E

bin\elasticsearch.bat

[http://localhost:9200](http://localhost:9200/)

<http://localhost:9200/_cat/indices?v>

<http://localhost:9200/new_index-2017.10.30?pretty>

K

bin\kibana.bat

[http://localhost:5601](http://localhost:5601/)

<https://www.elastic.co/webinars/getting-started-kibana>

<https://www.elastic.co/guide/en/kibana/current/getting-started.html>

L

bin\logstash.bat -f logstash.conf

KOPF

bin\cerebro.bat

[http://localhost:9000](http://localhost:9000/)

<https://github.com/lmenezes/cerebro>

<https://howtodoinjava.com/microservices/elk-stack-tutorial-example/>

APKA

java -jar target\elk-example-spring-boot-0.0.1-SNAPSHOT.jar

<http://localhost:8080/elk>.

Wizualizacje polazenia, podzialu info na skad, slupkowy ile razy przez component, slupkowy dlugosc trwania procesu ?, na jakies pole pie chart, wykres po slupkach, tendencji

- graf ilosc query na minute, mozna podzial na komponenty

- np. Podzial rzadan w czasie i podzial na statusy ile na jednostke

- mozna tabele agregacji, np. Ilosc uderzen na komponent

-metric – ilosc wszystkich uderzen, albo exceptionow , ilosc unique , przecietna wielkosc danych

- mozna w czym wartosc wyswietlana w management, np. Zamiast byte to MB

- w timelion wycinek czasu danych jest, mozna okreslic index – zrudlo danych

- mozna kilka wykresow, points , np. Rozlozenie czasow max i avg

- mozna dzielic przez share na dashboard

- w management mozna export – do

- mozna tabele ile nan adny url uderzen

- jak naloze filtr na dasboard i mam kilka zrodel powiazanych ,to wszystkie wyfiltruje

- np. Podzial na status, na status w lokalizacji – kolowe, rozklad na miapie, jakie linki dawaly itd.

- mozna tez np. Jaki ruch o konkretnej porze , mozna w management nowe script field i np. Wyciaganc godzine , potem mozemy agregowac po godzinie

<https://www.digitalocean.com/community/tutorials/how-to-use-kibana-dashboards-and-visualizations>

<https://howtodoinjava.com/spring/spring-boot/logging-application-properties/>

<https://www.timroes.de/2015/02/07/kibana-4-tutorial-part-3-visualize/>

ELASTIC

- klaster ma nody, w klastrze sa elastic search indexy

- indexy rozciagaja wiele nodes na odlamki (shards)

- shard to lucene index

- index sklada sie z segmentow – mini indexow

- segment ma wiele wartosci,

---

- kazde dodanie danych do elastic to dodanie do indexu, by moglo byc potem wyszukane, dodanie do indexu Lucene



\_source to dokument co byl indexowany, od \_ to metadane



- mamy czas zapytania, czy timed\_out, przeszukane shards, ...

- username:johnb – documents where username equal “johnb”

john\* – start with john then zero or more

john? – start with john and only one char

- Bardziej rozbudowane przez DSL

-Przy update Elastic wazna zasada

Minor version changes (from 2.X to 2.Y) support **rolling upgrades** (one node at a time), but major version updates (from 1.X to 2.X) **require full cluster restarts**.

- dobra dokumentacja ja update by nie bylo bledow (https://logz.io/learn/complete-guide-elk-stack/#Elasticsearch Cluster Setup & Upgrading )

--

Na prod wazna over lokacja shards, by potem po dodaniu noda mozna bylo shard przerzucic, bo ilosc shard na wejscie dla algorytmu indexowania wazna, jak chcemy nowy dodac, to reindex, co zabiera czas

<https://www.elastic.co/guide/en/elasticsearch/guide/current/overallocation.html>

- przy dodatkowym hardware – nodach, mozemy wiecej replica shards, co nam przyspiesza czytanie, gdy duze obicazenie mamy load balancing

- **elasticsearch.yml** - okresla do jakiego clustra nalezy i jaki typ node jest – master czy replica

- monitoring wydajnosci elastic <https://www.datadoghq.com/blog/monitor-elasticsearch-performance-metrics/>

- z default mamy 5 primary shards, 5 replica, ilosc primary nie moze byc zmieniona po stworzeniu indexu , po zmienie wymagany reindex, zmiana replica zasze OK

- master node zapewnia, ze replika shard nie jest na tym samym node co master shard

3 rodzaje node Elastic :

- Master Eligable, z default taki, jak pada inne takie wybieraja mastera z ME, master koordynuje zadania klastra jak dystrybucja shards na nodach czy dodaj/usun index , ME moze tez byc jako data node , ale mozna w konfig okreslic by nie byl (duze klastry, zapewnienie ze master obsluzy robote jak inny padnie, **node.data: false**)

- data node – z default kazdy, 3ma dane w formie shards,

- **Client nodes**: **node.master** and **node.data** to **false**, designed to act as a load balancer that helps **route indexing and search requests**.

Nie trzyma indexu, ale przesyla req do odpowiedniego data node , pomocne gdy duzy ruch, by sie mogly data node skupiac na danych

- dodawanie dokumentu

-- dodany do buffer, robiony segment, w fazie refresh index, dodawany segment do shard a buffer w pamieci czyszczony , po refresh dok jest wyszukiwalny,jak update dok to wyrzucenie i dodanie nowego, bo segment sie nie zmienia a shard z segmentow sie sklada

- czyli shard z wielu segmentow, ktore dodawane a potem laczone razem w background dla lepszej wydajnosci

- przy szukaniu przeszukany kazdy segment z shard

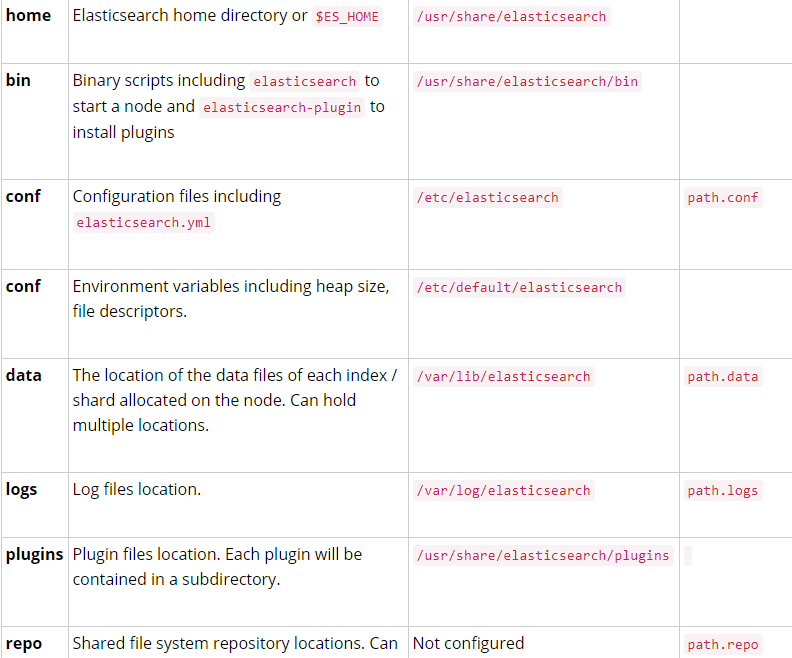
-- nowe dokumenty sa tez dodawane do shards translog wtedy co do buffera, to transaction log of operations. Co 30mins, lub jak zapelni sie (512MB), wywolany jest **FLUSH** , wtedy wszystkie segmenty z pamieci sa zapisane na dysku i log jest czyszczony , to chroni przed utrata danych, pomaga odtworzyc gdy padnie node. Log co 5 sec zapisywany badz po operacjach index, delete, update, bulk

--

Kazdy shard to lucene index, ma segmenty gdzie wewnetrzny index jest – dictionary, jak wyszukiwanie po kawalku slowa, to przez kazde slowo musi przejsc co wolne jest

<https://youtu.be/PpX7J-G2PEo> - lucene index

* dobrze data jako index, bo szukamy na okreslone dni, czyli w kilku indexach
* wybieramy w jakiej zone index, mozemy bardziej popularne na szybszych maszynach



curl 'localhost:9200/\_cat/health?v'

http://localhost:9200/<index>/<type>/[<id>]

jak baza/tabela/id

* na dole performance metrics :

<http://www.bogotobogo.com/Hadoop/ELK/ELK_Elastic_Search_Tutorial.php>

MySQL => Databases => Tables => Columns/Rows

Elasticsearch => Indices => Types => Documents with Properties

elasticsearch-head

<http://www.bogotobogo.com/Hadoop/ELK/ELK_ElasticSearch_Logstash_Kibana4.php>

dla monitorowania byl marvel, jest xpack

LOGSTASH

* Pobiera dane, obrabia do json, wrzuca do ES
* **grok** (see [patterns that are distributed with Logstash](https://github.com/elastic/logstash/tree/v1.4.1/patterns" \t "_blank)) filter.
* Moze byc wiele input, filtry, output plugin
* W konfiguracji output podaje index dla elasticsearch
* **rubydebug** codec will output our Logstash event data using the **ruby-awesome-print library**.

Pluginy :

<https://logz.io/blog/5-logstash-filter-plugins/>

- grok – pole def po kolei

- mutate – mutuje, zmienia typ, laczy,

- date – okresla date z pola, jak nie ma to data wlozenia do logu – data wpisu

- json – np. Okresla ktore pole json i je potem tak wyswietla

- kv – podzial logu na klucz wartosc

- dodatkowo mozna ruby tez

- pluginow b duzo, moze warto sie przyjrzec, czesc trzeba doinstalowac

<https://www.elastic.co/guide/en/logstash/current/output-plugins.html>

- date – co jako @timestamp, default czas wczytania, czyli stare logi jako jedno

--

po obrobieniu

https://stackoverflow.com/questions/33937936/how-to-parse-json-in-logstash-grok-from-a-text-file-line

kv, ruby

https://www.google.co.uk/search?q=logstash+dynamic+fields&ie=utf-8&oe=utf-8&client=firefox-b-ab&gfe\_rd=cr&dcr=0&ei=aFj3WYr0E-rBXrf9l\_gI

https://discuss.elastic.co/t/dynamic-fields/35309

https://www.elastic.co/guide/en/logstash/current/plugins-filters-kv.html

https://stackoverflow.com/questions/24605331/logstash-grok-filter-name-fields-dynamically

----------------

json{

source => "jsonLog"

target => "parsedJson"

remove\_field=>["jsonLog"]

}

1. Wybrane pola

filter {

grok {

match => { "message" => "%{TIMESTAMP\_ISO8601:timestamp} %{LOGLEVEL:log-level} \[%{DATA:class}\]:%{GREEDYDATA:jsonField}" }

}

json {

source => "jsonField"

target => "parsedJson"

remove\_field=>["jsonField"]

}

mutate {

add\_field => {

"messageID" => "%{parsedJson[messageID]}"

}

}

date {

match => [ "timestamp" , "yyyy-MM-dd\*HH:mm:ss.SSS" ]

}

}

1. Wszystkie pola

filter {

grok {

match => { "message" => "%{TIMESTAMP\_ISO8601:timestamp} %{LOGLEVEL:log-level} \[%{DATA:class}\]:{%{GREEDYDATA:jsonField}}" }

}

kv {

field\_split => ","

value\_split => ":"

}

date {

match => [ "timestamp" , "yyyy-MM-dd\*HH:mm:ss.SSS" ]

}

}

---

Mozna konwertowac na wybrany typ

mutate {convert =&gt; ["High", "float"]}

mutate {convert =&gt; ["Open", "float"]}

WARUNKI :

* Moze byc wiele plikow config, na rozne typy/indexy
* Potrzebne warunki wy konkretne typy obejmowal konkretny filter, zle olewal, a dany typ trafial na dany output (pod konkretny index elastic)

https://logz.io/learn/complete-guide-elk-stack/#A Logstash Tutorial

---

Można reload config :

<https://www.elastic.co/guide/en/logstash/current/reloading-config.html>

Można by beats strzelal

<https://www.elastic.co/guide/en/logstash/current/plugins-inputs-beats.html>

with filebeat

<https://www.elastic.co/guide/en/logstash/current/logstash-config-for-filebeat-modules.html>

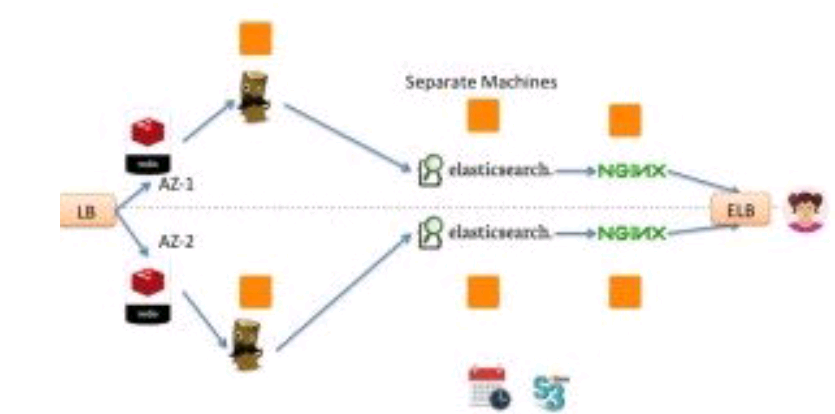
<http://www.bogotobogo.com/Hadoop/ELK/ELK_ElasticSearch_Logstash.php> - tu plugin

--

PLUGINS

https://logz.io/learn/complete-guide-elk-stack/#A Guide to Logstash Plugins

* Filter plugin na wejscie intersujacy
* Grok, mozna wlasne patterns, duzo zdefiniowanych
* Date mozna z pattern , match =&gt; [ "mydate", "MMM dd YYY HH:mm:ss", "MMM d YYY HH:mm:ss", "ISO8601" ]
* Redis I Kafka pluginy na wyjscie
* Format logow musi byc spojny !!!, ustalenie, ze pod indexem jeszcze type, ze dane pole ten sam typ wszedzie



- usuwanie indexow

Elasticsearch Curator to delete indices. We recommend having a cron job that automatically spawns Curator with the relevant parameters to delete any old indices, ensuring you don’t end up holding too much data. It is commonly required to save logs to S3 in a bucket for compliance, so you want to be sure to have a copy of the logs in their original format. Copying should be done before logs are parsed by Logstash.

- zwracac uwage na upgrade dobry

https://logz.io/learn/complete-guide-elk-stack/#How to Deploy the ELK Stack in Production

elasticsearch{

hosts => ["http://es01:9200", "http://es02:9200"]

}

Monitor performance , mozna razem z logami wyswietlic

https://logz.io/learn/complete-guide-elk-stack/#How to Use the ELK Stack to Monitor Performance

-----------------------------------------------------------------

Kazdy logstash ma swoj redis

- redisy i elastic na osobnych maszynkach – duzo pamieci , duze zuzycie JVM

KIBANA

<https://www.elastic.co/webinars/getting-started-kibana?baymax=rtp&elektra=docs&storm=top-video&iesrc=ctr>

* Jak przy szukaniu nie okreslimy pola to idzie po wszystkich
* W „”dokladne szukanie, bez tego nie case sensitive
* [\*] means any number of characters, and [?] means only one character
* Szukanie po polu

geoip.country\_name:Canada, name:chrome,name:Chrome, name:Chr\*,response:200,bytes:65,bytes:[65 TO \*],bytes:[65 TO 99]

* If you use {}, this means that the results are exclusive
* Using the **\_exists\_** prefix for a field will search the documents to see if the field exists
* AND, OR, NOT to samo co – i !
* Znaki specjalne : + – && || ! ( ) { } [ ] ^ ” ~ \* ? : \
* Proximity search, jak roznica o n liter, ale duze zuzycie pamieci , [categovi~2],

---

<https://www.timroes.de/2016/05/29/elasticsearch-kibana-queries-in-depth-tutorial/>

- szukanie w kibana

- po wlozeniu stringi na male litery, pociete i inverted index – z linkiem do pierowtnego dokumentu,

Mozna sort po ilosci wystapien

- mamy nanlized datased – czyli rozbite na slowa i unanalized, gdzie bedzie caly tekst, czyli w author:ben wyszuka w analized, bo wdrugim imie i nazwisko, wazne bez spacji po :, bo wtedy nie znajdzie

- jak szuka po kilka to w „”, „imie nazwisko”, nie znajdzie w unanalyzed – bo tam case sensitive, ale znajdzie w analyzed

- mamy wildcards, ? to 1, a \* dowolna ilosc znakow

- w „” wildcards nie dziala jak powinno, bo „imie nazwisko\*” nie znajdzie imie nazwiskoxxx, tylko kompletne dopasowanie

- uzywajac wildcards, zamienia wyszukiwanie na male litry, wiec nie wyszuka w unanalized gidze wielkie sa - earching for author:Doug\* will be converted to author:doug\*

- jak nie podam pola, szuka w \_all, czyli calym \_all:Douglas.

WIECEJ NA TEMAT SZUKANIA W KIBANA TU :

<https://www.timroes.de/2016/05/29/elasticsearch-kibana-queries-in-depth-tutorial/>

DODATKI

- XPACK

Zpac to dodatkowe alerty, monitoringi, grafy , dodane logowanie do kibana elastic/changeme

- mintoring – zdrowie klastru elastic i kabana,

Pod dashbord jest reportin button -> do pdf

-Jest logout

- to monitoring I reportin tool

- BEAT

, co sle logi z wielu maszyn do logstash, np. Filebeat do obrobki textowych i wysylania na logstash

- ALERTING

- warunki na elasticsearch danych

- monitoruje logi pod kontem warunkow

- sle notyfikacje przez email, Slack, Hipchat, PagerDuty

With docker

<https://logz.io/blog/migrating-to-docker/>

LINKI

Start : <https://blog.webkid.io/visualize-datasets-with-elk/>

-----------

Warto – na co uwazac (How to Use AWS Elasticsearch for Log Management)

* https://logz.io/learn/complete-guide-elk-stack/#How to Deploy the ELK Stack in Production

Architektura

* Redis by nie tracic danych jak indexowanie zwalnia odbieranie, zapewnia, ze nic nam nie ucieka

PRZYKLADY :

https://www.elastic.co/blog/getting-started-with-elk

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

https://www.google.pl/search?q=elasticsearch+cluster+openshift&oq=elastic+serch+cluster+op&aqs=chrome.1.69i57j0l4.7655j0j7&sourceid=chrome&ie=UTF-8

https://medium.com/@happymacaron/how-to-set-up-elasticsearch-on-openshift-405d0460c818

https://www.elastic.co/blog/found-logstash-openshift

https://www.google.pl/search?ei=LkAIWtXBAsXb6QSL8aXgCw&q=elsticsearch+cluster+docker+openshift&oq=elsticsearch+cluster+docker+openshift&gs\_l=psy-ab.3...343483.343483.0.343856.1.1.0.0.0.0.89.89.1.1.0....0...1.1.64.psy-ab..0.0.0....0.YPZl4fuM9eU

https://docs.openshift.com/container-platform/3.4/install\_config/aggregate\_logging.html#aggregated-elasticsearch

https://www.elastic.co/guide/en/elasticsearch/reference/current/docker.html

https://github.com/fabric8io/elasticsearch-cloud-kubernetes