

accepting (word
article).

focus n point

converging rays of light,
heat, waves of sound, meet;
centre of activity or
intensity; pl focuses, foci; v
adjust; cause to converge;
concentrate; a focal
pertaining to focus

Elasticsearch

@RoboNovotny | Eastcode Sessions | 24. 2. 2015

developerské kurzy?






hľadať podľa...

názov

popis kurzu

čas konania

lokácia



```
SELECT *  
FROM course  
WHERE  
    name = '%elastic%'
```




frázy

žolíky

našepkávanie

alternatívy

“Turbo Pascal”

Úvod do *

Visual_____

Phyton

A man with a long, dark beard and a hat made of twigs and branches is shown in a forest. He is holding a herring in his hand. The background consists of trees and foliage.

You must cut down
the mightiest tree in
the forest with a
herring!



Elasticsearch

Elasticsearch

- fulltextsearch server
 - distribuovatelný
 - škálovatelný
 - REST API
-
- Java
 - Apache License 2.0

setup

- `wget https://.../elasticsearch-2.2.0.zip`
- `tar xvzf elasticsearch-2.2.0.zip`
- `./bin/elasticsearch`
- `curl http://localhost:9200`

bežíme?

```
4. bash
whitehall:~ novotnyr$ curl http://localhost:9200
{
  "name" : "Sabreclaw",
  "cluster_name" : "elasticsearch",
  "version" : {
    "number" : "2.2.0",
    "build_hash" : "8ff36d139e16f8720f2947ef62c8167a888992fe",
    "build_timestamp" : "2016-01-27T13:32:39Z",
    "build_snapshot" : false,
    "lucene_version" : "5.4.1"
  },
  "tagline" : "You Know, for Search"
}
```

-



Sense (Beta)

offered by cheics

★★★★★ (98)

[Developer Tools](#)

49,232 users

OVERVIEW

REVIEWS

SUPPORT

RELATED

```
Server 192.168.231.50:9200/telemetry-dev History Help
6 GET _search
7 {
8   // All events actually have a doctype
9   query: {
10    "term": { "doctype": "event" }
11  }
12 }
13 GET _search
14 {
15   "explain": true,
16   "facets": {
17     "validity": {
18       "field": "k2"
19     }
20   },
21   "search_type": "explain",
22   "stats": true
23 }
24 //No get 49 users here
25 }
26 }
27 }
28 }
29 }
30 }
31 GET _search
32 }
```

```
{
  "book": 2,
  "timed_out": false,
  "_shards": {
    "total": 5,
    "successful": 5,
    "failed": 0
  },
  "hits": {
    "total": 123568,
    "max_score": 0.99996144,
    "hits": [
      {
        "_index": "telemetry-dev",
        "_type": "event",
        "_id": "924148433984d17fb3831acc251f1d4",
        "_score": 0.99996144,
        "_source": {
          "_rev": "1-2372ccb8e67563e4386da7aac238c43",
          "appId": "a1Ew0kayGV00eZbwRkxk48@yU4u6eE8",
          "doctype": "event",
          "_id": "924148433984d17fb3831acc251f1d4",
          "event": {
            "ts": 1487482918,
            "Type": "ValidateAccessTokenTelemetryEvent",
            "IpAddress": "157.56.164.154",
            "UserIdnToken": "3bc32978bbe4456ab37def"
          }
        }
      }
    ]
  }
}
```

LEFFINQWELLÚV PSACÍ STÚL

(KONKRETNÍ SITUACE)

KNIHY PŘIPRAVENÉ
KE STUDIU

ZDROJ SVĚTLA

PUPĚNÍ

KNIHY PŘESTUDOVÁNY

1

2

4

5

- index ~ databáza
- type ~ tabuľka
- document ~ riadok
- field ~ hodnota

indexovanie dokumentu

POST /skola/course

```
{  
  "code": "TP-1",  
  "name": "Introduction to Turbo Pascal",  
  "start": "2016-02-25",  
  "summary": "A brief introduction to Turbo  
Pascal syntax and principles"  
}
```


dokument je v indexe

```
{  
  "_index": "skola",  
  "_type": "course",  
  "_id": "AVMK_96YKqMeCvS5C2eN",  
  "_version": 1,  
  "_shards": {  
    "total": 2, "successful": 1, "failed": 0  
  },  
  "created": true  
}
```

dokument je v indexe

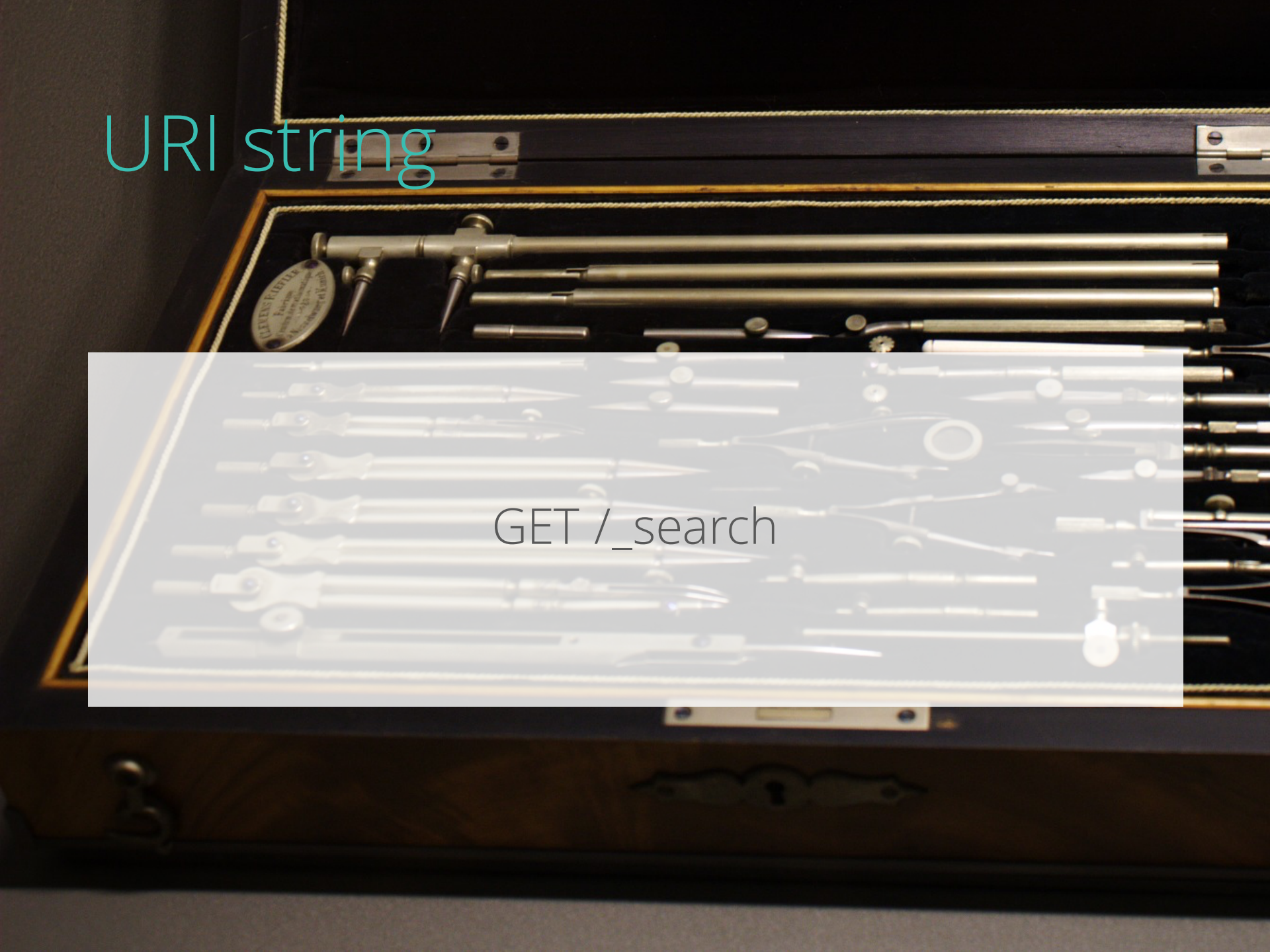
- automaticky pridelené ID
- automaticky odvodené dátové typy pre políčka
- automaticky pridelená verzia
- automatické políčko `_all` so všetkými textami



dopytovanie: ukážky

URI string

GET /_search



vyhl'adaj JavaScript

```
{  
  "query":{  
    "match":{  
      "name":"javascript"  
    }  
  }  
}
```

query dsl

```
{  
  "query": {  
    "match_all": {}  
  }  
}
```



dopyt

do akej miery vyhovuje dokument?

filter

spĺňa dokument podmienku?

vyhl'adaj JavaScript

```
{  
  "query":{  
    "match":{  
      "name":"javascript"  
    }  
  }  
}
```

úvody do nePascalu

```
{ "query": {  
  "bool": {  
    "must": {  
      "match": {  
        "name": "introduction" }  
      },  
    "must_not": {  
      "match": {  
        "name": "pascal" }  
      }  
    }  
  }  
}
```

autocomplete

```
{  
  "query":{  
    "match_phrase_prefix":{  
      "name":"intro"  
    }  
  }  
}
```

termové dopyty

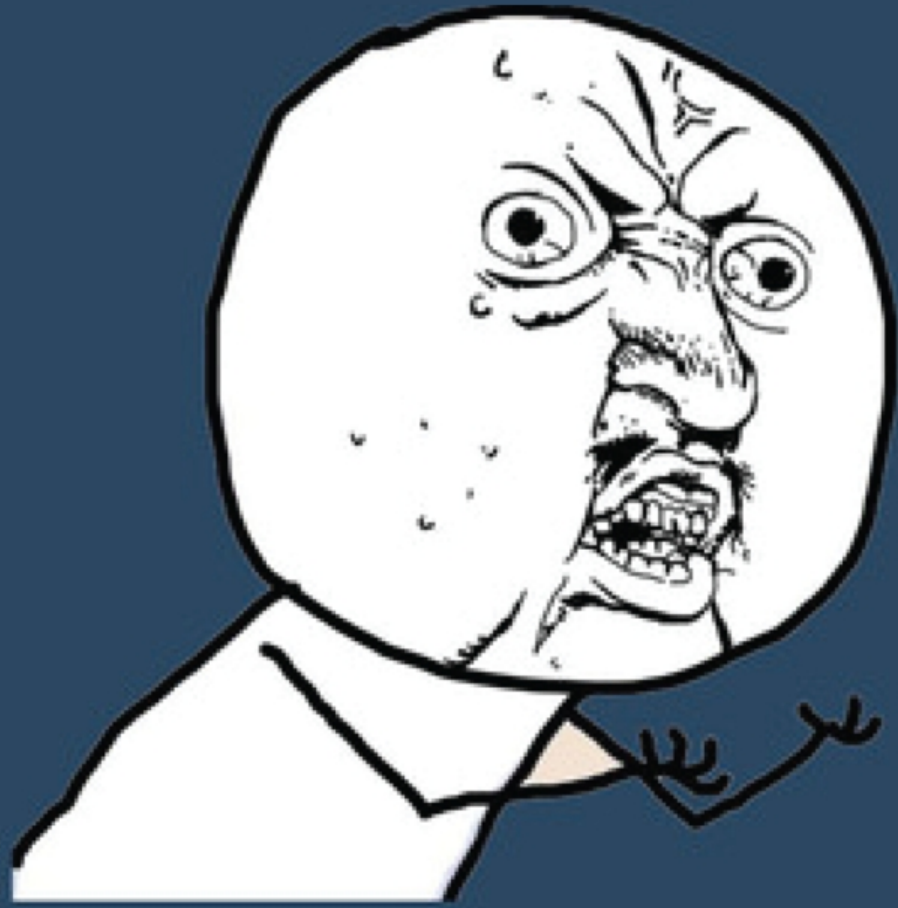
- filter na presnú zhodu s termom v indexe
- analýza sa ignoruje
- fungujú len s neanalyzovanými políčkami

termové dopyty

```
{  
  "query": {  
    "term": {  
      "code" : "PAS-1"  
    }  
  }  
}
```

“code”: “PAS-1”

```
"hits": {  
  "total": 0,  
  "max_score": null,  
  "hits": []  
}
```



Bol raz jeden
analyzér...

Artist



analyzátor textu

- odfiltruje zlé znaky
- rozseká na tokeny
- odstráni zlé tokeny

standard analyzer

A brief introduction to Turbo Pascal syntax and principles



[a](#) brief introduction to turbo Pascal syntax and principles

english analyzer

A brief introduction to Turbo Pascal syntax and principles



brief introduct turbo pascal syntax principl

analyzátor políček

PAS-1



"pas" 1

analyzátor políčok i dopytov

- fulltextové vyhľadávanie analyzuje
- presné zhody neanalyzujú

termové dopyty

```
{
  "query": {
    "term": {
      "code" : "PAS-1"
    }
  }
}
```

Hľadáme PAS-1, ale v indexe sú len tokeny PAS a 1



Bol raz jeden dátový typ...

Artist



mapovania

ES je bezschémový¹

¹ vedľajšie účinky konzultujte s vaším lekárom alebo lekárnikom

mapovania

- **schéma** dokumentov
 - štruktúra
 - dátové typy políček
- **metadáta** dokumentu
 - identifikátor
 - `_all`
 - `_source`

mapovania

- dátové typy políček sú automaticky uhádnuté

GET /skola/_mapping

mapovania

- dátové typy políček sú **automaticky** uhádnuté
 - **dynamic** mapping
- možno **customizovať**
 - pri vytváraní indexu
 - pri pridávaní políček

obmedzenia mapovaní

- existujúce mapovania sa nedajú zmeniť
- nedá sa zmeniť dátový typ políčka
- nedajú sa zmeniť metadáta dokumentu

reindex 😞

index + mapovanie

PUT skola

```
{  
  "mappings": {  
    "course": {  
      "properties": {  
        "code": {  
          "type": "string",  
          "index": "not_analyzed"}}}}}}}
```

highlighting

- priama podpora zvýrazňovania nájdených termov
- možnosť prispôbiť štýl

highlighting?

```
{"query": {  
  "match": {  
    "summary": "programming"  
  }  
},  
"highlight": {  
  "fields": {  
    "summary": {}  
  }  
}}
```

geolokácie

- priama podpora geolokácie
- hľadanie v okolí
- hľadanie v útvere
- triedenie podľa vzdialeností

hľadaj v okolí Košíc

```
{  
  "query": {  
    "geo_distance": {  
      "distance": "10km",  
      "location": {  
        "lat": 48.71,  
        "lon": 21.25  
      }  
    }  
  }  
}
```

tried podľa vzdialenosti

```
{  
  "sort" : {  
    "_geo_distance" : {  
      "location" : [21.25, 48.71],  
      "unit" : "km"  
    }  
  }  
}
```

a je toho viac...

- agregácie
- perkolácia
- verzie
- bulk API
- limitované joiny

škálovateľnosť



shard

- atomická jednotka škálovania v ES
- nesie časť indexu

- index začína s 5 shardami



shard

- počet shardov v indexe sa nedá meniť
- shardy možno presúvať medzi uzlami



replica

- shard možno replikovať
- replice musia bežať na samostatných uzloch
- potom zlepšujú čítanie
- prispievajú k failoveru

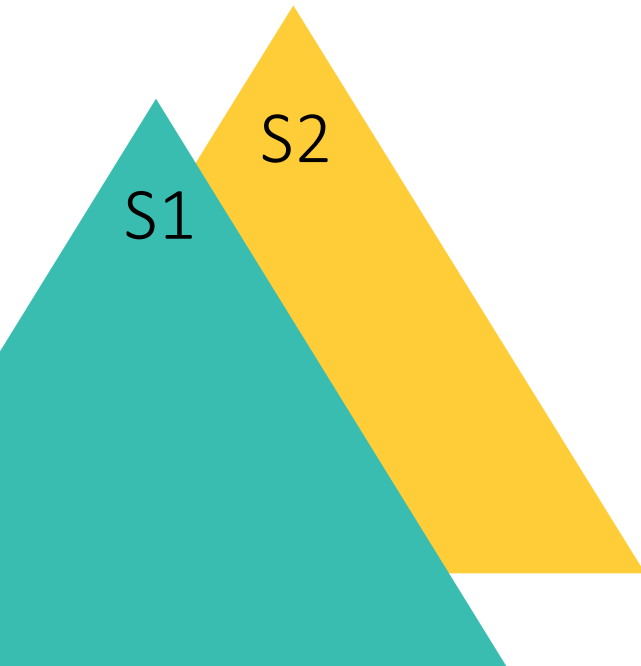
autoškálovanie

- nová inštancia Elasticsearchu sa automaticky pripojí do klastra
 - protokol [Zen Disco](#)
- shardy a replice sa automaticky rozložia

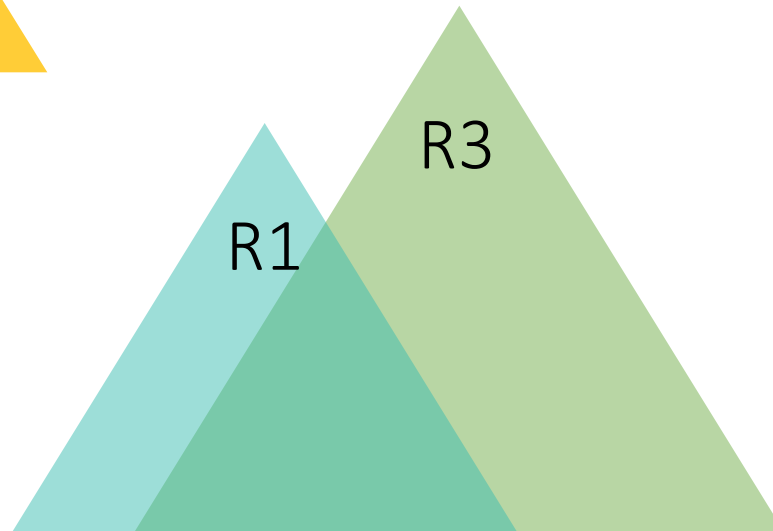
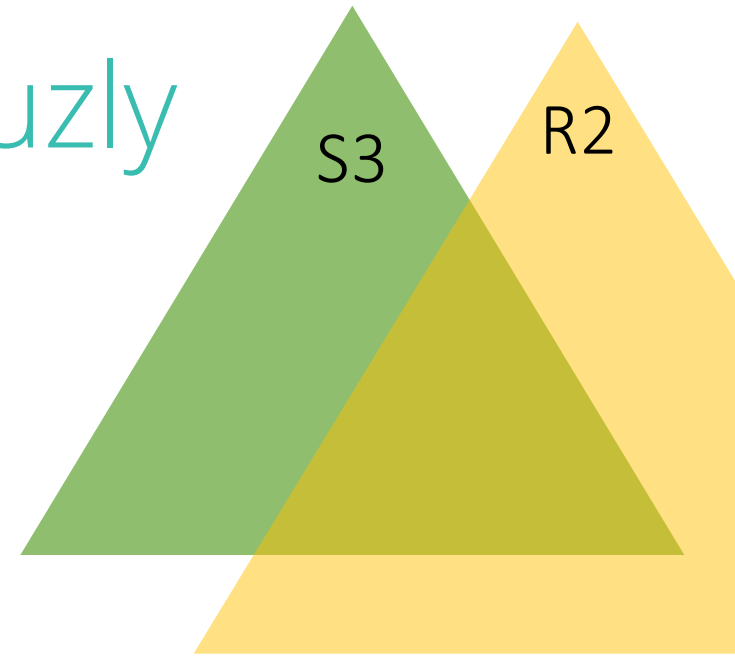
škálovanie?

- rovno na 3 stroje
 - 2 stroje nemajú failover
 - 2 uzlom hrozí split-brain

ukážkový scenár: 3 uzly



Môžeme prísť
o 2 uzly a dáta
sa nestratia



Škálovanie = hrajkanie

- počet shardov
- počet replík
- počet indexov
- organizácia typov



kam ďalej? ELK

- Elasticsearch : sklad
- Logstash : preprava
- Kibana : vizualizácia



otázky?