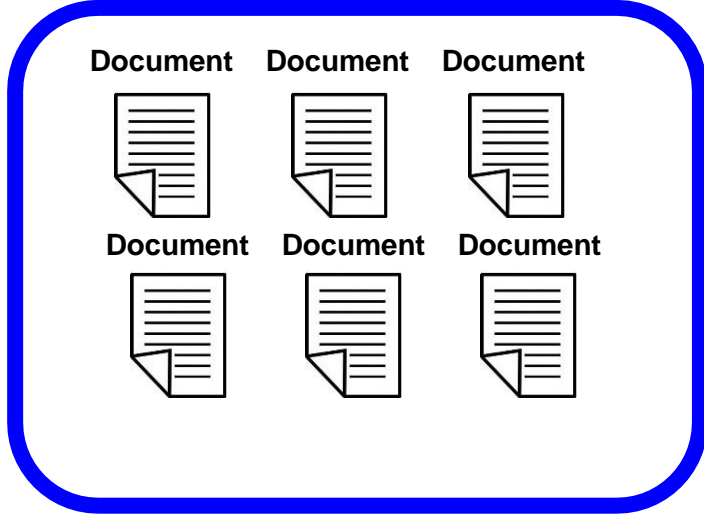


A screenshot of the Elastic UI interface. The top bar shows the "Elastic" logo and a search bar. Below the top bar, there are tabs for "Console", "Search Profiler", "Grok Debugger", and "Painless Lab". The "Console" tab is active, showing a search query: `GET kibana_sample_data_ecommerce/_search`. The "Search Results" panel on the right displays the JSON response, which is highlighted with an orange border. The JSON response includes a "hits" object with a "total" field and a "hits" array. The first hit in the array is a document from the "kibana\_sample\_data\_ecommerce" index, of type "\_doc", with ID "79WD1XYBy9gvFWLxZogX". The document's source is a JSON object containing category, currency, and customer information.

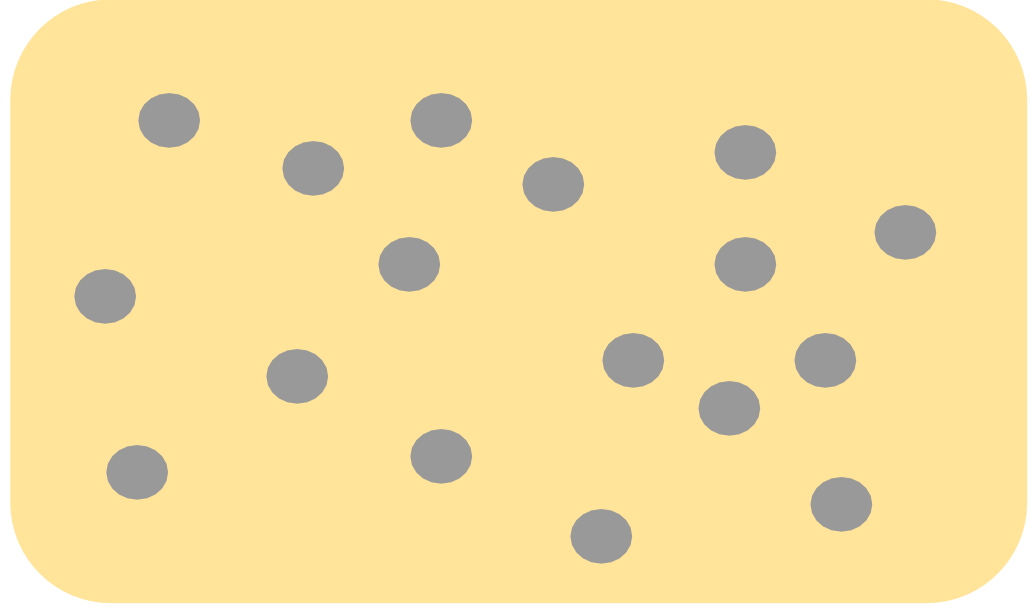
```
10+ "hits": {
11+   "total": {
12+     "value": 4675,
13+     "relation": "eq"
14+   },
15+   "max_score": 1.0,
16+   "hits": [
17+     {
18+       "_index":
19+         "kibana_sample_data_ecommerce",
20+       "_type": "_doc",
21+       "_id": "79WD1XYBy9gvFWLxZogX",
22+       "_score": 1.0,
23+       "_source": {
24+         "category": [
25+           "Men's Clothing"
26+         ],
27+         "currency": "EUR",
28+         "customer_first_name": "Eddie",
29+         "customer_full_name": "Eddie Underwood",
30+       }
31+     }
32+   ]
33+ }
```

# Bu iki görsel aynı anlamı ifade ediyor.

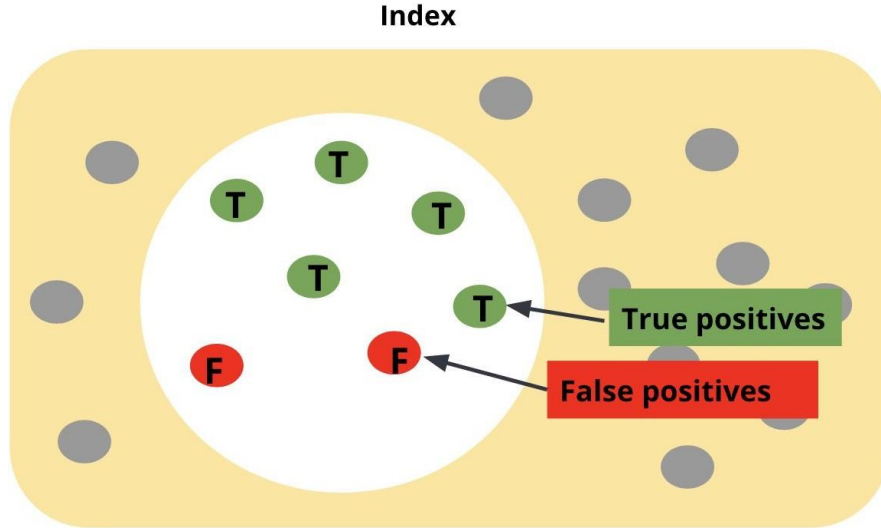
**Index**



**Index**



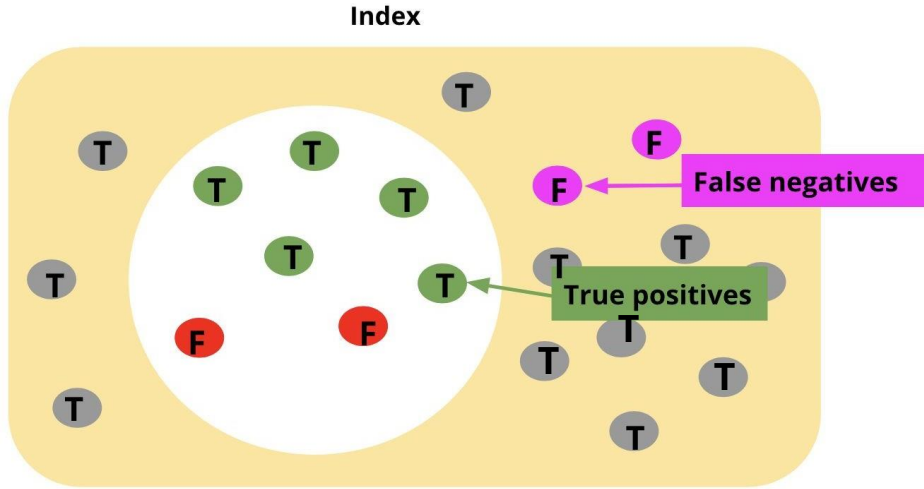
## Veride Kesinlik(Precision) Nedir ?



$$\text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}}$$

**Alınan verilerin hangi kısmı arama sorgusuyla gerçekten alakalı?**

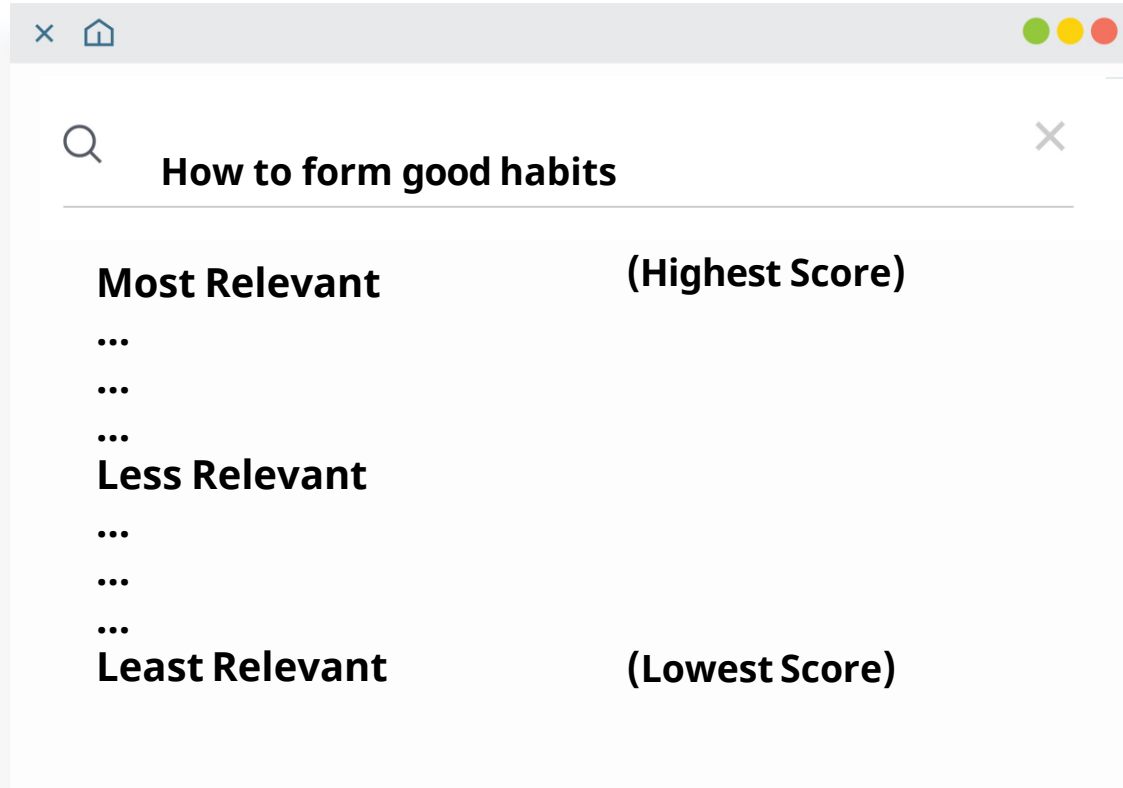
# Veride Duyarlilik(Recall) Nedir?



$$\text{Recall} = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}}$$

**İlgili verilerin hangi kısmı arama sonuçları olarak döndürülüyor?**

**Sonuçların sıralanması(Ranking):**  
en üstten başlayarak **en alakalı** sonuçlardan, **en az alakalıya**



## TF-IDF Nedir?

- **Term Frequency(TF)**

Arama, alanda ne kadar fazla görünürse, alan o kadar alakalı olur

- **Inverse Document Frequency(IDF)**

Arama, belgelerin tüm alt kümesinde ne kadar fazla görünürse, o kadar az alakalı olur.



**Terim Frekansı(TF),** her bir arama teriminin bir belgede kaç kez görüneceğini belirler.

The screenshot shows a search window titled "Search Terms". Below the title bar is a search bar containing the query "How to form good habits". To the right of the search bar is a red bracket labeled "Search Query". Below the search bar, there are two document snippets. The first snippet is for a document titled "Atomic Habits" by James Clear, and it shows the term frequency for "habits" as TF=4. The second snippet is for a document titled "The Mental Toughness Handbook" by Damon Zahariades, and it shows the term frequency for "habits" as TF=1.

```
{
  "title": "Atomic Habits",
  "author": "James Clear",
  "category": "self-help",
  "description": "No matter your goals, Atomic Habits offers a proven
    framework for improving every day. James clear, ... habits...habits
    ...habits" TF=4
}
```

```
{
  "title": "The Mental Toughness Handbook",
  "author": "Damon Zahariades",
  "category": "self-help",
  "description": "Imagine boldly facing any challenge that comes your way... 5
    daily habits you must embrace to strengthen your mind and harden your
    resolve. Why willpower and motivation are unreliable..." TF=1
}
```

**Bir belgede  
arama terimleri  
yüksek sıklıkta bulunursa,  
belgenin arama  
sorgusuyla  
daha alakalı  
olduğu kabul edilir.**

# Inverse Document Frequency(IDF) Nedir ?

IDF, belge kümesinde çok sık geçen terimlerin ağırlığını azaltır ve nadiren geçen terimlerin ağırlığını artırır !



How to form good habits

Bazı arama terimlerini içerebiliriz, ancak «iyi alışkanlıklar» oluşturmakla hiçbir ilgimiz yok !!

## Hits

How to form a meetup group



Good chicken recipe



How to form a band



Good times rolling



Good habits 101



Good habits are easy to master!

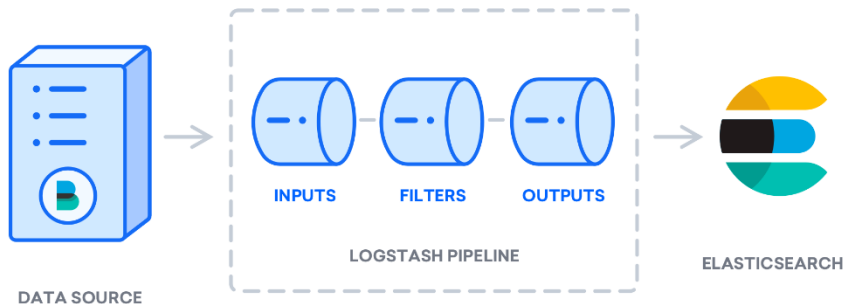


# Logstash



- Çeşitli kaynaklardan veri toplamanıza
- Anında dönüştürmenize ve istediğiniz hedefe göndermenize olanak tanıyan
- Açık kaynaklı, sunucu tarafı bir veri işleme hattıdır.







Genellikle Elasticsearch için bir veri hattı olarak kullanılır.



# The Beats Family

The Beats Family

All kinds of shippers for all kinds of data.

					
<b>Filebeat</b>	<b>Metricbeat</b>	<b>Packetbeat</b>	<b>Winlogbeat</b>	<b>Auditbeat</b>	<b>Heartbeat</b>
Log Files	Metrics	Network Data	Windows Event Logs	Audit Data	Uptime Monitoring

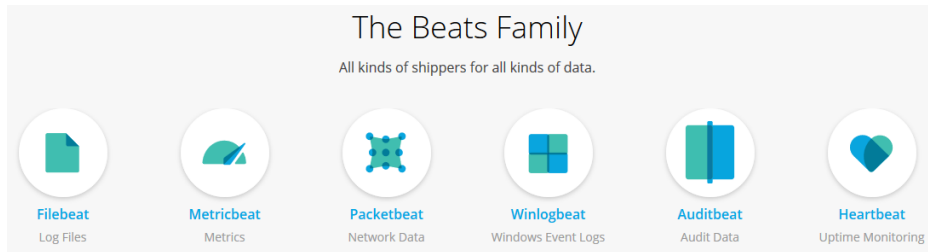
## Elasticsearch

Store | Search | Analyze



elasticsearch

# The Beats Family



- **Packetbeat** sunucularınız arasındaki ağ trafiğini izleyebilmemizi
- **Filebeat** sunucularınızdaki log dosyalarınızı takip edebilmemizi
- **Metricbeat** 'de periyodik olarak dış bir kaynaktan veri almamızı sağlar.

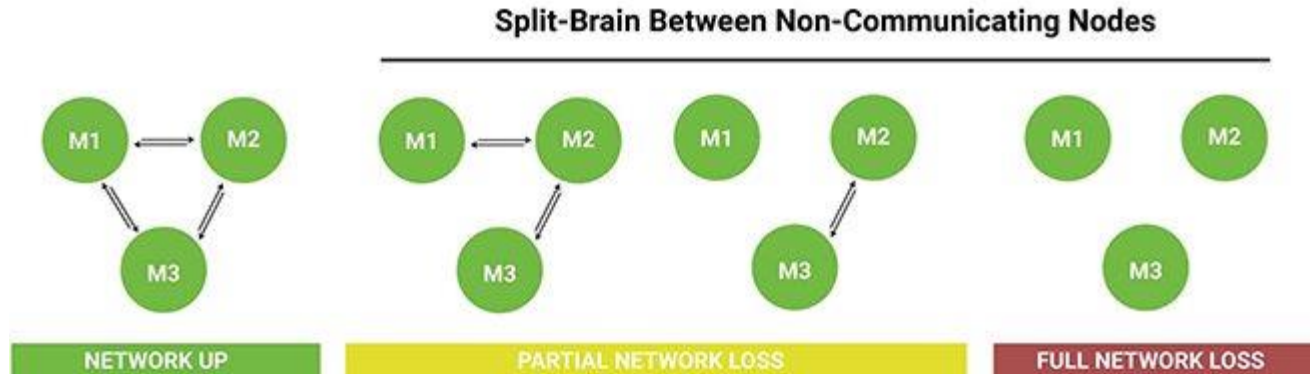
**Elasticsearch**

Store | Search | Analyze



elasticsearch

# Araştırma Ödevi: «Split Brain» problemi nedir?

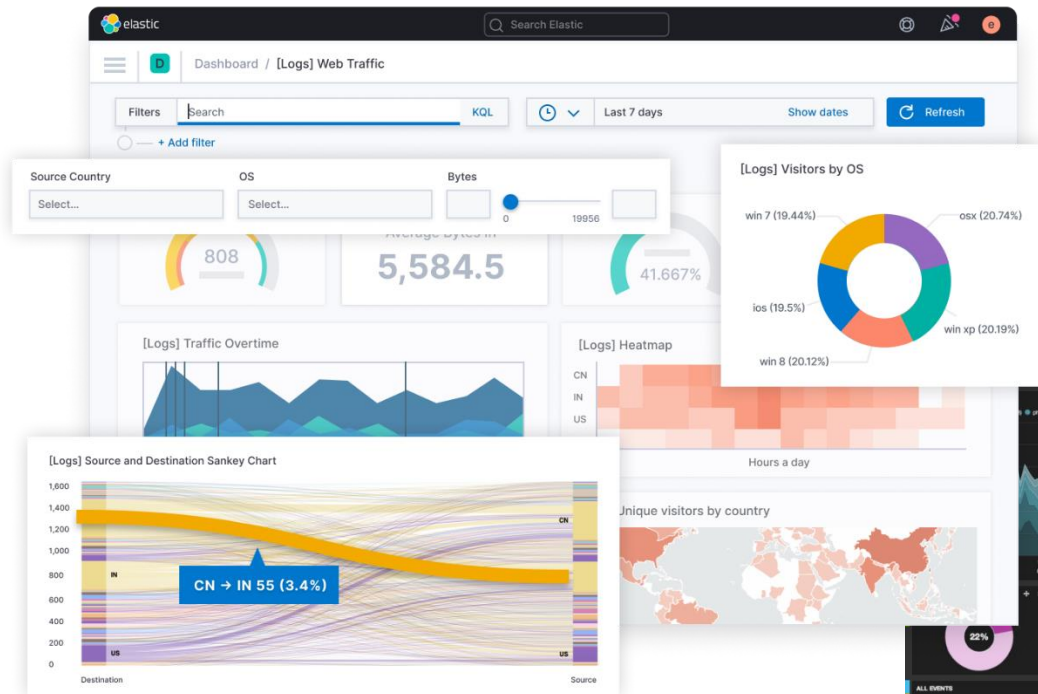


**Elasticsearch**

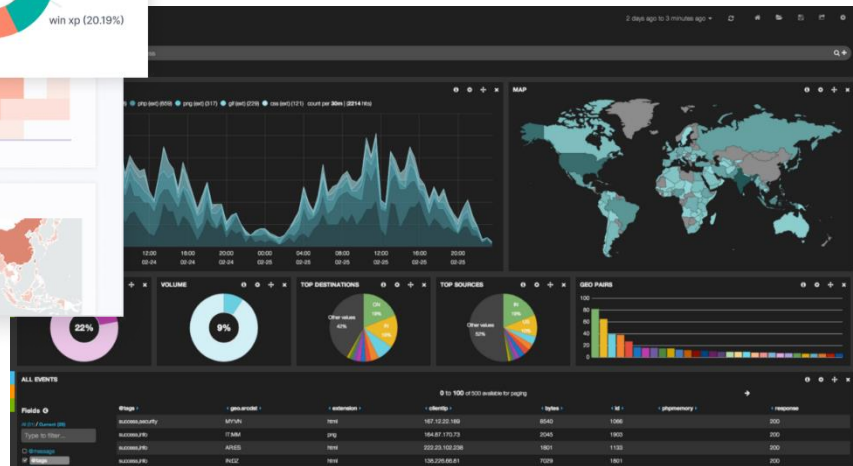
Store | Search | Analyze



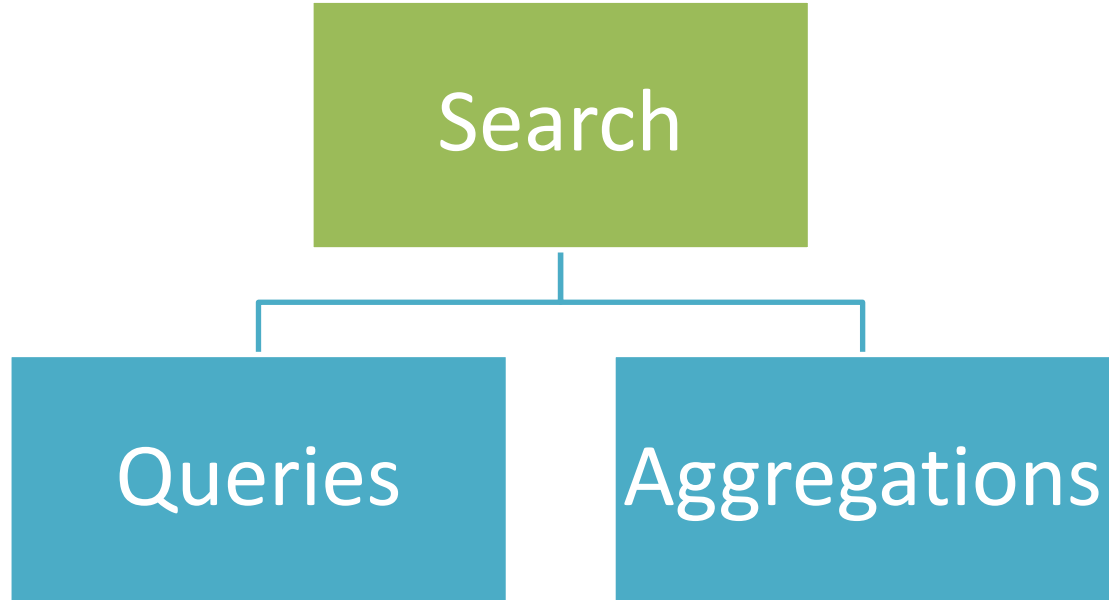
elasticsearch



# kibana



# Elasticsearch Temel Arama Sorgu Yapısı





# Queries

Syntax:

GET Enter\_name\_of\_the\_index\_here/\_search

Example:

GET news\_headlines/\_search

# Aggregations

Syntax:

```
GET Enter_name_of_the_index_here/_search
{
  "aggregations": {
    "Name your aggregation here": {
      "Specify aggregation type here": {
        "field": "Name the field you want to aggregate here",
        "size": State how many buckets you want returned here
      }
    }
  }
}
```

Example:

```
GET news_headlines/_search
{
  "aggregations": {
    "by_category": {
      "terms": {
        "field": "category",
        "size": 100
      }
    }
  }
}
```

# Aggregations

Example:

GET news\_headlines/\_search

```
{
  "aggregations": {
    "by_category": {
      "terms": {
        "field": "category",
        "size": 100
      }
    }
  }
}
```

```
1 {
2   "took" : 36,
3   "timed_out" : false,
4   "_shards" : {
5     "total" : 1,
6     "successful" : 1,
7     "skipped" : 0,
8     "failed" : 0
9   },
10  "hits" : { },
169 "aggregations" : {
170   "by_category" : {
171     "doc_count_error_upper_bound" : 0,
172     "sum_other_doc_count" : 80845,
173     "buckets" : [
174       {
175         "key" : "POLITICS",
176         "doc_count" : 32739
177       },
178       {
179         "key" : "WELLNESS",
180         "doc_count" : 17827
181       },
182       {
183         "key" : "ENTERTAINMENT",
184         "doc_count" : 16058
185       },
186       {
187         "key" : "TRAVEL",
188         "doc_count" : 9887
189       },
190       {
191         "key" : "STYLE & BEAUTY",
192         "doc_count" : 9649
193       },
194       {
195         "key" : "PARENTING",
196         "doc_count" : 8677

```

# Full Text Queries

Syntax:

```
GET Enter_name_of_index_here/_search
{
  "query": {
    "match": {
      "Specify the field you want to search": {
        "query": "Enter search terms"
      }
    }
  }
}
```

# Searching for a phrase

Example :

GET news\_headlines/\_search

```
{
  "query": {
    "match": {
      "headline": {
        "query": "Shape of you"
      }
    }
  }
}
```

```
{
  "hits" : {
    "total" : {
      "value" : 10000,
      "relation" : "gte"
    },
    "max_score" : 12.274778,
    "hits" : [
      {
        "_index" : "news_headlines",
        "_type" : "_doc",
        "_id" : "u9g9S3cBwsjPafpA2HGP",
        "_score" : 12.274778,
        "_source" : {
          "date" : "2012-08-30",
          "short_description" : "Get stronger. Practice wall squats (with back against the wall, lower your body until knees are at 90 degrees; hold for 30",
          "@timestamp" : "2012-08-30T00:00:00.000-06:00",
          "link" : "https://www.huffingtonpost.com/entry/fitness-test-women_us_5b9c2c91e4b03a1dcc7cda8e",
          "category" : "WELLNESS",
          "headline" : "Fitness Test: Are You In Shape?",
          "authors" : ""
        }
      },
      {
        "_index" : "news_headlines",
        "_type" : "_doc",
        "_id" : "u9g9S3cBwsjPafpA2HGP",
        "_score" : 12.274778,
        "_source" : {
          "date" : "2012-08-30",
          "short_description" : "Get stronger. Practice wall squats (with back against the wall, lower your body until knees are at 90 degrees; hold for 30",
          "@timestamp" : "2012-08-30T00:00:00.000-06:00",
          "link" : "https://www.huffingtonpost.com/entry/fitness-test-women_us_5b9c2c91e4b03a1dcc7cda8e",
          "category" : "WELLNESS",
          "headline" : "Fitness Test: Are You In Shape?",
          "authors" : ""
        }
      }
    ]
  }
}
```

# Searching for phrases using the match\_phrase query

Syntax :

```
GET Enter_name_of_index_here/_search
{
  "query": {
    "match_phrase": {
      "Specify the field you want to search": {
        "query": "Enter search terms"
      }
    }
  }
}
```

Example :

```
GET news_headlines/_search
{
  "query": {
    "match_phrase": {
      "headline": {
        "query": "Shape of You"
      }
    }
  }
}
```

# Searching for phrases using the match\_phrase query

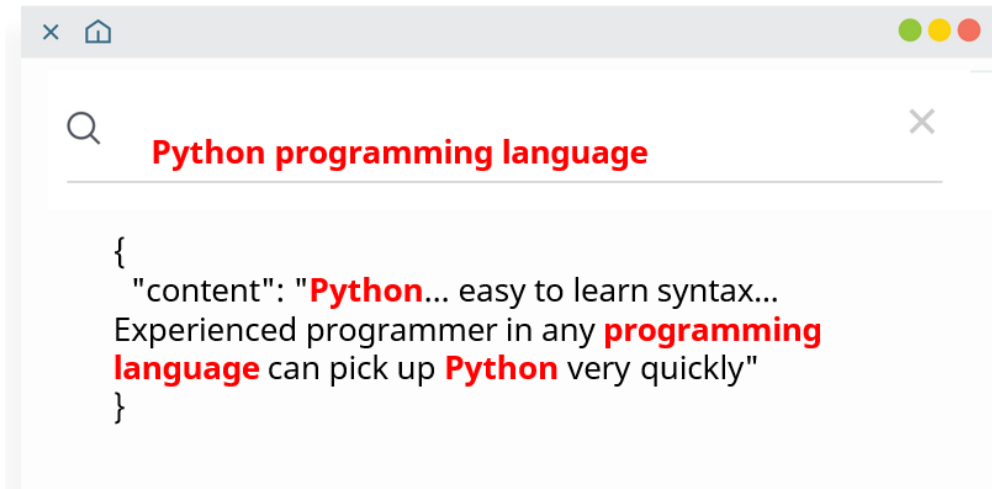
Example :

```
GET news_headlines/_search
{
  "query": {
    "match_phrase": {
      "headline": {
        "query": "Shape of You"
      }
    }
  }
}
```

```
10  "hits" : {
11    "total" : {
12      "value" : 3,
13      "relation" : "eq"
14    },
15    "max_score" : 12.074881,
16    "hits" : [
17      {
18        "_index" : "news_headlines",
19        "_type" : "_doc",
20        "_id" : "lMs7S3cBVGnaeHqj6RUZ",
21        "_score" : 12.074881,
22        "_source" : {
23          "date" : "2017-03-20",
24          "short_description" : "Puerto Rico's
                                Zion & Lennox are behind the new
                                version.",
25          "@timestamp" : "2017-03-20T00:00:00.000
                                -06:00",
26          "link" : "https://www.huffingtonpost
                                .com/entry/ed-sheerans-zion-lennox
                                -shape-of-you-latin
                                -remix_us_58d03b09e4b0be71dcf72c6f",
27          "category" : "LATINO VOICES",
28          "headline" : "Ed Sheeran's 'Shape Of
                                You' Gets An Unexpected Latin Remix",
29          "authors" : "Carolina Moreno"
30        }
31      },
```

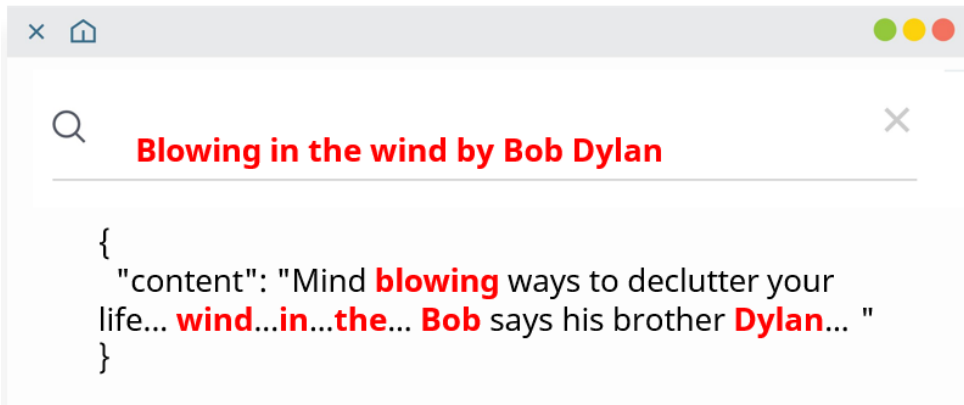
# match query

```
GET enter_name_of_index_here/_search
{
  "query": {
    "match": {
      "Specify the field you want to search":{
        "query":"Enter search terms"
      }
    }
  }
}
```



A search results window with a search bar containing the text "Python programming language". Below the search bar, a JSON object is displayed, showing a match for the field "content". The text in the JSON object is "Python... easy to learn syntax... Experienced programmer in any programming language can pick up Python very quickly". The words "Python", "programming language", and "Python" are highlighted in red.

```
{
  "content": "Python... easy to learn syntax...
Experienced programmer in any programming
language can pick up Python very quickly"
}
```



A search results window with a search bar containing the text "Blowing in the wind by Bob Dylan". Below the search bar, a JSON object is displayed, showing a match for the field "content". The text in the JSON object is "Mind blowing ways to declutter your life... wind...in...the... Bob says his brother Dylan... ". The words "blowing", "wind", "in", "the", "Bob", and "Dylan" are highlighted in red.

```
{
  "content": "Mind blowing ways to declutter your
life... wind...in...the... Bob says his brother Dylan... "
}
```



## match\_phrase query

```
GET Enter_the_name_of_the_index_here/_search
{
  "query": {
    "match_phrase": {
      "Specify the field you want to search over": "Enter the phrase you are searching for"
    }
  }
}
```

**Elasticsearch**

Store | Search | Analyze



elasticsearch

# TEŞEKKÜRLER

Meltem YILMAZ

Big Data Engineer

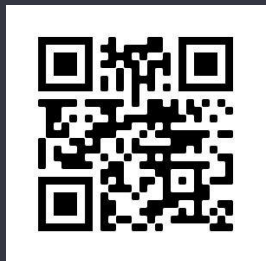
[meltemyilmaz@havel-san.com.tr](mailto:meltemyilmaz@havel-san.com.tr)

# Uygulamalı Pratikler

- Lisa Hjung - Beginner's Crash Course to Elastic Stack Series  
<https://github.com/LisaHJung/Beginners-Crash-Course-to-Elastic-Stack-Series-Table-of-Contents>
- Devnot Atölye - Elasticsearch, Kibana ve Logstash  
<https://www.youtube.com/watch?v=ZLQSCfXhpIM>
- **Getir**'de Elasticsearch  
<https://www.youtube.com/watch?v=O3-X6r-pJHk>
- Elasticsearch Mimarisi  
<http://cagataykiziltan.net/elasticsearch/>

# Connect with the Elastic Community

 Elastic  
meetups



[https://ela.st/  
amervirtual](https://ela.st/amervirtual)



Elastic Community  
Slack Workspace



<https://ela.st/slack>



YouTube  
Channel



[https://ela.st/  
community-  
youtube](https://ela.st/community-youtube)