

Jouons un peu avec ElasticSearch

Apéro Web Nancy – 25 septembre 2014

Thomas Jost – thomas@findspire.com – [@Schnouki](https://twitter.com/Schnouki)

ElasticSearch c'est quoi ?

- Une base de données NoSQL orientée documents
- ...distribuée, avec sharding et réplication automatique au sein du cluster
- Un moteur de recherche basé sur Lucene
 - Critères classiques : $A = B$, $A \leq B$, AND, OR, ...
 - Recherche plein-texte, avec des analyseurs aussi complexes qu'on veut
 - Recherche par critères géographiques

Un exemple pratique

Un soir de semaine... Il est 18h et Dave Loper, habitant de la belle ville de Nancy, sort du travail. Mais en ce beau mois de septembre 2014, il fait si beau et chaud que Dave a envie d'aller se désaltérer.

Il faut donc prendre des décisions. Quel(s) bar(s) ? Quelle(s) boisson(s) ?
Le choix est dur...

Dave décide alors de faire appel à l'outil ultime du développeur assoifié...

ElasticSearch !

Un petit jeu de données...

- Une centaines de bars et restaurants nancéiens (source : Google Places)
- Quelques milliers de sortes de bières (source : Brewery DB)
- Une liste de bières par bar (factice, générée aléatoirement)

À quoi ressemble une requête ?

- Pas de SQL mais un *query DSL* spécifique
- Dictionnaire JSON POSTé sur une URL
- Dedans : query, filter, agrégations, ...
 - *query* : tout type de recherche, scoring ; idéal pour des recherches plein-texte
 - *filter* : pas de scoring → performances++, résultats cachés automatiquement ; idéal pour des recherches booléennes ou sur termes exacts

C'est parti pour la démo...

Quelques mots sur Kibana

- ELK : ElasticSearch, Logstash, Kibana
- Logstash : collecte les logs et événements, les parse, les transforme si nécessaire, les stocke dans...
- ElasticSearch : stocke et indexe les logs et événements, afin de pouvoir faire des requêtes avec...
- Kibana : présente les données stockées sous forme de *dashboards* avec différents types de visualisations.

Deux exemples d'utilisation de Kibana

- Marvel : pour le monitoring de cluster ElasticSearch
- Les dashboards de Findspire

Pfiou, c'était long sa présentation ! Et y'avait quasiment pas d'images, j'ai rien compris 😞 Et puis c'est compliqué son machin ! Franchement qui a besoin de tout ça quand on a `mysql_query()` ? Il a même pas parlé de sécurité et de `mysql_real_escape_string()`, le noob ! Bon, c'est pas tout ça, mais j'ai soif moi...

Questions ?

Thomas Jost – thomas@findspire.com – [@Schnouki](https://twitter.com/Schnouki)

Merci pour votre attention !

Slides + démos + dataset disponibles demain sur schnouki.net

Bon, on va picoler ?