**Sujet 1 :**  Mettre en œuvre une base de données dont l'objectif est d'héberger les journaux (logs) d'équipements, logiciels , systèmes d'exploitation, équipements dans le but de permettre l'analyse à froid et à chaud . Faire une démonstration sur un exemple concret (Cybersécurité : tentative d' intrusion ou dysfonctionnement pour analyse et correction de bugs , …)

Recherche de la techno à préconiser :

* MongoDB
* ElasticSearch
* Cassandra
* Mysql
* Neo4j

Pour être performant dans le stockage de logs, une base de données doit pouvoir couvrir ces différents points :

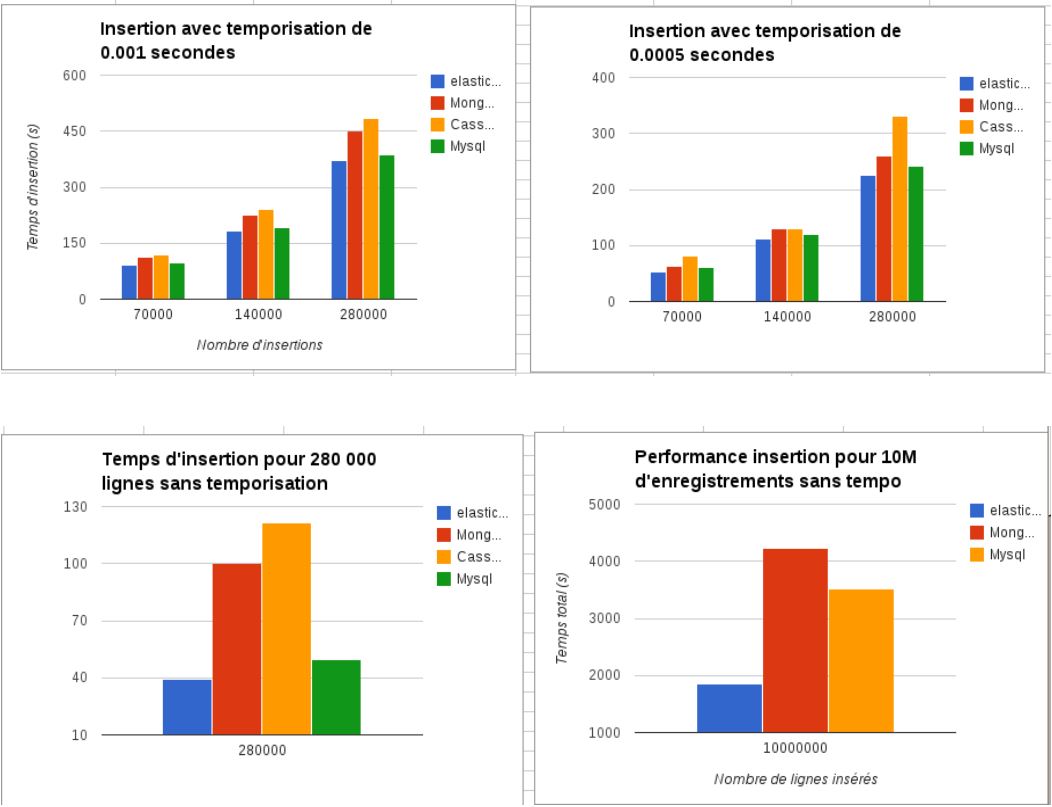
* Collection
* Transport
* Traitement
* Stockage
* Recherche
* Visualisation

MongoDB semble être la solution de base et répondre plutôt bien aux attentes, surtout en termes de stockage.

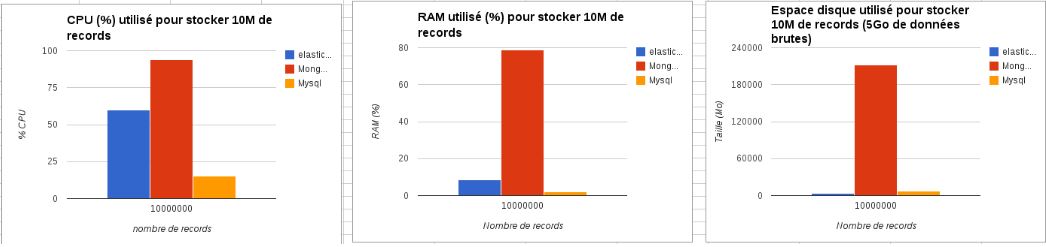
Un combo sort gagnant des recherches, cela implique **Logstash-Forwarder** pour la collecte et le transport, **Logstash et Riemann** pour le process, et **Kibana3** pour la visualisation.

Workbench: ElasticSearch – MongoDB – Cassandra – Mysql

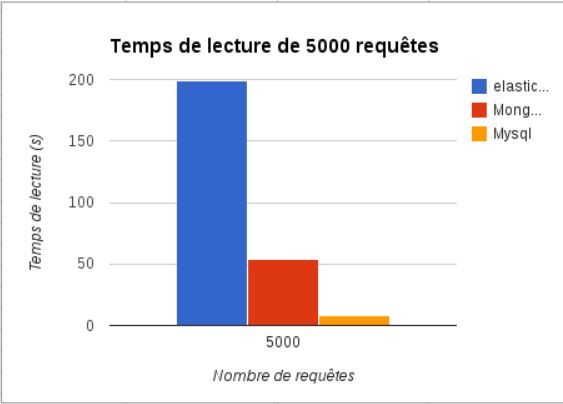
*Insertion*



*Consommation*



*Lecture*



Les tests ont été réalisé en essayant d’être cohérent pour un comparatif pertinent même si d’autres workbench dans d’autres circonstances et avec d’autres paramètres pourrait être très différents.

Au niveau des résultats, **Cassandra** sort gagnant avec des très bons temps de lecture et d’écriture tout en ayant une faible consommation. Malgré cela l’équipe qui a réalisé ces tests ont finalement décidé d’utiliser **ElasticSearch** pour leur projet, Les atouts sont la facilité de mise en place et les avantages liés au services annexes comme **Kibana**.

ELK Stack, MongoDB et Cassandra sont tous des outils populaires qui peuvent être utilisés pour la gestion de journaux. Voici une comparaison de la façon dont ces outils diffèrent:

* ELK Stack: ELK Stack est une plateforme d'analyse de journaux qui se compose d'Elasticsearch, de Logstash et de Kibana. Elasticsearch est utilisé pour stocker et indexer les journaux, Logstash est utilisé pour collecter et traiter les journaux, et Kibana est utilisé pour visualiser et analyser les journaux. ELK Stack est un outil puissant et flexible pour l'analyse de journaux et est largement utilisé dans une variété d'environnements.
* MongoDB: MongoDB est une base de données orientée documents conçue pour stocker et interroger de grands volumes de données. Il peut être utilisé pour stocker des journaux, mais il n'est pas spécifiquement conçu pour l'analyse de journaux. En revanche, il est plus souvent utilisé pour stocker et interroger des données pour des applications et d'autres fins.
* Cassandra: Cassandra est une base de données distribuée conçue pour stocker et interroger de grandes quantités de données sur de nombreux serveurs. Il peut être utilisé pour stocker des journaux, mais comme MongoDB, il n'est pas spécifiquement conçu pour l'analyse de journaux. Il est plus souvent utilisé pour stocker et interroger des données pour des applications et d'autres fins.

En général, ELK Stack est la meilleure option pour l'analyse de journaux, car il est spécifiquement conçu à cet effet. MongoDB et Cassandra sont des bases de données plus généralistes qui peuvent être utilisées pour stocker des journaux, mais elles ne proposent pas le même niveau de fonctionnalités d'analyse de journaux que ELK Stack.

**Cas d’usage :**

Dans notre cas d’usage les logs seront probablement stockés en termes de jours, voir mois, avec une nécessitée de pouvoir facilement et rapidement accéder aux données pour une analyse en temps réel.

Nous sommes responsables des postes dans une entreprise et nous avons besoin de pouvoir récupérer les informations importantes des machines pour prévenir les problèmes techniques, les défaillances et les failles potentielles de sécurité.

Dans ce contexte nous avons mit en place un système basé sur ELK (ElasticSearch, Kibana et Logstash) pour pouvoir récupérer les logs Windows des postes de travail de manière automatique et stockés dans une base.

Une fois dans la base, on peut distinguer deux types d’analyse :

Une analyse à chaud, c’est-à-dire des processus qui vont analyser chaque nouveau log pour y détecter des warnings, des erreurs ou d’autres choses qui peuvent nous servir.

Une deuxième analyse à froid, c’est-à-dire une analyse d’ensemble pour faire ressortir des informations pertinentes, des rapports détaillés de l’état des machines.

Tutoriel pour Elastic Cloud :

Elastic Cloud est une plate-forme basée sur le cloud pour le stockage et l'analyse de journaux et d'autres types de données. Voici un aperçu général de la configuration et de l'utilisation d'Elastic Cloud pour l'analyse des journaux :

1. Ouvrez un compte Elastic Cloud : accédez au site Web d'Elastic Cloud et créez un compte. Vous devrez fournir votre adresse e-mail et créer un mot de passe.
2. Créer un déploiement : un déploiement dans Elastic Cloud est un regroupement logique d'instances Elasticsearch et Kibana que vous pouvez utiliser pour stocker et analyser vos données. Pour créer un déploiement, accédez au tableau de bord Elastic Cloud et cliquez sur le bouton « Créer un déploiement ». Choisissez le nombre d'instances Elasticsearch et la quantité de stockage et de mémoire que vous souhaitez allouer à votre déploiement, puis cliquez sur le bouton "Créer un déploiement".
3. Configurer la collecte de journaux : pour collecter des journaux avec Elastic Cloud, vous pouvez utiliser Logstash ou Filebeat pour envoyer les journaux à votre déploiement. Pour le configurer, vous devrez installer et configurer Logstash ou Filebeat sur les serveurs ou appareils qui génèrent les journaux. Ensuite, vous devrez créer un fichier de configuration qui spécifie comment Logstash ou Filebeat doit envoyer les journaux à Elastic Cloud.
4. Visualisez et analysez les journaux dans Kibana : Une fois vos journaux collectés et stockés dans Elastic Cloud, vous pouvez utiliser Kibana pour les visualiser et les analyser. Kibana propose une gamme d'options de visualisation, notamment des graphiques linéaires, des graphiques à barres, des graphiques circulaires, etc. Vous pouvez utiliser ces visualisations pour comprendre vos journaux plus en détail et identifier des tendances ou des modèles.

**Progression :**

**ElasticSearch**, **kibana**, **logstash** et **filebeat** installés sur la machine distante, ports 9200 et 5601 ouverts, fichiers de config modifiés pour accepter les connexions remote via *10.8.2.35:5601*.

# Bibliographie

* Comparatif database: <https://eventuallycoding.com/2013/11/05/comparaison-de-moteurs-de-base-de-donnees-pour-du-stockage-de-logs/>
* Comparatif database: <https://dzone.com/articles/5-data-storages-for-better-log-management>
* Log monitoring: <https://www.elastic.co/fr/observability/log-monitoring>
* Tuto complet: <https://medium.com/@vitalypanukhin/elasticsearch-elk-stack-for-log-management-8f4e61a60239>
* Install ELK Ubuntu: <https://infosecwriteups.com/how-to-install-elastic-stack-on-ubuntu-22-04-lts-a2f1b00eced>