

Introduction et Historique

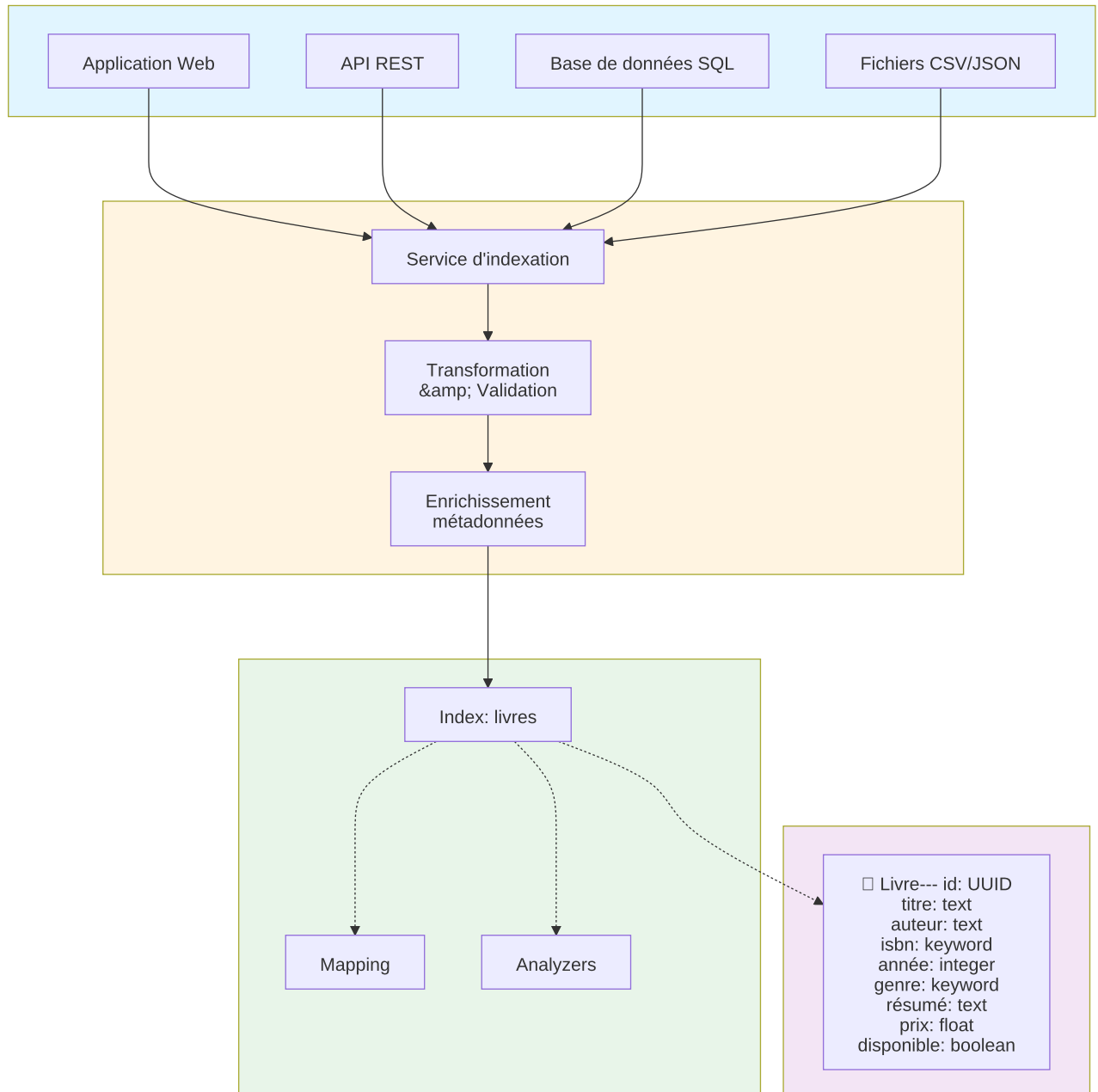
- Qu'est-ce qu'Elasticsearch ?
- Alternatives à Elasticsearch : Comparatif
 - Concurrents privés (propriétaires)
 - Alternatives libres (open source)
 - Écosystème d'Elasticsearch
 - Architecture Distribuée et Scalabilité
 - Qu'est-ce qu'une Architecture Distribuée ?
 - Les Avantages de cette Architecture
- Benchmarks
- L'Écosystème Elastic Stack
- Pourquoi Apprendre Elasticsearch ?
 - Pour les Développeurs
 - Pour les Data Analysts
 - Pour les Ops et DevOps
- Ressources Complémentaires



elasticsearch

Qu'est-ce qu'Elasticsearch ?

Elasticsearch est bien plus qu'un simple moteur de recherche. C'est une plateforme complète de recherche et d'analyse de données, développée en Java, qui permet d'indexer, de rechercher et d'analyser de gros volumes de données en temps quasi-réel.



- **Recherche textuelle**

- Utilisé par **Wikipedia** ou **GitHub** pour permettre des recherches rapides

- **Analyse de logs et observabilité**

- Au cœur de la stack **ELK (Elasticsearch, Logstash, Kibana)**.

- **Monitoring d'infrastructure**

- Utilisé avec **Metricbeat** et **Filebeat**

dans des millions de documents.

- Exemple : un utilisateur tape *"microservices Java"*, Elasticsearch renvoie les articles les plus pertinents en analysant le contenu, les titres et les balises.

- Exemple : un administrateur collecte les logs d'une application (erreurs, latence, appels API). Elasticsearch indexe ces logs pour permettre des recherches et des tableaux de bord dans Kibana :

- Rechercher `error AND payment`
- Voir le nombre d'erreurs par minute.

pour surveiller des serveurs.

- Exemple : détecter une surconsommation CPU ou mémoire sur un cluster Kubernetes. Elasticsearch stocke les métriques, Kibana affiche des courbes et déclenche des alertes.

• E-commerce

- Exemple : un site comme **Decathlon** ou **Zalando** utilise Elasticsearch pour ses filtres produits.
 - Requête : "chaussures de course homme, taille 43, prix < 100 €".
 - Réponse quasi instantanée avec tri, autocomplétion et suggestions.

• Systèmes de recommandation

- Utilisé dans les plateformes de contenu (films, musique, articles).
- Exemple : filtrer des vidéos "similaires à celles vues récemment", grâce aux agrégations et similarités de texte.

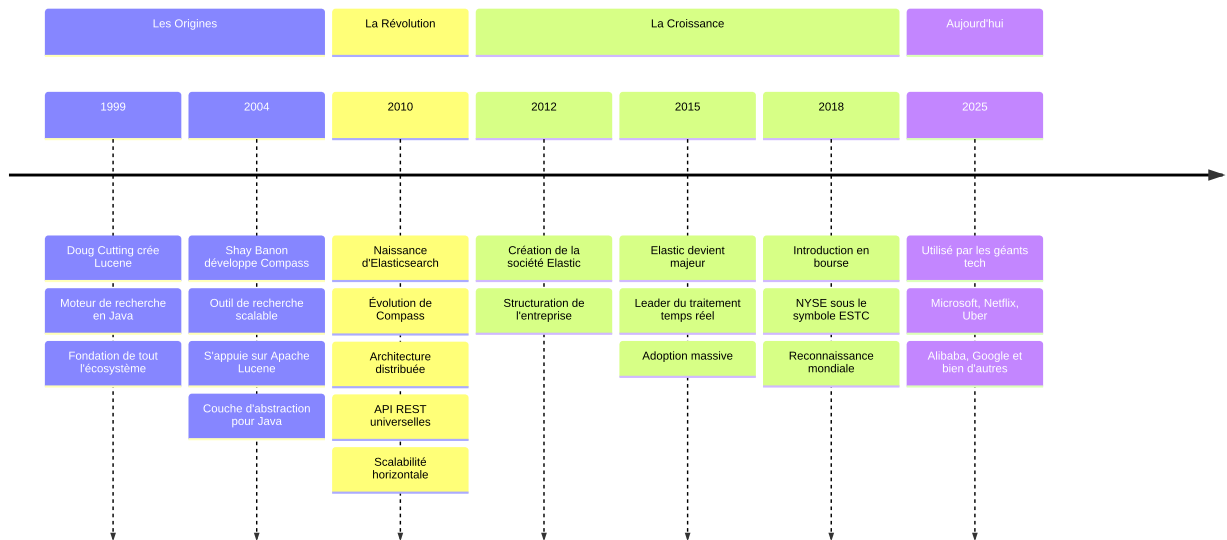
• Cybersécurité (SIEM)

- Exemple : un SOC centralise les alertes de pare-feux, antivirus, IDS. Elasticsearch indexe chaque événement pour corréliser les attaques (ex. même IP repérée sur plusieurs serveurs).

• Business Intelligence en temps réel


- Exemple : un tableau de bord de ventes qui affiche les commandes par région, catégorie ou canal, actualisé à la seconde.


L'Histoire d'Elasticsearch : De Compass à l'Entreprise Cotée en Bourse



Alternatives à Elasticsearch : Comparatif

Concurrents privés (propriétaires)

| Nom | Description | Avantages | Inconvénients |
|--|---|--|--|
|  algolia | Plateforme SaaS de recherche ultra-rapide spécialisée dans les expériences e-commerce et applications web. Offre des résultats en temps réel avec tolérance aux fautes de frappe. | <ul style="list-style-type: none"> Vitesse (< 50ms) Facilité d'implémentation Tolérance aux fautes intégrée Excellent pour e-commerce | <ul style="list-style-type: none"> Coût élevé à grande échelle Solution propriétaire fermée Tarification basée sur l'usage Moins flexible qu'Elasticsearch |
| splunk > | Plateforme d'analyse de données et monitoring spécialisée dans la conformité et la sécurité pour industries réglementées (finance, santé). | <ul style="list-style-type: none"> Sécurité et conformité robustes Excellent pour les logs Chiffrement et audit avancés Support entreprise | <ul style="list-style-type: none"> Prix très élevé Complexité d'implémentation Courbe d'apprentissage importante Surqualifié pour usage simple |

| Nom | Description | Avantages | Inconvénients |
|---|--|--|--|
|  coveo™ | Plateforme de recherche et personnalisation d'entreprise avec capacités IA et génération de réponses pour commerce et gestion des connaissances. | <ul style="list-style-type: none"> Gouvernance niveau entreprise Réponses génératives IA Intégrations d'entreprise Évolutif | <ul style="list-style-type: none"> Coût élevé Cycle d'implémentation long Moins orienté développeurs Complexe pour PME |
| Yext | Plateforme d'expérience numérique spécialisée dans la recherche géolocalisée et gestion de données de localisation précises. | <ul style="list-style-type: none"> Excellent pour recherche locale Gestion multi-emplacements Données de localisation précises Interface intuitive | <ul style="list-style-type: none"> Spécialisé géolocalisation uniquement Coût élevé Moins polyvalent Fonctionnalités limitées hors géo |

Alternatives libres (open source)

| Nom | Description | Avantages | Inconvénients |
|--------------------|--|---|---|
| Apache Solr | Plateforme de recherche d'entreprise mature construite sur Apache Lucene. Réputée pour sa fiabilité et flexibilité dans la recherche en texte intégral complexe. | <ul style="list-style-type: none"> Mature et éprouvé (Netflix, eBay) Hautement personnalisable Recherche texte puissante Gratuit et open source | <ul style="list-style-type: none"> Interface administrative datée Configuration complexe Moins orienté analytics Requiert expertise technique |
| OpenSearch | Fork open source d'Elasticsearch créé par AWS en 2021 sous licence Apache 2.0. Compatible avec l'API Elasticsearch et offrant les mêmes fonctionnalités de base. | <ul style="list-style-type: none"> Vraiment open source (Apache 2.0) Compatible API Elasticsearch Sécurité avancée gratuite Forte intégration AWS | <ul style="list-style-type: none"> Innovation plus lente Écosystème de plugins fragmenté Moins de fonctionnalités IA/vector Dépendant de AWS |

| Nom | Description | Avantages | Inconvénients |
|------------------|---|--|--|
| Typesense | Moteur de recherche open source moderne et léger avec recherche instantanée, tolérance aux fautes, recherche sémantique et vectorielle. | <ul style="list-style-type: none">• Ultra-rapide et léger• Simple à configurer• Auto hébergement économique• API développeur conviviale | <ul style="list-style-type: none">• Moins mature qu'Elasticsearch• Écosystème plus petit• Fonctionnalités limitées vs concurrents• Moins d'intégrations |

Ecosystème d'Elasticsearch

Architecture Distribuée et Scalabilité

Qu'est-ce qu'une Architecture Distribuée ?

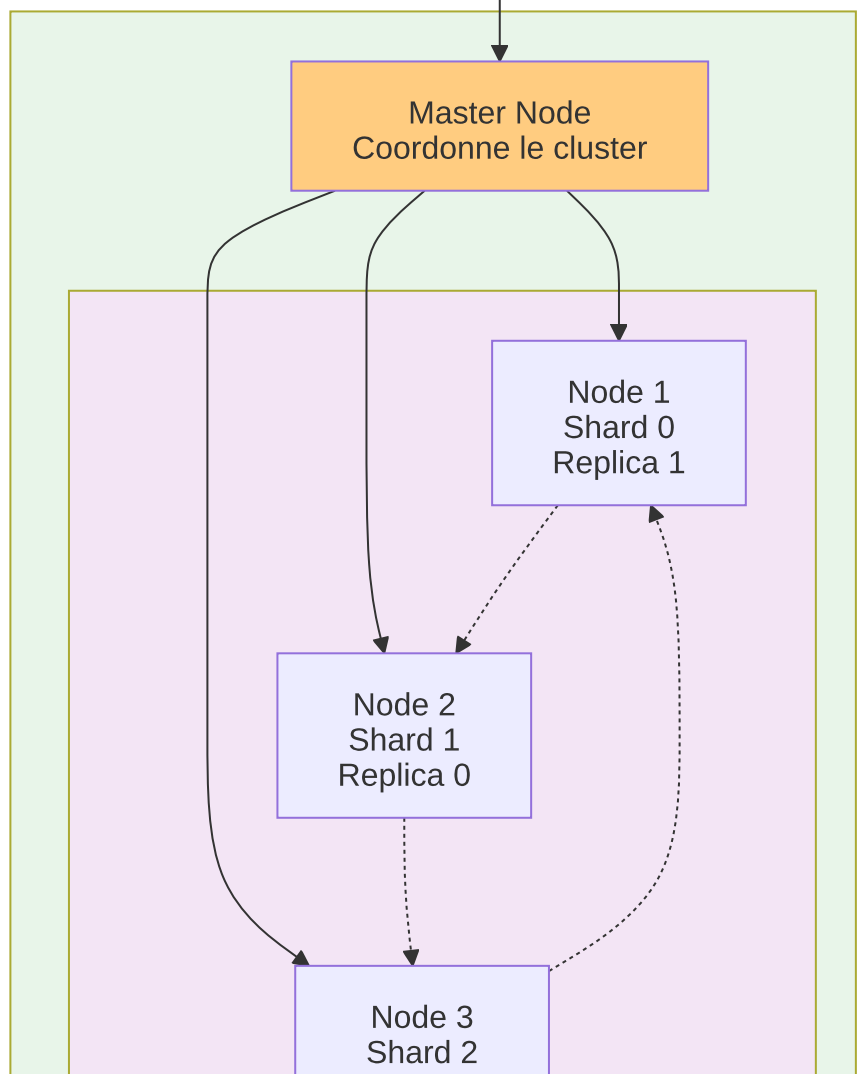
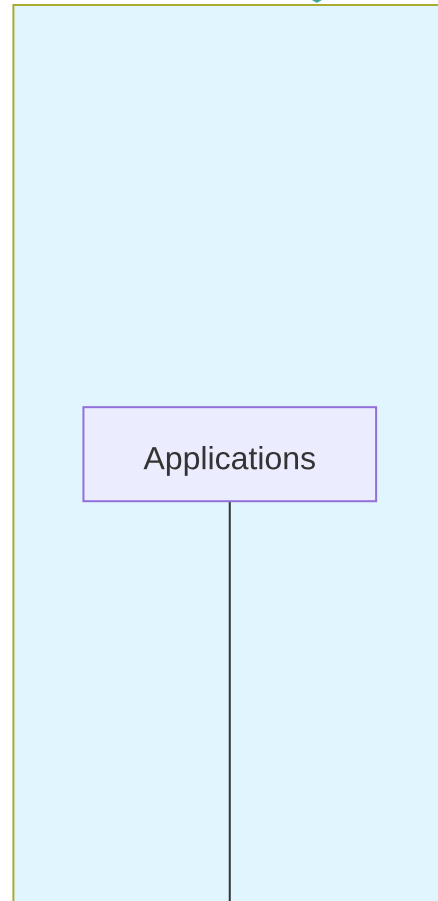
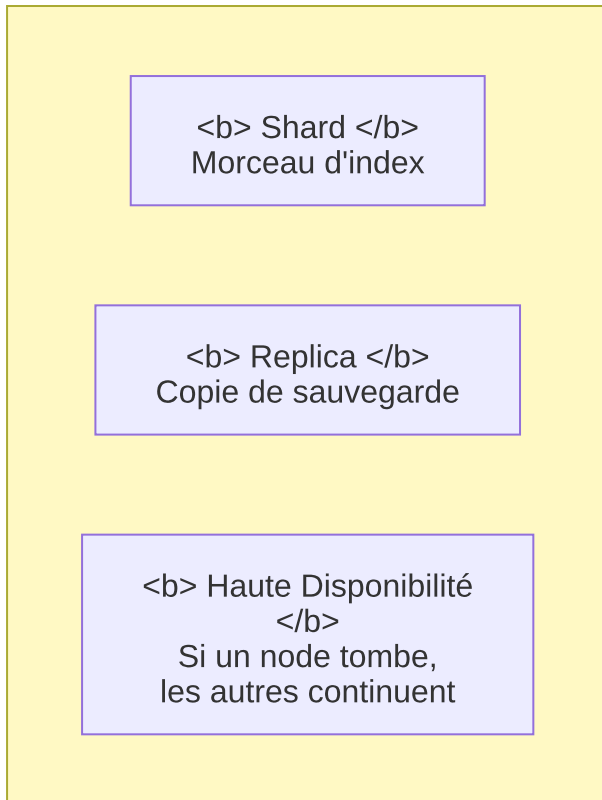
Elasticsearch utilise une architecture distribuée. Cela signifie que vos données peuvent être réparties sur plusieurs serveurs (appelés **nœuds**), formant un **cluster**.

Les Avantages de cette Architecture

Scalabilité horizontale : Vous pouvez simplement ajouter de nouveaux serveurs pour augmenter la capacité de stockage et de traitement, sans interruption de service.

Haute disponibilité : Si un serveur tombe en panne, les autres nœuds du cluster continuent de fonctionner. Vos données sont répliquées pour garantir qu'aucune perte ne survienne.

Performance : Les requêtes sont distribuées sur plusieurs nœuds, ce qui permet de traiter de gros volumes de données très rapidement.



Benchmarks

Matériel :

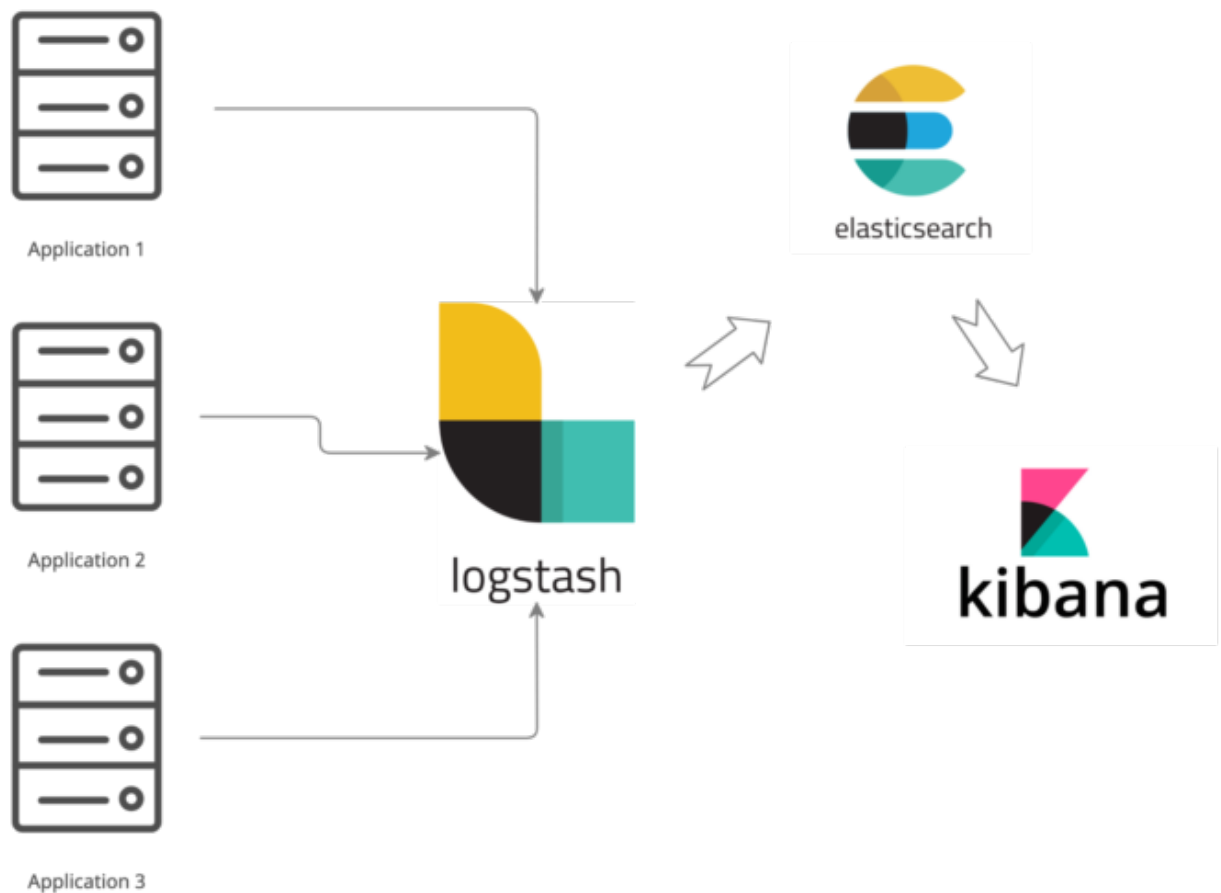
- nœuds de données : **4 vCPU + 16 GiB de RAM** chacun [alibabacloud.com+1](https://www.alibabacloud.com/help/en/es/product-overview/performance/?spm=a2c63.p38356.help-menu-57736.d_0_0_9.14f510cciRVXHR)
- nombre de nœuds : 3 [alibabacloud.com](https://www.alibabacloud.com/help/en/es/product-overview/performance/?spm=a2c63.p38356.help-menu-57736.d_0_0_9.14f510cciRVXHR)
- stockage : ESSD, PL1, 200 GiB par nœud

Un cluster 8.9 ingère ~ 219 531 documents par seconde (moyenne). (Logs de type HTTP)

https://www.alibabacloud.com/help/en/es/product-overview/performance/?spm=a2c63.p38356.help-menu-57736.d_0_0_9.14f510cciRVXHR

L'Écosystème Elastic Stack

Elasticsearch ne travaille pas seul. Il fait partie d'une suite d'outils appelée **Elastic Stack** (anciennement ELK Stack) :



- **Elasticsearch** : Le moteur de recherche et d'analyse
- **Logstash** : Un outil de collecte et de transformation de données
- **Kibana** : Une interface de visualisation pour explorer et visualiser les données
- **Beats** : Des collecteurs de données légers pour envoyer des informations vers Elasticsearch

Ensemble, ces outils forment une solution complète pour collecter, stocker, rechercher, analyser et visualiser vos données.

Pourquoi Apprendre Elasticsearch ?

Pour les Développeurs

- Amélioration des performances de recherche dans vos applications

Pour les Data Analysts

- Capacité à analyser de gros volumes de données rapidement
- Création de tableaux de bord interactifs avec Kibana
- Détection de tendances et d'anomalies en temps réel

Pour les Ops et DevOps

- Surveillance et monitoring d'infrastructures
 - Centralisation et analyse des logs
 - Détection proactive des problèmes
-

Ressources Complémentaires

- [Documentation officielle Elasticsearch](#)
- [Elastic Stack \(ELK\)](#)

