ElasticSearch

Plan

- Introduction
- Qu’est-ce que Elasticsearch?
- Des atouts indéniables
- Quelques problèmes à régler
- Des exemples concrets
- Les alternatives à ElasticSearch
- Conclusion

Introduction

- Première release février 2010 – Shay Banon
- 2014 – Levée de 104 millions de $
- Moteur de recherche open source basé sur Lucene

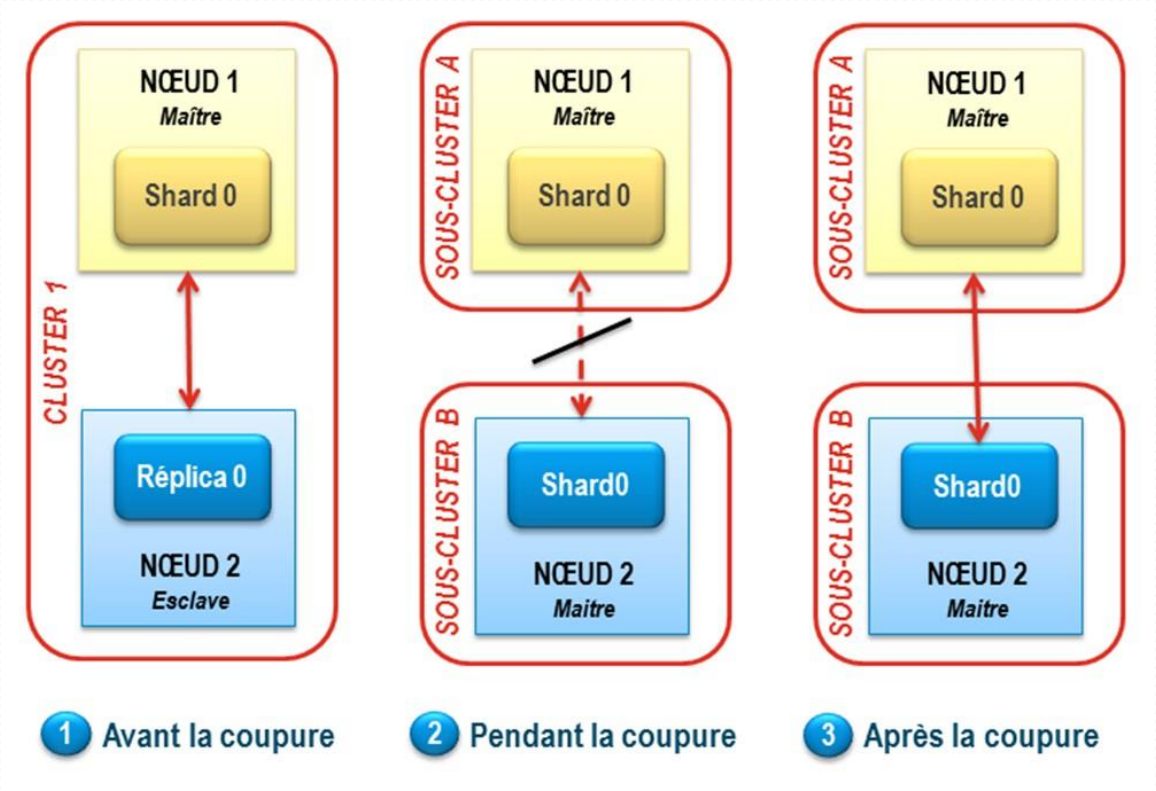
Qu’est ce que Elasticsearch ?

- Application modulaire
- Fonctionne en cluster
- Un moteur de recherche complet :
  - Recherche par mot clef
  - Recherche par combinaison de critères et de filtres
  - Tri et pagination des résultats
  - Gestion des synonymes
  - Navigation par facettes

Des atouts indéniables

- Beaucoup de release : un outil qui évolue vite
- Analyse des données en temps réel
- Robustesse
- Contrôle pour éviter les conflits
- Indexage customisable (JSON)
- API complète et facile à utiliser via le protocole http et le format JSON
- Injection de données via des Rivers (Wikipedia River, Twitter River…)
- Navigation par facet (terms, range…)

Quelques problèmes à régler



Des exemples concrets

- Maptimize : Afficher des millions de points sur une map
- Deezer : Gérer des millions de recherches et un catalogue avec 40 millions d'entrées
- GitHub: Indexer 8 millions de répertoires de codes

Les alternatives à ElasticSearch

- Apache Solr
- Sphinx
- Xapian

Conclusion

- Flexible, puissant et distribué, ElasticSearch permet les recherches et l'analyse en temps réel
- Facile à mettre en place
- Clients pour les principaux langages de programmation (Java, C++, Python…)

Sources

- http://www.technologies-ebusiness.com/solutions/les-problematiques-elasticsearch
- http://www.elasticsearch.org/
- http://blog.zenika.com/index.php?post/2012/11/14/Premiers-pas-avec-ElasticSearch-Partie-1