ENSET

elasticsearch logstash kibana

Introduction à :

ElasticSearch

Preparé par :

- Soukayna AYOUJJIL
- Ilham IDRISSI
- Essadeq El AAMIRI

Encadré par:

- M.Abdelmajid BOUSSELHAM

Année universitaire: 2021/2022

PLAN

01 Presentation d'ElasticSearch

02 Installation

03 concepts de bases

04 exploitation d'ElasticSearch

05 Travaux pratiques

1

Presentation d'ElasticSearch

C’est quoi Elasticsearch ?

elasticsearch

Elasticsearch est un moteur de recherche et d'analyse distribué et open-source basé sur Apache Lucene et développé en Java.

https://www.elastic.co/guide/en/elasticsearch/reference/current/getting-started.html

C’est quoi Elasticsearch ?

Stockage

Analyse

Recherche

Client

JSON

elasticsearch

Des données volumineux

2

Installation d'ElasticSearch

Guide d'installation d'Elasticsearch et Elsaticsearch-head

3

Les concepts clés d'ElasticSearch

Les conceptes logiques

Documents

Les documents sont l'unité de base des informations qui peuvent être indexées dans Elasticsearch, exprimées en JSON. On peut considérer un document comme une ligne (row) dans une base de données relationnelle, représentant une entité (entrée) donnée.

Chaque document possède un identifiant unique et un type de données.

Les conceptes logiques

Indices

Un index est une collection de documents qui ont des caractéristiques similaires. Un index est l'entité de plus haut niveau sur laquelle vous pouvez effectuer des requêtes dans Elasticsearch. On peut considérer que l'index est similaire à une base de données dans un système de base de données relationnelle. Tous les documents d'un index sont généralement liés logiquement. Un index est identifié par un nom qui est utilisé pour faire référence à l'index lors des opérations d'indexation, de recherche, de mise à jour et de suppression des documents qu'il contient.

Les conceptes logiques

Index inversé

C'est le mécanisme par lequel tous les moteurs de recherche fonctionnent. Il s'agit d'une structure de données qui stocke une correspondance entre le **contenu**, tel que des mots ou des chiffres, et son **emplacement** dans un document ou un ensemble de documents. En gros, il s'agit d'une structure de données de type hashmap qui vous dirige d'un mot vers un document. Un index inversé ne stocke pas directement les chaînes de caractères, mais divise chaque document en termes de recherche individuels (c'est-à-dire chaque mot), puis associe chaque terme de recherche aux documents dans lesquels il apparaît.

Les conceptes logiques

Index inversé

Documents 1 & 2

The bright blue butterfly hangs on the breeze

Under blue sky, in bright sunlight, one need no search around

ID	Term	Doccument
1	butterfly	1
2	blue	1,2
3	bright	1,2
4	retire	2
5	wind	2

Les composants Back-end

Cluster

Un cluster Elasticsearch est un groupe d'une ou plusieurs instances de nœuds qui sont connectées entre elles. La puissance d'un cluster Elasticsearch réside dans la distribution des tâches, la recherche et l'indexation, sur tous les nœuds du cluster.

Les composants Back-end

Node

Un nœud est un serveur unique qui fait partie d'un cluster. Un nœud stocke des données et participe aux capacités d'indexation et de recherche du cluster. Un nœud Elasticsearch peut être configuré de différentes manières :

Master Node, *Data Node*, *Client Node*

Les composants Back-end

Node

Master Node - Contrôle le cluster Elasticsearch et est responsable de toutes les opérations à l'échelle du cluster, comme la création/suppression d'un index et l'ajout/suppression de nœuds.

Data Node - Stocke les données et exécute les opérations liées aux données telles que la recherche et l'agrégation.

Client Node - Transmet les demandes de cluster au nœud maître et les demandes liées aux données aux nœuds de données.

Les composants Back-end

Shards

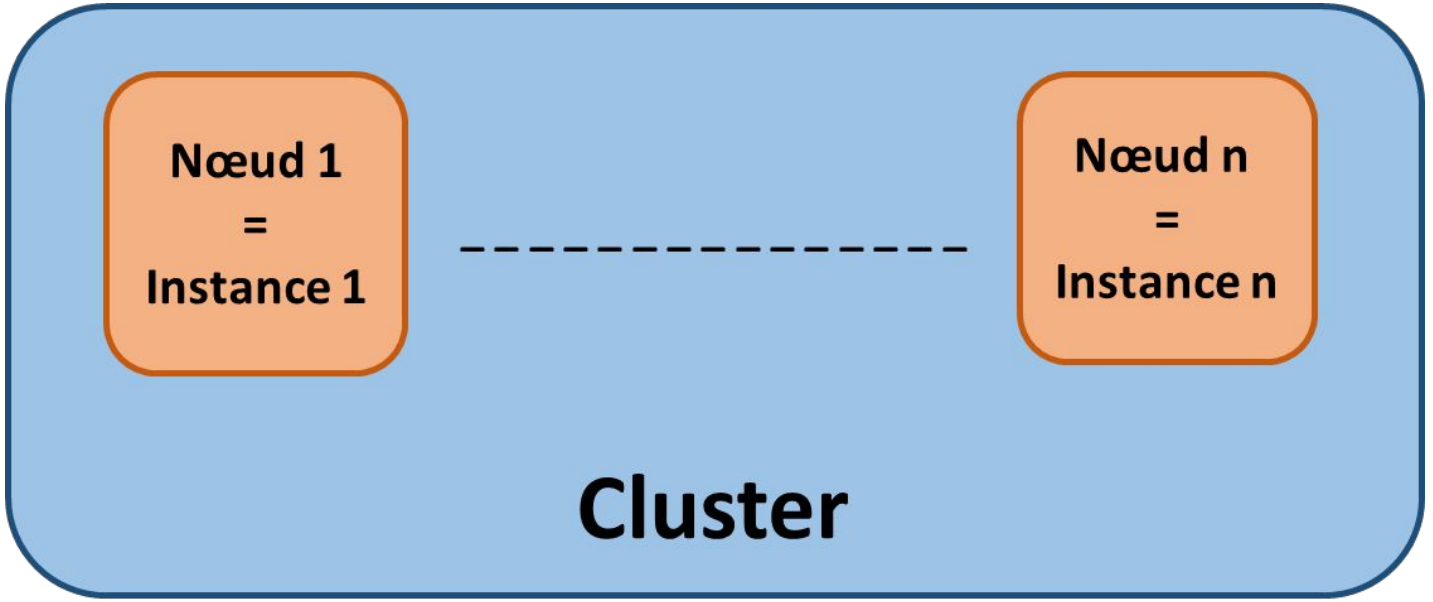
Elasticsearch offre la possibilité de subdiviser l'index en plusieurs morceaux appelés shards. Chaque shard est en soi un "index" entièrement fonctionnel et indépendant qui peut être hébergé sur n'importe quel nœud d'un cluster. En répartissant les documents d'un index sur plusieurs shards, et en répartissant ces shards sur plusieurs nœuds, Elasticsearch peut assurer la redondance, ce qui permet à la fois de se protéger contre les pannes matérielles et d'augmenter la capacité de requête lorsque des nœuds sont ajoutés à un cluster.

Les composants Back-end

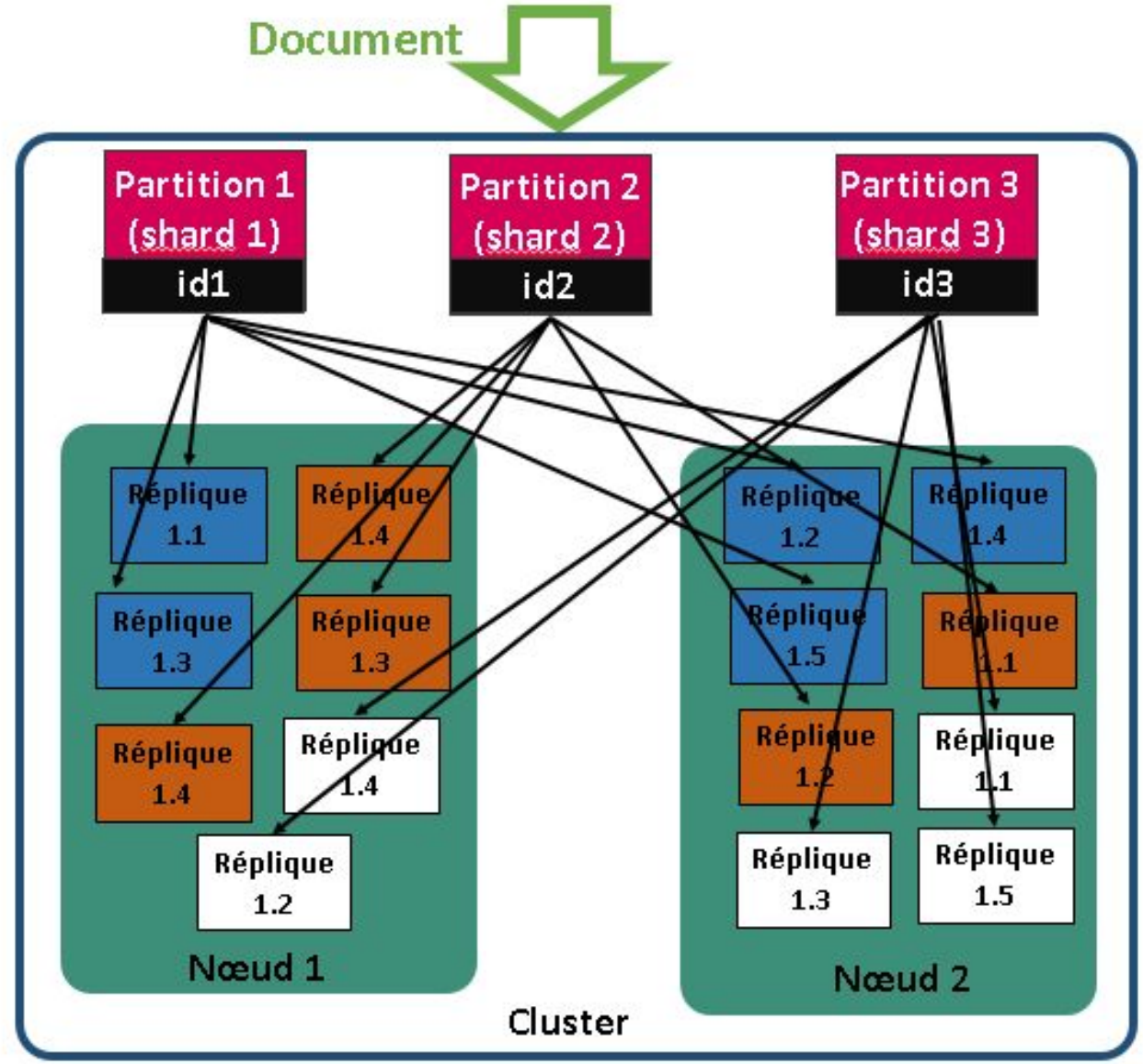
Replicas

Elasticsearch nous permet de faire une ou plusieurs copies des shards de notre index, appelées "replica shards" ou simplement "replicas". Fondamentalement, un replica shard est une copie d'un shard primaire. Chaque document dans un index appartient à un shard primaire. Les répliques fournissent des copies redondantes de nos données afin de nous protéger contre les pannes matérielles et d'augmenter la capacité de traitement des requêtes de lecture telles que la recherche ou la récupération d'un document.

Architecture d'ElasticSearch



Traitement des données



Qu'est-ce que Elastic Slack (ELK stack) ?

Elastic Slack est un écosystème complet d'outils open-source pour l'ingestion, l'enrichissement, le stockage, l'analyse et la visualisation de données. Outre Elasticsearch, les autres logiciels sont Logstash, Kibana et Beats.

https://www.elastic.co/what-is/elk-stack

Principaux cas d'utilisation

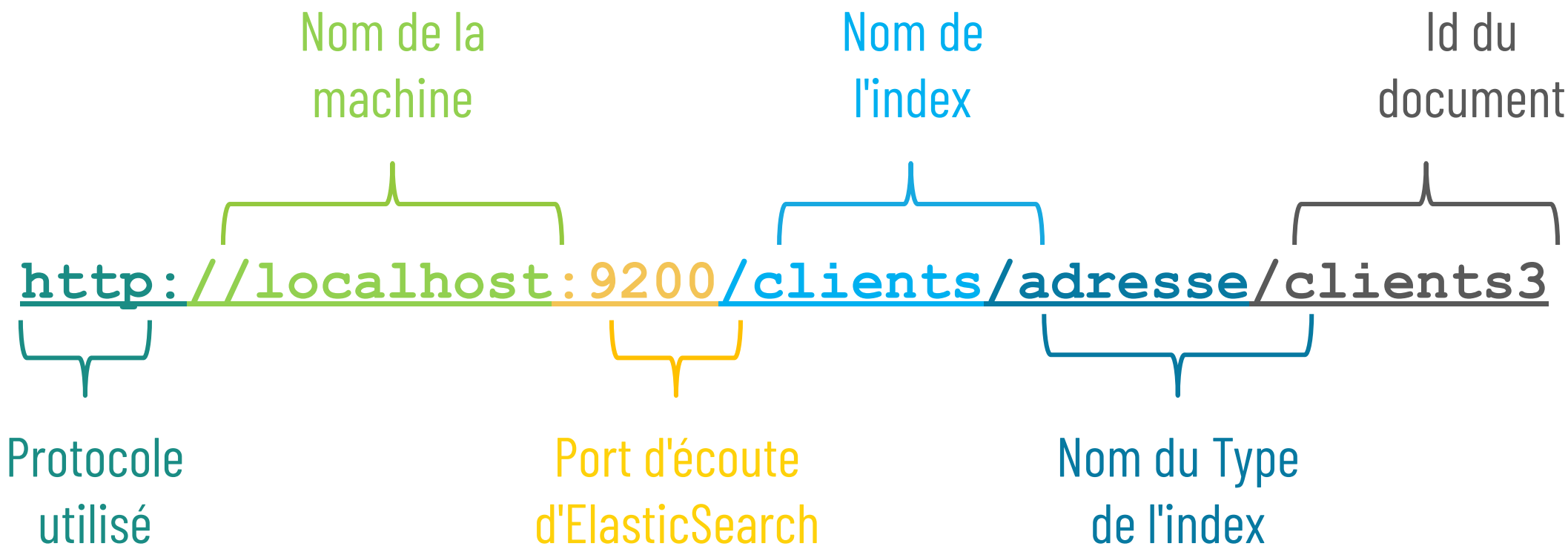
- Recherche dans les applications
- Recherche sur site Web
- Analyse de la sécurité (Security analytics)
- Business analytics

Qui utilise Elasticsearch ?

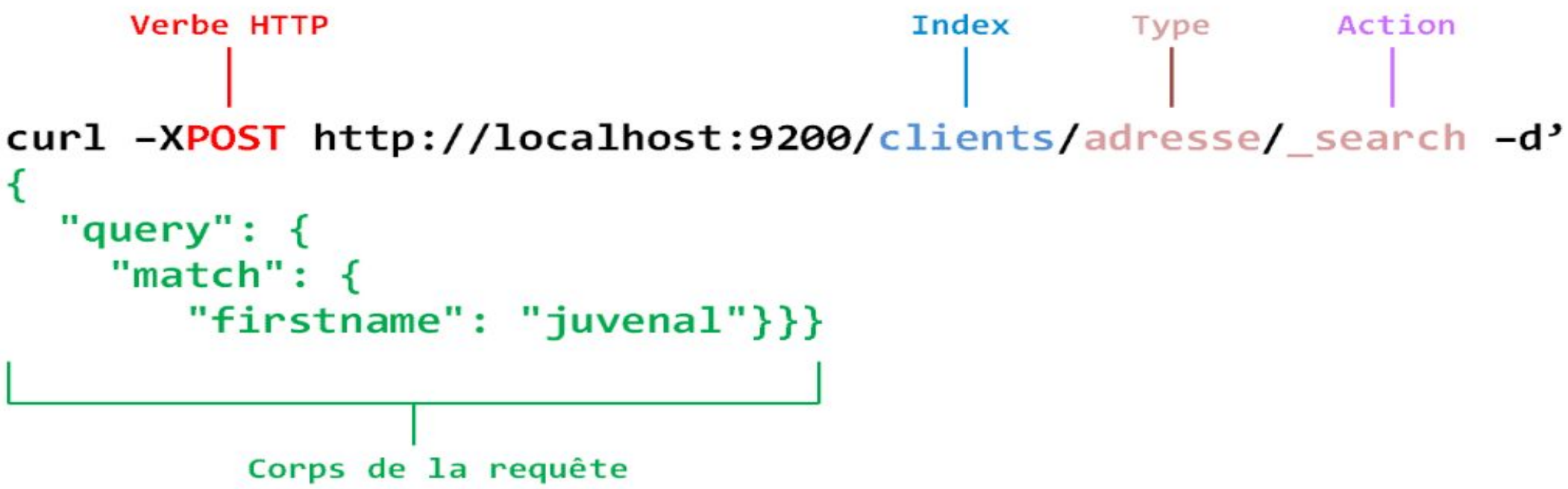
4

Exploitation d'ElasticSearch

1 structure de l'URI ElasticSearch

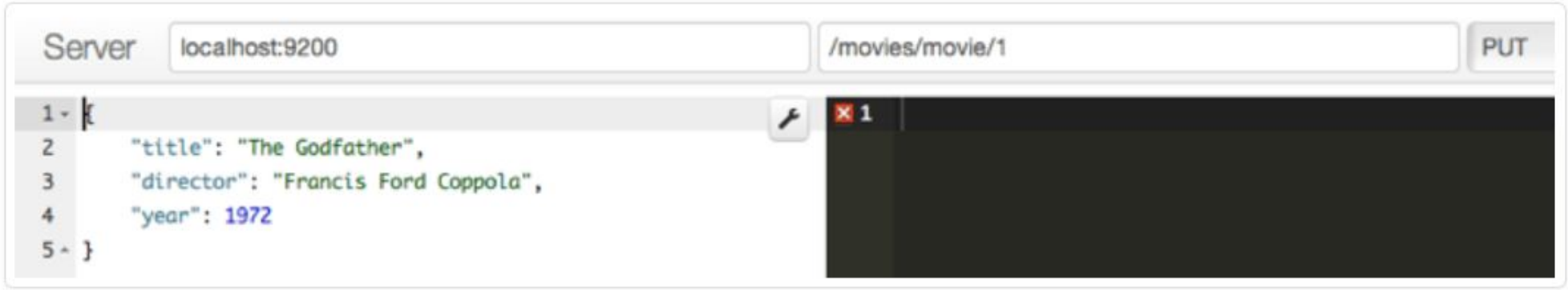


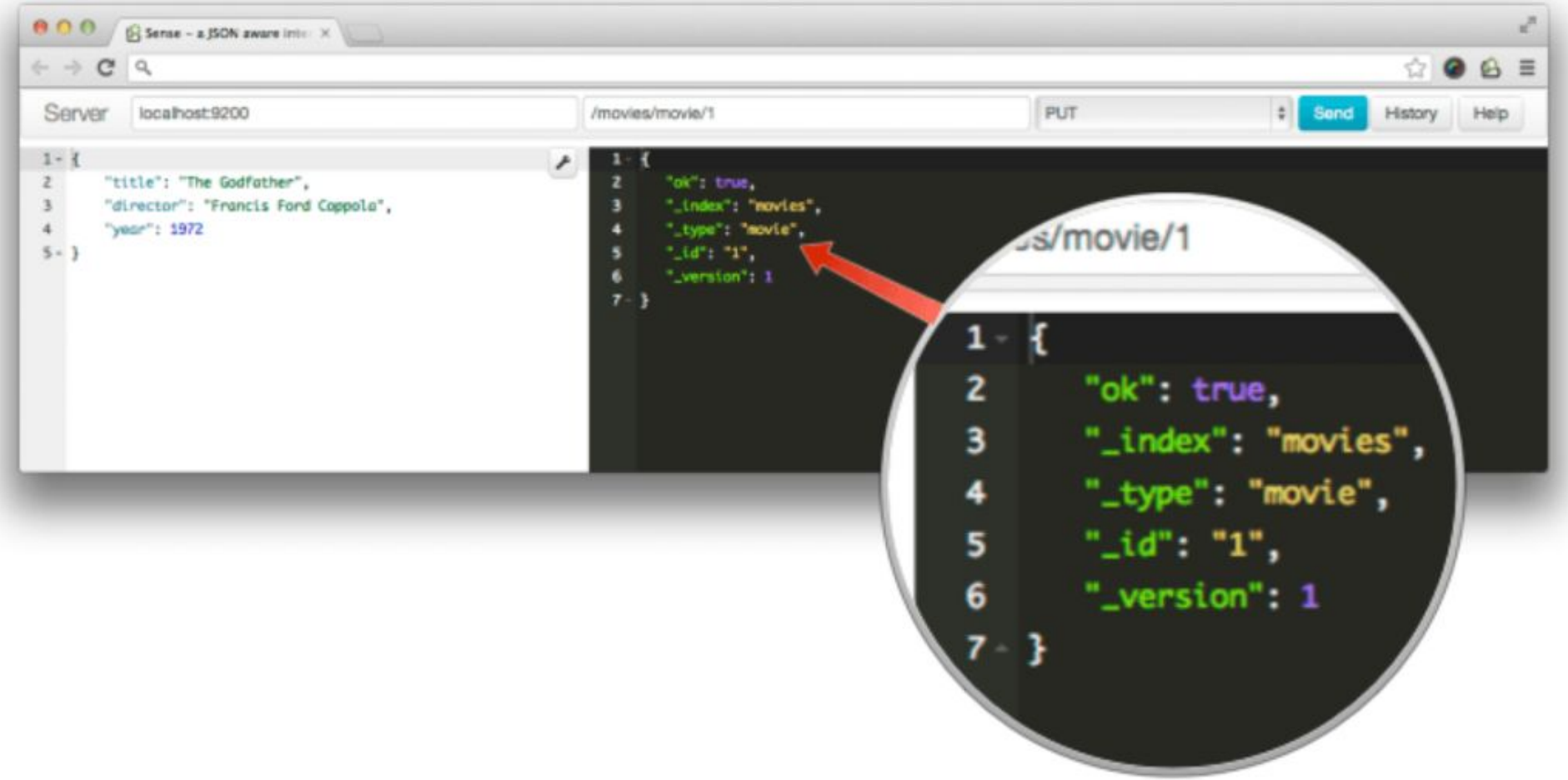
2 syntaxe des requêtes adressées à ElasticSearch via cURL



Ajouter un documents

```
curl -XPUT "http://localhost:9200/movies/movie/1" -d'
{
    "title": "The Godfather",
    "director": "Francis Ford Coppola",
    "year": 1972
}'
```




Server
localhost:9200
/movies/movie/1
PUT
Send
History
Help
{
"title": "The Godfather",
"director": "Francis Ford Coppola",
"year": 1972
}
{
"ok": true,
"_index": "movies",
"_type": "movie",
"_id": "1",
"_version": 1
}

Mise à jour d'un documents

```
curl -XPUT "http://localhost:9200/movies/movie/1" -d'
{
    "title": "The Godfather",
    "director": "Francis Ford Coppola",
    "year": 1972,
    "genres": ["Crime", "Drama"]
}'
```

Suppression de documents

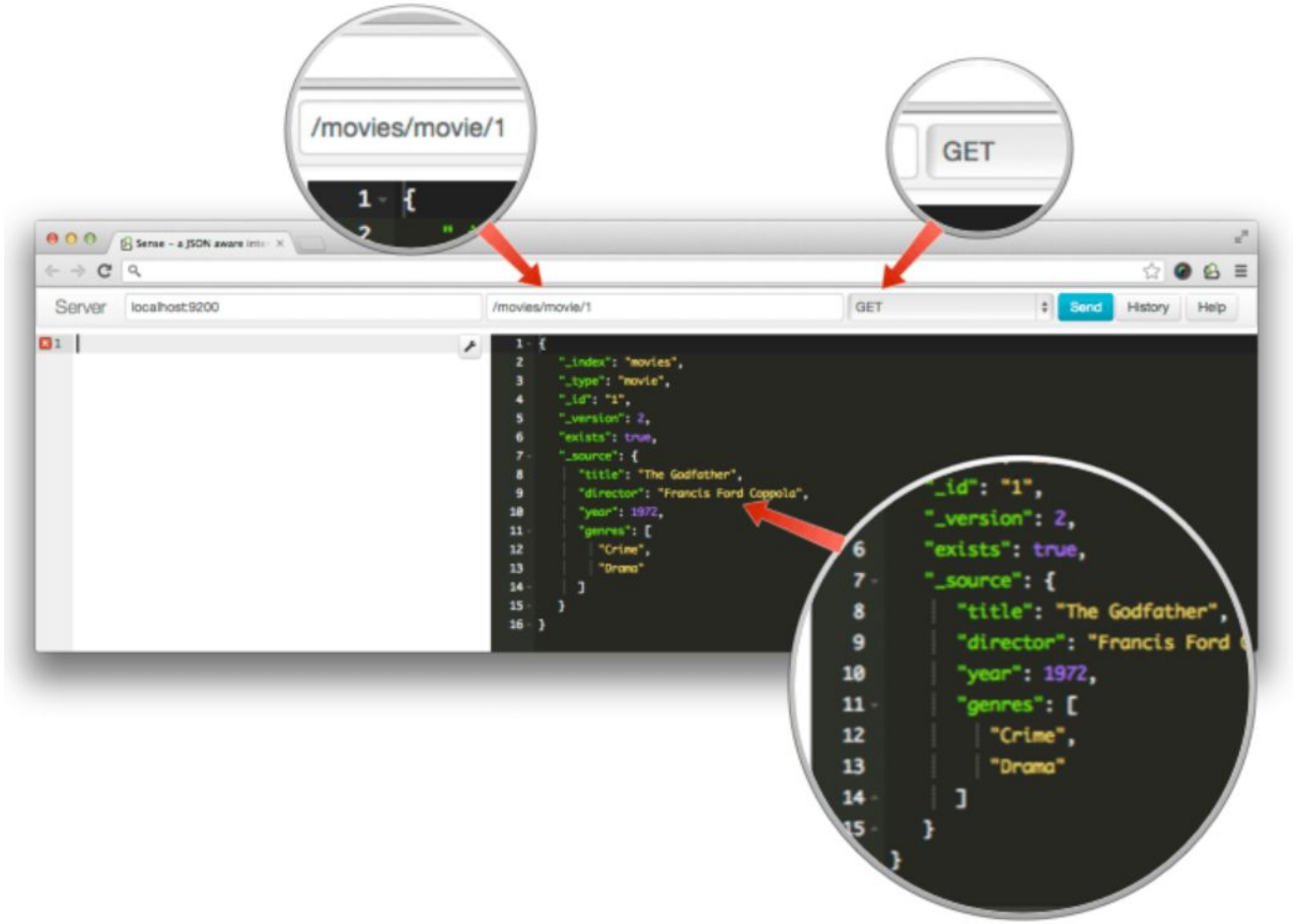
```
curl - XDELETE "http://localhost:9200/movies/movie/1" - d ''
```

/movies/movie/1 DELETE

```
{
    "ok": true,
    "found": true,
    "_index": "movies",
    "_type": "movie",
    "_id": "1",
    "_version": 3
}
```

Obtention par ID

```
curl - XGET "http://localhost:9200/movies/movie/1" - d ''
```



Recherche

- **Le point de terminaison _search**

Requêtes à une URL suivant ce modèle : <index>/<type>/_search

Afin de rechercher des films, nous pouvons faire des requêtes POST à l'une des URL suivantes :

- http://localhost:9200/_search : Recherche dans tous les index et tous les types.
- http://localhost:9200/movies/_search : Recherche parmi tous les types dans l'index des movies.
- http://localhost:9200/movies/movie/_search : Recherche explicitement des documents de type movie dans l'index des movies.

Recherche

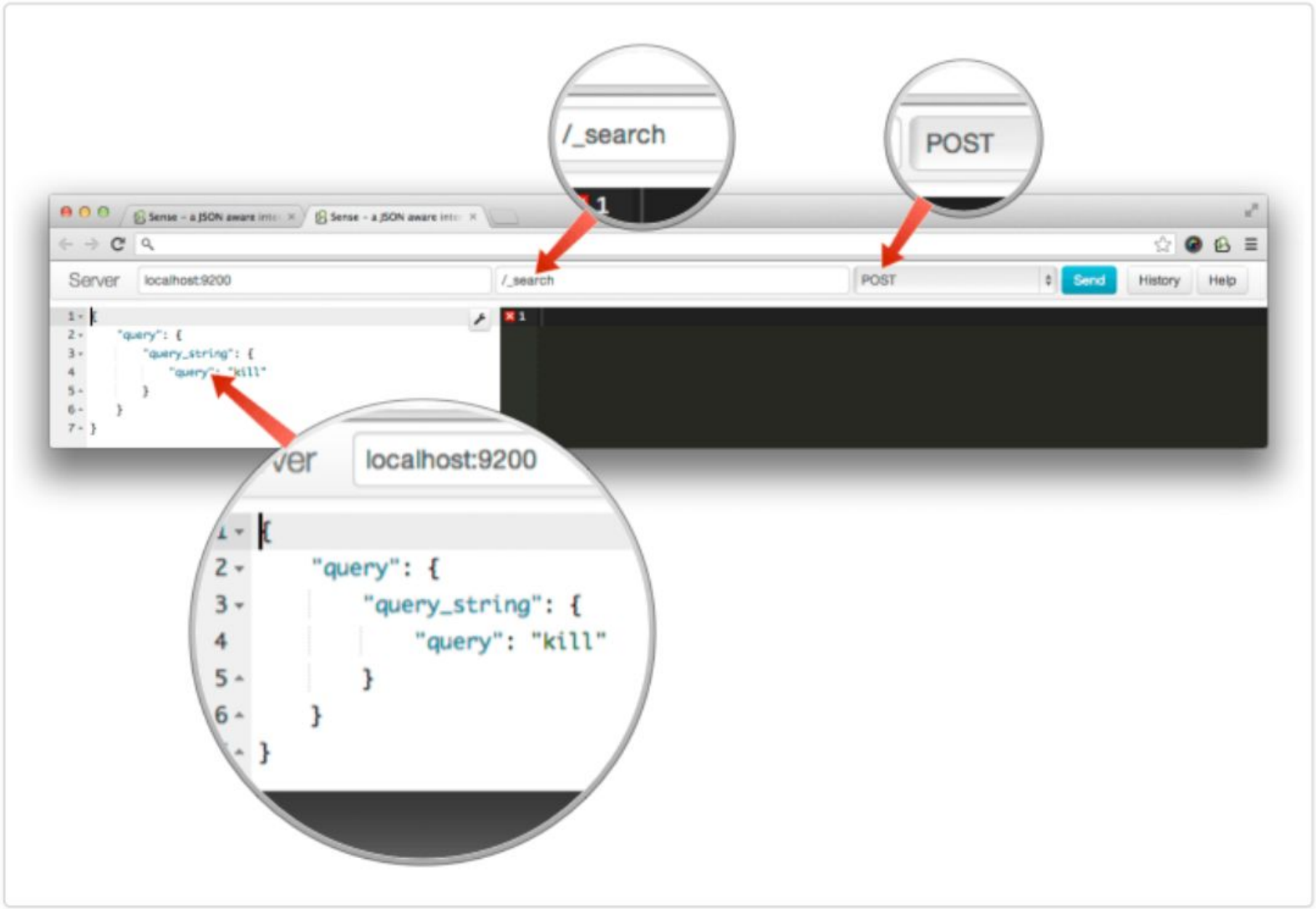
- **Corps de la requête de recherche et requête DSL d'ElasticSearch**

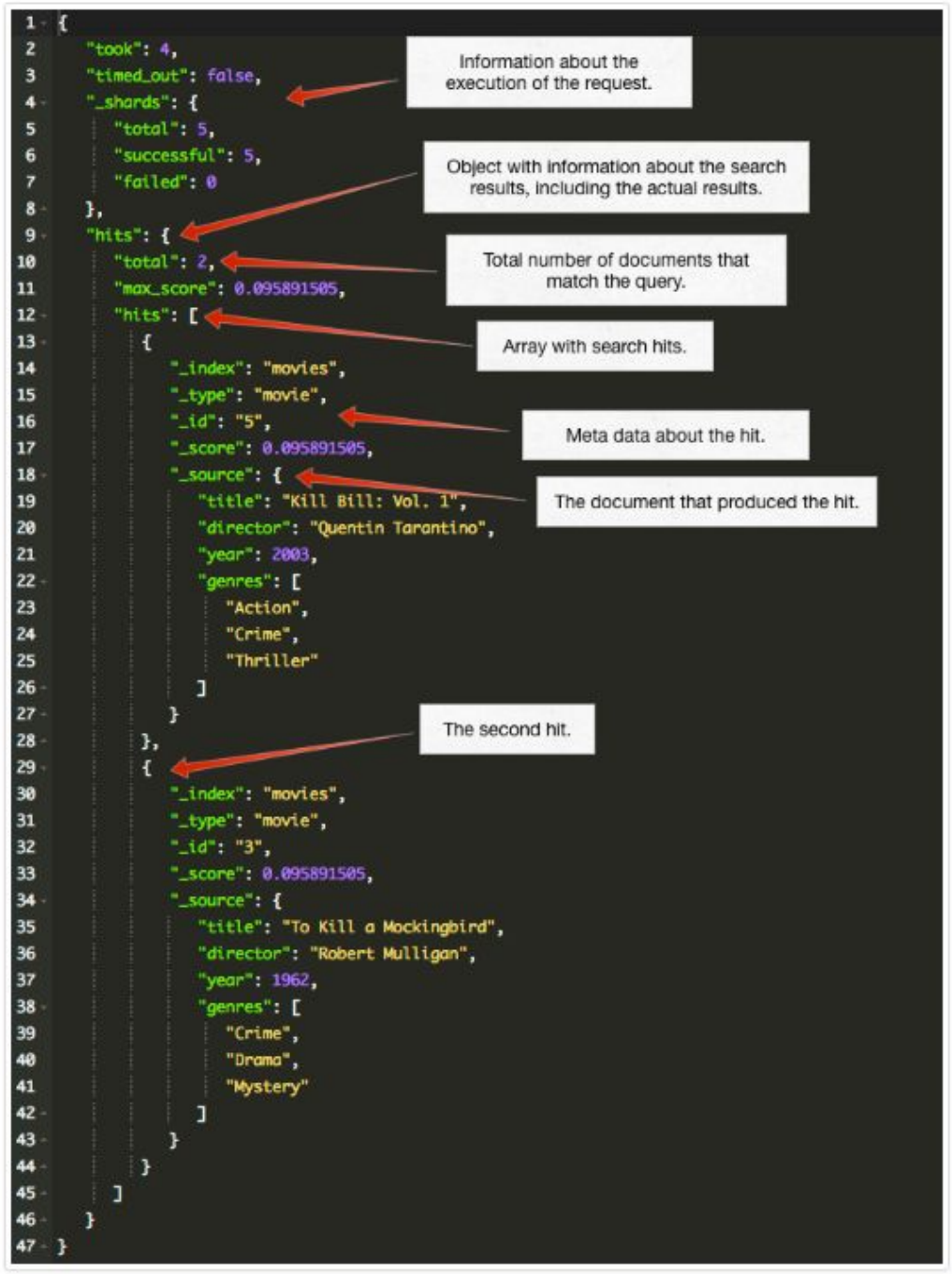
```
{
    "query": {
        //Query DSL here
    }
}
```

Recherche

- **Recherche de base en texte libre**

```
curl -XPOST "http://localhost:9200/_search" -d'
{
    "query": {
        "query_string": {
            "query": "kill"
        }
    }
}'
```


/_search
POST
Server
localhost:9200
/_search
POST
Send
History
Help
{
"query": {
"query_string": {
"query": "kill"
}
}
}
localhost:9200


Information about the execution of the request.
Object with information about the search results, including the actual results.
Total number of documents that match the query.
Array with search hits.
Meta data about the hit.
The document that produced the hit.
The second hit.

Plus d'informations

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

5

Travaux pratiques

Merci

Pour votre attention

Introduction à :

ElasticSearch

Preparé par :

- Soukayna AYOUJJIL
- Ilham IDRISSI
- Essadeq El AAMIRI

Encadré par:

- M.Abdelmajid BOUSSELHAM

Année universitaire: 2021/2022